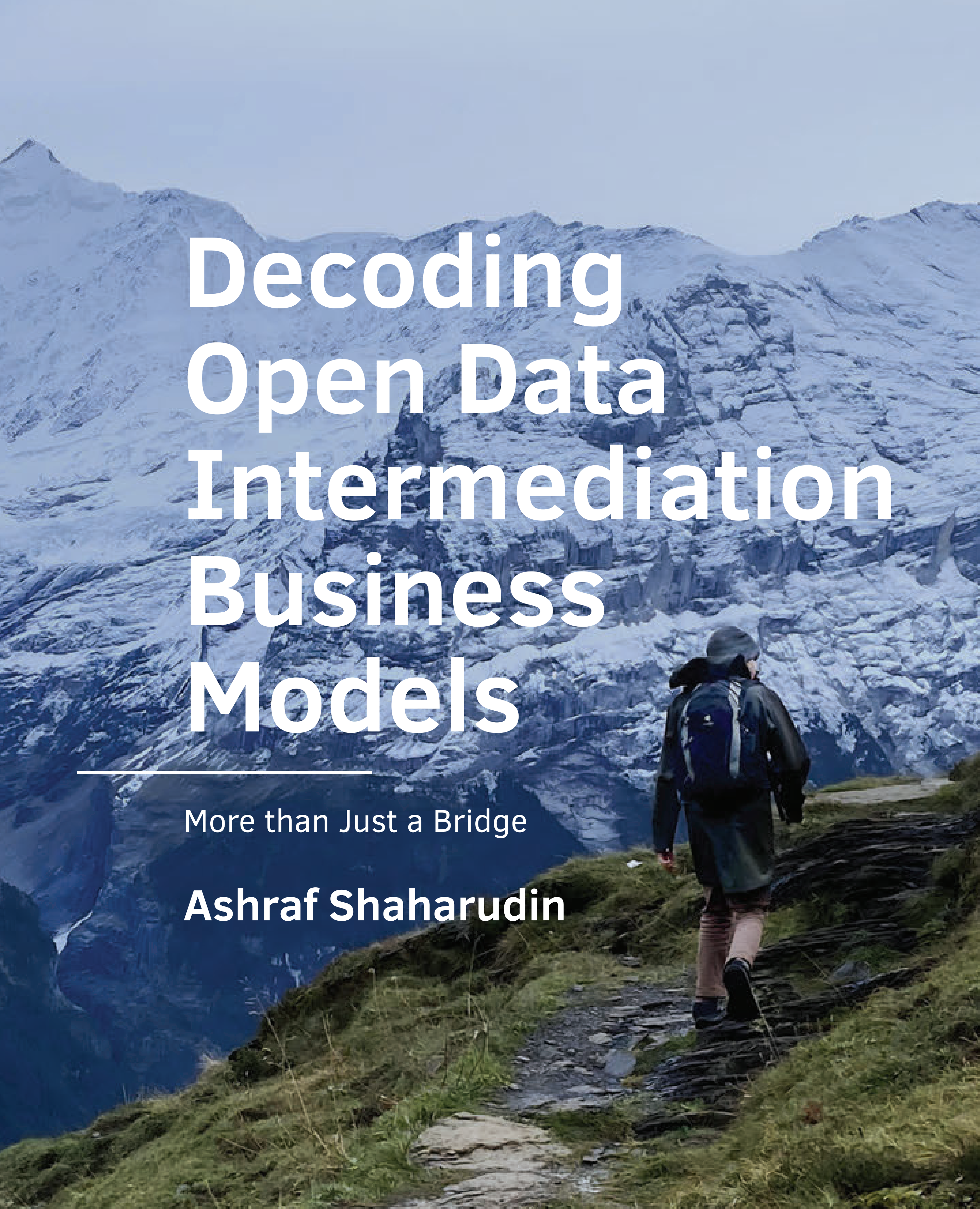


Decoding Open Data Intermediation Business Models

More than Just a Bridge

Ashraf Shaharudin

Decoding Open Data Intermediation Business Models

More than Just a Bridge

Ashraf Shaharudin

TOC

A+BE | Architecture and the Built Environment | TU Delft BK

25#22

Design | Sirene Ontwerpers, Véro Crickx

ISBN 978-94-6518-186-8
ISSN 2212-3202

This dissertation is open access at https://doi.org/10.71690/abe.2025.22

Decoding Open Data Intermediation Business Models

More than Just a Bridge

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, prof.dr.ir. T.H.J.J. van der Hagen
chair of the Board for Doctorates
to be defended publicly on
Thursday 4 December 2025 at 15:00 o'clock

by

Ahmad Ashraf bin AHMAD SHAHARUDIN
Master of Economics, Universiti Putra Malaysia, Malaysia
born in Johor Bahru, Malaysia

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof.dr.ir. M.F.W.H.A. Janssen	Delft University of Technology, promotor
Dr.ir. B. van Loenen	Delft University of Technology, promotor

Independent members:

Prof.dr.ir. P.J.M. van Oosterom	Delft University of Technology
Prof.dr. S.M. Coetzee	United Nations University, Dresden, Germany
Dr. P.A. Johnson	University of Waterloo, Canada
Dr. P. Mooney	Maynooth University, Ireland
Dr. A.M.G. Zuiderwijk	Delft University of Technology
Prof.dr.ing. S. Nijhuis	Delft University of Technology, reserve member

The research for this dissertation was funded through the ‘Towards a Sustainable Open Data ECOsystem’ (ODECO) project. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955569. The opinions expressed in this document reflect only the author’s view and in no way reflect the European Commission’s opinions. The European Commission is not responsible for any use that may be made of the information it contains.

Acknowledgements

Ever since I first watched Lupita Nyong'o deliver her heartfelt Oscars acceptance speech in 2014, after winning the Best Supporting Actress for her performance in '12 Years a Slave', I have always imagined quoting her when it was my turn to give an important acknowledgement: 'It doesn't escape me for one moment that so much joy in my life is thanks to so much pain in someone else's'.

My parents, *Ibu Ayah*, who endured so much hardship in life, only to ensure their children have a good life. Thank you for always supporting and encouraging me to chase my dreams. For putting my interests above yours. Only God can truly reward you. My little brothers, I admire the way you navigate life with all its challenges, and I am proud of you. Thank you for cheering on your big brother as he explores the world far from home. To my late *Nenek Johor*, who gave me some money before I flew to the Netherlands even though you probably needed it more than I did, I am sad that you did not get to see me complete my PhD, but your love motivated me.

My teachers, whose every small act has brought me to where I am today. Countless of you to thank. But some especially deserve special mention. Thank you to Cikgu Rohayu and Cikgu Amal, whose encouragement for me to join a speech (*pidato*) contest when I was 13 profoundly shaped me and boosted my confidence. I then went on to become a school debater, and that was what I was known for all my school life. As a naturally introverted person, I thus learned that I can fake it till I make it—a skill that has proved necessary in academia, especially when I need to share my work. To my late Teacher Norsitah, thank you for building my confidence in speaking English and for everything else. I always remember your first words when you saw me visiting you in your final days: 'My favourite'. You are my favourite.

My friends—in Australia, Malaysia, and the Netherlands—with whom I share my happy and sad moments. Thank you for putting up with me. For always reminding me that I am not alone, even when some days can be lonely. For practical help when I need it. And for seeing me for who I am and not what I have.

My mentors, whose generosity and sincerity have helped me blossom not only as a researcher but also as a person. How lucky I am. Dr Rachel Gong, who meticulously reviewed and edited my work at KRI, and helped me learn to write better. Thank you for your (life) advice and encouragement, even when we no longer work together.

And most certainly, thank you for writing such an incredibly thoughtful recommendation letter, which I believe had a significant part in helping me land my PhD position. Dr Bastiaan van Loenen and Prof Marijn Janssen, I cannot ask for a better team of PhD supervisors. Thank you for believing in me and giving me the space to follow my curiosity. Thank you for your time and energy in reviewing my work and asking critical questions. I have always had a sense of direction during my PhD—thanks to you both. To Bastiaan especially, I truly enjoyed our countless conversations about academia, life, and the world. Thank you also for always having my back.

My officemate, Dr Stefano Calzati; my other KCOD research group colleagues, Dr Frederika Welle Donker and Dr Hendrik Ploeger; and my ODECO project colleagues—through all the meetings and chats we had over lunches, coffees, and dinners—you have made my PhD journey so much more fulfilling and enriching. And fun!

While learning the ropes in the first year of my PhD, I watched Julie Goodwin—the first season's winner of MasterChef Australia—return to compete in the show's 14th season (2022). In the first episode, she shared an inspiring message: 'It's not about what you know now, but what you're willing to take on board from all the judges (read: supervisors) and fellow contestants (read: other colleagues). It's about what you are willing to try and how you are willing to grow'. That was exactly what I did—or at least tried my best to do. Ultimately, in the battle of completing a PhD, where the only person you compete with is yourself, I believe I won.

Not to forget all the interviewees who generously shared their time, experiences, and insights with me, and all the reviewers whose constructive feedback has greatly improved my work. My gratitude also to my secondment hosts: Jan Willem van Eck at Esri Nederland, and Prof Joep Crompvoets at KU Leuven.

To all of you, thank you so much. *Heel erg bedankt*. In Malay, we say *terima kasih*, which is translated word for word as 'receive love' (*terima* = receive, *kasih* = love). So, from the bottom of my heart, please receive my love.

While writing this acknowledgement, it never slipped my mind that I was born with immense privilege. I am fully aware that so many people around the world have never had even a fraction of the opportunities I have been given. Many are even deprived of basic necessities to live, like food, shelter, and safety—especially today in places like Palestine, Sudan, Congo, Myanmar, and Ukraine, to name just a few. There is so much human suffering in this world, inflicted by the hands of fellow humans, sustained by the apathy of others. With such privilege that I have, comes responsibility: a duty to humanity and a commitment to contribute to a better and more just world for everyone, however limited my power may be. May I always honour this.

Time is a funny thing. Ten years ago, I received my Bachelor's degree in Australia. As I pondered my next steps in life then, I encountered a powerful poem, 'The Will to Live' by celebrated Tunisian poet Abu al-Qasim al-Shabbi (1909 – 1934). I vividly remember translating it into Malay and posting it on my now-defunct blog. Fast forward, in May 2025, I visited Wereldmuseum Amsterdam when the exhibition 'Poetry of the People' was running, and I came across the poem again. What a beautiful coincidence! As I once more found myself reflecting on my next steps in life, I include a snippet of the poem below, as a reminder to myself and anyone reading this that life is what we make of it and that—again quoting Lupita Nyong'o—'no matter where you're from, your dreams are valid'.

When people one day desire life,
Then destiny must respond,
Night must be dispelled,
And chains must be broken.
If a yearning for life does not embrace you,
You will fade away and perish.
Beware O you who are not passionate!
Beware of the strike of the triumphant void!

[Arabic to English translation of the poem by Mohammad Fakhreddine.]

Ashraf Shaharudin
August 2025

Contents

List of Tables

List of Figures

List of Acronyms

AHN	*Actueel Hoogtebestand Nederland* (Current Dutch Elevation)
ANT	actor-network theory
API	Application Programming Interface
AWS	Amazon Web Services
BAG	*Basisregistratie Adressen en Gebouwen* (Dutch basic registration of addresses and buildings)
BKG	*Bundesamt für Kartographie und Geodäsie* (German Federal Agency for Cartography and Geodesy)
CSO	civil society organisation
DWG	OSMF Data Working Group
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable
HIF	Humanitarian Innovation Fund
HOT	Humanitarian OpenStreetMap Team
IFRC	International Federation of the Red Cross/Red Crescent (IFRC),
IGN (Belgium)	*L'Institut géographique national* (National Geographic Institute); same as NGI (Belgium)
IGN (France)	*Institut national de l'information géographique et forestière* (National Institute of Geographic and Forest Information)
IGN (Spain)	*Instituto Geográfico Nacional* (National Geographic Institute)
INSPIRE	Infrastructure for Spatial Information in the European Community Directive
IOER	*Leibniz-Institut für ökologische Raumentwicklung* (Leibniz Institute of Ecological Urban and Regional Development)
ISTAT	*Istituto Nazionale di Statistica* (Italian National Institute of Statistics)
IT	information technology
Kadaster	The Netherlands' Cadastre, Land Registry and Mapping Agency
LCCWG	Local Chapters and Communities Working Group
ML	machine learning
MSF	*Médecins Sans Frontières* (Doctors Without Borders)
NGI (Belgium)	*Nationaal Geografisch Instituut* (National Geographic Institute); same as IGN (Belgium)
NGO	non-governmental organisation
NPO	non-profit organisation
OAM	OpenAerialMap
ODbL	Open Database Licence
ODE	open data ecosystem
OEG	OSM Organised Editing Guidelines

>>>

TOC

OGC	Open Geospatial Consortium
OKF	Open Knowledge Foundation
OSGeo	Open Source Geospatial Foundation
OSM	OpenStreetMap
OSMF	OpenStreetMap Foundation
OWG	OSMF Operations Working Group
POI	Point of interest
RBV	resource-based view
RQ	research question
SDFI	*Styrelsen for Dataforsyning og Infrastruktur* (Danish Agency for Data Supply and Infrastructure)
SLR	systematic literature review
STS	science and technology studies
TIGER	Topologically Integrated Geographic Encoding and Referencing
TU Delft	Delft University of Technology
UAE	United Arab Emirates
UCL	University College London
UK	United Kingdom
UN	United Nations
UNDP	United Nations Development Programme
US	United States of America
USAID	United States Agency for International Development
VRIN	valuable, rare, imperfectly imitable, and non-substitutable
WCK	World Central Kitchen
WFP	UN World Food Programme

Summary

Open data intermediaries are crucial for the sustainability (i.e., long-term durability) of the open data ecosystem (ODE). They enhance the access to and the (re-) use of open data and connect other open data actors. Additionally, open data intermediaries play a role in mitigating information asymmetry between actors. However, despite the importance of open data intermediaries in the ODE having been widely acknowledged in research and practice, studies on open data intermediation business models are limited. This knowledge is essential to better understand the role of open data intermediaries within the ODE and provide recommendations to develop their business models in such a way that they support the overall sustainability of the ODE. Thus, the objective of this dissertation is to understand how open data intermediation business models can support a sustainable ODE. The objective was addressed through four research questions (RQs).

The first research question (RQ1) asked: What are open data intermediaries? This research question tackled the lack of consensus on what constitutes open data intermediaries. Past literature offers various, and sometimes, conflicting definitions of open data intermediaries. Some also use the term open data intermediaries without adequately defining them. Thus, to clarify the obscurity of open data intermediaries, a systematic literature review was conducted to offer a common definition. This was done by compiling and analysing the existing definitions of open data intermediaries in the literature, as well as the activities, types of actors, and objectives of open data intermediaries. Eventually, this dissertation proposes a definition of open data intermediaries as ‘third-party actors who provide specialised resources and capabilities to (i) enhance the supply, flow, and/or use of open data and/or (ii) strengthen the relationships among various open data actors’. This definition is used throughout this dissertation.

The second research question (RQ2) asked: What are potential contributions of open data intermediaries in addressing challenges in an open data ecosystem? To better situate open data intermediaries within the ODE, it is necessary to look into what they (can) contribute to other actors. This can help refine the understanding of open data intermediaries’ characteristics and their relationships with other ODE actors. Most studies have focused solely on examining the current activities of open data intermediaries. While this baseline understanding is valuable, by also exploring what they can potentially contribute, RQ2 also aims to identify gaps in the ODE that

TOC

open data intermediaries may be able to close or narrow. Through semi-structured interviews and a validation exercise with various stakeholders, this dissertation mapped the potential contributions of open data intermediaries to various challenges in the ODE. Several additional insights were also drawn. Notably, findings from this research question break the misconception that open data intermediaries are merely a 'bridge' between open data providers and end-users. Instead, an open data provider or user could benefit from the contributions of multiple open data intermediaries simultaneously, in parallel and/or sequentially. This also implies that the beneficiary of an open data intermediary's contributions could be another open data intermediary.

The third research question (RQ3) asked: What are archetypes of open data intermediation business models? Open data intermediaries exist in various shapes and forms. Hence, RQ3 identified the common archetypes of open data intermediation business models. Such knowledge is necessary before any business model recommendations can be prescribed, as different business model archetypes may need to consider different aspects. Past literature was limited in the number of open data intermediaries surveyed and/or in the comprehensiveness of the business model described. Using the machine learning technique of K-means clustering and based on a survey of 190 open data intermediaries, this dissertation identifies nine business model archetypes. Each archetype is described based on its value proposition (i.e., what the organisation offers), value creation (i.e., what resources and activities are deployed), and value capture (i.e., how the organisation is compensated) dimensions. The archetypes identified are:

1. collaborative open data platform,
2. paid self-service data delivery,
3. personalised open data service,
4. interactive app with other complementary products,
5. open data repository funded by sponsorship,
6. one-stop package around an (augmented) open data platform/repository,
7. single-purpose app,
8. interactive app without complementary products, and
9. open data advocacy.

The fourth and overarching research question of this dissertation (RQ4) asked: What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem? RQ4 was addressed through two single-case studies representing two different business model archetypes, namely Esri, representing the one-stop package around an (augmented) open data platform/repository archetype, and OpenStreetMap (OSM), representing the collaborative open data platform archetype. Six considerations that apply to both archetypes were identified:

1. Ensure consistency in how members of the organisation or community view the (potentially multifaceted) organisational identity, especially as the community or organisation expands or becomes more decentralised.
2. Consider offering new or non-traditional solutions instead of being entrenched in traditional paradigms.
3. Invest in the adoption of new and emerging technology for the software infrastructure (e.g., the application of AI for metadata recommendations).
4. Offer products or services that simplify, as much as possible, the process of supplying and using open data.
5. Stimulate potential multistakeholder collaborations (e.g., via projects or events).
6. Offer consultancy or training services, including through the formal education sector (e.g., schools or universities).

Eleven considerations that apply specifically to the one-stop package archetype were identified:

1. Ensure that open data intermediation services offered are fittingly integrated with the existing core products or services.
2. Offer open data intermediation services that are consistent with the organisational identity (i.e., does not involve a significant shift from the core business).
3. Offer diverse complementary products or services that leverage open data.
4. Offer services that minimise open data-associated risks that customers have to deal with.
5. Offer customised data, services, or projects catering to local needs.
6. Offer open-source software, at least partially as part of the larger product suite (e.g., through a freemium model).
7. Facilitate feedback on open data through a structured mechanism.
8. Ensure that the unique resources or position are not leveraged in ways that unfairly stifle the growth of other actors (the winner-takes-it-all situation), e.g., by committing to the development of broadly adoptable open standards and technical interoperability.
9. Showcase the value of open data.
10. Advocate for the release of open data from non-public sectors.
11. Invest in open data-based collaborations.

Fourteen considerations that apply specifically to the collaborative open data platform were identified:

1. Foster a healthy and constructive community of contributors.
2. Protect the overall interests of the contributors by implementing transparent (decision-making) processes and putting in place a mechanism to prevent potential takeover or hijacking.
3. Invest in enhancing the visibility and reach of the organisation or community.
4. Ensure a clear, efficient, civility-focused communication mechanism.
5. Facilitate and encourage the development of (especially open-source) tools around the platform (i.e., federated architecture).
6. Leverage a self-correcting mechanism by cultivating a 'team-minded' culture and having a clear process for dispute resolution.
7. Support local chapters or communities in engaging with local organisations (e.g., local governments or NGOs).
8. Transform disagreements and conflicts into opportunities for enhancing governance mechanisms and technical development.
9. Allow diverse types of data to be contributed to accommodate broad and diverse use cases.
10. Take initiatives to include the perspectives of indirect impact recipients (i.e., the non-users), e.g., by working with on-the-ground NGOs and prioritising their needs.
11. Choose a reciprocal licence that does not impose (many) restrictions on the end-use (i.e., derivative) work.
12. Proactively recruit new data contributors among typically marginalised or disadvantaged groups and nurture a welcoming environment to retain their participation.
13. Meaningfully consider feedback related to the technological development from non-technical expert members.
14. Cultivate a culture of deliberating ethical responsibilities in the contribution and use of data, beyond providing general ethics guidelines.

These considerations can be undertaken by existing or potential open data intermediaries in developing or innovating their business model. Notably, the aspects recommended highlight the importance of the social dimension (e.g., organisational identity, engagement with the users, and community management) and the legal dimension (e.g., the choice of licence), apart from the technological dimension, in the development of open data intermediation business models.

Furthermore, this dissertation also offers several theoretical reflections on the (sustainable) ODE concept. Among others, this dissertation reaffirms the need to distinguish a *role* from an *actor* in the ODE. Various types of organisations or entities can be open data intermediaries and, in particular, they do not only exist outside of the public sector as some have implied. This dissertation, through the case of OSM, also highlights that an ODE where open data from the non-government sector is circulating widely (i.e., not only open government data) already exists. However, there is a risk that open government data and open non-government data are growing in silos, due to limited legal and/or technical interoperability and the diverging priorities of different actors (e.g., government organisations prioritising government data infrastructure and neglecting the opportunities to collaboratively contribute to infrastructure initiated outside the public sector). Additionally, this dissertation emphasises the importance of considering the entire technology stack to the sustainability of the ODE. It is not only the characteristics of the open data (coverage, quality, format, license, etc.) that are important to the sustainability of the ODE, but also how it is disseminated and re-used, including the software used.

Looking forward to future research, while the focus of this dissertation has been on the business models, policy aspects that could steer open data intermediaries to act in ways that support the sustainability of the ODE should also be studied, such as investigating the viability and benefits of extending the open data legislation to non-government open data intermediaries. Additionally, open data intermediaries' hidden and potentially undesirable impacts on other ODE actors deserve attention, taking into account that these impacts may evolve over the course of their business model development (i.e., nascent versus mature and potentially dominant stage).

In conclusion, this dissertation advances the understanding of open data intermediaries, their position within the ODE, and their business models. In particular, this dissertation theoretically contributed to the definition of open data intermediaries, potential contributions of open data intermediaries, and the archetypes of open data intermediation business models. It also identified practical aspects to consider in developing open data intermediation business models that contribute to a sustainable ODE, ultimately enhancing the generation of open data value. This value, in turn, can be leveraged to foster innovation, promote economic well-being, and address pressing social and environmental challenges.

Samenvatting

Open data-intermediairs zijn cruciaal voor de toekomstbestendigheid van een Open Data Ecosysteem (ODE). Zij kunnen de toegang tot en het gebruik van open data verbeteren en andere open data-actoren met elkaar verbinden. Daarnaast kunnen open data-intermediairs een rol spelen bij het verminderen van een mogelijke informatie-asymmetrie tussen actoren. Ondanks het belang van open data-intermediairs in de ODE is er nauwelijks onderzoek gedaan naar bedrijfsmodellen voor open data intermediatie. Deze kennis is echter essentieel om de rol van open data intermediairs binnen ODE beter te begrijpen en aanbevelingen te kunnen doen om hun bedrijfsmodellen zodanig te optimaliseren dat ze ook de toekomstbestendigheid van de ODE als geheel vergroten. Het doel van dit proefschrift is om te begrijpen hoe bedrijfsmodellen voor open data intermediatie de toekomstbestendigheid van ODE kunnen bevorderen. Dit is door middel van vier onderzoeksvragen (RQ's) onderzocht.

De eerste onderzoeksvraag (RQ1) luidde: Wat zijn open data intermediairs? De literatuur is niet eenduidig en soms tegenstrijdig in de definitie van open data intermediairs. In sommige publicaties wordt de term open data intermediairs zelfs niet gedefinieerd. Een systematische literatuurstudie is uitgevoerd om tot een eenduidige definitie te komen. Hierbij zijn ook de activiteiten, soorten actoren en doelstellingen van open data intermediairs meegenomen. Dit resulteerde in de volgende definitie van open data-intermediairs: 'derde partijen die gespecialiseerde middelen en capaciteiten leveren om (i) het aanbod, de stroom en/of het gebruik van open data te verbeteren en/of (ii) de relaties tussen verschillende belanghebbenden bij open data te versterken'. Deze definitie wordt in dit proefschrift aangehouden.

De tweede onderzoeksvraag (RQ2) luidde: Wat zijn de potentiële bijdragen van open data-intermediairs aan een open data-ecosysteem? Hiervoor is onderzocht wat zij (kunnen) doen om andere actoren te faciliteren in hun open data activiteiten. Dit geeft inzicht in de kenmerken van open data-intermediairs en hun relaties met andere ODE-actoren. De meeste onderzoeken hebben zich uitsluitend gericht op de huidige activiteiten van open data-intermediairs. Het onderliggende onderzoek kijkt, via RQ2, naast de huidige activiteiten ook naar wat open data-intermediairs potentieel kunnen bijdragen, en onderzoekt eventuele hiaten in de ODE die open data-intermediairs mogelijk kunnen dichten of verkleinen. Door middel van semigestructureerde interviews met diverse belanghebbenden, heeft dit proefschrift de potentiële bijdragen van open data-intermediairs aan een ODE in kaart gebracht. Uit het onderzoek wordt duidelijk dat open data-intermediairs niet alleen een ‘brug’ vormen tussen aanbieders van open data en hun eindgebruikers. Een aanbieder of gebruiker van open data kan namelijk gelijktijdig, parallel en/of sequentieel profiteren van de bijdragen van meerdere open data-intermediairs. De begunstigde van de activiteiten van een open data intermediair kan dus evengoed een andere open data intermediair zijn.

De derde onderzoeksvraag (RQ3) luidde: Wat zijn archetypen van bedrijfsmodellen voor open data intermediatie? Open data-intermediairs kunnen vele verschillende bedrijfsmodellen hebben. Daarom richtte RQ3 zich op het onderzoeken van soortgelijke archetypen van bedrijfsmodellen voor open data-intermediatie. Op basis hiervan kunnen er aanbevelingen voor bedrijfsmodellen voor verschillende archetypen worden gedaan. De bestaande literatuur heeft gekeken naar een zeer beperkt aantal open data-intermediairs en/of richtte zich slechts op een aantal onderdelen van een bedrijfsmodel. In dit onderzoek zijn de bedrijfsmodellen van 190 open data-intermediairs met behulp van de K-means-clustering-machine-learning-techniek geclusterd en geanalyseerd. Dit resulteerde in negen archetypen van bedrijfsmodellen. Elk archetype wordt beschreven op basis van de dimensies van de waarde propositie (d.w.z. wat de organisatie biedt), waarde creatie (d.w.z. welke middelen en activiteiten worden ingezet) en waarde vastlegging (d.w.z. hoe de organisatie wordt gecompenseerd). De geïdentificeerde archetypen zijn:

1. een collaboratief open dataplatform,
2. betaalde selfservice-datalevering,
3. gepersonaliseerde open dataservice,
4. een interactieve app met andere aanvullende producten,
5. een open data repository die wordt gefinancierd door sponsoring,
6. een totaalpakket rond een (uitgebreid) open data platform/repository,
7. een app voor één doel,
8. een interactieve app zonder aanvullende producten, en
9. open data belangenbehartiging.

De vierde onderzoeksvraag van dit proefschrift (RQ4) luidde: Welke aspecten moeten in overweging worden genomen bij de ontwikkeling van bedrijfsmodellen voor open data-intermediatie die een duurzaam open data-ecosysteem ondersteunen? RQ4 is beantwoord aan de hand van twee casestudies die twee verschillende archetypen van bedrijfsmodellen vertegenwoordigen: Esri, dat het archetype van een totaalpakket rond een (uitgebreid) open dataplatform/ repository vertegenwoordigt, en OpenStreetMap (OSM), dat het archetype van een collaboratief open dataplatform vertegenwoordigt. Zes principes bleken op beide archetypen van toepassing:

1. Zorg voor consistentie in hoe leden van de organisatie of gemeenschap de (potentieel veelzijdige) organisatorische identiteit zien, vooral naarmate de organisatie of gemeenschap groeit of meer gedecentraliseerd wordt.
2. Overweeg om nieuwe of niet-traditionele oplossingen aan te bieden in plaats van vast te houden aan traditionele producten.
3. Investeer en omarm nieuwe en opkomende technologie voor de software-infrastructuur (bijv. de toepassing van AI voor metadata-aanbevelingen).
4. Bied producten of diensten aan die het proces van het leveren en gebruiken van open data zoveel mogelijk vereenvoudigen.
5. Stimuleer potentiële samenwerkingen tussen meerdere belanghebbenden (bijv. via projecten of evenementen).
6. Bied consultancy- of trainingsdiensten aan, bijvoorbeeld via het formele onderwijs (bijv. scholen of universiteiten).

Er werden elf overwegingen geïdentificeerd die specifiek van toepassing zijn op het archetype van het totaalpakket:

1. Zorg ervoor dat de aangeboden open data intemediatiediensten goed geïntegreerd zijn met de bestaande kernproducten of -diensten.
2. Bied open data intermediatiediensten aan die consistent zijn met de identiteit van de organisatie (d.w.z. geen significante verschuiving ten opzichte van de kernactiviteiten).
3. Bied diverse complementaire producten of diensten aan die gebruikmaken van open data.
4. Bied diensten aan die de risico’s die klanten lopen met betrekking tot open data minimaliseren.
5. Bied data, diensten of projecten op maat aan die aansluiten op lokale behoeften.
6. Bied opensourcesoftware aan, ten minste gedeeltelijk als onderdeel van de bredere productsuite (bijv. via een freemium-model).
7. Maak het mogelijk dat feedback op open data op een gestructureerde manier kan worden gegeven.
8. Zorg ervoor dat de unieke middelen of positie niet worden benut op manieren die de groei van andere actoren oneerlijk belemmeren (de ‘winner takes it all’-situatie), bijvoorbeeld door je in te zetten voor de ontwikkeling van breed toepasbare open standaarden en technische interoperabiliteit.
9. Laat de waarde van open data zien.
10. Pleit voor het vrijgeven van open data vanuit niet-publieke sectoren.
11. Investeer in samenwerkingen op basis van open data.

Er werden veertien overwegingen geïdentificeerd die specifiek van toepassing zijn op het collaboratief open dataplatform:

1. Bevorder een gezonde en constructieve gemeenschap van mensen en organisaties die bijdragen aan het platform.
2. Bescherm de algemene belangen van de mensen en organisaties die bijdragen aan het platform door een transparante (besluitvormings-) processen te implementeren en een mechanisme in te stellen om een mogelijke (vijandige) overname te voorkomen.
3. Investeer in het vergroten van de zichtbaarheid en het bereik van de organisatie of gemeenschap.
4. Zorg voor een duidelijk, efficiënt en constructief communicatiemechanisme.
5. Faciliteer en moedig de ontwikkeling aan van (met name open-source) tools rond het platform (d.w.z. federatieve architectuur).
6. Maak gebruik van een zelfcorrigerend mechanisme door een ‘teamgerichte’ cultuur te cultiveren en een duidelijk proces voor geschillenbeslechting te hanteren.
7. Ondersteun lokale afdelingen of gemeenschappen bij het samenwerken met lokale organisaties (bijv. lokale overheden of ngo’s).
8. Maak van meningsverschillen en conflicten kansen voor het verbeteren van besturingsmechanismen en technische ontwikkeling.
9. Sta toe dat diverse soorten data worden bijgedragen om tegemoet te komen aan de vele verschillende behoeften van gebruikers.
10. Zorg er proactief voor om ook het perspectief van actoren die indirect baat bij de activiteiten van de organisatie hebben (d.w.z. niet-gebruikers) te betrekken, bijvoorbeeld door samen te werken met ngo’s ter plaatse en hun behoeften te prioriteren.
11. Kies een wederkerige licentie die niet (veel) beperkingen oplegt aan het eindgebruik (d.w.z. afgeleid werk).
12. Werf proactief nieuwe databijdragers waaronder gemarginaliseerde of achtergestelde groepen en creëer een gastvrije omgeving om ze blijvend aan de organisatie te verbinden.
13. Neem feedback over de technologische ontwikkeling van niet-technische experts op een zinvolle manier in overweging.
14. Ontwikkel een cultuur waarin ethische verantwoordelijkheden bij het aanleveren en het gebruik van data worden overwogen, en die verder gaat dan het verstrekken van algemene ethische richtlijnen.

Deze nieuwe kennis kan door bestaande of potentiële nieuwe open data-intermediairs in overweging worden genomen bij het ontwikkelen van hun bedrijfsmodel. Hierbij is het vooral van belang om naast de technologische dimensie, ook de sociale dimensie (bijvoorbeeld organisatorische identiteit, betrokkenheid bij de gebruikers en gemeenschapsmanagement) en de juridische dimensie (bijvoorbeeld de keuze van de licentie), mee te nemen bij de ontwikkeling van bedrijfsmodellen voor open data intermediatie.

Verder biedt dit proefschrift ook reflecties op het concept van een (duurzame) ODE. Het bevestigt onder andere de noodzaak om onderscheid te maken tussen rollen en actoren in een ODE. Verschillende soorten organisaties of entiteiten kunnen open data-intermediairs zijn of worden; dus ook organisaties uit de publieke sector. Dit proefschrift benadrukt ook, aan de hand van de casus OSM, dat er al een ODE bestaat waar met name open data van buiten de overheid breed circuleert. Er is echter wel een risico dat open overheidsdata en open niet-overheidsdata zich onafhankelijk van elkaar ontwikkelen (bijvoorbeeld overheidsorganisaties die prioriteit geven aan het ontwikkelen van de overheidsdata-infrastructuur en niet de mogelijkheden benutten om samen bij te dragen aan infrastructuur die buiten de publieke sector is geïnitieerd). Daarnaast toont dit proefschrift het belang aan van de rol van de technologie voor de duurzaamheid van de ODE. Hiervoor zijn niet alleen de kenmerken van de open data (dekking, kwaliteit, formaat, licentie, etc.) van belang, maar ook de manier waarop de data verspreid en hergebruikt worden, inclusief de gebruikte software.

Toekomstig onderzoek kan beleidsaspecten bestuderen die open data-intermediairs ertoe kunnen aanzetten om te handelen op een manier die de toekomstbestendigheid van de ODE ondersteunt. Dit omvat bijvoorbeeld het onderzoeken van de haalbaarheid en voordelen van het uitbreiden van de open data wetgeving naar open data intermediairs van buiten de overheid. Daarnaast verdienen de verborgen en mogelijk ongewenste effecten van open data-intermediairs op andere ODE-actoren aandacht, rekening houdend met het feit dat deze effecten kunnen evolueren gedurende de ontwikkeling van hun bedrijfsmodel (d.w.z. van opkomende versus volwassen en potentieel dominante fase).

Dit proefschrift heeft zich gericht op slechts twee casestudies die betrekking hebben op twee van de negen archetypen van het bedrijfsmodel van open data-intermediatie. Uitbreiding van het onderzoek naar de andere archetypen is daarom aanbevelenswaardig, omdat deze mogelijk tot andere conclusies leiden. Bovendien zijn de twee gekozen casestudies, Esri en OSM, beide gevestigde open data-intermediairs uit hetzelfde geografische informatie domein. Of de bevindingen ook van toepassing zijn op open data intermediairs van dezelfde archetypes in een ander domein moet blijken uit vervolgonderzoek. Echter ook binnen het geografische informatie domein bestaan er zeer veel verschillende soorten open data en actoren. Vervolgonderzoek moet aantonen of de bevindingen uit de twee cases ook van toepassing zijn op andere cases uit het geografische informatie domein. Zulk onderzoek kan bijvoorbeeld betrekking hebben op open data-intermediairs die zich bezighouden met lucht- en straatbeelden en historische archiefkaarten.

Concluderend vergroot dit proefschrift het inzicht in open data-intermediairs, hun positie binnen de ODE en hun bedrijfsmodellen. Dit proefschrift heeft met name theoretisch bijgedragen aan de definitie van open data-intermediairs, de mogelijke bijdragen van open data-intermediairs aan de toekomstbestendigheid van een ODE en ontwikkelde negen archetypen van bedrijfsmodellen voor open data-intermediatie. Het identificeerde ook praktische aspecten die in overweging moeten worden genomen bij de ontwikkeling van bedrijfsmodellen voor open data-intermediatie die bijdragen aan een toekomstbestendige ODE. Deze inzichten kunnen open data-intermediairs helpen hun rol binnen de ODE te versterken, wat uiteindelijk de waardecreatie van open data ten goede komt. Deze waarde kan vervolgens worden ingezet om innovatie te stimuleren, het economisch welzijn te bevorderen en urgente sociale en ecologische uitdagingen aan te pakken.

1 Introduction

A substantial portion of this chapter was previously published in
– Shaharudin, A., van Loenen, B., & Janssen, M. (2023). Towards a Common Definition of Open Data Intermediaries. *Digital Government: Research and Practice*, 4(2), 6:1-6:21. https://doi.org/10.1145/3585537
– Shaharudin, A., van Loenen, B., & Janssen, M. (2024). Exploring the contributions of open data intermediaries for a sustainable open data ecosystem. *Data & Policy*, *6*, e56. https://doi.org/10.1017/dap.2024.63
– Shaharudin, A., van Loenen, B., & Janssen, M. (2025). Developing an Open Data Intermediation Business Model: Insights From the Case of Esri. *Transactions in GIS*, 29(1), e13304. https://doi.org/10.1111/tgis.13304
and submitted for journal consideration as
– Shaharudin, A., van Loenen, B., & Janssen, M. (n.d.). *Business model archetypes of open data intermediaries: Empirical insights from practice*. Manuscript submitted for publication.

In today's digital era, data is everywhere, generated constantly by individuals, organisations, and sensors, at an unprecedented scale. This vast and growing availability of data holds great potential for innovation and societal well-being. It is instrumental in addressing complex local and global challenges, from improving public health and urban planning to generating new products and services. However, the value of data can be better realised when it is used widely by diverse actors. Three decades ago, McLaughlin & Nichols (1994, p. 67) wrote that

'If goals such as sustainable development are to be reached, then private individuals, citizen groups, and all levels of public and private sector organisations require timely access to a wide variety of databases [and] […] the capability of integrating that data horizontally [i.e., across types] and vertically [i.e., across administrative boundaries]'.

Open data plays a crucial role in enhancing value generation from data. It facilitates the use of data by those without the resources and capabilities to collect or purchase data themselves. Thus, open data allows not only governments and large corporations but also smaller actors such as startups and civil society organisations (CSOs) to collectively harness and materialise the value of the data. Consequently, open data allows a broader engagement and re-use of data, encouraging the production of new insights and innovation for the benefit of the society as a whole.

1.1 Background: Context, motivation, and terminologies

1.1.1 Open data

The International Open Data Charter (IODC) defines open data as ‘digital data that is made available with the technical and legal characteristics necessary for it to be freely used, re-used, and redistributed by anyone, anytime, anywhere’ (IODC, n.d.). Similarly, the Open Knowledge Foundation (OKF) defines open data as ‘data that can be freely used, modified, and shared by anyone for any purpose’ (OKF, n.d.-b). Open data have four requirements according to OKF: the data must be (1) ‘in the public domain or provided under an open license’, (2) ‘provided as a whole and at no more than a reasonable one-time reproduction cost, and should be downloadable via the internet without charge’, (3) ‘provided in a form readily processable by a computer and where the individual elements of the work can be easily accessed and modified’, and (4) ‘provided in an open format […] which places no restrictions, monetary or otherwise, upon its use and can be fully processed with at least one free/libre/open-source software tool’.

Open data is also defined legally or administratively. For example, the European Union (EU) open data and the re-use of public sector information directive 2019 (known as the Open Data directive) defines open data as ‘data in an open format that can be freely used, re-used and shared by anyone for any purpose’ (Directive (EU) 2019/1024). The United States (US) Open, Public, Electronic, and Necessary Government Data Act or OPEN Government Data Act defines ‘open Government data asset’ as ‘a data asset maintained by the Federal Government that is (A) machine readable; (B) available in an open format; (C) not encumbered by restrictions that would impede use or re-use; and (D) based on an underlying open standard that is maintained by a standards organisation’ (United States Congress, 2019).

While the various definitions described are consistent with each other, this dissertation favours and adopts the definition by IODC due to its succinctness. It thus should be emphasised that data shared on a case-by-case basis or with a select group of users is not considered open data. In other words, open data is only one type of data sharing, and not all types of data sharing are in the form of open data. While the importance and relevance of other types of data sharing are certainly recognised, this dissertation focuses on open data because of its peculiar characteristics in practice, namely its reusability by anyone for any purpose. It should also be emphasised that open data can be used to develop new data products and services offered under a proprietary format or for a fee. Furthermore, this dissertation focuses on open data *beyond* open government data, including open data provided by companies, non-governmental organisations (NGOs) or individuals, while acknowledging that much of the progress in terms of policy and research so far has been on open government data.

Open data promises various benefits, including stimulating innovation, enhancing public sector processes, improving disaster responses, addressing public health crises, facilitating environmental protection, catalysing citizen participation and engagement, and improving transparency and accountability (Adaktylou et al., 2020; Biljecki et al., 2021; Brovelli & Coetzee, 2021; Degbelo, 2022; Janssen et al., 2012; Mooney et al., 2021; Zhu et al., 2019). For instance, open data provided by public transportation providers has been used to develop urban mobility apps (Citymapper, 2025), open data collected by volunteers of the Humanitarian OpenStreetMap Team (HOT) has been used in disaster and humanitarian crisis responses (HOT, 2018c), and open data provided by the Global Forest Watch has been used by environmental advocates and policymakers in the Philippines to support the enactment of the National Mangrove Forests Conservation and Rehabilitation Act (Camero, 2015) as well as by indigenous groups in Peru to track and report forest loss in the Amazon (Moloney, 2021). Additionally, the implementation of an open procurement data policy in the EU has led to more competitive bidding, which has been attributed to increased scrutiny by non-governmental organisations and investigative journalists, as well as learning by national procurement regulators (Duguay et al., 2023). In terms of economic value, based on an econometric analysis, Koski (2011) found that firms operating in countries where public sector geographic information provided for free or at maximum marginal costs grew about 15% per annum on average. Another more recent econometric analysis showed that the higher the level of open data in a country, the higher its quality and scale of entrepreneurial activities, measured by the Global Entrepreneurship Index (GEI) (Huber et al., 2022).

1.1.2 Open data ecosystem

Despite its promises, generating value from open data is not without hurdles (Coetzee et al., 2020; Johnson et al., 2017; Johnson & Varga, 2022). For example, processing open data may involve laborious work, especially if the data comes from multiple sources in different models and formats (Aydinoglu & Bilgin, 2015). It may also involve technical skills and complementary assets that are not at the disposal of open data users (Okamoto, 2016; Temiz et al., 2022). Open data users may also have to deal with poor or inconsistent open data quality (Benitez-Paez et al., 2018; Quarati et al., 2021; Welle Donker & van Loenen, 2017).

Based on empirical studies of the United Kingdom (UK) government's open data initiatives and the International Aid Transparency Initiative, Davies (2011) argued that successful value generation from open data relies on more than just the dataset; it also relies on the 'mobilisation of a wide range of technical, social and political resources, and on interventions [...] [to] support coordination of activity around datasets' (Davies, 2011, p. 1). Thus, he advocated for the *open data ecosystem* (ODE) as an analytical lens through which 'the emergent, autonomous and self-organising components' are 'linked together in local and global feedback loops and developing according to local specialisations and adaptation rather than top-down design' (Davies, 2011, p. 3). According to Davies, the *ecosystem* metaphor is useful in describing the interrelations of open data actors and devising interventions to increase cooperation and interaction among actors, just as the metaphor is used in other fields such as economics and political science. This dual function of the metaphor was also emphasised by Harrison et al. (2012, p. 907): 'Although ecosystems are naturally occurring phenomena and the metaphor may be applied to any existing socio-technical domain, they can also be seeded, modelled, developed, managed, that is, intentionally cultivated for the purpose of achieving a managerial and policy vision'. Even though the conceptualisation of the ODE by Davies (2011) is more specifically referring to open government data, it is still relevant in thinking about the ODE more broadly beyond open government data.

Csáki (2019) succinctly defined the ODE as a '*way of looking* [emphasis added] at how participating actors and groups create shared meaning and generate value around open data and how the structural properties of their interactions shape this process, which in turn enables or constrains the growth and health of the ecosystem itself' (Csáki, 2019, p. 19). This definition evokes similarities with how actor-network theory (ANT) is described by one of its co-founders, which is 'a disparate family of material-semiotic *tools, sensibilities and methods of analysis* [emphasis added] that treat everything in the social and natural worlds as a *continuously generated effect of the webs of relations* [emphasis added] within which they are located' (Law 2008, p. 141). Thus, in this dissertation, the ODE is viewed as a specific application of ANT, rooted in the larger field of science and technology studies (STS). From this perspective, the ODE is not an isolated nor necessarily novel analytical approach to a socio-technical network but one that can draw theoretically and empirically grounded insights from other types of networks, including information systems, see e.g., Díaz Andrade & Urquhart (2010), Doolin & Lowe (2002), Nehemia-Maletzky et al. (2018), Pandey et al. (2022), and Walsham (1997). This dissertation thus adopts the definition of the ODE by Csáki (2019) that emphasises its utility as an analytical lens.

A fundamental notion underpinning ANT is the participation of humans and non-humans in a network, including technological artefacts (in the case of open data, e.g., data portals, application programming interfaces (APIs), and data standards), not only passively as a resource or constraint, but also actively influencing the dynamic interactions in the network (Callon & Law, 1997). This mirrors the conceptual elements of the data ecosystem identified by Oliveira and Lóscio (2018), namely resources, roles, actors, and relationships. Although their conceptualisation is for the data ecosystem beyond open data, it is still useful in formulating the ODE. According to Oliveira & Lóscio (2018), resources can be datasets, data-based software, and hardware, which may be exchanged individually or in combination, through relationship transactions. A role is a function of an actor in the ecosystem. Actors are autonomous entities, such as businesses, public institutions, and individuals, serving one or more specific roles. Relationships are the interactions among actors in the ecosystem. Notably, Oliveira and Lóscio (2018) made a distinction between roles and actors: actors are not wedded to a particular role.

Poikola et al. (2011) drew attention to the autonomy yet interdependency of actors in the open government data ecosystem, which is also relevant to the broader ODE. As they put it, the 'ecosystem evokes an image of well-being of the entity and, on the other hand, fulfilling one's own needs through the richness and vitality of the ecosystem' (Poikola et al., 2011, p. 14). They also highlighted the ever-changing nature of an ecosystem: 'With the ecosystem, we wish to highlight not only the technological systems and institutionalised organisations, but also the living, *dynamically changing network of interaction* [emphasis added]' (Poikola et al., 2011, p. 14). This hybridity of self-interested yet interdependent actors is likewise an essential notion of ANT, asserting that 'every stable social arrangement is simultaneously a point (an individual) and a network (a collective)' (Callon and Law 1997, p. 165).

This dissertation uses the singular form to refer to the ODE instead of the plural form, i.e., ODEs (unless when direct quoting from other sources). Even though there could be multiple subsets or disconnected networks of open data actors, this particular terminological choice is meant to emphasise the approach taken in this dissertation, where the ODE is used as an analytical lens instead of necessarily referring to a tangible ontic structure.

1.1.3 Sustainable open data ecosystem

One of the empirical focuses of ANT is ‘to trace and explain the processes whereby *relatively stable* [emphasis added] networks of aligned interests are created and maintained, or alternatively to examine why such networks fail to establish themselves’ (Walsham 1997, p. 469). In the same manner, researchers are investigating factors that contribute to a sustainable ODE, where stable networks of open data actors’ aligned interests are created and maintained (adopting the words of Walsham). According to the Oxford Dictionary (Oxford, n.d.), the word ‘sustainable’ is defined as ‘that can continue or be continued for a long time’. Since the ODE can be understood from both descriptive and normative perspectives, the term ‘sustainable ODE’ thus refers specifically to the normative scenario of an ODE that can continue to function and grow for a long term.

TABLE 1.1 Features of a sustainable ODE by van Loenen et al. (2021), adopted as the starting point of the dissertation

Feature	Description
User-driven	• Open data is provided based on the needs of diverse users, including in terms of what and how data should be provided. • Context: Current open data provision is considered to be mainly supplier-driven, where the data supplied does not satisfactorily meet the needs of diverse users. Meanwhile, there are often limited mechanisms in place for open data providers to understand and deliver those needs. A sustainable ODE is understood by van Loenen et al. (2021) as one that fosters the open data demand-supply matching.
Circular	• All actors mutually create and capture value; for example, open data re-users provide value-added data back to the providers. • Context: Currently, open data value flow is considered linear, where most open data users merely benefit from open data without contributing value back to the open data providers (who have to bear the costs of providing open data). This situation is deemed to be an unfair distribution of value and may discourage further/future provision of open data. Thus, a sustainable ODE is considered by van Loenen et al. (2021) as one where actors simultaneously contribute and extract value to/from other open data actors.
Inclusive	• Diverse actors contribute open data (i.e., not only governments) and participate in the ecosystem decision-making processes (within their capacity). • Context: Current research and practice focus mainly on governments as open data providers and others, especially companies, as merely open data users, overlooking many potential datasets that can be made open. Similarly, open data decision-making processes (e.g., around open data policy) are often confined to within government organisations. A sustainable ODE is considered by van Loenen et al. (2021) as one that is not limited to open government data but includes open non-government data and one where diverse types of actors can be involved in decision-making processes.
Skills-based	• Application of appropriate data skills and competencies to optimise value generation from open data. • Context: Much of the current open data use is hindered by a widespread lack of digital and data skills, undermining the value generation from open data. A sustainable ODE is considered by van Loenen et al. (2021) to emphasise data literacy and skills development.

Notably, van Loenen et al. (2021) suggested the key features of a sustainable ODE, which can be summarised as user-driven (open data supply matches the demands of users of different types and domains), circular (all actors mutually create and capture value), inclusive (all actors, not only government organisations, are incentivised to contribute open data and participate in related decision-making processes), and skills-based (appropriate data skills and competencies are applied). These features are not mutually exclusive. In this dissertation, the normative framework of a sustainable ODE by van Loenen et al. (2021) is used as a starting point, since it provides a concise frame of reference with which to work (Table 1.1).

1.1.4 Open data intermediaries

Various expectations for open data to be 'actionable data' (Gutierrez & Landa, 2021) cannot be fulfilled by open data providers or end-users alone. Chattapadhyay (2014) argued that a vital aspect of the ODE lies in the efficient circulation of resources (e.g., software, technical skills, funding), with open data intermediaries serving a pivotal role in this regard. The role of open data intermediaries is instrumental in the access to and use of open data (González-Zapata & Heeks, 2015; Neves et al., 2020) and in connecting open data actors (Mayer-Schönberger & Zappia, 2011; Yoon et al., 2018). Furthermore, since the ODE emphasises the self-organisation of actors (Davies, 2011; Oliveira & Lóscio, 2018), open data intermediaries are crucial in mitigating information asymmetry between actors in the ODE.

In the same vein, the intermediary is a crucial concept in ANT. ANT differentiates intermediaries, a messenger that 'transports meaning or force without transformation' from mediators that 'transform, translate, distort, and modify the meaning or the elements they are supposed to carry' (Latour 2007, p. 39). However, this dissertation's conceptualisation of open data intermediaries does not make such a distinction (i.e., both ANT's intermediaries and mediators are termed 'open data intermediaries' in the ODE) since this is how the term is conventionally used in the literature as an umbrella term for heterogeneous types. Regardless, ANT highlights the importance of intermediaries/mediators as match-makers for situations in which 'people, goods and services are brought together' (Goodchild and Ferrari 2024, p. 107).

Various types of actors serve the role of open data intermediaries, including CSOs (Meng et al., 2019; Sangiambut & Sieber, 2017), companies (Andrason & van Schalkwyk, 2017; Germano et al., 2016), the media (Enaholo & Dina, 2020; Johnson & Greene, 2017), researchers (Corbett et al., 2020; S. Park & Gil-Garcia, 2017), and government organisations (Hablé, 2019; Meijer & Potjer, 2018). They undertake various tasks, deploying different types of resources based on their specialisation. Examples of open data intermediaries include software providers that process and integrate open data in their software (e.g., Citymapper and Esri), crowdsourcing platforms that compile and facilitate the re-use of open data from various contributors (e.g., OpenStreetMap and Wikidata), and portal providers integrating data from different open data sources (e.g., Global Forest Watch and Humanitarian Data Exchange). Actors referred to as open data intermediaries do not necessarily undertake activities solely related to open data, even though intermediating open data is part of their activities. For example, a company can sell non-open data-based products while offering open data intermediation services.

As a point of clarification, open data intermediaries should certainly include *open data* in their activities, despite their end product not necessarily being free and open to everyone; the same emphasis was notably made by van Schalkwyk et al. (2016, p. 12). Although this may seem obvious, being involved with data shared on a case-by-case basis through individual arrangements and not with open data (conventionally defined), as in the case of certain forms of data collaboratives (Susha et al., 2017) or boundary organisations (Perkmann & Schildt, 2015), does not make such actors open data intermediaries. As similarly highlighted by scholars of ANT (Klecuń, 2004; Lee & Hassard, 1999), a delimitation based on concrete practices and functioning for empirical studies is necessary to avoid dealing with an ‘endless chain of association’ (Müller 2015, p. 30).

Moreover, open data intermediaries are *not* to be confused with the *data intermediaries* established within the EU Data Governance Act framework (Regulation (EU) 2022/868 (Data Governance Act), 2022). As written (e.g., Item 10 of the Preamble), this Act does not apply to open data that falls under the EU Open Data Directive. Moreover, data intermediaries in the Data Governance Act are not allowed to use the data they intermediate (e.g., by developing products based on the data) and can only facilitate data sharing between parties. In practice, this cannot be applied to open data intermediaries, which can freely use the open data they deal with (even for financial profit), because the data is already open.

At the beginning of this dissertation journey, it was apparent that there were various definitions of open data intermediaries in the literature, with some potentially at odds with one another. For example, definitions by Janssen & Zuiderwijk (2014) and Johnson & Greene (2017) regard open data intermediaries as those that are actively involved in the processing of open data, whereas the definition by Robinson & Mather (2017) considers them as those that connect community members with open data and do not necessarily process the data themselves. Hence, the first research question of this dissertation addresses this conundrum by proposing a common definition of open data intermediaries (Chapter 2). From thereon, the proposed definition is used throughout the dissertation. Based on Chapter 2, open data intermediaries are defined as 'third-party actors who provide specialised resources and capabilities to (i) enhance the supply, flow, and/or use of open data and/or (ii) strengthen the relationships among various open data actors'.

1.1.5 Open data intermediation business model

The term *business model* is an elusive concept with numerous definitions and interpretations in the literature (see the review by Andreini and Bettinelli (2017)). For example, in the context of electronic business models, Timmers (1998) defined a business model as 'an architecture for the product, service and information flows, including a description of the various business actors and their roles; and a description of the potential benefits for the various business actors; and a description of the sources of revenues'. Also in the same context, Dubosson-Torbay et al. (2002) defined a business model as 'the architecture of a firm and its network of partners for creating, marketing and delivering value and relationship capital to one or several segments of customers in order to generate profitable and sustainable revenue streams'. More recently, Snihur & Markman (2023) described a business model as 'a blueprint that outlines how an organisation creates value, generates revenue, delivers offerings, and even interacts with its direct stakeholders (employees, customers, suppliers) and indirect stakeholders (rivals, regulators, community)'.

Notwithstanding the variations in specific definitions and terminologies, typical conceptualisations of a business model consist of at least three general dimensions (Afuah, 2018; Andreini & Bettinelli, 2017; Teece, 2010; Voigt et al., 2017): (i) value proposition (potential benefits for the consumers); (ii) value creation (resources and activities deployed by organisations to deliver the value proposition); (iii) value capture (compensation, not necessarily in monetary form, to the organisations offering the value) (Figure 1.1). The literature on data-related business models, such as big data business models (Acciarini et al., 2023), data marketplaces (Bergman et al., 2022), and platform business models (Täuscher & Laudien, 2018), took on a similar conceptualisation of business models as described. Hence, following its typical conceptualisation, this dissertation defines a business model as a framework making sense of what values an organisation offers (value proposition), how it offers such values, including the activities and resources involved (value creation), and why it offers them, particularly what benefits does it gain (value capture).

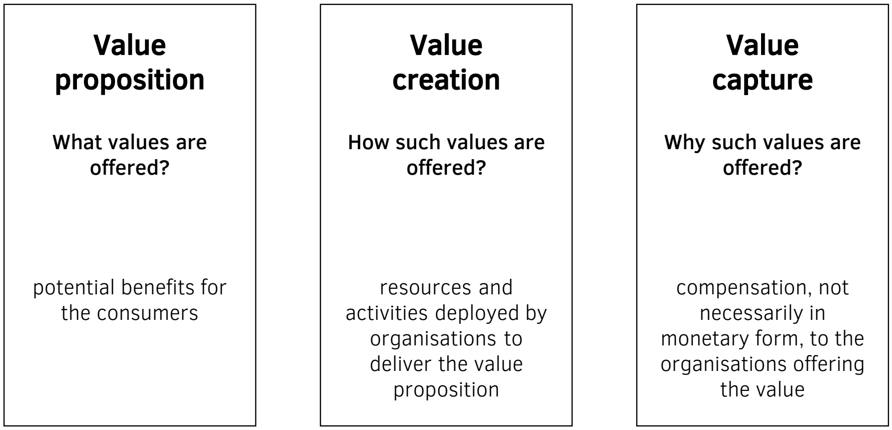


FIG. 1.1 Three dimensions of a business model adopted in the dissertation

Both for-profit and non-profit organisations (NPOs) require business models (Bocquet et al., 2020; Maguire, 2009). The latter need revenue to support operational costs, even though they do not generate profits (in economic terms, profits equal revenues minus costs). Additionally, since actors referred to as open data intermediaries do not exclusively engage in activities related to open data, the term *open data intermediation business model* refers specifically to elements or components of their business model that include open data.

Business models are crucial to the success of an organisation since they clarify how organisations (should) operate (Magretta, 2002). To ensure long-term sustainability, an organisation must select suitable business model(s) based on the present circumstances, execute them well, develop and strengthen the organisation's dynamic capabilities, and modify the business models effectively and in a timely manner when an opportunity or threat arises (DaSilva & Trkman, 2014). Business models also encapsulate the relationships between an organisation and other actors (Johannesson, 2007; Lambert & Davidson, 2013), such as open data providers and users, in the case of open data intermediaries. Thus, business models can serve as strategic tools for exploring new markets or opportunities and strengthening existing relationships or forging new ones (Wieland et al., 2017). Business models are often characterised as the framework linking an organisation's long-term strategy with its (micro-level) business processes (Di Valentin et al., 2012; Spencer, 2013; Veit et al., 2014).

TOC

1.2 Problem statement

Despite the importance of open data intermediaries in the ODE having been widely acknowledged in research and practice (see e.g., Carolan (2016), Davies & Perini (2016), Dove et al. (2023) and Publications Office of the European Union (2023)), studies on open data intermediation business models are limited in number and scope. This poses a problem since business model design and innovation are often associated with organisations' performance and longevity (DaSilva & Trkman, 2014; Kesting & Günzel-Jensen, 2015; Peric et al., 2017). Besides, business models also clarify the relationships between an organisation and other stakeholders (Lambert & Davidson, 2013). Thus, it is crucial to understand how aspects of open data intermediation business models affect other ODE actors and be developed to support the sustainability of the ODE. Such understanding is currently missing in the academic literature.

The critical literature review by Wiener et al. (2020) on data business models indicated that organisations (especially large ones) tend to exhibit a high level of vertical integration, where they simultaneously take on the role of data users, data suppliers, and data facilitators. Nevertheless, the authors argued that going forward, more organisations would rely on external parties to fully leverage data as they may not possess all the capabilities needed internally. This resonates with the recognition of open data intermediaries' growing importance in supporting open data users and providers, as well as other open data intermediaries, for specific activities.

As third-party actors, open data intermediaries' business models rest on offering products/services to open data providers, users, or other intermediaries. Since open data is already available free of charge under an open license, open data intermediaries do not capture value solely by facilitating the transaction of data from one party to another (as in the case of data intermediaries dealing with private or proprietary data). Instead, open data intermediaries primarily gain profits or non-monetary benefits by adding value to existing open data (e.g., by integrating open data from various sources or improving the quality of open data) or the processes linked to their supply and re-use (e.g., by curating relevant open data for specific purposes or facilitating open data actors' interactions).

Conceptualising business models as a modular unit of analysis enables researchers and practitioners to evaluate them based on their merit and investigate the factors contributing to their successful adoption and ability to achieve organisations' goals (Lambert & Davidson, 2013; Seddon et al., 2004). Many researchers and practitioners have been interested in exploring innovative new business models as

TOC

alternatives to existing (or mainstream) ones (e.g., Chesbrough, 2010; Geissdoerfer et al., 2020; Rossi et al., 2022). In order to explore and develop new models or even refine existing models, having thorough knowledge of the current business models is essential. Such knowledge is currently lacking for open data intermediaries.

1.3 Objective and research questions

The overarching objective of this dissertation is

- **To understand how open data intermediation business models can support a sustainable open data ecosystem.**

Accordingly, the overarching research question (RQ) of this dissertation is

- **What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem?**

To that end, four sub-research questions are addressed in this dissertation:

- **RQ1: What are open data intermediaries?**

Description: Various definitions and conceptualisations of open data intermediaries were described in the academic literature. Some contradict each other, while some are narrower than others. Concurrently, there is no established definition of open data intermediaries in policy documents that can be relied upon. Before delving deeper into studying open data intermediaries and their business models, a clear understanding of what constitutes open data intermediaries must first be established to ensure knowledge about them is built on top of a consistent vocabulary.

One may ask why this dissertation chooses to define *open data intermediaries* instead of *open data intermediation*. This decision is because simply defining the latter, which is an activity, obscures the fact that actors who perform open data intermediation do not necessarily deal exclusively with open data. However, their overarching motivation and interest may significantly influence their open data intermediation (including why they get involved with open data in the first place). Moreover, defining open data intermediaries also triggers a better understanding or further interrogation of the relationships between actors who perform open data intermediation with other ODE actors.

Methodology: This RQ was tackled through a systematic literature review (SLR) to (i) compile existing definitions of open data intermediaries, (ii) identify the wide range of actor types, tasks, and objectives of open data intermediaries, and (iii) propose a common definition of open data intermediaries based on (i) and (ii). Chapter 2 addresses RQ1.

- **RQ2: What are potential contributions of open data intermediaries in addressing challenges in an open data ecosystem?**

  Description: To better situate open data intermediaries within the ODE, it is necessary to look into what they (can) contribute to other actors. This can help refine the understanding of open data intermediaries' characteristics and their relationships with other ODE actors. Most studies have focused solely on examining the current activities of open data intermediaries. While this foundational understanding is certainly valuable, by also exploring what they can potentially contribute, this RQ also aims to identify gaps in the ODE that open data intermediaries may be able to close or narrow.

  Methodology: This RQ was tackled by exploring the connections between challenges in the ODE and potential contributions by open data intermediaries to address them. Towards that end, a two-stage methodology was employed. In Stage 1, data was gathered through semi-structured interviews. From these interviews, challenges in the ODE and the potential contributions of open data intermediaries were derived. In Stage 2, the links between the individual potential contributions of open data intermediaries and specific challenges in the ODE that they can address were explored. These links were validated with organisations interviewed and additional open data practitioners and researchers. Chapter 3 tackles RQ2.

- **RQ3: What are archetypes of open data intermediation business models?**

  Description: Since there can be an endless number of business models, each with its unique elements and arrangements of value proposition, value creation, and value capture, this RQ aims to identify common archetypes of open data intermediation business models. Such identification is necessary to consider more targeted, appropriate, and relevant recommendations in developing open data intermediation business models, beyond generic recommendations. Only two past studies have explicitly investigated open data intermediation business models. Janssen and Zuiderwijk (2014) identified six archetypes based on their study of 12 cases in the Netherlands. However, all six were described only in terms of the value proposition dimension (i.e., the type of product). On the other hand, based on seven cases in Brazil, Germano et al. (2016) identified only three archetypes, which are very limited.

Methodology: The RQ was answered through a four-stage methodology. First, an initial codebook consisting of categories and elements of open data intermediation business models was developed through an SLR. Second, relevant qualitative data was collected from 190 samples of existing open data intermediaries facilitated by the initial codebook developed. The codebook was iteratively modified based on the learning throughout the data-gathering process. Third, K-means clustering was employed to group the business models of the sample cases. Fourth, the K-means clustering results were interpreted, and the archetypes were identified. Chapter 4 addresses RQ3.

- RQ4 (also the overarching RQ): **What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem?**

  Description: Open data intermediaries serve an instrumental role in the circulation of resources in the ODE. Thus, their business models influence the sustainability of the ODE. RQ4, which is also the overarching research question of this dissertation, aims to recommend aspects to consider in developing open data intermediation business models that support sustainable ODE.

  Methodology: Since there are multiple open data intermediation business model archetypes (identified in RQ3), RQ4 was addressed specifically for two archetypes, each through an in-depth single-case study presented in Chapters 5 and 6. A single-case study allows for a deep contextualised understanding of the case in question through 'thick' descriptions, which is difficult to achieve through a multiple-case study (Dyer & Wilkins, 1991; Siggelkow, 2007).

  The decision to study two archetypes was guided by feasibility considerations within the scope of a doctoral dissertation. Nevertheless, rather than selecting the archetypes in advance, this dissertation first selected the case studies based on the following reasoning: This dissertation opted to concentrate on open geospatial data to address RQ4. This is because, as Gray (2014) pointed out in his work on the genealogy of open data, the geospatial domain had dealt with various issues and controversies related to open data years before they were encountered in other domains. Furthermore, four out of six thematic categories of the EU Open Data Directive's high-value datasets constitute (geo)spatial datasets, thereby affirming the importance of such data. In the geospatial domain, Esri and OpenStreetMap (OSM) are two notable open data intermediaries; hence, they are chosen as the case studies for this dissertation. They represent two different business model archetypes.

1.4 Dissertation outline

The organisation of this dissertation is shown in Table 1.2. RQ1 is addressed in Chapter 2 and RQ2 in Chapter 3. The two chapters constitute Part A of the dissertation, which focuses on clarifying the baseline understanding of open data intermediaries, including their position and contributions in the ODE. RQ3 is addressed in Chapter 4, which forms Part B of the dissertation, characterising different archetypes of open data intermediation business models. RQ4 is addressed in Chapters 5 and 6, composing Part C of the dissertation that aims to recommend aspects to consider in developing open data intermediation business models that support the sustainability of the ODE. By analogy, the progression from Part A to Parts B and C can be understood as a shift from a global, 'forest-level' perspective, to more detailed 'tree-level' views. Lastly, Part D, which consists of Chapter 7, brings all the insights and knowledge developed in this dissertation together, (re)framing and reflecting on the present understanding about open data intermediaries and the ODE.

TABLE 1.2 Dissertation outline

Part	Chapter	Research question
PART A: Understanding open data intermediaries	**Chapter 2**: Defining open data intermediaries	**RQ1**: What are open data intermediaries?
	Chapter 3: Contributions of open data intermediaries in addressing challenges in an open data ecosystem	**RQ2**: What are potential contributions of open data intermediaries in addressing challenges in an open data ecosystem?
PART B: Characterising archetypes of open data intermediation business models	**Chapter 4**: Archetypes of open data intermediation business models	**RQ3**: What are archetypes of open data intermediation business models?
PART C: Delving deeper into selected open data intermediation business models	**Chapter 5**: The case of Esri	**RQ4 (also overarching RQ)**: What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem?
	Chapter 6: The case of OpenStreetMap (OSM)	
PART D: (Re)framing the big picture	**Chapter 7**: Discussion and conclusion	

1.5 Scope and limitations

This dissertation studies open data intermediation business models within a particular theoretical understanding of a (sustainable) open data ecosystem (ODE) as previously described. Such understanding is largely inspired by the actor-network theory (ANT) and guided by the conceptualisation of a sustainable ODE by van Loenen et al. (2021). Dealing with loaded terms and a rich theory, some limitations of this dissertation are thus specified:

On sustainability: 'Sustainability' is a value-laden term. In both academic and public discussions, 'sustainability' is often spoken in terms of the environmental dimension. While the environmental impacts of data production and storage warrant attention (Jarvenpaa & Essén, 2023; Lucivero, 2020), they fall outside the scope of this dissertation. Instead, this dissertation focuses on the creation and maintenance of ODE actors' relationships, following a similar line of inquiry in ANT about how networks become and stay stable over time. Investigating the environmental aspect of the ODE would necessitate a different research trajectory and collecting different types of data, such as on land, water, and energy use.

On ecosystem: According to Krivý (2023), by emphasising the complex and self-organising processes, the 'ecosystem' metaphor, as it is used in various domains ranging from economics to information systems, prioritises adaptation instead of challenging the status quo based on the underlying assumption of the immutability of the natural order. In the context of this dissertation, where the study is centred around the notion of *sustainability* of the ecosystem, one may argue that the mental model of this dissertation is precisely of an adaptation, to ensure long-lasting relationships of open data actors, as opposed to challenging them. However, while the line of argument of Krivý (2023) deserves further investigation, the ecosystem metaphor in this dissertation is used to trace out the relationships among different open data actors while recognising the agency of each actor, and thus, can also be leveraged to disrupt the status quo. In other words, this dissertation does not presuppose that an ecosystem cannot or should not be challenged; however, the interdependency among different (concurrently autonomous) actors means that an actor would likely have to work with other actors to achieve certain objectives.

On ANT: ANT is useful in theoretically describing the ODE and understanding the role of open data intermediaries within the ecosystem. It provides a potent initial frame of reference as previously shown. However, while there are other key concepts in ANT, such as translation, immutable mobile, inscription, and black box (Walsham, 1997), they are not applied in this dissertation, which focuses on the business model development aspects. Furthermore, shortcomings of ANT are also recognised, including the lack of an explanatory framework for causality, the tendency to overlook expressions of power, the inability to account for social structures, and the attribution of agency to non-humans (Tummons, 2021). However, since the application of ANT in this dissertation is limited to foregrounding the diverse open data-related resources, roles, actors, and relationships (i.e., as a lens) for the purpose of studying open data intermediation business models, without underplaying the common criticisms of ANT, they are, however, not directly relevant in this dissertation's context.

PART A

Understanding open data intermediaries

The globe

Chapters 2 and 3 of the dissertation focus on understanding what constitutes open data intermediaries and their positions and roles in the ODE. Against the diverse, sometimes obscure, interpretations of open data intermediaries, Chapter 2 offers a common definition as a starting point, facilitating further research based on a consistent understanding of what they are. This was achieved through a systematic literature review whose goals are not only to compile existing definitions already put forth, but also to interrogate the types of actors, tasks, and objectives of open data intermediaries discussed in the literature.

Next, to situate open data intermediaries within the broader ODE, Chapter 3 studies how they can potentially contribute to addressing a myriad of challenges in the ODE. Moving from only engaging with literature in Chapter 2 to have a sense of what open data intermediaries are, Chapter 3 involves interviews with practitioners and asking more specific questions around what the contributions of open data intermediaries for other actors in the ODE may be. Findings from Chapter 3 reveal that open data intermediaries are not merely a bridge or a middleperson between open data providers and end-users but much more.

2 Defining open data intermediaries

The core content of this chapter was previously published as Shaharudin, A., van Loenen, B., & Janssen, M. (2023). Towards a Common Definition of Open Data Intermediaries. *Digital Government: Research and Practice*, *4*(2), 6:1-6:21. https://doi.org/10.1145/3585537.

2.1 Introduction

This chapter addresses the RQ1: What are open data intermediaries? Various definitions of open data intermediaries can be found in the literature. Several definitions differ significantly; for example, definitions by Janssen & Zuiderwijk (2014) and Johnson & Greene (2017) regard open data intermediaries as those that are actively involved in the processing of open data, whereas the definition by Robinson & Mather (2017) considers them as those that connect community members with open data and do not necessarily process the data themselves. These definitions can benefit from harmonisation to ensure a shared understanding of open data intermediaries among researchers and practitioners. The absence of a common definition may lead to a divergence of understanding of what open data intermediaries constitute. Therefore, this chapter reviews the existing definitions in the literature and synthesises them. No study has previously compiled the various definitions of open data intermediaries and harmonised them into a common definition. Hence, this chapter fills the said gap.

The organisation of this chapter is as follows. In Section 2.2, the research methodology is presented, including data gathering and analysis methods. In Section 2.3, the findings from the first round of data gathering are presented, of which the goal is to gather all definitions of open data intermediaries in the literature. The definitions gathered were broken down into basic components, and another round of data gathering and analysis was conducted to substantiate and ascertain the basic components identified. The findings and analysis from the second round of data gathering and analysis are presented in Section 2.4. Section 2.5 develops a common definition of open data intermediaries based on the findings from the two rounds of data gathering and analysis. Section 2.6 discusses the definition proposed. Section 2.7 elaborates the contributions of this chapter.

2.2 Research methodology

2.2.1 Overview

A systematic literature review (SLR) was employed as it is a robust method to gather what the existing literature says about a topic (Okoli & Schabram, 2010). It is more rigorous than a non-systematic literature review, as it involves following clearly defined protocols and transparent reporting, which allows replications (Thomé et al., 2016).

The eight steps of the SLR process by Xiao & Watson (2019) was followed. First, the problem to addressed from the SLR was formulated. In this case, the question to be addressed is, What is the definition of open data intermediaries in the literature? Second, the review protocol was developed and validated. Third, the literature was searched. Fourth, the literature was screened for inclusion by reviewing the title and abstract. Fifth, the quality of each publication was assessed by reviewing the full text. Sixth, relevant data from the literature was extracted. Seventh, the data was analysed. Lastly, the findings were reported. The abovementioned steps are detailed in the subsequent sections.

2.2.2 Literature search

Relevant publications were searched in four databases, namely Scopus, Web of Science (WoS), Google Scholar, and Open Access Theses and Dissertations (OATD). Publications up to June 1, 2022 are included in the search (no start year was set)[1]. The search terms used were 'open data intermediaries', 'open data intermediary', 'open data intermediation', 'open data infomediaries', 'open data infomediary', 'open government data intermediaries', 'open government data intermediary', 'open government data intermediation', 'open government data infomediaries', and 'open government data infomediary'. The terms 'infomediaries' and 'infomediary' were included in the searches as the initial literature scanning shows that they are occasionally used as a synonym to or a type of open data intermediary. In addition, the term 'intermediation' was also included to capture literature that uses the said term instead of 'intermediary'; while linguistically, the former is a participle whereas the latter is a subject, both terms would point to the equally relevant literature as far as this study's goal is concerned. Although the scope of this dissertation is not limited to open government data but a broad range of open data (including open non-government data), the term 'open government data' was also included in the search terms since the academic sub-area of open government data has gained tremendous interest in the past, resulting in much literature in this area.

The search strategy for each database, including the search query and the number of publications found, is shown in Appendix A. Note that Google Scholar only allows terms searched either in the title or in the whole publication. Because the latter gives an unmanageable number of publications, which was about 16,900 publications, only searches based on the title was done for Google Scholar, whereas in the title and abstract for the other databases.

1 Although the research for this chapter was conducted in 2022, hence, only the literature published by mid-June 2022 was included, the findings, particularly the common definition of open data intermediaries proposed remain relevant. This can be attributed to the broad nature of the definition, serving as an umbrella definition applicable to diverse kinds of open data intermediaries. The relevance of the definition was reassured based on the in-depth case studies conducted in Chapters 5 and 6. Moreover, since the publication of this chapter in Shaharudin et al. (2023), several research papers have adopted the proposed definition.

2.2.3 Literature filtering

In total, there were 176 publications compiled from the four databases searched. Duplicated publications, publications with no authors' information, and inaccessible publications were removed in the first filtering stage, resulting in 101 publications. Next, irrelevant publications (publications that are not about open data) based on the title and abstract and publications in a non-English language were removed, leaving 59 publications. Based on the content of each publication, nine of them were found irrelevant to the study's objective: seven publications do not describe anything informational about open data intermediaries except referring to them in passing, one publication is an engineering article about a novel method to integrate information from multiple systems, and one publication is a two-page conference paper with the research method vaguely described and the findings section of only one paragraph. In the end, 50 publications were selected[2]. Figure 2.1 visualises the filtering stages.

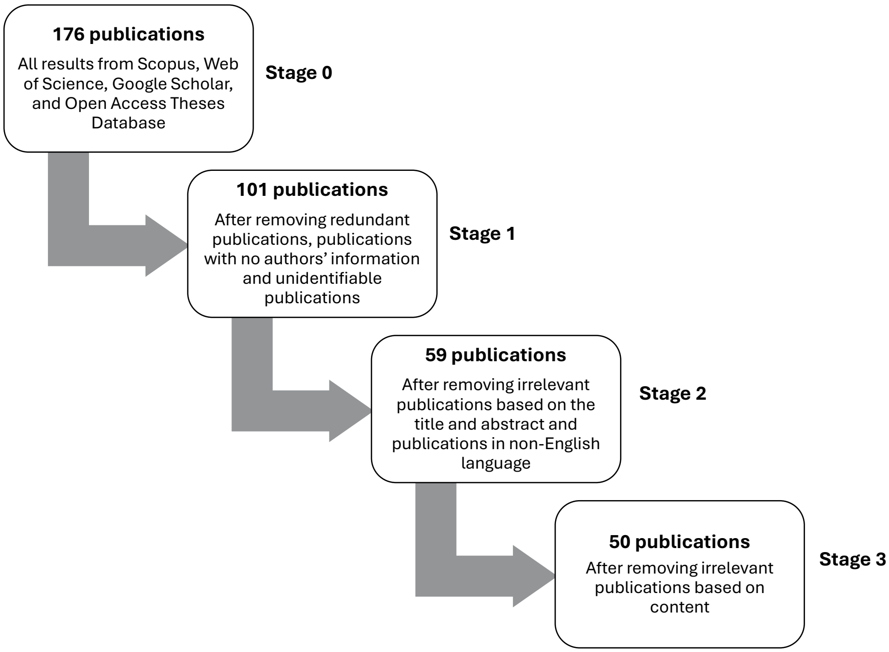


FIG. 2.1 Filtering stages of literature

2 Note that van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) and van Schalkwyk, Cañares, et al. (2016) were counted as one since the former, which is a non-academic report, is republished as the latter in an academic journal.

Most of the publications were journal articles (23) , 13 were conference papers, four were book chapters, seven were dissertations, two were reports, and one was a working paper. Almost all of those publications employed qualitative methods except one that used a quantitative method and four that employed a combination of qualitative and quantitative methods. The earliest publication was from 2011 and the largest share of publications in the pool (12) was from 2017.

2.2.4 Data gathering and analysis

Inductive coding was conducted to gather data from the literature. Inductive coding allows 'research findings to emerge from frequent, dominant or significant themes inherent in raw data' (Thomas, 2006). With inductive coding, the process started with limited preconceived ideas about open data intermediaries. On the contrary, coding was done based on what was written in the literature. There are five key features of inductive coding (Thomas, 2006), namely, (i) the code that is tagged to the raw text, (ii) the code description, (iii) text or data associated with the code, (iv) links between codes (in this case, the links across different publications were studied to see how they cross-reference each other) and, (v) the type of model in which the code is embedded (in this case, the research method and context of each publication were recorded).

To develop a common definition of open data intermediaries, two rounds of data gathering and analysis were conducted. In the first round, existing definitions were gathered from the literature pool. These definitions were analysed by breaking them down into basic components. Next, relevant data categories that capture each of the basic components was identified and a second round of data gathering and analysis were conducted based on the same literature pool. The goal of this second round was to substantiate and ascertain the best description for each of the basic components of the definitions from the first round. Finally, the most appropriate description for each of the basic components was stitched together to produce a common definition of open data intermediaries.

2.3 Findings: Definitions of open data intermediaries in the literature

2.3.1 Compilation of definitions

There were 12 definitions of open data intermediaries found in the literature surveyed. Appendix B shows the list of the definitions, their source, and publications that adopted or were inspired by the respective definition. 'Adopt' here means that the publication followed entirely the definition provided by the source publication, whereas 'inspired by' means that the publication built on the definition in the source publication in order to propose a new definition.

Although all of the 50 publications reviewed discussed open data intermediaries, it was not simply assumed that every publication attempted to define open data intermediaries based on what the publication associated to open data intermediaries. This was to avoid misrepresenting the viewpoints of the authors by taking their texts out of context. Therefore, unless the publication explicitly mentions something along the line of 'open data intermediaries are …' or 'open data intermediaries are defined as …', the publication was not assumed to define open data intermediaries.

Based on the literature pool, the first attempt to define open data intermediaries was made by Chattapadhyay (2014, p. 362), who defined them as 'organisations that share data for its access, consumption and re-usage (including re-sharing) by other organisations and individuals'. The author further clarified three points, namely, (i) 'sharing of open data by such organisations can either be done on a commercial or a non-commercial basis'; (i) 'shared data can either be primary (collected by the organisation concerned) or secondary (sourced from an external creator) in nature'; and (iii) 'the data intermediary organisation may or may not add value to the data before sharing it further' (Chattapadhyay, 2014, p. 362). Enaholo (2017, p. 96) built on Chattapadhyay (2014) to define open data intermediaries as 'those who operate within the open data ecosystem by means of their contribution, in one way or the other, to the supply of open data by governments as well as to the demand for such data by citizens', which goes beyond sharing open data as defined by Chattapadhyay (2014).

In 2015, da Silva Craveiro & Albano (2015, p. 226) defined open data intermediaries as 'all the players (in an individual way or representatives of governments and social organisations), who are involved with public data that are released in an open format. They may or may not make use of technological, legal or structural artifacts in their activities. In making use of open data, the intermediaries aggregate value to the data to ensure that they can be understood more easily (and hence have a greater value) [by] third parties after their intervention'. Meanwhile, González-Zapata & Heeks (2015, p. 4) defined open government data intermediaries as 'all actors that assist OGD [open government data] initiatives by bridging the barriers that separate public sector data producers and civil society data consumers'. They emphasised that open government data intermediaries have a two-way relationship, with the government on the supply side and with the civil society on the demand side.

Van Schalkwyk et al. (2015, p. 7) defined an open data intermediary as 'an agent (i) positioned at some point in a data supply chain that incorporates an open dataset, (ii) positioned between two agents in the supply chain, and (iii) facilitates the use of open data that may otherwise not have been the case'. Van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) noted that an open data intermediary may neither supply nor access open data but facilitates the flow of data. To distinguish open data intermediaries from internet intermediaries such as internet service providers, van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) emphasised the degree of 'agency' of actors in fulfilling the function of intermediating open. In this regard, according to them, internet service providers are not considered open data intermediaries as they do not execute a high degree of involvement in intermediating open data. Note that in the following year, van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) was republished as van Schalkwyk, Cañares, et al. (2016); the former is a report of a project, while the latter is an article in an academic journal.

In the literature pool, six publications adopted the definition offered by van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) or van Schalkwyk, Cañares, et al. (2016): Andrason & van Schalkwyk (2017), da Silva Craveiro & Albano (2017), den Haan (2018), Enaholo & Dina (2020), Maail (2017), and Yoon et al. (2018). Interestingly, in the same year van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) was published, da Silva Craveiro & Albano (2015) came up with their own definition of open data intermediaries but later, in da Silva Craveiro & Albano (2017), they adopted the definition by van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) instead of reiterating their own.

According to Brugger et al. (2016, p. 222), open government data intermediaries are ‘actors who bridge gaps between data producers (governments) and data users (civil society) in that they supply essential resources and capabilities necessary to turn government data into development actions and results’. Meng (2016, p. xi) defined them as ‘actor[s] that bridge the gap between marginalized groups and OGD [open government data] by facilitating physical access, technical capacity, and value for use of information’ whereas Schrock & Shaffer (2017, p. 2) as ‘actors that translate, use, or otherwise mediate communication using data produced by or for government’. Meanwhile, Balvert & van Maanen (2019, p. 133) defined them as ‘the in-between actor standing between a government and a citizen in the process of data communication’.

A term that is occasionally used as a synonym to or a subgroup of open data intermediaries is infomediaries. Janssen & Zuiderwijk (2014, p. 695) considered infomediaries as those involved in ‘the handling of information between information providers and consumers’. This definition was adopted by Sangiambut & Sieber (2017). Johnson & Greene (2017, p. 10) defined infomediaries as ‘specific categories of open data users who extract, aggregate, and transform data, altering it into a format that is seen as valuable, beneficial, and, most importantly, usable to the general public’. Gao & Janssen (2022) adopted the definition of infomediaries by Johnson & Greene (2017). Meanwhile, Robinson & Mather (2017, p. 31) defined a *civic* infomediary as ‘a person or organisation that connects community members with open data so that public value can be derived from the data’.

Based on the compilation, it can be observed that some definitions are rather different from each other and thus may result in conceptual confusion about open data intermediaries. For example, while several definitions highlight open data intermediaries’ function in the use of open data (da Silva Craveiro & Albano, 2015; Johnson & Greene, 2017; van Schalkwyk, Chattapadhyay, Cañares, et al., 2015), the definition by Enaholo (2017) emphasises their function in the supply *and* use. Furthermore, half of the definitions compiled are specifically for open government data intermediaries (Balvert & van Maanen, 2019; Brugger et al., 2016; Enaholo, 2017; González-Zapata & Heeks, 2015; Meng, 2016; Schrock & Shaffer, 2017), thereby limiting their applicability to broader open data intermediaries. Meanwhile, two from the other six definitions (Johnson & Greene, 2017; Robinson & Mather, 2017) are limited to the context of public value generation (thus, not inclusive to private value generation).

2.3.2 Breakdown of the definitions

Inspired by the 5W1H questions method (what, who, where, when, why, and how), derived from the *Septem Circumstantiae* (elements of circumstances) from the field of philosophy (Sloan, 2010), the elements in the 12 definitions gathered were categorised into the who, what, where, and why, that are called basic components (Table 2.1). Specifically,

- The who: Who are the actors of open data intermediaries?
- The what: What do open data intermediaries do?
- The where: Where are open data intermediaries located in the open data lifecycle?
- The why: Why are open data intermediaries needed?

For *the where*, the open data lifecycle model introduced by van Veenstra & van den Broek (2015) was referred. Open data lifecycle is ‘a conceptualisation of the process and practices around handling data, starting from its creation, through the provision of open data to its use by various parties’ (Charalabidis et al., 2018, p. 12). While there are several open data lifecycle models in the literature (Attard et al., 2016; Charalabidis, Zuiderwijk, Alexopoulos, Janssen, Höchtl, et al., 2018; Sutherland & Cook, 2017), the model by van Veenstra & van den Broek (2015) was chosen because it concisely integrates the activities of both data providers and users in one lifecycle, instead of separate lifecycles. There are five stages in the open data lifecycle model by van Veenstra & van den Broek (2015), namely (i) identification: setting the open data strategy and selecting the data; (ii) preparation: setting requirements for data publication, modelling and describing data, converting data to a machine-readable format, linking data, and storing data; (iii) publication: publication of data and metadata; (iv) re-use: exploiting published data; and (v) evaluation: assessing the value of open data, and monitoring and improving data.

Naturally, based on the 5W1H, one may ask, do the definitions not describe *the when* and *the how*? *The when*, which one may likely put as ‘When do open data intermediaries carry out their tasks?’ is similar to *the where*, which is, ‘Where are open data intermediaries located in the open data lifecycle?’ Meanwhile, from the definitions compiled, it is rather difficult to differentiate *the how*, which one may likely put as ‘How do open data intermediaries do what they do?’ from *the what* which is, ‘What do open data intermediaries do?’. For the said reasons, in this study, *the when* is considered equivalent to *the where,* and *the how* to *the what*.

TABLE 2.1 The breakdown of open data intermediaries' definitions gathered from the literature

No.	Source	The who	The what	The where	The why
1.	Chattapadhyay (2014)	organisations	share data for its access, consumption and re-usage (including re-sharing) by other organisations and individuals		
2.	Janssen & Zuiderwijk (2014)		the handling of information	between information providers and consumers	
3.	da Silva Craveiro & Albano (2015)	all the players (in an individual way or representatives of governments and social organisations), who are involved with public data that are released in an open format	they [i.e., open data intermediaries] may or may not make use of technological, legal or structural artifacts in their activities		to ensure that they [i.e., data] can be understood more easily (and hence have a greater value) [by] third parties after their intervention
			in making use of open data, the intermediaries aggregate value to the data		
4.	González-Zapata & Heeks (2015)	all actors	assist OGD [open government data] initiatives		bridging the barriers that separate public sector data producers and civil society data consumers
5.	van Schalkwyk, Cañares, et al. (2016) and van Schalkwyk, Chattapadhyay, Cañares, et al. (2015)	an agent		positioned at some point in a data supply chain that incorporates an open dataset	facilitates the use of open data that may otherwise not have been the case
				positioned between two agents in the supply chain	
6.	Brugger et al. (2016)	actors	bridge gaps	between data producers (governments) and data users (civil society)	to turn government data into development actions and results
			they supply essential resources and capabilities necessary		
7.	Meng (2016)	an actor	bridges the gap by facilitating physical access, technical capacity, and value for use of information	between marginalized groups and OGD [open government data]	

>>>

TOC

TABLE 2.1 The breakdown of open data intermediaries' definitions gathered from the literature

No.	Source	The who	The what	The where	The why
8.	Schrock & Shaffer (2017)	extra-institutional actors	translate, use, or otherwise mediate communication using data produced by or for government		
9.	Enaholo (2017)	those who operate within the open data ecosystem			by means of their contribution, in one way or the other, to the supply of open data by governments as well as to the demand for such data by citizens
10.	Johnson & Greene (2017)	specific categories of open data users	extract, aggregate, and transform data		altering it [i.e., data] into a format that is seen as valuable, beneficial, and, most important, usable to the general public
11.	Robinson & Mather (2017)	a person or organisation	connects community members with open data		so that public value can be derived from the data
12.	Balvert & van Maanen (2019)	the in-between actor		standing between a government and a citizen in the process of data communication	

Note: Definitions are taken in verbatim from the source, but are arranged based on the four basic components.

2.4 Findings: Basic components of the definitions compiled

2.4.1 Overview

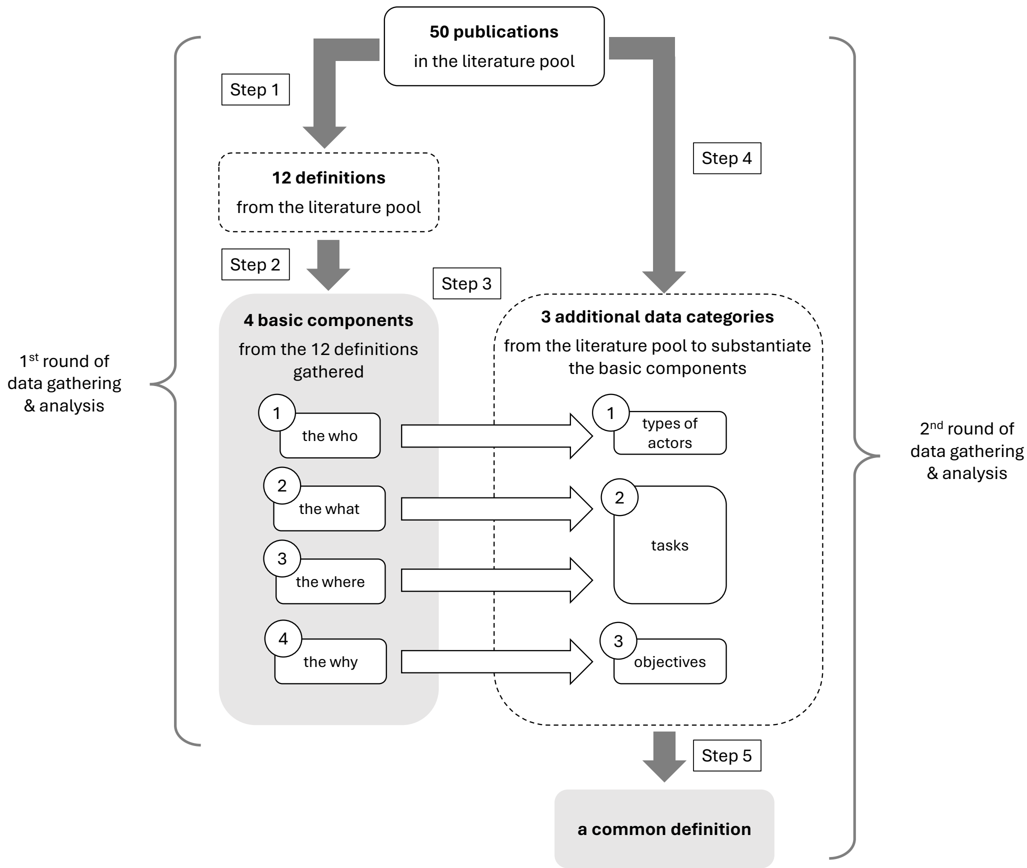


FIG. 2.2 The development of a common definition based on two rounds of data gathering and analysis

TOC

As the goal of this chapter is to develop a *common* definition of open data intermediaries, the basic components of the definitions (i.e., *the who*, *the what*, *the where*, and *the why*) were studied further in order to substantiate and ascertain the best description for each component and develop a definition of open data intermediaries that is encompassing.

A second round of data gathering and analysis was conducted from the same literature pool. *The who* was captured by looking into the types of actors of open data intermediaries, *the what* and *the where* were captured by looking at their tasks, and *the why* was captured by looking at the objectives of open data intermediaries. In short, three more data categories (types of actors, tasks, objectives) were gathered from the 50 publications in the literature pool to substantiate the four basic components of open data intermediaries' definitions. Figure 2.2 visualises the steps.

The following subsections describe (i) the types of actors of open data intermediaries (Section 2.4.2) to answer *the who*, (ii) the tasks of open data intermediaries (Section 2.4.3) to answer *the what* and *the where*, and (iii) the objectives of open data intermediaries (Section 2.4.4) to answer *the why*, based on the second round of systematic literature review.

2.4.2 Types of actors: the who

Compilation of types of actors

Six types of open data intermediary actors were identified in the literature (Table 2.2). One of them are civil society organisations (CSOs). Several publications described some civic application or platform developers as open data intermediary CSOs (Germano et al., 2016; González-Zapata & Heeks, 2015; Johnson & Greene, 2017; Kassen, 2018; Meijer & Potjer, 2018; Meng et al., 2019; Sangiambut & Sieber, 2017), whereas another group of publications identified some advocacy groups as open data intermediary CSOs (Brugger et al., 2016; Cañares, 2014; Enaholo, 2017; Enaholo & Dina, 2020).

TABLE 2.2 Types of actors of open data intermediaries in the literature

No.	Type of actor	Description	Sources
1.	**Civil society organisations**	Non-profit and non-governmental organisations that offer services based on open data or advocate for open data	Mayer-Schönberger & Zappia (2011), Cañares, (2014), González-Zapata & Heeks (2015), Brugger et al. (2016), Germano et al. (2016), Meng (2016), Enaholo (2017), Glassey (2017), Johnson & Greene (2017), Sangiambut & Sieber (2017), Kassen (2018), Meijer & Potjer (2018), Hablé (2019), Meng et al. (2019), Enaholo & Dina (2020), Žuffová (2020), Navalkha (2021), Reggi & Dawes (2022)
2.	**Entrepreneurs/ businesses**	For-profit entrepreneurs and companies that provide products and services based on open data	Cañares (2014), Janssen & Zuiderwijk (2014), Germano et al. (2016), Andrason & van Schalkwyk (2017), Glassey (2017), Johnson & Greene (2017), Sangiambut & Sieber (2017), Corbett et al. (2018), Kassen (2018), Hablé (2019)
3.	**The media**	Media organisations and individual journalists who use open data to produce news stories	Cañares (2014), Baack (2015), Brugger et al. (2016), Meng (2016), Johnson & Greene (2017), Corbett et al. (2018), Hablé (2019), Meng et al. (2019), Enaholo & Dina (2020), Žuffová (2020), Reggi & Dawes (2022)
4.	**Public organisations**	Public organisations, including public libraries and public research institutions, that help others make use of open data	Janssen & Zuiderwijk (2014), Chan et al. (2016), Johnson & Greene (2017), Robinson & Mather (2017), D. Kim (2018), Meijer & Potjer (2018), Hablé (2019), Reggi & Dawes (2022)
5.	**Researchers**	Researchers in universities or research institutions who use and distribute open data in research activities	Meng (2016), Johnson & Greene (2017), S. Park & Gil-Garcia (2017), Corbett et al. (2018), D. Kim (2018), Meijer & Potjer (2018), Hablé (2019), Meng et al. (2019), Navalkha (2021)
6.	**Multi-partner**	Collaborative organisations of different types, such as public-private partnerships that facilitate the supply and use of open data	Hielkema & Hongisto (2013), Meijer & Potjer (2018)

Entrepreneurs and businesses are another type of open data intermediary actors identified in the literature. They use open data to offer web or mobile applications (Germano et al., 2016; Kassen, 2018; Sangiambut & Sieber, 2017) and (advisory) services (Andrason & van Schalkwyk, 2017). Several publications described media as another type of open data intermediaries. In some, they are called data journalists (Enaholo & Dina, 2020; Johnson & Greene, 2017). The resources and competencies provided by data journalists that act as open data intermediaries mainly involve transforming open data into digestible information in the form of news stories (Enaholo & Dina, 2020).

Another type of open data intermediary actors described in the literature are public organisations, including government organisations (Chan et al., 2016; Johnson & Greene, 2017; Meijer & Potjer, 2018; Reggi & Dawes, 2022), public libraries (Robinson & Mather, 2017), and public research organisations (Meijer & Potjer, 2018). Several publications found that some of them perform the role of both open data providers and intermediaries (Johnson & Greene, 2017; Reggi & Dawes, 2022). Other types of open data intermediary actors are researchers (Corbett et al., 2018; Johnson & Greene, 2017; Meijer & Potjer, 2018; Meng et al., 2019; S. Park & Gil-Garcia, 2017) and multi-partner open data intermediaries such as public-private partnership organisations (Meijer & Potjer, 2018) and living labs formed by universities, development agencies, and private companies (Hielkema & Hongisto, 2013).

The who: Who are the actors of open data intermediaries?

Based on the literature, there are various types of open data intermediary actors. They are not necessarily organisations as some of them are individuals, such as entrepreneurs, individual journalists, and researchers. There are also multi-partner open data intermediaries. Although they are mostly users of open data, the literature also identified open data intermediaries among those that advocate for open data or facilitate its supply and use.

Therefore, to capture the multifaceted types of open data intermediary actors, several literature (González-Zapata & Heeks, 2015; Robinson & Mather, 2017; van Schalkwyk, Chattapadhyay, Cañares, et al., 2015) refer to *the who* in their definitions generically as ‘actors’, ‘an agent’, or ‘a person or organisation’. Nevertheless, based on the analysis, open data intermediaries mostly take care of the interest of other open data actors, distinguishing them from solely open data providers or end-users. For example, open data intermediaries process open data to deliver products and services benefitting other open data actors, and not only for their own internal use. Similarly, open data intermediaries that facilitate access to open data, support users to access open data, but they are not the original open data providers themselves.

Accordingly, *the who* in defining open data intermediaries can be more accurately described as ‘third-party actors’. Based on Oxford Dictionary, a third party is ‘a person who is involved in a situation in addition to the two main people involved’ (Oxford University Press, n.d.).

2.4.3 Tasks of open data intermediaries: the what and the where

Compilation of tasks

Fourteen tasks of open data intermediaries were gathered from the literature (Table 2.3). Typically, multiple tasks are carried out simultaneously by an open data intermediary. One of the most popular tasks of open data intermediaries is compiling data from various sources. They then publish the data on their platform and/or use the data to offer products and services to users. For example, Aclímate Colombia compiled data on commercial crops, station-level daily weather data, and data related to crop yield from various open data sources and made it all conveniently downloadable through its one-stop platform (Young & Verhulst, 2017).

Another task of open data intermediaries is building data capacity of data users (e.g., citizens, community organisations, journalists) as well as data providers. For example, BudgIT, a civil society organisation (CSO) in Nigeria, organised training sessions and workshops for journalists and individuals to engage with open data (Enaholo, 2017). Others organised open data-related events (da Silva Craveiro & Albano, 2017; Robinson & Mather, 2017) or worked closely with open data providers to improve the process of publishing open data (Meng et al., 2019).

Some open data intermediaries also augmented data by combining open data (from multiple sources) with non-open data. For example, Farmerline, a company based in Accra, Ghana, combined open meteorological data sourced from the Ministry of Food and Agriculture with the data that the company collected (e.g., weekly market prices) and purchased from non-open sources (e.g., weather forecast data) in order to provide advisory services to farmers (Andrason & van Schalkwyk, 2017). Another example is Geonext, a geodetic engineering company that combined its own data with open data to offer data services to clients (den Haan, 2018).

Contextualising data is another task of open data intermediaries described in the literature, which involves adding relevant and specific context to the data for it to be relatable and meaningful to a targeted audience. In the literature, this task is often carried out by journalists and CSOs. For example, CSOs in Argentina, Mexico, and Uruguay that championed budget transparency, translated budget data into spending stories for the data to be easily understood by the public (da Silva Craveiro & Albano, 2017). Meanwhile, some journalists used open data for their news reports (Brugger et al., 2016; Enaholo & Dina, 2020). Another task of open data intermediaries is curating data. The vast amount of open datasets that potential users may need to sift through may deter some from using them (Yoon et al., 2018). Hence, some open data intermediaries, for example, curated datasets based on specific geographic areas (Yoon et al., 2018) or communities' needs (Chan et al., 2016).

Several open data intermediaries developed products and services, typically in the form of web and mobile applications. For example, an open data intermediary in the city of Edmonton, Canada combined the city's open data and data from other sources to develop an application that simplifies the process of finding a home (Corbett et al., 2018). Another example is a company named IntellinQ that built a spatial database management software to help users to access open datasets from various sources and reorganise them (den Haan, 2018). There were also CSOs that develop data tools to facilitate community organisations to aggregate and integrate data that is of their interest (Yoon et al., 2018). Some open data intermediaries offered advisory services, for example, Farmerline and Esoko, which used open data to provide recommendations to farmers in Ghana (Andrason & van Schalkwyk, 2017).

Some open data intermediaries interpreted data to transform it into accessible information for their audience. For example, some CSOs interpret complex data into more easily understandable information on socio-economic topics such as health and education as well as using the interpreted data for advocacy work (Enaholo, 2017). A related task to interpreting data is visualising data of which open data is represented in charts, maps, and other visual forms (Dumpawar, 2015; Enaholo & Dina, 2020; Meng, 2016). Some open data intermediaries were involved in validating the quality of open data, which addresses the 'inaccuracy, incompleteness and obsolescence' of data (Dumpawar, 2015). Some open data intermediaries also demanded open data. This task includes identifying specific datasets that should be made open based on local needs (González-Zapata & Heeks, 2015; Corbett et al., 2018) and advocating for open data policies to be adopted (Enaholo & Dina, 2020).

Other tasks of open data intermediaries found in the literature are facilitating actors' interactions either through direct networking or open data events (Chan et al., 2016; den Haan, 2018; Dumpawar, 2015; Juell-Skielse et al., 2014), channelling feedback regarding data or issues identified based on data to the relevant actors (Chan et al., 2016; den Haan, 2018; Enaholo, 2017), improving the technical openness of data such as by transforming data into a machine-readable format (den Haan, 2018; Maail, 2017; Meng et al., 2019; Navalkha, 2021), and assisting data providers in identifying potential risks in publishing certain datasets as open data (Davies & Edwards, 2012).

TABLE 2.3 Tasks of open data intermediaries in the literature

No.	Task	Description	Source
1.	**Compile data**	Collect open data from multiple sources	Dumpawar (2015), González-Zapata & Heeks (2015), van Schalkwyk, Chattapadhyay, Cañares, et al. (2015), Chan et al. (2016), Andrason & van Schalkwyk (2017), Young & Verhulst (2017), Corbett et al. (2018), den Haan (2018), D. Kim (2018), Meijer & Potjer (2018), Yoon et al. (2018), Hablé (2019), Meng et al. (2019), Enaholo & Dina (2020), Žuffová (2020), Navalkha (2021), Reggi & Dawes (2022)
2.	**Build data capacity**	Organise training sessions, workshops, hackathons, and other open data-related events and engage with actors to improve open data practices	Davies & Edwards (2012), da Silva Craveiro & Albano (2017), Enaholo (2017), Maail (2017), Robinson & Mather (2017), Yoon et al. (2018), Meng et al. (2019), Reggi & Dawes (2022)
3.	**Augment data**	Integrate open data from multiple sources or with non-open data	Davies & Edwards (2012), Dumpawar (2015), Andrason & van Schalkwyk (2017), Young & Verhulst (2017), Corbett et al. (2018), den Haan (2018), Yoon et al. (2018), Gao & Janssen (2022)
4.	**Contextualise data**	Add relevant and specific context to the data to enhance its relatability and meaning for the targeted audience	Dumpawar (2015), González-Zapata & Heeks (2015), Brugger et al. (2016), Meng (2016), da Silva Craveiro & Albano (2017), den Haan (2018), Enaholo & Dina (2020), Navalkha (2021)
5.	**Curate data**	Select and reorganise open data based on its relevance and needs for the targeted audience	Davies & Edwards (2012), Dumpawar (2015), Chan et al. (2016), Andrason & van Schalkwyk (2017), den Haan (2018), Yoon et al. (2018), Gao & Janssen (2022)
6.	**Develop products and services**	Use open data to offer or complement products and services such as web-based and mobile applications and advisory services	González-Zapata & Heeks (2015), Chan et al. (2016), Meng (2016), Andrason & van Schalkwyk (2017), Corbett et al. (2018), den Haan (2018), D. Kim (2018), Yoon et al., (2018), Hablé (2019), Yoon & Copeland (2020), Enaholo & Dina (2020), Gebka et al. (2021), Navalkha (2021)
7.	**Interpret data**	Transform data into more digestible information for the targeted audience	Dumpawar (2015), van Schalkwyk, Chattapadhyay, Cañares, et al. (2015), Meng (2016), Enaholo (2017), Corbett et al. (2018), den Haan (2018), D. Kim (2018), Yoon et al. (2018), Žuffová (2020), Navalkha (2021)
8.	**Validate data**	Check, update, and rectify open data in terms of its accuracy, completeness, and timeliness	Dumpawar (2015), González-Zapata & Heeks (2015), Corbett et al. (2018), den Haan (2018), D. Kim (2018), Enaholo & Dina (2020)

>>>

TABLE 2.3 Tasks of open data intermediaries in the literature

No.	Task	Description	Source
9.	**Demand open data**	Identify datasets that should be made open data or advocate for the general adoption of open data policy	González-Zapata & Heeks (2015), Enaholo (2017), Corbett et al. (2018), D. Kim (2018), Meng et al. (2019), Enaholo & Dina (2020)
10.	**Visualise data**	Represent open data in charts, maps, and other visual forms to improve its understandability for the targeted audience	Dumpawar (2015), Brugger et al. (2016), Meng (2016), Enaholo (2017), den Haan (2018), Enaholo & Dina (2020), Navalkha (2021)
11.	**Facilitate actors' interactions**	Connect open data actors through direct engagements or events	Juell-Skielse et al. (2014), Dumpawar (2015), Chan et al. (2016), Meng (2016), den Haan (2018), D. Kim (2018)
12.	**Channel feedback**	Channel feedback regarding open data quality or other issues identified based on open data to the relevant actors	Chan et al. (2016), Enaholo (2017), den Haan (2018), Hablé (2019), Navalkha (2021)
13.	**Improve technical openness of data**	Enhance the technical openness of data such as by converting it into a machine-readable format	Meng (2016), Maail (2017), den Haan (2018), Meng et al. (2019), Navalkha (2021)
14.	**Identify risks of opening data**	Assist data providers in identifying potential risks in making particular datasets open	Davies & Edwards (2012)

The what: What do open data intermediaries do?

Based on the literature, open data intermediaries carried out a wide range of tasks, deploying various types of resources and capabilities. Most of the tasks involved direct processing of open data such as collecting, augmenting, contextualising, visualising data, and developing products and services with open data. However, some tasks did not necessarily require them to directly process open data, for example, building data capacity, facilitating actors' interactions, and channelling feedback.

In this regard, it is not straightforward to describe *the what* in defining open data intermediaries without potentially excluding certain tasks, an issue several past definitions grappled with. As summarised by van Schalkwyk, Chattapadhyay, Cañares, et al. (2015), open data intermediaries deploy various types of capital, including economic capital (e.g., financial resources), cultural capital (e.g., knowledge of local custom), social capital (e.g., existing networking with other actors), symbolic capital (e.g., well-regarded position in society), and/or technical capital (e.g., data processing skills). Different open data intermediaries offered different types of resources and capabilities according to their specialisation. These resources and capabilities are often not at the disposal of most lay users. Therefore, adapting the words of Brugger et al. (2016), *the what* in defining open data intermediaries can be concisely described as 'provide specialised resources and capabilities'.

The where: Where are open data intermediaries located in the open data lifecycle?

The different tasks of open data intermediaries found in the literature were assigned to the relevant open data lifecycle stage(s) (Table 2.4). It is clear that open data intermediaries performed tasks at various stages of the open data lifecycle. In fact, two tasks, namely facilitating actors' interactions and building data capacity, fall in multiple stages. For this reason, the approach taken by several publications (Chattapadhyay, 2014; da Silva Craveiro & Albano, 2017; Enaholo, 2017; González-Zapata & Heeks, 2015; Johnson & Greene, 2017; Robinson & Mather, 2017; Schrock & Shaffer, 2017) that are agnostic about *the where* in defining open data intermediaries is thus adapted.

TABLE 2.4 Tasks of open data intermediaries based on the open data lifecycle stages

No.	Stage of open data lifecycle	Tasks of open data intermediaries in the literature
1.	**Identification**	demand open data, facilitate actors' interactions, identify risks of opening data
2.	**Preparation**	compile data, build data capacity, augment data
3.	**Publication**	curate data, improve technical openness of data
4.	**Re-use**	build data capacity, contextualize data, develop products and services, interpret data, visualise data, facilitate actors' interactions
5.	**Evaluation**	validate data, channel feedback

2.4.4 Objectives of open data intermediaries: the why

Compilation of objectives

Seven objectives of open data intermediaries were found in the literature (Table 2.5). Many publications described facilitating the use of open data as an objective of open data intermediaries. Barriers to open data use call for interventions by open data intermediaries to make open data more usable and impactful through various ways, such as by directly working with community organisations to understand their data needs (Yoon et al., 2018), building the community's data skills (Yoon et al., 2018), and simplifying complex data into more digestible information (Enaholo & Dina, 2020). Another objective of open data intermediaries highlighted in the literature is to increase the access to open data by end-users. This objective is closely related to the objective of facilitating the use of open data. However, the emphasis of this objective is to advocate for non-open data to be provided as open data in appropriate formats (Enaholo, 2017; Maail, 2017; Thakuriah et al., 2017).

TABLE 2.5 Objectives of open data intermediaries from the literature surveyed

No.	Objective	Description	Sources
1.	**Facilitate the use of open data**	Help overcome the barriers to open data use, including socioeconomic gaps, the lack of awareness, and the lack of data skills by users	Chattapadhyay (2014), Maail (2017), Robinson & Mather (2017), Thakuriah et al. (2017), Yoon et al. (2018), Meng et al. (2019), Yoon & Copeland (2020), Enaholo & Dina (2020), Gebka et al. (2021), Reggi & Dawes (2022)
2.	**Increase the availability of and accessibility to open data**	Advocate for the provision of non-open data as open data and in formats that users need	Chattapadhyay (2014), Baack (2015), Meng (2016), van Schalkwyk, Willmers, et al. (2016), Enaholo (2017), Maail (2017), Thakuriah et al. (2017), D. Kim (2018), Yoon et al. (2018), Meng et al. (2019), Gebka et al. (2021)
3.	**Close the feedback loop**	Channel feedback between open data providers and users (or governments and citizens)	Hielkema & Hongisto (2013), Frank & Waddell (2014), Meng (2016), Enaholo (2017), Maail (2017), Sangiambut & Sieber (2017), Gebka et al. (2021)
4.	**Provide services**	Provide value-added services around open data	Davies & Edwards (2012), Frank & Waddell (2014), Andrason & van Schalkwyk (2017), Glassey (2017), Sangiambut & Sieber (2017), Hablé (2019), Meng et al. (2019), Navalkha (2021)
5.	**Bring actors together**	Connect open data providers, (potential) users and other actors	Mayer-Schönberger & Zappia (2011), Hielkema & Hongisto (2013), Juell-Skielse et al. (2014), Andrason & van Schalkwyk (2017), Maail (2017), Yoon et al. (2018)
6.	**Enhance trust between actors**	Build and maintain trust between diverse open data actors	Andrason & van Schalkwyk (2017), Johnson & Greene (2017), Maail (2017), D. Kim (2018)
7.	**Improve open data practices**	Assist open data providers in publishing open data and performing open data initiatives	Maail (2017), S. Park & Gil-Garcia (2017), Meng et al. (2019)

The literature also described the objective of open data intermediaries of closing the feedback loop between open data providers and users (Gebka et al., 2021; Hielkema & Hongisto, 2013; Sangiambut & Sieber, 2017). There were open data intermediaries created specifically by governments to collect feedback, such as the Toronto Cycling App which was developed by a company commissioned by the City of Toronto. It offered cycling-related information based on Toronto’s open data of which app users have the option to share their cycling trips for the city council to improve the city’s cycling network infrastructure (Sangiambut & Sieber, 2017).

Providing value-added services around open data is also one of the objectives of open data intermediaries gathered from the literature. The services are aimed, for example, to enhance the economic benefits (e.g., to improve farm profit) (Andrason & van Schalkwyk, 2017) and improve day-to-day activities (e.g., guiding travel decisions) (Frank & Waddell, 2014; Sangiambut & Sieber, 2017). The literature also described bringing actors together as another objective of open data intermediaries. Some open data intermediaries aimed to establish relationships between open data providers with potential users (Maail, 2017), actors in an economic market (e.g., farmers and distributors) (Andrason & van Schalkwyk, 2017), and developers that use open data (Hielkema & Hongisto, 2013; Juell-Skielse et al., 2014). Other objectives of open data intermediaries in the literature are to enhance trust between open data actors (Andrason & van Schalkwyk, 2017; Johnson & Greene, 2017; Maail, 2017) and to improve open data practices by assisting (prospective) open data providers in publishing open data and performing open data initiatives (Maail, 2017; Meng et al., 2019; S. Park & Gil-Garcia, 2017).

The why: Why are open data intermediaries needed?

In general, the objectives of open data intermediaries can be grouped into two: (i) to enhance the supply, flow, and/or use of open data; and (ii) to strengthen the relationships among various open data actors. While the first objective involves the active processing of open data by open data intermediaries, the second objective does not necessarily imply so. Open data intermediaries that contribute towards the second objective facilitate relationships building among open data actors. Table 2.6 groups the various objectives of open data intermediaries in the literature into the two general objectives.

TABLE 2.6 Categorisation of the objectives of open data intermediaries

No.	General objective	Objective found in the literature
1.	**Enhance the supply, flow, and/or use of open data**	facilitate use, increase the availability of and accessibility to open data, close the feedback loop, provide services, improve open data practices
2.	**Strengthen the relationships among various open data actors**	bring actors together, enhance trust between actors

Both groups of objectives were described in some ways in several existing definitions (Brugger et al., 2016; da Silva Craveiro & Albano, 2015; Enaholo, 2017; González-Zapata & Heeks, 2015; Johnson & Greene, 2017; Robinson & Mather, 2017; van Schalkwyk, Chattapadhyay, Cañares, et al., 2015). However, the two were not mentioned together in any of the definitions. Thus, this dissertation proposes stating the two together in describing *the why* in defining open data intermediaries.

2.5 Proposing a common definition of open data intermediaries

Ultimately, stitching together *the who*, *the what*, *the where*, and *the why* (Figure 2.3), the following common definition of open data intermediaries is developed: third-party actors that provide specialised resources and capabilities to (i) enhance the supply, flow, and/or use of open data and/or (ii) strengthen the relationships among various open data actors.

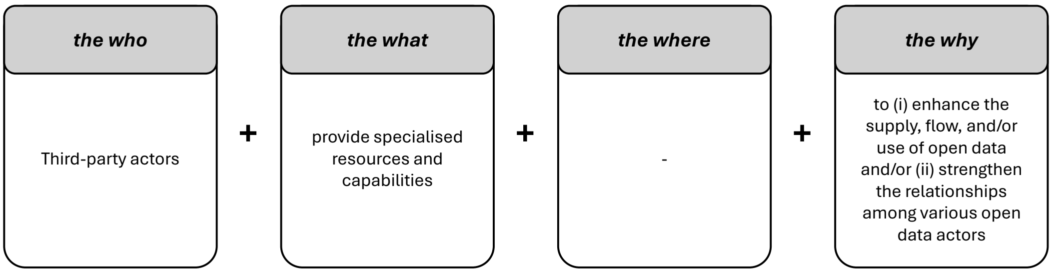


FIG. 2.3 Stitching together the who, the what, the where, and the why for a common definition of open data intermediaries

2.6 Discussion

Definitions have often been the starting point for intellectual inquiry since the time of Socrates (Flowerdew, 1992). According to Kant, as summarised by Beck (1956, p. 180), to define means ‘to present the complete concept of a thing within its limits and in its primary or original character’. According to Edwards (1967) as cited by Flowerdew (1992, p. 204), there are six rules for definitions, namely: (i) ‘a definition should give the essence or nature of the thing defined, rather than its accidental properties’; (ii) ‘a definition should give the genus and differentia of the thing defined’; (iii) ‘one should not define by synonyms’; (iv) ‘a definition should be concise’; (v) ‘one should not define by metaphors’; and (vi) ‘one should not define by negative terms or by correlative terms’. LeMay et al. (2017) noted that a definition ‘should include all of those things that fall under it and exclude all of those things that do not’.

The proposed common definition was assessed according to these six rules. The proposed definition describes the essence of open data intermediaries and not their accidental properties (rule i). Rule (ii) is not applicable in this research context because the definition proposed is not an *intensional* definition (i.e., a definition that specifies the necessary and sufficient conditions for a thing) (Cook, 2009). Instead, it is a theoretical definition, which is a definition that 'function[s] as proposals to see or interpret some phenomenon in a certain way' and 'since proposals have no truth value, neither do theoretical definitions' (Hurley, 2014).

The proposed definition was not defined using synonyms (rule iii). The conciseness of the definition was also thoughtfully considered (rule iv). In particular, although the approach by van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) who described *the where* as 'positioned at some point in a data supply chain' may also be considered, it is not the most concise approach for a definition because the term 'data supply chain' itself may need to be further explained. Likewise, if 'open data lifecycle' were to be included in the proposed common definition, one may ask which model of open data lifecycle is referred to. Hence, the proposed definition silences *the where* since doing so does not change the essence of the definition whereas the alternative may raise more questions and render the proposed definition to be less succinct.

The proposed definition does not contain metaphors (rule v), negative terms or correlative terms (rule vi). Overall, the proposed definition abides by the rules of Edwards (1967). The definition is also inclusive enough while excluding actors that do not play the function of open data intermediation, thus aligning with LeMay et al. (2017).

Apart from being more succinct than the previously widely adopted definition by van Schalkwyk, Chattapadhyay, Caňares, et al. (2015), the proposed definition is also more broadly applicable to diverse open data intermediaries, as it accounts for their wider range of objectives beyond merely 'facilitat[ing] the use of open data that may otherwise have not been the case', as stated in the former. Furthermore, the proposed definition is not limited to open government data intermediaries or to the generation of public value.

There may be studies that looked into open data actors performing intermediation functions but did not label these actors as open data intermediaries. While this implies a limitation in the literature search of this chapter, it also reaffirms the need for a common definition.

2.7 Conclusion

This chapter proposes the following common definition of open data intermediaries: third-party actors who provide specialised resources and capabilities to (i) enhance the supply, flow, and/or use of open data and/or (ii) strengthen the relationships among various open data actors. This common definition can be used by researchers and practitioners to mutually identify open data intermediaries and build knowledge about them on top of a mutual understanding of constitute open data intermediaries.

In the process of developing the definition, various types of actors, tasks, and objectives of open data intermediaries were compiled, which clarified them further. This compilation shows the diversity of open data intermediaries, which matters when designing related policies or business models. Besides, it can also support ODE actors to account for ‘Who is currently doing what?’ and by extension ‘What else needs to be done?’. From such understanding, the expectations on open data intermediaries can be better identified.

3 Contributions of open data intermediaries in addressing challenges in the open data ecosystem

The core content of this chapter was previously published as Shaharudin, A., van Loenen, B., & Janssen, M. (2024). Exploring the contributions of open data intermediaries for a sustainable open data ecosystem. *Data & Policy*, *6*, e56. https://doi.org/10.1017/dap.2024.63.

3.1 Introduction

Various challenges exist in preparing, disseminating, processing, and reusing open data (Conradie & Choenni, 2014; Johnson et al., 2017; Sugg, 2022). These include those faced by open data providers, such as scattered data management (Linders, 2013; Ma & Lam, 2019), resource constraints (Janssen et al., 2012; Nikiforova & Zuiderwijk, 2022), and the lack of a convincing business case for publishing open data (Barry & Bannister, 2014; Cahlikova & Mabillard, 2020), and those faced by open data users, such as poor or inconsistent open data quality (Crusoe et al., 2019; Huijboom & Van den Broek, 2011), underdeveloped data skills (Janssen et al., 2012; Zuiderwijk & de Reuver, 2021), and limited complementary technologies for using open data (Temiz et al., 2022; Zuiderwijk et al., 2012). More recent studies seem to reveal that similar challenges to those identified a decade ago persist (to illustrate this, studies published before/in 2013 and after 2013 were cited for each of the aforementioned examples). Therefore, it is imperative to investigate how to address these challenges. As such, this chapter asks what the contributions of open data intermediaries may be in this regard.

From an ecosystem perspective, all actors, not limited to open data providers, can and should contribute to addressing challenges in the ODE. Actors in the ODE are individually and collectively affected by each other; hence, solving issues in the ODE is not a 'one-person job'. This understanding lays the groundwork for the RQ2: What are potential contributions of open data intermediaries in addressing challenges in an open data ecosystem? Research on the potential contributions of open data intermediaries to addressing persistent challenges in the ODE remains lacking. Most studies have only investigated their existing activities. While this baseline understanding is undoubtedly important, the next necessary step is to explore what (more) could they do for the benefit of the entire ODE. One pathway toward this end is to explore the connections between challenges in the ODE and potential contributions by open data intermediaries to address them.

The structure of this chapter is as follows. The next section, Section 3.2 presents the research methodology. Section 3.2 presents the findings. Section 3.4 reflects on the findings. Lastly, Section 3.5 concludes the chapter.

3.2 Research methodology

This chapter focuses on the EU context due to the shared legal regime concerning open data, which enables a more contextual interpretation of the findings. The EU Open Data and the Re-use of Public Sector Information Directive (Open Data Directive) was published in 2019, and EU countries were required to transpose this directive into national laws, regulations, and administrative provisions by July 2021 (Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-Use of Public Sector Information, 2019). This directive is a recast to the earlier Re-use of Public Sector Information (PSI) Directive 2003, and it establishes a set of minimum rules (including technical aspects) governing the re-use of data held by public bodies and 'public undertakings' (including non-public entities operating in the water, energy, transport, and postal services sectors), as well as research data. The directive also lists specific high-value datasets across six thematic categories, four of which constitute spatial datasets (geospatial, earth observation and environment, meteorological, and mobility), while the remaining two are statistics and companies and company ownership. EU countries that fail to comply with this directive may face legal consequences brought by the EC, including heavy financial penalties. In fact, the EC referred several countries to the Court of Justice of the EU last year due to non-compliance with the Open Data Directive (European Commission, 2023). This legal framework makes open data in the EU context rather specific since many organisations are legally obligated to publish and facilitate the re-use of open data.

This study adopted a qualitative research approach. It does not seek to offer generalisable quantitative sample-to-population findings nor suggest any causal relationships. Instead, it aims to pave the way for new areas of inquiry and interventions regarding the potential contributions of open data intermediaries for a sustainable ODE. There are two stages in the methodology. In Stage 1, data was gathered through semi-structured interviews involving ten interviewees from eight organisations representing open data providers, intermediaries, and a data standard body. From these interviews, challenges in the ODE and the potential contributions of open data intermediaries were derived (Section 3.2.1). In Stage 2, the links between the individual potential contributions of open data intermediaries and specific challenges in the ODE that they can address were explored. These links were validated with four organisations interviewed in Stage 1 and nine additional open data practitioners and researchers (Section 3.2.2). Guided by the ecosystem perspective, where actors are interrelated, representatives of diverse roles (open data providers, intermediaries, users, and researchers) were involved in Stages 1 and 2.

3.2.1 Stage 1: Semi-structured interviews

Semi-structured interviews with ten experts from eight organisations based in Denmark, Germany, the Netherlands, and Spain, representing open data providers, intermediaries, and a data standard body were conducted between May and July 2023 (Table 3.1). These organisations were selected based on purposive sampling, in which the aim was to obtain insights from organisations that could provide information on the challenges in ODE and the potential contributions of open data intermediaries. Six out of the eight organisations are spatial data organisations since four out of six thematic categories of the high-value datasets listed under the EU Open Data Directive framework constitute spatial datasets.

The individual interviewees were involved in formulating open data legislation at the national and EU levels, fulfilling the requirements of open data legislation, and coordinating open data stakeholders' engagements. Seven of the ten interviewees had at least ten years of experience in dealing with open data. The minimum open data experience among all of the interviewees was five years. All interviews were conducted online except for two, of which one was conducted in person and another one in writing.

Questions across four areas were asked: (i) the background of the interviewee and their organisation in relation to open data; (ii) their perception of the value of open data and its benefits and costs to their own organisation; (iii) their perceptions of the sustainability of the ODE, challenges in the ODE, and potential remedies; (iv) their perceptions of the current and potential contributions of open data intermediaries. However, the questions were not exactly the same for each interview since they were semi-structured. In semi-structured interviews, interviewers ask practical, customised, open-ended questions in a conversation-like manner without expecting specific answers, thereby allowing interviewees to speak freely and fully (Magnusson & Marecek, 2015). Since the ODE is a rather abstract concept, its description as 'a network of interdependent yet self-interested open data actors' were offered to interviewees to guide the interviews. The interview guide can be found Appendix C and can also be found here (with de-identified interview transcripts and the informed consent form template): https://doi.org/10.4121/d7dd11e0-7c6c-49db-946a-ffe71520f8fd.v1. As an additional privacy protection measure, the order of the de-identified transcripts in the repository differs from that in Table 3.1.

TABLE 3.1 List of interviewee organisation

No.	Organisation	Country	Role related to open data
1.	Agency for Data Supply and Infrastructure (SDFI)	Denmark	Provider & intermediary
2.	Federal Agency for Cartography and Geodesy (BKG)	Germany	Provider & intermediary
3.	Berlin Senate Chancellery	Germany	Intermediary
4.	Cadastre, Land Registry and Mapping Agency (Kadaster)	Netherlands	Provider & intermediary
5.	Geonovum	Netherlands	Standard body
6.	National Centre for Geographic Information (CNIG)	Spain	Provider
7.	Red.es (a public business entity under the Ministry of Economic Affairs and Digital Transformation)	Spain	Intermediary
8.	Department of Urban Planning, Environment and Mobility, Madrid City Council	Spain	Provider

The interview data was coded inductively (Linneberg & Korsgaard, 2019) with the aid of ATLAS.ti software. There are two categories of codes: (i) challenges in the ODE and (ii) potential contributions of open data intermediaries. The coding results can be found here: https://doi.org/10.4121/d7dd11e0-7c6c-49db-946a-ffe71520f8fd.v1. The inductive coding was adopted in order to derive the challenges in the ODE and the potential contributions of open data intermediaries close to the context the interviewees referred to, without being cognitively restricted by preconceived vocabularies. This contextual understanding is crucial in making informed decisions when linking the two to avoid speculating at a very abstract level.

3.2.2 Stage 2: Analysis and validation

In Stage 2, the individual contributions of open data intermediaries were linked to the challenges in the ODE based on the data derived in Stage 1. To validate the analysis, feedback from all of the interviewees was requested; four out of eight organisations responded with their input (organisations 1, 6, 7, 8 in Table 3.1). In addition, feedback from nine other open data practitioners and researchers were also obtained (Table 3.2). In total, 13 inputs were received during the validation exercise in February 2024. Eleven of those were done in writing, where the initial version of Table 3.5 (with the description for each item) was shared and validators edited it by suggesting new links between the challenges of the ODE and potential contributions of open data intermediaries where they considered appropriate and removing initial links where they considered otherwise. They also added new challenges in the ODE and the potential contributions of open data intermediaries that were not listed in the initial version but deemed relevant. The other two validations (with organisation 1 in Table 3.1 and organisation 6 in Table 3.2) were performed through online meetings, where the table was discussed with the validators.

Based on the input of the 13 validators, the individual link suggestions were evaluated. Each suggestion was considered by speculating a case where the suggestion could be materialised. If there is such a case, the suggestion would be accepted unless there was a counterargument against it (i.e., the contribution unlikely addresses the challenge). Where appropriate, the new challenges or contributions proposed by validators were merged (e.g., *develop common standards* proposed by a validator as one of the potential contributions was merged with *transform data into open standards*). New links that were not directly suggested by validators but are appropriate and consistent with the general insights and reasoning offered by validators were also added. Finally, the mapping of potential contributions of open data intermediaries and the ODE challenges that they could address was generated (Table 3.5).

Most of the experts engaged in this study (in Stages 1 and 2) are from the public sector, which may represent a limitation. However, these experts have rich experience in open data accumulated over many years in various capacities, including in coordinating open data stakeholder engagements. The organisations they represented are not only open data providers but several are also open data intermediaries and users. This study does not claim to offer exhaustive lists of ODE challenges and the potential contributions of open data intermediaries (if that is even possible), nor does it seek to establish causal relationships.

TABLE 3.2 List of additional validators' organisations (aside from those interviewed)

No.	Organisation	Country	Role related to open data
1.	**Aalborg University**	Denmark	Researchers (x 2) on open data use
2.	**National Institute of Geographic and Forestry Information (IGN)**	France	Provider & intermediary
3.	**Leibniz Institute of Ecological Urban and Regional Development (IOER)**	Germany	User
4.	**HuffPost**	Greece	User and researcher on open data use
5.	**Italian National Institute of Statistics (ISTAT)**	Italy	Provider
6.	**Dataninja**	Italy	Intermediary
7.	**Delft University of Technology (TU Delft)**	Netherlands	(i) Researcher on open data governance & (ii) Researcher who uses open data

3.3 Findings

Section 4.1 discusses the challenges in the ODE, while Section 4.2 discusses the potential contributions of open data intermediaries and connects them with the challenges they may address in the ODE. The challenges in the ODE and the potential contributions of open data intermediaries were gathered from the interviews. Then, the connections proposed between the two were validated with input from some of the interviewees and additional experts.

3.3.1 Challenges in the ODE

Table 3.3 presents challenges in the ODE identified from the interviews. An attempt was made to categorise them based on the key features of a sustainable ODE suggested by van Loenen et al. (2021), namely user-driven, circular, inclusive, and skills-based (refer Table 1.1 for the description of the features). This form of categorisation helps to show which challenges threaten which specific features of a sustainable ODE. However, some challenges were found to not fitting well into any of the four proposed features. Those that do not fit are either very foundational challenges around data management systems (categorised as *foundational*) or those associated with broader political factors beyond open data (categorised as *broad*). A sample quote from the interviews for each of the challenges is presented in Appendix D. Challenges indicated by an asterisk (*) were suggested by validators.

TABLE 3.3 Challenges in the ODE identified from the interviews

ID	Challenges in the ODE	Description
User-driven		
C01	**Different data/metadata standards**	Open data are provided under different data/metadata standards.
C02	**Different open data licenses**	Open data are provided under different licenses (e.g., CC0, CC BY, CC BY-SA, OdBL, OGL).
C03	**Siloed open data domains (e.g., across sectors)**	Open data domains (e.g., across sectors) exist in silos, impeding cross-domain data integration.
C04	**High technical threshold to use open data**	Using open data may require complex technical know-how or tooling by users (including experts on a particular domain, but not the other).
C05	**Unfulfilled user needs**	Open data users' needs do not match what and how open data are provided.
C06	**Limited feedback from lay users (i.e., non-expert users)**	Limited feedback from lay users (i.e., non-expert users) on open data initiatives.
Circular		
C07	**Loss of open data providers' revenue**	Open data providers lost revenue from data that they previously charged users for.
C08	**Limited value return from data re-use**	Open data providers are not compensated well for the costs they bear to develop and maintain open data initiatives.
C09	**Limited use case visibility**	Limited visibility of open data use cases (which open data are used, by whom, for what purposes, and how are they used).
Inclusive		
C10	**Limited open data from non-government sectors**	Limited open data are provided by sectors other than the government sector.
C11	**Lack of incentives for voluntarily publishing open data**	Open data providers (including government agencies) have limited incentives for voluntarily publishing open data without legal force.
C12	**Requiring viable business models**	Running initiatives or businesses based on open data requires a viable business model.
C13	**Overlooked non-government open data**	Open data provided by organisations apart from the government (e.g., businesses and non-profit organisations) may not be visible to open data users.
C14	**Practical constraints in multistakeholder engagement**	Open data stakeholders face practical constraints to conducting multistakeholder engagements, such as a lack of time, resources, and commitment.
C15	**Lack of data awareness***	Some segments of society have a limited general understanding of the role of data, different types of data, the application and value of open data, ethical issues around data, and privacy aspects.
Skills-based		
C16	**Limited knowledge of open data providers**	Certain open data providers lack knowledge regarding the legal requirements and best practices for publishing open data.
C17	**Limited knowledge of (potential) open data users**	Potential or current open data users have limited knowledge regarding open data.

>>>

TABLE 3.3 Challenges in the ODE identified from the interviews

ID	Challenges in the ODE	Description
Foundational		
C18	**Poor open data quality***	Certain types of open data have low or inconsistent quality.
C19	**Incurring maintenance costs for open data provider**	Maintaining the infrastructure and human resources to publish and update open data and engage stakeholders incurs costs for data providers.
C20	**Incurring development costs for open data provider**	Developing the infrastructure and human resources to publish and update open data and engage stakeholders incurs costs for data providers.
C21	**Technical difficulties in establishing open data management systems**	Certain open data providers face significant technical bottlenecks or difficulties in establishing open data management systems.
C22	**Complex and/or rigid open data standards to comply with**	Some open data standards are too complex and/or rigid for open data providers to comply with.
C23	**Heterogeneous data administration**	Open data are provided by different data administrations that operate under different systems/jurisdictions.
C24	**Privacy concerns**	Open data providers have to address privacy concerns that may not be easily identifiable and addressable.
Broad		
C25	**Inflexible/unclear government-market boundary**	Open data providers among government agencies are uncertain or restricted in terms of the extent to which they can offer value-added products using their open data.
C26	**Reliance on a political agenda**	Some open data initiatives (not only by government agencies) rely on the political agendas of their funders.
C27	**Inflexible governance/law (esp. with evolving technology)**	Laws and bureaucracy related to open data publication may take a long time to change, thus being unable to keep up with rapidly evolving technologies and data applications.

*Note: * were suggested by validators*

Challenges related to user-drivenness: Open data users have to deal with different open datasets published under various standards and licenses. As highlighted by an interviewee, the EU Directive on Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) states that spatial data published by national mapping agencies in the EU should follow ISO standards. However, generic administrative open data are rarely published in those standards; instead, they are published in other standards (e.g., DCAT). Simultaneously, to use and possibly integrate open data from different sources for specific purposes, users must check the conditions of the relevant data licenses and whether they are compatible. This task is not necessarily straightforward, especially for users with limited legal literacy. These technical and legal interoperability issues are exacerbated by siloed open data domains (geospatial, demographic statistics, etc.).

Using open data may require complex technical know-how and tooling, which particularly impacts non-specialist users (including experts on a particular domain but not on others). Simultaneously, with the growing (potential) applications of open data, the diverse needs of users (e.g., in terms of the data and the format they require) are not entirely met. While some open data providers have taken the initiative to seek user feedback, limited input is obtained from non-expert users, especially among small and medium enterprises and individual citizens.

Challenges related to circularity: Open data providers that previously sold their data lost (one of) their source(s) of income upon providing it as open data. In the EU, many organisations are legally required to publish open data even though it would not have been in line with their managerial or strategic decisions. However, open data providers receive limited value returns (particularly monetary) from the open data they provide. Meanwhile, open data use cases are not always visible. Thus, open data providers cannot fully assess the value of their open data to decide what data they (do not) need to provide and prove its value to seek (more) funding.

Challenges related to inclusivity: Non-government sectors (i.e., the private and civil sectors) provide limited open data, which is partly attributed to the lack of incentives to voluntarily publish open data. As noted by at least two interviewees, the legal obligation in the EU entails a major push for open data; however, this mainly affects the public sector at present. Running initiatives based on open data (as providers, users, and intermediaries) requires viable business models, which remain underdeveloped. One interviewee also noted that open data from non-government sectors are available in some cases but simply not visible enough. Multistakeholder engagements are necessary to encourage non-government sectors to contribute open data and participate in decision-making related to open data. However, individual actors face practical constraints in leading such engagements due to limited resources, time, and commitment.

Challenges related to skills: Some existing or potential open data providers have limited knowledge of the best practices for publishing open data. In the context of the EU, certain open data providers struggle to meet the requirements set by the law due to the lack of technical expertise in the organisation. Conversely, potential or current open data users may have limited knowledge and skills related to using open data in a meaningful manner.

Foundational challenges: Publishing open data comes with considerable costs for providers linked to developing and maintaining the relevant infrastructure and expertise. This is especially the case for organisations that had underdeveloped data management systems before they had/decided to publish open data. In the development stage, open data providers must deal with major technical undertakings to build open data infrastructure and processes. Additionally, some open data providers are legally required to publish data in specific data models that are rather complex, which was noted by at least two interviewees. Certain open data providers may also need to integrate data from different administrations before publishing it as open data, which involves significant coordination efforts. Furthermore, as the volume of open data grows, privacy concerns may become more challenging to address.

Broad challenges: Certain challenges in the ODE are an extension of broader political factors beyond open data alone, particularly surrounding policies around market competition, overarching digital strategies, and technological governance. Different governments and societies may have different preferences and approaches to these aspects, which ultimately affect the ODE. In the EU, decisions in these areas are also negotiated and made at the supranational level.

3.3.2 Potential contributions of open data intermediaries

Table 3.4 presents the potential contributions of open data intermediaries derived from the interviews. The contributions were categorised into technical, non-technical, and combination contributions. Technical contributions may include the direct processing of open data and developing tools to facilitate its supply or use. Non-technical contributions do not necessarily involve directly processing open data or developing tools but are more geared towards relationship building, stakeholder engagement, and financial support. Combination contributions are those that involve a mix of technical and non-technical activities. Moreover, these potential contributions may be interrelated. Some of them are not necessarily new; for instance, fostering collaborations is a known contribution of open data intermediaries, as described in Chapter 2. Table 3.5 links the potential contributions of open data intermediaries with the ODE challenges they may address. A sample quote from the interviews for each of the potential contributions is presented in Appendix E. Potential contributions indicated by an asterisk (*) were suggested by validators.

TABLE 3.4 Potential contributions of open data intermediaries

ID	Potential contributions of open data intermediaries	Description
Technical		
P01	**Implement federated architecture**	Develop and maintain federated architecture open data systems.
P02	**Integrate data (e.g., across sectors/ administration)**	Integrate open data from different sectors and administrations.
P03	**Transform data into open standards (esp. web standards)**	Transform open data in certain standards into open standards to proliferate its re-use.
P04	**Customise data (based on use cases)**	Transform open data into formats familiar to certain professionals, domains, or use cases.
P05	**Offer process automation**	Offer applications/software that provide process automation in data publication and/or use.
P06	**Develop open-source tooling**	Develop open-source tooling to be used by open data providers or users.
P07	**Provide direct technical services***	Provide direct technical services to open data providers or users (beyond offering tooling or pre-processed data).
P08	**Offer freemium data platform***	Offer a one-stop platform for data providers to provide open data together with their paid data.
Non-technical		
P09	**Foster public-private collaboration**	Initiate and coordinate collaboration between the government and businesses.
P10	**Foster public-civic collaboration**	Initiate and coordinate engagement between the government and non-governmental organisations (NGOs).
P11	**Implement multistakeholder collaboration**	Initiate and coordinate the collaboration between the government, businesses, and NGOs.
P12	**Perform open data advocacy***	Conduct advocacy/lobbying activities related to open data issues.
P13	**Invest in open data-based civic tech**	Invest in businesses or initiatives that offer open data-based applications or programs that benefit the community, in line with societal goals.
P14	**Showcase open data value**	Demonstrate open data use cases and the value of open data.
P15	**Promote open non-governmental data**	Identify open data from organisations outside the government sector and promote their re-use.
Combination		
P16	**Provide consultancy**	Advise and help open data providers and/or users with aspects related to the publication and use of open data.
P17	**Streamline cross-administrative processes**	Advise and coordinate efforts to streamline data management processes across different administrations.
P18	**Facilitate internal re-use of open data**	Identify and recommend the internal re-use of open data.
P19	**Facilitate feedback on open data**	Help users provide feedback to open data providers.
P20	**Provide education on data literacy and skills***	Provide training and other education activities to improve data literacy and skills.

*Note: * were suggested by validators*

TABLE 3.5 Potential contributions of open data intermediaries and the ODE challenges they may address

Potential contributions of open data intermediaries (column) Challenges in the ODE (row)	Technical contributions								Non-technical contributions							Combination contributions				
	P01: Implement federated architecture	P02: Integrate data	P03: Transform data into open standards	P04: Customise data	P05: Offer process automation	P06: Develop open-source tooling	P07: Provide direct technical services	P08: Offer freemium data platform	P09: Foster public-private collaboration	P10: Foster public-civic collaboration	P11: Implement multistakeholder collaboration	P12: Perform open data advocacy	P13: Invest in open data-based civic tech	P14: Showcase open data value	P15: Promote open non-governmental data	P16: Provide consultancy	P17: Streamline cross-administrative processes	P18: Facilitate internal re-use of open data	P19: Facilitate feedback on open data	P20: Provide education on data literacy and skills
User-drivenness challenges																				
C01: Different data/metadata standards	X	X	X	X					X		X						X			
C02: Different open data licenses	X										X									
C03: Siloed open data domains (e.g., across sectors)	X	X	X	X	X				X		X						X			
C04: High technical threshold for users		X	X	X	X	X	X						X			X				
C05: Unfulfilled user needs		X	X	X	X				X	X	X				X	X			X	
C06: Limited feedback from lay users (i.e., non-expert users)									X	X	X					X			X	
Circularity challenges																				
C07: Loss of open data providers' revenue												X	X	X						
C08: Limited value return from data re-use													X	X			X	X	X	
C09: Limited use case visibility									X	X	X			X		X	X		X	
Inclusivity challenges																				
C10: Limited open data from non-government sectors								X	X	X	X		X		X					
C11: Lack of incentives for voluntarily publishing open data								X	X	X	X		X	X	X			X		
C12: Requiring viable business model						X			X	X	X		X	X		X				
C13: Overlooked non-government open data									X	X	X			X	X					
C14: Practical constraints in multistakeholder engagement									X	X	X					X	X		X	
C15: Lack of data awareness										X	X	X		X	X	X				X

>>>

TABLE 3.5 Potential contributions of open data intermediaries and the ODE challenges they may address

Potential contributions of open data intermediaries (column) Challenges in the ODE (row)	Technical contributions								Non-technical contributions							Combination contributions				
	P01: Implement federated architecture	P02: Integrate data	P03: Transform data into open standards	P04: Customise data	P05: Offer process automation	P06: Develop open-source tooling	P07: Provide direct technical services	P08: Offer freemium data platform	P09: Foster public-private collaboration	P10: Foster public-civic collaboration	P11: Implement multistakeholder collaboration	P12: Perform open data advocacy	P13: Invest in open data-based civic tech	P14: Showcase open data value	P15: Promote open non-governmental data	P16: Provide consultancy	P17: Streamline cross-administrative processes	P18: Facilitate internal re-use of open data	P19: Facilitate feedback on open data	P20: Provide education on data literacy and skills
Skills-based challenges																				
C16: Limited knowledge of open data providers		X		X	X		X									X	X	X	X	X
C17: Limited knowledge of (potential) open data users		X	X	X	X		X								X	X		X		X
Foundational challenges																				
C18: Poor open data quality							X									X			X	
C19: Incurring maintenance costs for open data provider	X				X	X			X		X			X		X	X			
C20: Incurring development costs for open data provider	X				X	X			X		X			X		X	X			
C21: Technical difficulties in establishing data management systems	X				X		X									X	X			
C22: Complex and/or rigid open data standards to comply with	X				X		X									X				
C23: Heterogeneous data administration	X	X	X		X						X					X	X			
C24: Privacy concerns					X					X	X					X			X	X
Broad challenges																				
C25: Inflexible/unclear government-market boundary												X				X				
C26: Reliance on a political agenda									X	X	X	X		X		X	X	X	X	
C27: Inflexible governance/law (esp. with evolving technology)			X	X							X					X				

Technical contributions: By implementing federated architecture, open data intermediaries may help address user-driven challenges (different data standards, different open data licenses, and siloed open data domains) and the foundational challenges of data management systems. Integrating data may also help resolve issues of different data standards, siloed open data domains, the high technical thresholds for non-specialist users, and unfulfilled user needs (due to fragmented data). It may also help address the limited skills of data providers and users in combining different datasets and the problem of open data being provided by heterogeneous administrations. Transforming open data into open standards, especially web standards that are already used widely across many domains, may address issues similar to integrating data. It may also help to overcome issues where data are published according to standards set by the law, but the standards are not adaptive to changing technology and user needs. Likewise, customising data based on the common use cases in a specific industry or domain may also help to address the same problem of misaligned legal development vis-à-vis technological progress. It may also address some challenges related to skills and user-drivenness (different data standards, siloed open data domains, high technical threshold, and unfulfilled user needs) but not foundational challenges.

Open data intermediaries may also offer automation for certain open data publishing and use processes. This may help to overcome challenges around user-drivenness (siloed open data domains, high technical threshold, and unfulfilled user needs), skills, and foundational challenges (including by enhancing privacy protection). Offering open-source tooling may reduce the high technical threshold for non-specialist users through an affordable means. This would enable the implementation of certain business models by open data actors that do not have the technical expertise and proprietary in-house tooling or financial capacity to acquire such resources. Open data intermediaries may also provide direct technical services in certain parts of open data processing (instead of providing tooling and pre-processed data). This may resolve challenges linked to the high technical threshold for using open data, limited skills among open data providers and users, and several foundational challenges (poor open data quality, technical difficulties in establishing data management systems, and complex data standards to comply with). Additionally, open data intermediaries may offer a one-stop platform for providers to offer their open data together with paid data, thereby encouraging (non-government) actors to provide open data through the freemium business model (addressing inclusivity challenges).

Non-technical contributions: By fostering public-private, public-civic, and multistakeholder collaboration, open data intermediaries may help to address inclusivity, user-drivenness, circularity, and broad foundational challenges. While most links between these three contributions and their challenges are quite straightforward, the connections between public-private/multistakeholder collaborations and the challenges of development/maintenance costs incurred by the open data providers may not be very clear. The idea is that the public, private, and civil sectors may pool resources and jointly coordinate efforts to build and maintain technical infrastructure to publish open data (whether the open data are from the public, private, or civil sectors or a combination of them), thereby reducing the costs incurred for a single organisation. This form of collaboration may work if all parties are interested in publishing specific open data, such as data showing the locations of electric vehicle charging stations. Open data intermediaries may facilitate this type of collaboration.

Open data intermediaries may undertake open data advocacy, among other activities, to raise awareness of the loss of open data providers' revenue, and lobby for additional funding based on the socioeconomic value of open data. Open data intermediaries may also invest in and provide business model design support to civic technology companies producing applications that reduce the technical threshold for (lay) users to benefit from open data. They may return a share of the profits to open data providers (e.g., through joint ventures) and encourage more open data releases. Furthermore, open data intermediaries may contribute to showcasing the (critical) value of open data to address circularity, inclusivity, foundational, and broad political challenges. By promoting open non-government data, open data intermediaries may help address most of the inclusivity challenges as well as unfulfilled user needs (user-drivenness) and the limited knowledge of (potential) users (skills-based).

Combination contributions: Open data intermediaries may provide various types of consultancy to open data providers and users, such as those related to technical, managerial, economic, and political aspects. The technical and managerial types of consultancy may help open data providers and users to overcome user-drivenness, skills-based, circularity, and foundational challenges. The economic and political types of consultancy may help to address the challenges faced by open data providers and users linked to requiring viable business models and mitigating broad political challenges. Open data intermediaries may also help streamline cross-administrative data management processes to tackle various issues across all dimensions, particularly foundational issues.

Open data intermediaries may facilitate the internal re-use of open data to address shortcomings of limited value return from open data, lack of incentives for publishing open data voluntarily, reliance on a political agenda, and limited knowledge of open data providers and users. For this contribution, the open data intermediary may likely be a unit within the same organisation as the open data provider to have comprehensive knowledge of the data produced and used by the organisation. Open data intermediaries may also facilitate feedback on open data, which may tackle some challenges across all dimensions. This facilitation of feedback may involve technical aspects (e.g., introducing a feedback feature in a data platform) or non-technical aspects (e.g., organising user group meetings). Additionally, open data intermediaries may provide education on data literacy and skills. Apart from tackling skills-based challenges, this education may aim to improve general public data awareness, such as on different types of data, the potential applications of open data, ethical ways of using data, and privacy aspects.

3.4 Discussion

This study affirms the role of open data intermediaries in addressing challenges in the ODE. The findings show that not only could they help the ecosystem strengthen the four features of a sustainable ODE as proposed by van Loenen et al. (2021), but they could also help address foundational issues around data management systems and mitigate the broad political factors impacting the ODE. However, it is important to note that these potential contributions addressing the various challenges in the ODE would not magically materialise. To reiterate, ODE actors are deemed to be self-interested and open data intermediaries require internal incentives (e.g., through viable business models) and/or extrinsic incentives (e.g., through policies and regulations) to drive them to offer those contributions. Therefore, this study emphasises further inquiry into the business models of open data intermediaries and other external conditions of the ODE that could encourage open data intermediaries to act in ways that support the sustainability of the ODE, in which their interests and other actors are aligned.

This study highlights the need to distinguish a *role* from an *actor* in the ecosystem, as elucidated by Oliveira and Lóscio (2018) and Shaharudin et al. (2023). Diverse ODE actors can undertake open data intermediation, including public organisations, for-profit companies, and CSOs. It is especially worth emphasising that open data intermediaries do not only exist outside of the public sector, as some have implied (Balvert & van Maanen, 2019; Schrock & Shaffer, 2017). Several public organisations involved in this study, such as SDFI (Denmark), BKG (Germany), and IGN (France), have identified some of their tasks as being those of open data intermediaries. One of the interviewees mentioned,

‘We do not necessarily produce the data, but we [are] intermediaries ourselves. So we get the data from others, in particular from the official mapping agencies [redacted], and we process that data, combine it, and provide it’.

Similarly, during the validation exercise, a public organisation representative noted how they provide a paid service to help smaller public agencies fulfil the technical data requirements set by the law. Having said that, what a government organisation can do as an open data intermediary may differ from what a for-profit company or a civil society organisation can do in that role. This is due to the different legal obligations, societal expectations, resources, and other factors that these organisations have. Hence, the contributions of diverse types of open data intermediaries that are possible in practice are also different.

The findings also suggest that an open data provider or user could benefit from the contributions of multiple open data intermediaries simultaneously, in parallel and/or sequentially. As van Schalkwyk et al. (2016, p. 22) argued, ‘No single intermediary is likely to possess all the types of capital required to unlock the full [open data] value’. This also implies that the beneficiary of an open data intermediary’s contributions could also be another open data intermediary. This insight highlights what Oliveira and Lóscio (2018) emphasised: value is not created in a chain but instead in a network in which actors can participate in multiple networks. A chain refers to linear pathways where value is transferred from one actor to the next in a linear sequence, whereas a network refers to complex, non-linear pathways where value is transferred from and to multiple actors in various directions. This also means that there could be multiple orderings of open data providers, intermediaries, and users in the ecosystem. Thus, apart from provider-intermediary-user relationships, they could also take the form of provider-intermediary-intermediary-user or provider-parallel intermediaries-user relationships, among others. Thus, open data intermediaries are not merely the ‘bridge’ between open data providers and users.

Some of the challenges of the ODE identified from the interviews are very foundational issues around open data management systems or related to broader political factors and do not necessarily fit into the four features of a sustainable ODE suggested by van Loenen et al. (2021). This implies that the four features are inadequate to determine the sustainability of the ODE. Thus, additional layers of criteria may be necessary and deserve future attention. At the very least, the four features noted by van Loenen et al. (2021) may have to be clarified or refined to readily incorporate those foundational and broader political issues around open data. Some proposals for ODE sustainability assessment framework have been made in the past (Vancauwenberghe, 2018; Welle Donker & van Loenen, 2017) but they were largely limited to the context of open government data and, more importantly, the object of assessment is usually the data, whereas the ODE centres the relationships between actors and the value flows instead of the data per se.

The findings on the challenges of the ODE and the potential contributions of open data intermediaries call the boundaries of the ODE into question. Notably, the ODE challenges identified include those associated with external factors beyond just open data, and the contributions of open data intermediaries may involve many more activities than directly handling open data. Drawing from ANT, Latour (2007, p. 29) asserted that 'it's not the sociologist's duty to decide *in advance* [emphasis added] and in the member's stead what the social world is made of' since 'social aggregates are not the object of ostensive definition—like mugs and cats and chairs that can be pointed at by the index finger—but only of a performative definition' (p. 34). Latour emphasised tracing connections 'instead of being constantly bogged down in the impossible task of deciding once and for all what is the right unit of analysis' (p. 34). In other words, the boundaries of the ODE do not have to (or cannot) be defined beforehand. However, for a specific assessment, research inquiry, or intervention, one should identify relevant actors, trace their associations, and analyse their interactions, which would involve making and remaking boundaries (Harrison et al., 2012; Lee & Hassard, 1999). Similarly, van Schalkwyk et al. (2015) argued that to determine whether an actor can be considered an open data intermediary, one should assess its 'degree of agency' in fulfilling the open data intermediation functions. Having said that, striking a balance between being pragmatic and reductionist is undoubtedly something that open data researchers and practitioners would have to constantly grapple with.

3.5 Conclusion

This chapter has explored the potential contributions of open data intermediaries in addressing various challenges in the ODE. Through interviews and a validation exercise conducted with 19 individuals from 15 organisations in the EU, it has been shown that open data intermediaries could help overcome challenges that are detrimental to ODE sustainability through various technical and non-technical contributions. The contributions of open data intermediaries identified would not automatically resolve the challenges in the ODE linked to them by default; instead, they have to be designed for that purpose. As previously stressed, actors in the ODE are self-interested; hence, they require intrinsic incentives (e.g., through viable business models) and/or external conditions (e.g., through policies and regulations) that drive them toward acting in a particular manner. Hence, research into business models and policies relevant to open data intermediaries is necessary while considering the diverse types of actors serving the role of open data intermediaries (e.g., government organisations, companies, and CSOs). Notably, different types of actors would require different sets of incentives.

PART B

Characterising archetypes of open data intermediation business model

The forest

Chapter 4 moves another step closer to the heart of this dissertation, unravelling business models of open data intermediaries. Since it is evident from Part A that open data intermediaries may also perform other activities that are not directly related to open data, the term ‘open data intermediation business model’ refers to the aspects of the business model that are specifically relevant to open data activities. Just like there are different archetypes of forest (e.g., tropical rainforest, mangrove forest, dry forest, and coniferous forest), there are distinguishable archetypes of open data intermediation business models exist in practice. While individual business models within the same archetype may vary in their details, they share a common set of core components. Identifying these archetypes are necessary for this dissertation in order to propose more specific and targeted, instead of generic, recommendations in developing open data intermediation business models that support a sustainable ODE.

4 Archetypes of open data intermediation business models

The core content of this chapter has been submitted for journal consideration as Shaharudin, A., van Loenen, B., & Janssen, M. (n.d.). *Business model archetypes of open data intermediaries: Empirical insights from practice.* Manuscript submitted for publication.

4.1 Introduction

Most studies on open data business models did not focus specifically on open data intermediaries; instead, they focused on the business models of open data actors or re-users in general (Charalabidis, Zuiderwijk, Alexopoulos, Janssen, Lampoltshammer, et al., 2018; Ferro & Pizzamiglio, 2023; Garatzogianni et al., 2017; Magalhaes et al., 2014; Magalhaes & Roseira, 2020). Others have studied the business models of data intermediaries that do not specifically deal with open data (Micheli et al., 2020, 2023; Schweihoff et al., 2024; Susha et al., 2020); thus, obscuring or missing the peculiarities of open data intermediaries that deal with data that are already reusable free of charge under an open license. Meanwhile, the few studies specifically focus on open data intermediation business models (Germano et al., 2016; Janssen & Zuiderwijk, 2014) fell short of integrating the key business model dimensions (value proposition, value creation, and value capture), resulting in an incomplete overview. Those studies are also outdated, considering the advancement of open data and the emergence of many more open data intermediaries over the past decade.

Therefore, this chapter addresses the RQ3: What are archetypes of open data intermediation business model? An archetype is ‘a typical example of something, or the original model of something from which others are copied’ (Cambridge Dictionary, n.d.). Although business models of the same archetype may exist in multiple variations, all have similar core components (Sterk et al., 2024). Additionally, throughout this chapters, the terms organisations instead of companies and consumers instead of customers are used more commonly (unless in specific cases) to be semantically consistent across different types of open data intermediary organisations.

The structure of this chapter is as follows. Section 4.2 presents the multimethod approach adopted, consisting of the SLR, data collection from real-world 190 open data intermediaries, and the unsupervised machine learning (ML) K-means clustering method. Section 4.3 presents the identified open data intermediaries’ business model archetypes. Section 4.4 discusses the study’s empirical, theoretical and practical implications. It also elaborates on the limitations of the study. Lastly, Section 4.5 concludes the chapter.

4.2 Research methodology

Various methodologies have been employed to identify business model archetypes in the extant literature, such as thematic analysis (Chirumalla et al., 2024; Magalhaes & Roseira, 2020), case studies (Janssen & Zuiderwijk, 2014; Reinhardt et al., 2020; Susha et al., 2020), and an SLR (Trapp & Kanbach, 2021). Nevertheless, many studies do not rely on a single method, but adopt a multimethod approach. For example, to identify business models of 3D printer manufacturers, Holzmann et al., (2019) conducted a thematic analysis, followed by a two-step cluster analysis. Likewise, Urban et al. (2018) identified airline business model archetypes by filling out the elements of Osterwalder and Pigneur's (2010) business model canvas based on various sources and then employed a two-step cluster analysis. Weking et al. (2020) employed a literature review and Ward's hierarchical cluster analysis to discover blockchain-based business model archetypes. The same approach was adopted by Duparc et al. (2022) to identify open-source business model archetypes, and Sterk et al. (2024) for connected car. Lüdeke-Freund et al. (2019) identified business model patterns in the circular economy through a literature review and a morphological analysis. Overall, depending on the business models studied, data availability, and expected details of the business model archetypes, different studies employed different (combinations of) methods, as there is no one size fits all.

Much like in the literature, this study employed a multimethod approach over four sequential stages. First, an initial codebook consisting of categories and elements of open data intermediation business models was developed through an SLR. Second, relevant qualitative data was collected from 190 samples of existing open data intermediaries based on the abductive approach facilitated by the initial codebook developed. The codebook was then modified based on the learning throughout the data-gathering process. Third, an unsupervised ML technique, namely K-means clustering, was employed to group the business models of the sample cases. During this stage, several categories was deselected based on the calibration of the K-means clustering. K-means clustering was chosen instead of Ward's hierarchical clustering, which was used in some studies because the former is more appropriate for large datasets. Besides, Ward's clustering is commonly used to identify archetypes of which a hierarchical structure is important (e.g., gene expression), which is not expected in the present study since every category is considered as potentially key to determining a cluster. Fourth, the K-means clustering results were interpreted and open data intermediaries' business model archetypes were identified. The term *clusters* refer to the K-means output before the interpretation, whilst *archetypes* refer to the final business model groups after interpreting the K-means output. Figure 4.1 illustrates the four stages of the methodology employed in this study with fruits as the analogy.

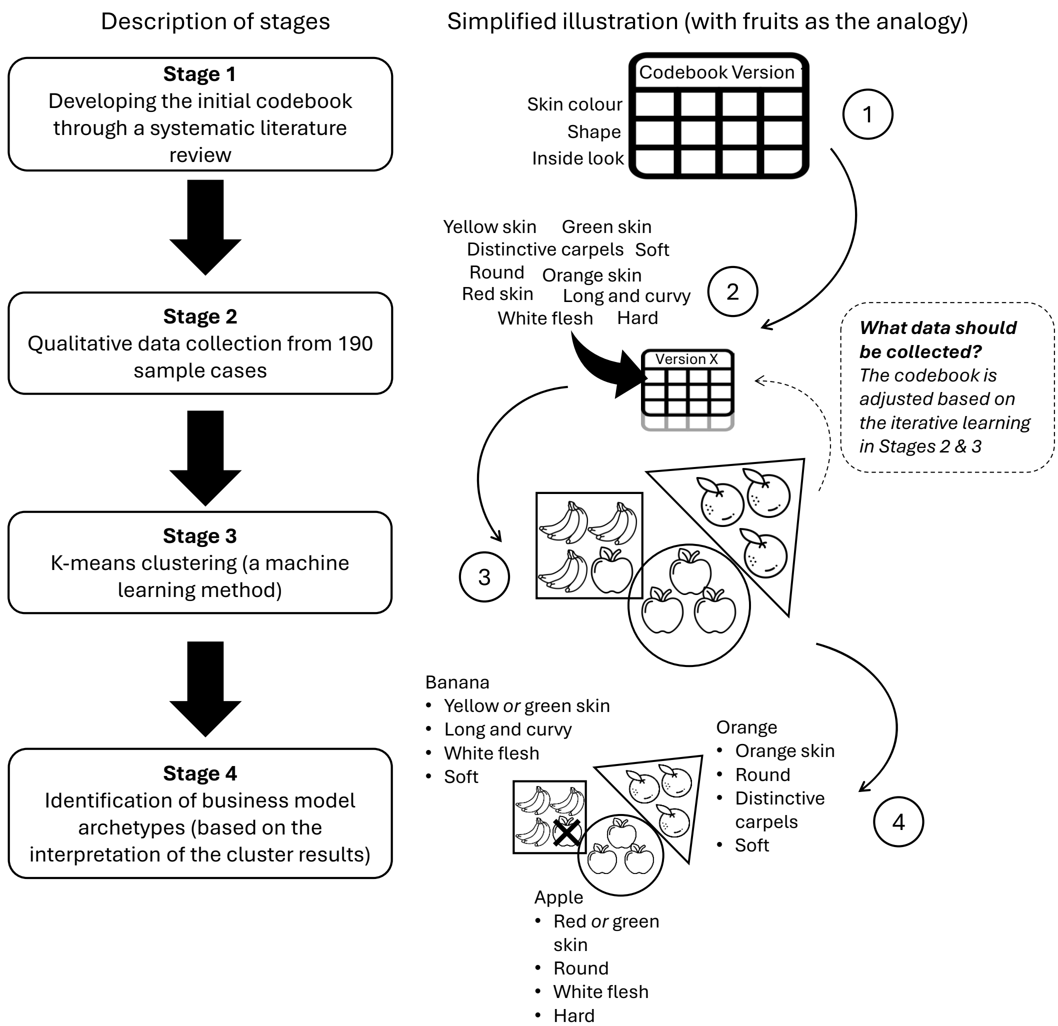


FIG. 4.1 Four stages of the research methodology

4.2.1 Stage 1: Developing the initial codebook through an SLR

In Stage 1, the initial codebook was developed to initiate data collection from the sample cases in Stage 2. An SLR was conducted to establish the initial codebook consisting of the categories and elements of open data intermediation business models. *Categories* refer to the characteristic groups of the business models. Each category consists of *elements*, which are the identifiers of specific characteristics. The SLR steps proposed by Xiao and Watson (2019) were followed. Relevant publications were searched in the Scopus and Web of Science databases, using the combination of keywords presented in Table 4.1. Notably, the search was not limited to *intermediaries* as it was learned from the literature screening that several publications on the business models of other open data roles are still appropriate for the purpose of this study. The term *open government data* was also included as a search term since it is a subset of open data. Apart from the term *business model*, different search terms that may capture relevant literature were also considered.

TABLE 4.1 Search terms for the SLR (30 combinations)

Boolean operator	OR
AND	open data, open government data
	business model, revenue, value proposition, value creation, value capture, value architecture, value network, finance, profit, business format, enterprise model, model of business, business plan, business strategy, business opportunity

Following the literature filtering (Appendix F), ten relevant publications were found. Two more publications were included in the literature pool through backward citation: Al-Debei and Avison (2010) and Osterwalder and Pigneur (2010). Ultimately, 12 publications were used to develop the initial codebook. Table 4.2 presents the relevant publications. *Publications relevant to categories* refer to the publications used to develop the categories of the codebook, and vice versa for *publications relevant to elements*. Several publications are relevant to both categories and elements. The publications were synthesised to develop the initial codebook (Appendix G).

TABLE 4.2 Relevant publications for the initial codebook

Publications relevant to *categories*	Publications relevant to *elements*
Ahmadi Zeleti et al. (2016); Al-Debei and Avison (2010); Kamariotou and Kitsios (2022); Osterwalder and Pigneur (2010); Yu (2016)	Ahmadi Zeleti et al. (2016); Janssen and Zuiderwijk (2014); Kitchin et al. (2015); Lindman et al. (2016); Magalhaes and Roseira (2020); Osterwalder and Pigneur (2010); Schroeder (2016); Smith et al. (2016); Welle Donker and van Loenen (2016); Yu (2016)

4.2.2 Stage 2: Data collection from 190 sample cases

Through purposive sampling, 190 samples of open data intermediary products/services were identified from the open data use cases compiled by data.europa.eu (the official portal for European data). Only use cases that represent products/services offered by an open data intermediary, following the definition in Chapter 2 were selected. Since open data intermediary organisations may perform activities beyond open data and may offer non-open data-based products and/or services (e.g., Nasdaq offers other products/services besides Nasdaq Data Link), the data collection was anchored to the open data intermediation products/services they offer instead of the organisations as a whole. Certainly, there are also cases where the products/services offered represent the entire organisation (e.g., OpenStreetMap).

Qualitative data from the sample cases was collected through a desk survey following the abductive approach facilitated by the initial codebook developed in Stage 1. This process occurred between January and April 2023. Figure 4.2 presents the number of sample cases according to the type of organisation, geographical area in which they operate (or that their products/services are accessible), and sector (following the categorisation by data.europa.eu). Most of the sample cases are companies (103), followed by non-governmental organisations (NGOs) or NPOs (34) and public organisations (26). More than half of them operate in Europe (118). In terms of sectoral categories, the sample cases were gathered following their corresponding population shares in the portal. However, it is worth noting that sectoral categorisation is not a rigid demarcation; for instance, a use case categorised in the transport sector may also fit in the regions and cities sector.

The codebook was modified based on the learning throughout the data collection process, which involved several iterations (Appendix G). The goal of such iterative modification is to strike a balance between collecting meaningful and sufficient data to capture the salient business model characteristics on the one hand, and avoiding excessive detail (noise) that obscures the salient characteristics on the other. The codebook iteration also ensures the consistency of data gathered across the sample cases. In particular, the elements within the *type of main open data-based product* were expanded, and an extra piece of information was added to each of the elements by indicating whether the type of product is data-to-data (D-D), data-to-information (D-INFO), data-to-knowledge (D-K), or support services (SU) products, partly inspired by the data-information-knowledge-wisdom (DIKW) hierarchy (Ackoff, 1989; Rowley, 2007). This helps to ensure consistency in the data collection process across the sample cases.

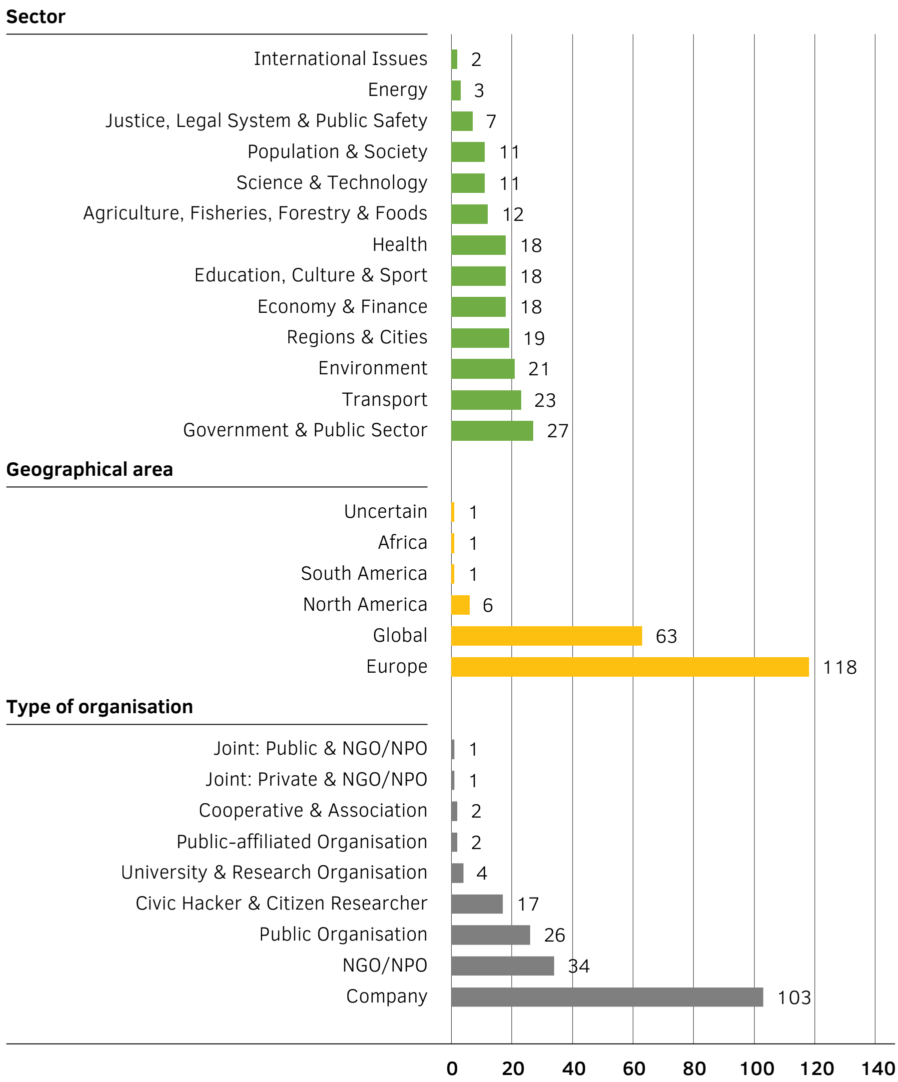


FIG. 4.2 Number of sample cases based on the type of organisation, geographical area, and sector

New categories were also added in the value dimension: *source of data*; *product components*; *other open data-based products*; *non-open data-based products*; *linking of other product(s) to the main products*. The new categories were added based on the acknowledgement that non-open data may be combined with open data to offer certain products/services, some products/services may consist of multiple modular units of products/services, and open data intermediaries may offer multiple open data-based products and/or non-open data-based products simultaneously. Furthermore, some expansion, contraction, or modification of the elements was done within the categories of *channel*, *consumer segment*, *critical partner* (substituting *key partners*), *critical resource* (substituting *key resources*), *customer relationship*, *cost structure,* and *main revenue stream*. The category of key activities was substituted with the critical stage of the open data lifecycle, with elements derived from van Veenstra and van den Broek (2015). Additionally, a new category was added under the value dimension, namely *source of revenue,* which indicates whether the revenue is derived solely from (augmented) open data or also from other sources.

The above codebook modifications were based on what was learned from the data itself (instead of based on the literature, as in Stage 1). To provide an illustrative example by using the analogy of fruits from Figure 4.1, when it was discovered that the feel of the fruits (hard, soft) is also a useful category for identifying what fruits are there, data based on this category was then added. On the other hand, when it was discovered that the smell of fruits is challenging to capture, this category was then dropped from the data collection process.

Since the data collected is qualitative instead of quantitative, the data collection process inevitably involved some interpretation and best efforts to capture the business model elements (e.g., in determining the *critical* stage of the open data lifecycle involved). During the data collection process, certain categories were left blank for certain sample cases where the data was uncertain. At the end of Stage 2, a dataset consisting of the business model elements of 190 open data intermediaries was developed. This dataset would be further modified in Stage 3.

4.2.3 Stage 3: K-means clustering

K-means clustering was used to facilitate the identification of the open data intermediaries' business model archetypes based on the dataset developed in Stage 2. The Orange Data Mining software package was utilised for this step. K-means is an unsupervised ML technique used to group *n* objects, each with measurements of *p* variables, into *K* clusters (Steinley, 2006). In this study, the objects are the individual sample cases (i.e., $n = 190$), and *p* is the number of categories in the codebook (hereafter, they are simply called *categories*). The goal of K-means is to minimise within-cluster variances. The objective function of the K-means algorithm can be expressed as:

$$Z = \sum_{k=1}^{K} \sum_{x_{ij} \in C_k} \left\| x_{ij} - \mu_k \right\|^2$$

where,

Z is the objective function to be minimised,
K is the number of clusters,
C_k are the data points assigned to the *k*-th cluster,
X_{ij} is the data point of the *j*-th variable for the *i*-th object (where $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$),
μ_k is the centroid (mean) of the *k*-th cluster, and
$\|x - \mu_k\|^2$ is the squared Euclidean distance between data point *x* and centroid.

Since the dataset consists of categorical data, they are converted into numerical data with one feature per value before the clustering procedure. K-means clustering requires initialising the number of clusters (i.e., *K*) *a priori*. Two strategies were employed to determine the optimal *K* based on two decision criteria. The first strategy is to experiment with a range value of *K*, while the second strategy is to experiment with deselecting several categories to minimise overspecification. To decide whether the optimal *K* has been reached, the *K* with the highest silhouette scores within a range of *K* values was selected, and the clustering output was ensured to make human sense (e.g., by observing whether rough commonalities among several sample cases' business models can be observed within the same cluster). A silhouette score (ranging from -1 to +1) assesses whether an object matches its designated cluster (the higher the score, the better it fits) (Rousseeuw, 1987).

This calibration process to find the optimal K is thus iterative. Ultimately, four categories were deselected (*channel*, *critical partner*, *critical resource*, and *cost structure*), resulting in the selection of 12 categories for the clustering (i.e., $p = 12$) (Table 4.3). These categories were deselected either because they correlate highly with another category and are thus considered redundant (*channel* with the *type of main open data-based products*), or they are highly subjective and speculative (*critical partner, critical resource,* and *cost structure*). Ultimately, based on the calibration process, the optimal K was determined as 10.

Table 4.3 presents the evolution of the codebook from Stage 1 to Stage 3 (the full version is shown in Appendix G). Figure 4.3 present the silhouette plot (cases with negative silhouette scores may have been assigned to the wrong cluster). Figure 4.4 shows the multidimensional scaling (MDS) of the clusters. MDS visualises multidimensional data (in this study's case, 12 dimensions) into a two-dimensional Cartesian space; therefore, some visible overlaps between K-means clusters are expected.

TABLE 4.3 Codebook categories (and the number of respective elements) by the end of stages 1, 2, and 3

Codebook, end of Stage 1	Codebook, end of Stage 2	Codebook, end of Stage 3
Value proposition		
Type of open data products (13)	Type of main open data-based product (15)	Type of main open data-based product (15)
	Source of data (2)	Source of data (2)
	Product components (2)	Product components (2)
	Other open data-based products (2)	Other open data-based products (2)
	Non-open data-based product (2)	Non-open data-based product (2)
	Link of other product(s) to the main open data-based product (4)	Link of other product(s) to the main open data-based product (4)
Offering (13)	Offering (13)	Offering (13)
Channel (2)	Channel (3)	*Deselected*
Consumer segment (4)	Consumer segment (6)	Consumer segment (6)
Value creation		
Key partners (*open coding*)	Critical partner (16)	*Deselected*
Key activities (*open coding*)	Critical stage of the open data lifecycle (5)	Critical stage of the open data lifecycle (5)
Key resources (*open coding*)	Critical resources (other than financial) (13)	*Deselected*
Customer relationship (6)	Customer relationship (4)	Customer relationship (4)
Value capture		
Cost structure (4)	Cost structure (2)	*Deselected*
Revenue streams (15)	Main revenue stream (13)	Main revenue stream (13)
	Source of revenue (3)	Source of revenue (3)

Note: Refer to Appendix G for the full version. Appendix H provides definitions for each codebook element by the end of Stage 3.

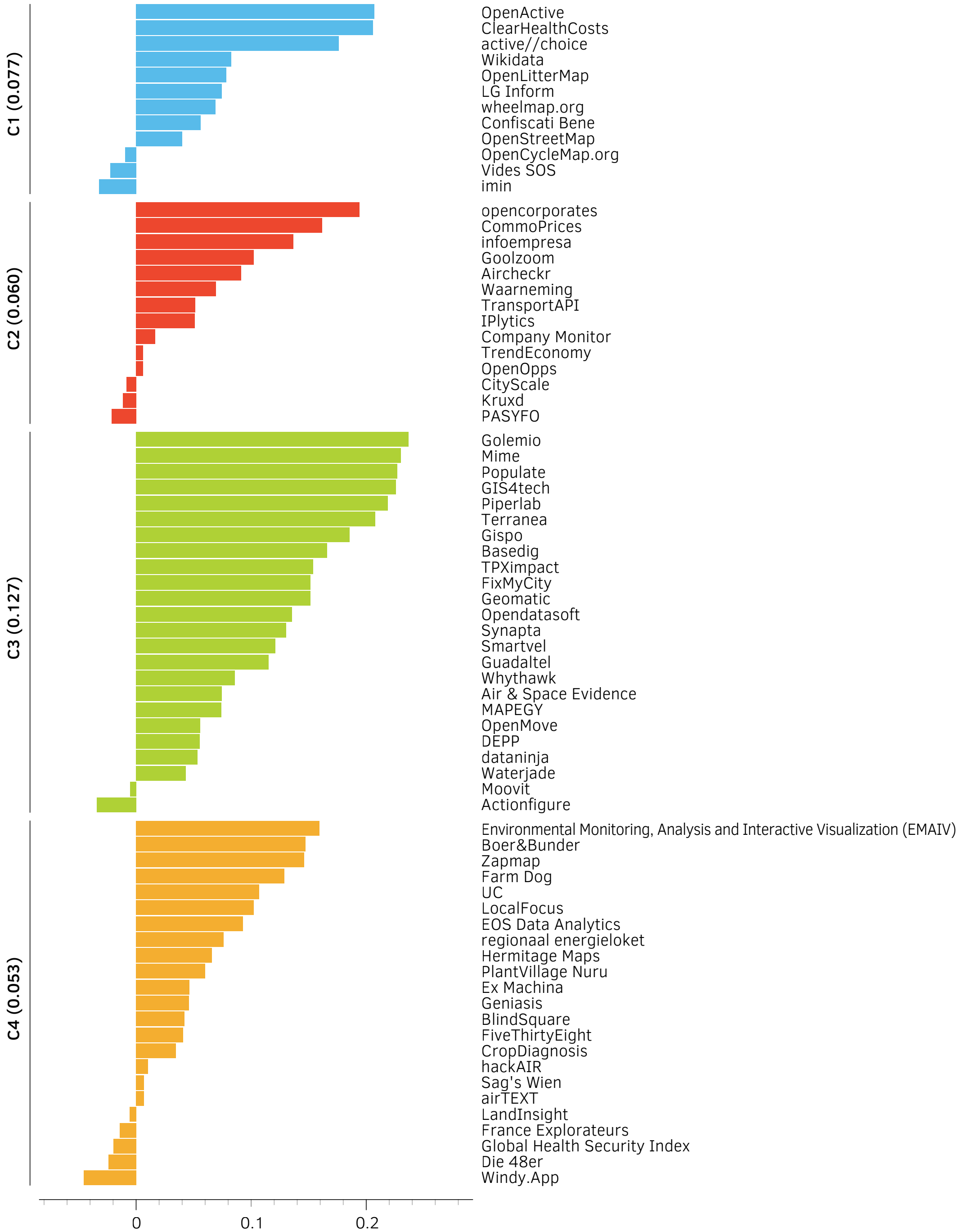


FIG. 4.3 Silhouette plot across clusters

>>>

TOC

FIG. 4.3 Silhouette plot across clusters

>>>

TOC

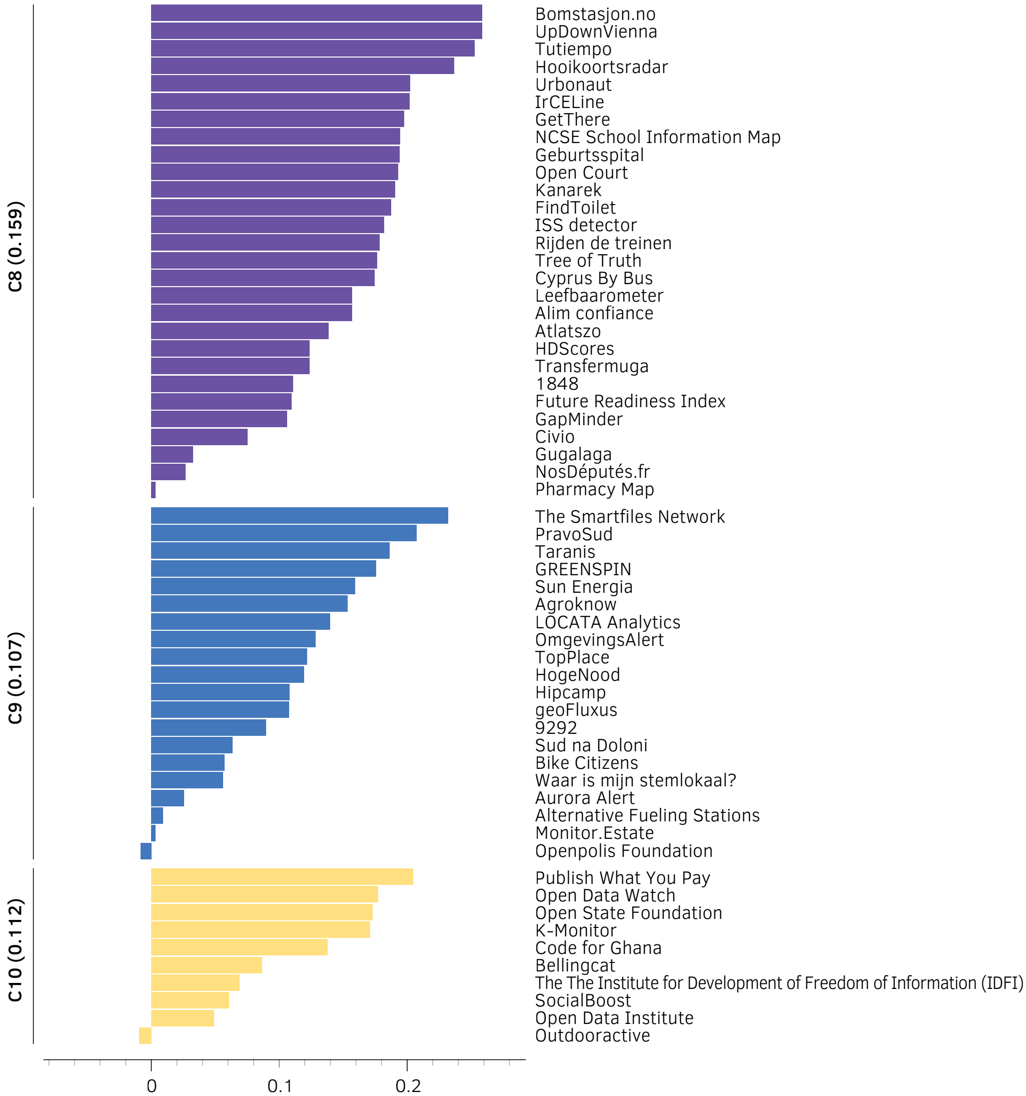


FIG. 4.3 Silhouette plot across clusters

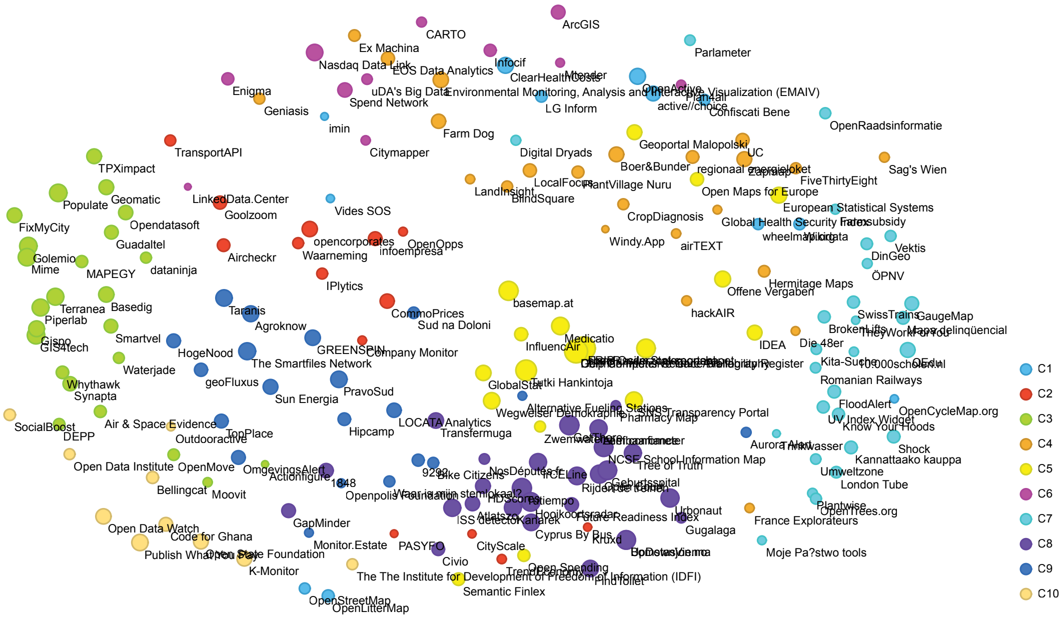


FIG. 4.4 Multidimensional scaling (MDS) of the clusters

4.2.4 Stage 4: Identification of business model archetypes

K-means clustering was used as an aid for identifying business model archetypes; the clustering results are not the final findings per se. In Stage 4, the clustering output of Stage 3 was interpreted, and the business model archetypes were identified. This was done by first identifying the categories within each cluster, of which most sample cases (at least 70% of the total cases) share the same element; they are called *defining categories*. Additionally, some similar elements within a category were grouped on a case-by-case basis, and consequently, the respective category was also considered a defining category. For example, if most sample cases within a cluster have the elements of either *crowdfunding*, *sponsorship: public*, or *sponsorship: private*, the cluster's *main revenue stream* category was considered a defining category since all three elements are based on external contributions (as opposed to transactions). Therefore, in this sense, the defining category is both a quantitative and qualitative indicator. Defining categories underscores the salient characteristics of sample cases within every cluster, which helps identify and describe the business model archetypes.

Appendix I presents the sample cases within each cluster (only cases with positive salient scores of at least 0.04 were listed) and their elements (with D indicating defining categories). The clusters are labelled from C1 to C10. Within the same cluster, the number of known sample cases across categories may differ because categories that were uncertain about were left blank during the data collection (Stage 2). Not all sample cases grouped within a particular cluster are meant to be in that cluster, since inaccurate cluster assignments are expected with ML. However, the objective of this study was to identify the archetypes that exist. Thus, inaccurate cluster assignments are a minor issue if most cases within each cluster exhibit commonalities.

4.3 Findings

Nine open data intermediaries' business model archetypes were identified. Two clusters (C7 and C8) from the K-means clustering were determined to be similar based on the defining variables (Appendix I); hence, they were combined to become a single archetype. Table 4.4 presents the archetypes, their salient characteristics, and several examples.

TABLE 4.4 Open data intermediaries' business model archetypes and salient characteristics

ID	Name	Salient characteristics based on value dimensions	Example	Cluster
A1	**Collaborative open data platform**	**Value proposition:** Open data platform freely available for both open data providers and users. **Value creation:** The critical open data stage is preparation, and the consumer relationship is collaborative (co-creation or community-based). **Value capture:** Funded by external contribution (crowdfunding or sponsorship)	Wikidata, Confiscati Bene, OpenStreetMap	C1
A2	**Paid self-service data delivery**	**Value proposition:** Augmented open data (i.e., in combination with non-open data) delivered via various types of products to data users. **Value creation:** The critical open data stage is preparation, and the consumer relationship is self-service. **Value capture:** Revenue generated from augmented open data through freemium or subscription models.	Opencorporates, Goolzoom, TransportAPI	C2

>>>

TOC

TABLE 4.4 Open data intermediaries' business model archetypes and salient characteristics

ID	Name	Salient characteristics based on value dimensions	Example	Cluster
A3	**Personalised open data service**	**Value proposition:** Multiple service units based on augmented open data, providing personalised services to open data providers and users. **Value creation:** The consumer relationship is personal assistance. **Value capture:** Revenue typically generated through service delivery.	FixMyCity, Opendatasoft, dataninja	C3
A4	**Interactive app with other complementary products**	**Value proposition:** Interactive app with other complementary products. **Value creation:** The critical open data stage is re-use, and the consumer relationship is self-service. **Value capture:** Revenue generated mainly from (augmented) open data via various means such as subscription fees, app sales, and sponsorship. Complementary products may enhance the benefit, visibility, or appeal of the interactive app.	Boer&Bunder, Zapmap, Local-Focus	C4
A5	**Open data repository funded by sponsorship**	**Value proposition:** Open data repository mainly targeted at generic open data re-users and is free. **Value creation:** The critical open data stage is the preparation, and the consumer relationship is self-service. **Value capture:** Funded by public or private sponsorship.	Tutki Hankinto-ja, FRIS Onder-zoeksportaal, basemap.at	C5
A6	**One-stop package around an (augmented) open data platform/ repository**	**Value proposition:** Multiple product units with complementary products, centred around a restricted data platform/repository based on augmented open data. The target consumers are typically (but not necessarily) highly skilled data users and providers. **Value creation:** The critical open data stage is the preparation, and the consumer relationship is self-service. **Value capture**: Revenue generated through subscription fees or software sales.	Nasdaq Data Link, ArcGIS, Enigma	C6
A7	**Single-purpose app**	**Value proposition:** Single-purpose app based on open data, targeting generic data users. **Value creation:** The critical open data stage is re-use, and the consumer relationship is self-service. **Value capture:** Various means of revenue generation, such as cross-subsidy and sponsorship.	QEdu, FloodAlert, SwissTrains	C7 and C8

>>>

TOC

TABLE 4.4 Open data intermediaries' business model archetypes and salient characteristics

ID	Name	Salient characteristics based on value dimensions	Example	Cluster
A8	**Interactive app without complementary products**	**Value proposition:** Interactive app without other complementary products. **Value creation:** The critical open data stage is re-use, and the consumer relationship is self-service or personal assistance. **Value capture:** Various means of revenue generation, such as subscription fees, brokerage, or app sales.	Taranis, geo-Fluxus, 9292	C9
A9	**Open data advocacy**	**Value proposition:** Multiple units of open data advocacy, campaigning, or lobbying services. **Value creation:** Various critical stages of the open data lifecycle and various forms of consumer relationships. **Value capture:** Mainly funded via external contributions (sponsorship or crowdfunding), but in some cases, through service delivery.	Publish What You Pay, Open State Foundation, Open Data Institute	C10

Note: Definitions for business model elements in every value dimension, including types of products (e.g., data platform, data repository, single-purpose app), stages of the open data lifecycle (e.g., preparation, re-use), and revenue streams (e.g., sponsorship, service delivery) are provided in Appendix H.

Archetype A1 is a *collaborative open data platform*. Archetype A1 offers a free open data platform for anyone to contribute and use open data. Since the platform is free, this business model captures value through external contributions (i.e., through crowdfunding or sponsorship, instead of market transactions). Value is created collaboratively through co-creation (where a lead body facilitates the contribution and use of open data) or through community-based organising (where all members, at least theoretically, have more or less equal opportunity to influence how open data is contributed and used).

Archetype A2 is *paid self-service data delivery*. The core value proposition of archetype A2 lies in the augmentation of open data with non-open data, delivered through various means, such as data repositories, APIs, or direct transfers. The augmented data are not offered for free, with revenue typically being captured via freemium or subscription models.

Archetype A3 is a *personalised open data service*. This business model helps open data providers or users with their open data-related activities. Since the services offered are personalised, the business model typically consists of multiple service units instead of a single product to cater to the diverse needs of open data providers and users. The consumer relationship is based on personal assistance, and revenue is obtained through service delivery.

Archetype A4 is an *interactive app with other complementary products*. This type of app allows dynamic interactions instead of static ones. It may be based on entirely open data or in combination with non-open data. At the heart of this business model are other open data-based or non-open data-based complementary products (e.g., other apps, data platforms, and advisory services). In this regard, this business model does not rely entirely on a single product to generate revenue. The complementary products enhance the benefit, visibility, or appeal of the interactive app. Revenue is obtained through various means, such as subscription fees, app sales, and sponsorship.

Archetype A5 is an *open data repository funded by sponsorship*. Compared to a data platform, which is two-sided and where multiple suppliers and users can offer and use the data on/from it, a data repository is one-sided, and only a limited number of suppliers can provide data on it. This archetype is relatively straightforward. It is funded by public or private sponsorship. The critical open data stage is the preparation, and the consumer relationship is self-service.

Archetype A6 is a *one-stop package around an (augmented) open data platform/ repository*. This archetype is a one-stop package with modular service units (e.g., various data analysis, visualisation, and dissemination tools) built around an open data platform/repository with (augmented) open data, where data users/providers can select the needed functionalities. The study's analysis shows that it typically targets highly skilled data users and providers in professional domains. The critical open data stage is the preparation, and the consumer relationship is self-service. Revenue is generated through subscription fees or software sales.

Archetype A7 is a *single-purpose app*. These apps are typically only based on open data (i.e., not in combination with non-open data) and have limited functionalities. The critical open data stage is re-use, and the consumer relationship is self-service. The app may be free or come at a cost. The revenue is obtained through various means, such as cross-subsidy, a freemium model, subscription fees, or sponsorship. The app may also be developed by volunteers, for which no funds were collected.

Archetype A8 is an *interactive app without complementary products*. The slight difference between archetypes A8 and A4 is the absence of complementary products. Archetype A8 may be self-sufficient or sufficiently viable without complementary products. Moreover, it was discovered that archetype A4 typically relies on augmented open data (i.e., in combination with non-open data), whereas archetype A8 relies solely on open data. The critical open data stage is re-use, and the consumer relationship is self-service or personal assistance. Revenue is generated via various means, such as subscription fees, brokerage, or app sales.

Lastly, archetype A9 is *open data advocacy*. The value offered through this archetype is advocacy or lobbying for open data policies, provision, and re-use through engagement with various open data stakeholders. Funding is mainly obtained via external contributions (sponsorship or crowdfunding); however, in some cases, revenue is generated through service delivery.

The typical types of organisations for each archetype were also analysed based on the number of sample cases (with a silhouette score of at least 0.04) within each archetype (Table 5). This insight does not rule out the possible adoption of a particular archetype for certain types of organisations; instead, it merely represents the common adoption currently in practice. The following was discovered:

- Archetype A1 primarily consists of companies and NGOs/NPOs.
- Archetypes A2, A3, A4, A6 and A8 mostly comprise companies.
- Archetype A5 mainly consists of public organisations.
- Archetype A7 comprises companies, civic hackers/citizen researchers, NGOs/NPOs, and public organisations.
- Archetype A9 consists of NGOs/NPOs.
- The most popular archetype for civic hackers/citizen researchers is archetype A7.
- The most popular archetypes for companies are archetypes A3, A7 and A8.
- The most popular archetypes for NGO/NPOs are A7 and A9.
- The most popular archetype for public organisations is A5.

TABLE 4.5 Type of organisations in each archetype

Type of organisation	A1	A2	A3	A4	A5	A6	A7	A8	A9
Civic hackers/ citizen researchers	1	0	0	0	1	0	12	0	0
Company	4	7	22	12	1	5	15	15	0
Cooperative/ association	0	0	0	0	0	0	1	0	0
Joint - Public and NGO/NPO	0	0	0	0	0	0	0	1	0
NGO/NPO	3	1	0	1	3	0	8	0	9
Public organisation	0	0	0	1	11	0	7	0	0
Public-affiliated organisation	1	0	0	0	1	0	0	0	0
University/ research organisation	0	0	0	0	3	0	0	0	0

4.4 Discussion

In this section, the empirical, theoretical, and practical implications of the findings are discussed, and the study's limitations are elaborated.

4.4.1 Empirical implications

The contribution of this study is affirmed by comparing the empirical findings with open data business model archetypes in the literature (not only open data intermediaries specifically, but also other open data roles generally). Notably, archetype A1 was not previously identified, whereas archetypes A2 to A9 are more refined in comparison to the business model archetypes described in the literature, especially in terms of expounding elements in all three business model value dimensions (i.e., value proposition, value creation, and value capture). Below, the specific comparisons made for each archetype are described.

Archetype A1 is a *collaborative open data platform*. This business model was not captured in the existing literature reviewed. Although the *data platform* business model identified by Magalhaes and Roseira (2020) may seem similar to archetype A1, the former is described as having 'a higher level of interactivity, thereby offering users the ability to effectively explore open government datasets' (Magalhaes and Roseira, 2020, p. 7). This description lacks detail in terms of the value creation and value capture dimensions, is limited to government data, and is also fundamentally different from the bottom-up nature of archetype A1. The *increasing quality through participation* business model identified by Ferro and Pizzamiglio (2023) may come close to archetype A1 even though the latter is not limited to quality enhancement but also includes the contribution in terms of the data itself, standards, and governance aspects.

Archetype A2 is *paid self-service data delivery*. The *data refining* and *data structuring* business models identified by Magalhaes and Roseira (2020), the *integrators* identified by Magalhaes et al. (2014), and the *information aggregators* noted by Janssen and Zuiderwijk (2014) come close to archetype A2. However, although the four business models described in the literature highlight the open data pre-processing value proposition, the data involved is solely (or at least mainly) open data. This contrasts with archetype A2, which foregrounds the combination of open data with non-open data.

Archetype A3 is a *personalised open data service*. This archetype is comparable to the *enablers* identified by Magalhaes et al. (2014) but is more specific in that the personalised relationship with consumers is at the heart of it. Thus, for archetype A3, the value is captured by open data intermediaries through service delivery fees. In contrast, the *enablers* noted by Magalhaes et al. (2014) are rather generic.

Archetype A4 is an *interactive app with other complementary products*. This archetype is a specific subset of the *interactive apps* business model identified by Janssen and Zuiderwijk (2014) and Magalhaes and Roseira (2020). The complementary products offered in archetype A4 support open data intermediaries to improve the benefit, visibility, or appeal of the interactive app.

Archetype A5 is an *open data repository funded by sponsorship*. This archetype is similar to the *open data repositories* identified by Janssen and Zuiderwijk (2014). While the two are identical, archetype A5 accentuates that the value capture is mainly based on sponsorship.

Archetype A6 is a *one-stop package around an (augmented) open data platform/ repository*. This archetype resembles the *service platforms* identified by Janssen and Zuiderwijk (2014). Nevertheless, archetype A6 described is more refined in clarifying subscription fees and software sales as the typical means for value capture.

Archetype A7 is a *single-purpose app*. This archetype is the same as the *single-purpose apps* identified by Janssen and Zuiderwijk (2014) and Magalhaes and Roseira (2020).

Archetype A8 is an *interactive app without complementary products*. Like archetype A4, this archetype is also a specific subset of the *interactive apps* business model identified by Janssen and Zuiderwijk (2014) and Magalhaes and Roseira (2020).

Archetype A9 is *open data advocacy*. This archetype is similar to the *advocacy* (for open data providers) and *consultancy* (for open data re-users) business models identified by Magalhaes and Roseira (2020). However, this study's findings suggest that the two business models identified by Magalhaes and Roseira (2020) typically come together as a single archetype (A9) in practice, i.e., open data intermediaries adopting this business model work simultaneously with providers, users, and other relevant stakeholders.

4.4.2 Theoretical implications

The findings of this study offer a broader overview (i.e., nine archetypes) and more detailed account (i.e., across the value proposition, value creation, and value capture dimensions) of existing open data intermediaries' business model archetypes. This knowledge is missing from the literature. Past studies (Germano et al., 2016; Janssen & Zuiderwijk, 2014; Magalhaes & Roseira, 2020) fall short of integrating the three key business model dimensions, even though business model scholars generally consider the three dimensions as the foundation of a business model (Afuah, 2018; Andreini & Bettinelli, 2017; Teece, 2010; Voigt et al., 2017). This study tackled such shortcomings. Furthermore, this study is based on a large number of cases across many countries, whereas previous studies are based on only a handful of cases and/or a single country.

This study's findings offer the groundwork for further inquiries into how open data can be better integrated into business models (Gurin, 2014), factors that could contribute to the success of such business models (Lambert & Davidson, 2013) and the conditions that could support such business models to sustainably deliver value to other open data stakeholders (Hossain et al., 2016; Jetzek et al., 2019; van Veenstra & van den Broek, 2013). Further research may also investigate whether certain business model archetypes are feasible for certain types of organisations (e.g., can open data intermediaries from public organisations adopt a *personalised open data service* even though this study showed no public organisations from the sample cases adopt that archetype). Furthermore, through the development of the codebook, this study offers categories and elements to consider in the research and development of open data intermediation business models.

This study shows that open data intermediation business models may be based on integrating open data with non-open data. In this case, open data is a crucial component of the business model, where the products offered would not be viable only with non-open data. This affirms the potential of open data in multiplying the creation of new products by serving as the basic data that organisations add their non-open data on. This also challenges the claim that the value of open data is 'often meager' (p. 1) if the evaluation is solely based on organisations that have 'participated in open data events' and/or 'had received government funding for open data projects' (p. 3) (Temiz et al., 2022), since the value that open data enables may not be obvious in the end-products or solely generated through open data events or projects.

TOC

Furthermore, this study found that many open data intermediaries also offer other open data-based or non-open data-based products. For certain archetypes (A1, A4 and A6), these other products are complementary to the main open data products. For archetype A6 in particular, the main source of revenue is in fact from the other products instead of the main open data products. This shows that open data intermediation business models do not have to rely on generating revenue solely or mainly from open data products. In addition, it was also confirmed that contributions of open data intermediaries can happen at various stages of the open data lifecycle (i.e., identification, preparation, publication, re-use, and evaluation). Therefore, open data intermediaries are not merely the 'bridge' between open data providers and users, as some (Brugger et al., 2016; Meng, 2016) have implied. This means that open data intermediaries can offer a lot more potential value propositions, consistent with the outlook by Wiener et al. (2020) that moving forward, organisations would rely more on external parties to maximise the potential of data, as they may lack the necessary internal capabilities.

This study's findings show that while companies employ many different business model archetypes, public organisations and NGOs/NPOs only rely on a handful of archetypes. Moreover, the few archetypes that they adopt are mostly based on sponsorship. This calls into question the sustainability and the innovativeness of such organisations in serving the role of open data intermediaries. Hence, the development of new and innovative business models specifically for public organisations and NGOs/NPOs deserves more attention.

The business model archetypes of open data intermediaries discovered are starkly different from those of data intermediaries that mainly deal with non-open data, as recently identified by Schweihoff et al. (2024). They found nine patterns of data intermediation services, and all, except one, offer services related to data control, consent management, or identity management. Neither of these aspects is particularly focal in the open data intermediation business model in the context of this study. The more crucial aspect is gaining benefits from intermediating data that is already freely reusable by everyone under an open license (i.e., not requiring registration, consent, or identity verification), while carving a space or maintaining a competitive advantage in the respective sector or market. This ascertains the peculiarity of open data intermediaries compared to generic data intermediaries. Having said that, certainly, organisations that operate as open data intermediaries can also intermediate non-open data simultaneously, i.e., employing multiple business models.

4.4.3 Practical implications

For practice, this study's findings offer insights for existing and potential open data intermediaries on the business model they can adopt and, indirectly, encourage more private and public value exploitation from open data (Robinson & Mather, 2022; van Loenen et al., 2021). Such knowledge may be particularly illuminating to public organisations and NGOs/NPOs since they currently only employ a limited number of business model archetypes. At the same time, their role as open data intermediaries is needed to ensure public interests are accounted for in open data value generation. The findings can also help open data intermediaries and other actors discern their value network configuration (i.e., how they are interrelated) (Lindman et al., 2014), which can be used to explore new opportunities and forge new relationships. Meanwhile, policymakers can use the findings to support policy design related to open data intermediaries, which is consistent with the call by Meijer et al. (2014) for governments to acknowledge the heterogeneous open data incentives, practices, and consequences.

Furthermore, the codebook that has been developed in the process of identifying the archetypes is in and of itself useful to practitioners as a morphological box for designing and experimenting with new open data intermediation business models, by mixing and matching different elements across different categories. In other words, the codebook has the potential to be turned into the equivalent of Osterwalder & Pigneur's (2010) business model canvas, but more granular and specific for open data intermediaries.

For potential funders of open data intermediaries, either from the public or private sector, that are assessing the viability of businesses or projects proposed by open data intermediaries, the findings provide an overview of the existing business models for reference. They can look for similar open data intermediaries within a particular archetype and identify critical aspects to consider while making funding decisions, including potential competitors. Such knowledge may give funders more confidence to fund open data intermediaries and support the emergence and growth of open data innovation.

TOC

4.4.4 Study limitations

This study's findings are limited by the data used. First, the sample cases are based on use cases gathered from data.europa.eu and thus predominantly (but not only) operate in Europe. Thus, further research could investigate whether other existing open data intermediaries' business model archetypes were not captured from this study's sample cases. Second, the qualitative data gathered from the sample cases relied on some interpretation and best effort. Notably, there might be overlooked aspects that could offer more insights into the business models of open data intermediaries. Despite this, this study only aimed to provide a bird's-eye view of the existing business model archetypes. Further research, especially based on qualitative methodologies such as in-depth case studies, can be used to investigate open data intermediation business models more deeply and capture missing nuances. Third, this study is informed by existing real-world open data intermediation business models. However, business models evolve, following technological, market, and regulatory conditions (de Reuver et al., 2009; Şimşek et al., 2022). This study could not predict future business models.

4.5 Conclusion

This chapter addresses the knowledge gap on the existing business model archetypes of open data intermediaries. This gap has limited our understanding of the conditions and potential innovations required to perform the roles of open data intermediaries sustainably. Through a multimethod approach based on 190 sample cases of open data intermediaries from various countries, nine business model archetypes were identified: collaborative open data platform; premium self-service data delivery; personalised open data service; interactive app with other complementary products; open data repository funded by sponsorship; one-stop package around an (augmented) open data platform/ repository; single-purpose app; interactive app without complementary products; open data advocacy. For each of these archetypes, identified its value proposition, value creation and value capture dimensions (three dimensions considered to be fundamental for a business model) were identified. Thus, the findings are based on a large number of cases while also being detailed in the description of each archetype.

The methodology that this study employed in and of itself can be helpful for researchers to systematically identify business model archetypes. Furthermore, the codebook that have been developed in the process of identifying the archetypes has the potential to be turned into a more detailed and tailored equivalent of Osterwalder & Pigneur's (2010) business model canvas, specifically for open data intermediaries. The findings from this chapter can support further research into the conditions that contribute to the success of open data intermediation business models and the development of new ones. They could also inspire (potential) open data intermediaries regarding the business models to adopt and policymakers in designing policies based on a good understanding of open data intermediation business models in practice. Ultimately, this can help optimise private or public value generation from open data.

PART C

Delving deeper into selected open data intermediation business models

The trees

Chapters 5 and 6 investigate in more detailed 2 out of 9 open data intermediation business model archetypes identified in Part B through two in-depth case studies; akin to looking at individual trees within forests. These chapters aim to recommend aspects to consider in developing open data intermediation business models that support sustainable ODE. The cases studied are Esri in Chapter 5, representing the one-stop package around an (augmented) open data platform/repository, and OpenStreetMap (OSM) in Chapter 6, representing the collaborative open data platform.

Both Esri and OSM are well-established entities that play an important role in open data intermediation in the geospatial data domain. Four out of six thematic categories of the EU Open Data Directive's high-value datasets constitute (geo)spatial datasets, thereby affirming the importance of such data. Besides, as Gray (2014) pointed out in his work on the genealogy of open data, the geospatial domain had dealt with various issues and controversies related to open data years before they were encountered in other domains. By focusing on a specific domain through the selection of the two case studies, the role of open data intermediaries can be understood more deeply by having a close familiarity with the context in which they operate.

5 The case of Esri

The core content of this chapter was previously published as Shaharudin, A., van Loenen, B., & Janssen, M. (2025). Developing an Open Data Intermediation Business Model: Insights From the Case of Esri. *Transactions in GIS*, *29*(1), e13304. https://doi.org/10.1111/tgis.13304.

5.1 Introduction

This chapter addresses the RQ4: What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem? Since there are nine open data intermediation business model archetypes identified in the previous chapter, this present chapter answers the research question specifically for the archetype of one-stop package around an (augmented) open data platform/ repository (archetype A6 in Chapter 4). Towards that end, the case of Esri is selected. The company has been involved in (open) data intermediation[3] since the 1990s, and its software package, ArcGIS, leads the global market share of geographic information system (GIS) software. Hence, Esri stands to offer illuminating insights for answering the said research question.

This chapter tackles three specific research objectives: (1) to detail Esri's open data intermediation business model, (2) to consider the current strengths and weaknesses, and potential opportunities and threats, of Esri's open data intermediation business model to the ODE, and (3) to formulate aspects to consider in developing an open data intermediation business model that can contribute to the sustainability of the ODE, specifically for the one-stop package around an (augmented) open data platform/ repository.

3 The term 'open' is placed in parentheses because, before 2010, the term 'open data' (or even the term 'open access', which was more common then) was not apparent in the descriptions of any Esri's data-related services. Although the company had been facilitating the access and reuse of free-of-charge data since the late 1990s, the legal rights of the data were not clearly defined.

The remainder of this chapter is organised as follows. Section 5.2 delves into a selection of relevant organisational and management theories that provide vantage points into Esri's business model and its implications to other ODE actors. Section 5.3 presents the methodology of this chapter. Section 5.4 briefly describes Esri and the evolution of its (open) data intermediation. Section 5.5 presents Esri's open data intermediation business model (Objective 1). Section 5.6 presents the analysis of the strengths, weaknesses, opportunities, and threats of Esri's open data intermediation business model to the ODE (Objective 2). Section 5.7 formulates the aspects to consider in developing an open data intermediation business model that supports the sustainability of the ODE (Objective 3). Finally, Section 5.8 concludes the chapter.

5.2 Conceptual framework

Open data intermediation business models exist in various shapes and forms (Janssen & Zuiderwijk, 2014; Magalhaes & Roseira, 2020). Esri's open data intermediation business model represents the one-stop package around an (augmented) open data platform/ repository archetype. This archetype is described as offering multiple product/service units around a (restricted) data platform/ repository based on (augmented) open data (Chapter 4). Augmented open data refers to open data that is enhanced by combining it with non-open data. Revenue for this archetype is generated through subscription fees or software sales. Other examples of this archetype are CARTO, Nasdaq Data Link, and Enigma.

Several strategic management and organisational theories and frameworks may offer guidance or explanations for business model design and innovation. One is the resource-based view (RBV) (DaSilva & Trkman, 2014; Hedman & Kalling, 2003), which postulates that organisations can maintain a sustained competitive advantage by leveraging their valuable, rare, imperfectly imitable, and non-substitutable (VRIN) resources (Barney, 1991). Resources can take the form of an organisation's assets, capabilities, organisational attributes, brand, and knowledge, among other examples (Barney, 1991). Information technology (IT) capabilities, including IT infrastructure and skills, can also be VRIN resources (Seddon, 2014). The literature on RBV has also expanded to consider inter-organisational networks, partnerships, and social capital as a type of VRIN resource (Eisenhardt & Schoonhoven, 1996; Gulati et al., 2000; Lavie, 2006; Yi et al., 2022).

Another relevant concept related to the business model is that of the value driver (Amit & Zott, 2001; Leppänen et al., 2023; Spieth et al., 2019; Visnjic et al., 2017). Value drivers are broad dimensions of attributes leveraged to attract and retain customers. Amit & Zott (2001) identified four value drivers, namely novelty (e.g., new content, new structures, and new participants), lock-in (high switching costs and positive network externalities), complementarity (e.g., between products and services, between technologies, and between activities), and efficiency (e.g., low transaction costs, simplicity, speed). These value drivers are not mutually exclusive. Visnjic et al. (2017) proposed the fifth value driver, accountability, where an organisation can help manage or eliminate risks and internalise 'unmanageable' risks of its customers.

Organisational identity theory may also explain or guide business model design and innovation (Bojovic et al., 2020; Kohtamäki et al., 2019; Snihur, 2016). The theory suggests that organisational identity (i.e., 'who we are as an organisation') informs strategic and organisational decisions (Gioia et al., 2013; Kohtamäki et al., 2019), including the business model. An organisational identity consists of three characteristics: central (some features are believed to be fundamentally core to the organisational identity, and are thus deliberately preserved), enduring or continuous (the identity is deemed stable over time, not necessarily from the eyes' of outsiders but from the perspective of the members), and distinctive (where the organisation sees itself simultaneously similar to some desirable referent group, such as an industry, but also notably different from members of the group) (Albert & Whetten, 2003; Gioia et al., 2013; Whetten, 2006). Organisations involved in IT businesses are particularly confronted with the need to maintain their well-established identity on the one hand, and adapt to the rapidly changing environment on the other (Wang et al., 2013).

5.3 Research methodology

This chapter and Chapter 6 employ the single-case study methodology (R. K. Yin, 2018). Case studies are used to derive new insights into topics for which existing studies and evidence are scarce (Gerring, 2006), as in the topic of the open data intermediation business model. A single-case study, in particular, is deemed appropriate when the case is remarkably revelatory or exemplar (Eisenhardt & Graebner, 2007; Siggelkow, 2007; R. K. Yin, 2018). A single-case study allows for a deep contextualised understanding of the case in question through 'thick' descriptions, that may be difficult to achieve through a multiple-case study (Dyer & Wilkins, 1991; Siggelkow, 2007).

Esri is a market leader in GIS and has long been an open data intermediary. The highly elaborate ways in which Esri offers, creates, and captures value from or with open data and the prevalence of Esri products in the geospatial domain warrant its investigation through a single-case study. Esri distributes its software outside of the US through local companies called distributors. While Esri Inc. (its parent company in the US) devises the overarching global mission regarding open data for the multinational entity, each distributor devises and implements its specific local strategies. Hence, to more fully understand how Esri plays a role as an open data intermediary, data was gathered not only from Esri Inc., but also from five of its distributors, namely Esri Germany, Esri Netherlands, Esri Spain, Esri United Kingdom (UK), and Geoinfo Denmark (Esri's distributor in Denmark). These five distributors were selected because they operate in countries with a considerable level of open geospatial data, driven by both the EU Directive on the Re-use of Public Sector Information (Directive 2003/98/EC) that came into force in 2003 (recast in 2019 as the EU Open Data Directive) and the EU Directive for Infrastructure for Spatial Information in the European Community (INSPIRE) (Directive 2007/2/EC) that came into force in 2007 (Haarsma, 2012; Minghini et al., 2021; van Loenen & Grothe, 2014).

Data from 27 semi-structured interviews was gathered involving 29 participants (between April 2023 and April 2024) and publicly accessible sources, such as materials on websites (including archival materials via Wayback Machine), videos, and audio podcasts (cited accordingly). The interviewees include a top-level senior executive from Esri Inc. and representatives from the above-listed five distributors. Moreover, representatives from eight geospatial data providers (including key persons from national mapping agencies, a city council, and OpenStreetMap (OSM) Foundation), ten Esri products users (from research institutions and the industry), a key person from a national geospatial data coordination organisation, and a key person from the Open Source Geospatial Foundation (OSGeo) were also interviewed. OSM is an open geodatabase project updated and maintained by a community of volunteers, and the OSM Foundation is a not-for-profit organisation that supports the OSM project (OSM Foundation, 2024h). OSGeo is a not-for-profit organisation that fosters the adoption of open source geospatial software (OSGeo, n.d.). The de-identified interview transcripts are available at https://doi.org/10.4121/f86d0e4c-851f-4378-a1bc-41210235ad61.

TABLE 5.1 Identification of interviewees

Interviewee ID	Organisation	Country	Role
01-Esri-A	Esri distributor	A	Consultant
02-Esri-B	Esri distributor	B	Content manager
03-Esri-C	Esri distributor	C	Content manager
04-Esri-C	Esri distributor	C	Content manager
05-Esri-D	Esri distributor	D	Content manager
06-Esri-D	Esri distributor	D	Content manager
07-Esri-E	Esri distributor	E	Content manager
08-Esri-E	Esri distributor	E	Marketing manager
09-Esri-O	Esri Inc.	N/R	C-level executive
10-StO-C	National geospatial data coordination organisation	C	Manager
11-Prov-B	Data provider (municipal)	B	Geospatial data manager
12-Prov-E	Data provider (national)	E	Geospatial data consultant
13-Prov-E	Data provider (national)	E	Geospatial data manager
14-Prov-D	Data provider (national)	D	Geospatial data manager
15-Prov-D	Data provider (national)	D	Geospatial data manager
16-Prov-C	Data provider (national)	C	Geospatial data manager
17-Prov-B	Data provider (national)	B	Geospatial data manager
18-Prov-O	OSM Foundation	N/R	Board member
19-OSG-O	OSGeo	N/R	Chair of a local chapter
20-User-E	Research institution	E	Esri user
21-User-E	Company	E	Esri user
22-User-A	Research institution	A	Esri user
23-User-D	Research institution	D	Esri user
24-User-C	Company	C	Esri user
25-User-C	Company	C	Esri user
26-User-C	Research institution	C	Esri user
27-User-D	Research institution	D	Esri user
28-User-C	Research institution	C	Esri user
29-User-C	Research institution	C	Esri user

Note: N/R means not relevant, as these interviewees represented a more global perspective.

Table 5.1 presents the identification (ID) of the interviewees to facilitate the presentation of the findings. The names of the interviewees' organisations and countries were omitted to avoid the re-identification of the interviewees, since the geospatial data community in some countries is close-knit. Throughout this chapter, depending on circumstances (e.g., for brevity or whether the statements cited may be deemed controversial), either the name of the organisation or the ID of the interviewees was used, but never both at the same time.

The data collected was analysed based on the abductive approach (Dubois & Gadde, 2002), which seeks to overcome the disengagement between theory and reality through systematic combining, i.e., going back and forth between the theories, data, and analysis. The strengths, weaknesses, opportunities, and threats of Esri's open data intermediation business model were considered from the overall ODE point of view, which were then used to synthesise aspects to consider in developing open data intermediation business models. The theories and frameworks discussed in Section 5.2 (i.e., RBV, value drivers, and organisational identity) as well as the (provisional) features of a sustainable ODE (Section 1.1.3) and the potential contributions of open data intermediaries identified in Chapter 3 are used to guide the identification of the aspects to consider.

Note that while this dissertation is inspired by the SWOT (Strengths-Weaknesses-Opportunities-Threats) analysis that is popular in management studies (Helms & Nixon, 2010), the analysis is based on the temporal dimension (current strengths and weaknesses, and potential opportunities and threats) rather than internal versus external dimensions (i.e., the manner in which the SWOT analysis is conventionally used). Hence, the analysis in this chapter is not referred to as a SWOT analysis[4].

Furthermore, open data intermediation business model exists and can be designed in various shapes and forms (Chapter 4), and Esri's business model represents only a specific archetype (i.e., one-stop package around an (augmented) open data platform/ repository). Hence, not all insights from the Esri case may be directly transferable to all archetypes of open data intermediation business models, but they may still set the groundwork for theorising other archetypes.

4 The same to be said for the analysis of the OSM case study in Chapter 6.

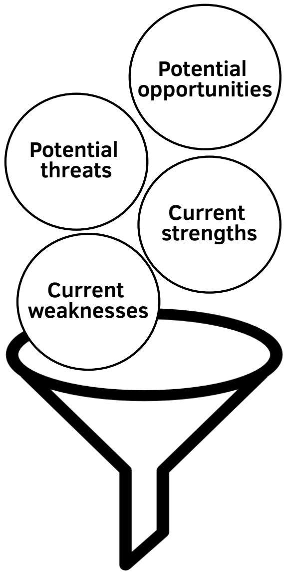


FIG. 5.1 Process of synthesising aspects to consider in developing open data intermediation business models based on the case study

5.4 Background: Esri and its evolution as an (open) data intermediary

5.4.1 Esri: The company's background

Esri is a multinational GIS software company headquartered in Redlands, California. It is a global market leader in GIS, with its software suite, ArcGIS, used by over 350,000 organisations, including 90% of Fortune 100 companies, over two-thirds of Fortune 500 companies, many national governments, approximately 30,000 cities and local governments, and roughly 12,000 nonprofit organisations (Esri, n.d.-a). The software is used in over 60 industries, including banking, retail, transportation, utilities, government, and health and human services (Geospatial Media and Communications, 2018). Esri has an annual revenue of over $1.3 billion (Hoffman, 2021). In 2018, Forbes valued the company at $5.5 billion (Daniel, 2018).

The company was founded in 1969, making it one of the oldest software companies, older even than SAP (1972), Microsoft (1975), Apple (1976), and Oracle (1977). It was founded as Environmental Systems Research Institute Inc. (ESRI) by Jack Dangermond, a landscape architect, and his wife, Laura Dangermond, a social scientist. They both worked at the Harvard Laboratory of Computer Graphics and Spatial Analysis, where the early development of computer map-making took place. Before commercialising its GIS software, the company started as a consulting firm helping land use planners and resource managers make informed decisions based on computer mapping and spatial analysis. Unlike most software companies, Esri has always been privately held – without outside investors, liquidity events, or stock options (Esri, n.d.-a; Hoffman, 2021).

Esri distributes its software outside the US through locally-owned companies called distributors. They have exclusive rights to distribute ArcGIS in their countries. They have been either independent from Esri Inc. from the start, or Esri Inc. made a small investment in them (Hoffman, 2021). Besides reselling ArcGIS, Esri distributors support local customers by addressing their specific queries and needs, as well as by facilitating the broad local development of GIS in the country. Esri Inc. and its distributors, while operating individually, often collaborate and exchange ideas about marketing, business development, and technology support (Hoffman, 2021).

ArcGIS products run on desktops, mobile devices, and the cloud (Esri, n.d.-e). ArcGIS Pro, ArcGIS Online, and ArcGIS Enterprise are some of the company's most popular products. ArcGIS Pro is a desktop GIS application that supports data visualisation, advanced analysis, and authoritative data maintenance in 2D, 3D, and 4D. ArcGIS Online is a cloud-based software for creating and sharing interactive web maps. ArcGIS Enterprise is a complete mapping and data management server software used to create maps, and analyse and share data. It can be deployed on-premises or in the cloud.

5.4.2 The evolution of Esri's (open) data intermediation

The term 'open' is placed in parentheses because, before 2010, the term 'open data' (or even the term 'open access', which was more common then) was not apparent in the descriptions of any Esri's data-related services. Although the company had been facilitating the access and re-use of free-of-charge data since the late 1990s, the legal rights of the data were not clearly defined (at least, based on accessible archival materials). Only in 2010 did the Community Maps Program state that organisations seeking to contribute data through the programme must provide 'royalty-free redistribution' data at no cost (Esri, 2010a), which reflects the concept of open data (despite the term itself having not been explicitly used). Besides, Esri also intermediates non-open data through its services, although such activities are not the focus of this dissertation.

Esri has long recognised that data is at the heart of GIS application and innovation and, thus, crucial for its products' continuous growth and relevance (Esri Events, 2016). Figure 5.2 shows the archived webpage of Esri captured in November 1996 that reads, 'Just as a car won't run without gasoline, a GIS without data has no information' (Esri, 1996). In the early 1990s, when geospatial data was shared as files on CDs and other media, such as FTP (file transfer protocol), Esri offered best practices and data models for efficient data sharing (Esri Events, 2016). Later, in 1996, only a few years after the launch of the World Wide Web, Esri introduced the ArcData Publishing Program, where users could download hundreds of ready-to-use datasets from Esri and other companies, some free of charge (Esri, 1996) (Figure 5.2). Esri also introduced Data Hound in 1998, a search engine that brought users to external websites offering freely downloadable data compatible with Esri software (Esri, 1998) (Figure 5.3).

In June 2000, Esri introduced Geography Network (Figure 5.4), a website from which to access, share, and download geographic content from around the world. While most of the content was freely downloadable or at least viewable, some were commercial. Whenever a commercial map was viewed or a commercial dataset downloaded, a charge was recorded in the Geography Network e-commerce system, and Esri would bill users and pay content providers (Dempsey, 2000; Esri, 2000). The Geography Network was an enhanced consolidation of Esri's previous data-related services (ArcData and Data Hound). Data Hound was discontinued by 2001 and ArcData by 2003 (Esri, 2001, 2003).

TOC

Additional offerings of the Geography Network included the Live Map Services, where several map services[5] were available on a subscription basis, and the Map Exchange, where users could share static map images on the website (Geography Network, 2000b). The Geography Network was retired in December 2009 as the services it provided were then incorporated into a web-based ArcGIS Online software package launched in the same year (Geography Network, 2009).

In 2010, Esri introduced the Community Maps Program (Esri, 2010a), where ArcGIS organisational users can share their local data to improve the suite of basemaps[6] created and hosted by Esri. Since then, the company has accepted, processed, and published hundreds of millions of vector data (e.g., roads, buildings, addresses) and tens of millions of square kilometres of raster data (e.g., imagery, digital elevation models) (Esri, n.d.-d; Kensok, 2020b). The basemaps also include OSM data (Kensok, 2020a).

In 2014, just about a year after US President Obama signed an executive order to make open data the new default for government information (Obama, 2013), Esri introduced ArcGIS Open Data as part of ArcGIS Online, to make it easier for data providers to publish open data (Claessens, 2016; Esri, n.d.-c). ArcGIS Open Data was later rebranded as ArcGIS Hub. Also in 2014, Esri launched Living Atlas of the World (Berry, 2024; Esri, n.d.-b), in which Esri Inc. and its distributors actively curate geospatial information (maps, apps, data layers) and imagery (e.g., on demographics, landscape, and transportation), beyond basemaps (Esri, 2014).

5 Map services contain one or more data layers, depending on the map's purpose.

6 Basemaps are reference maps on which one overlays data from other layers and visualises graphic information.

The Wayback Machine - https://web.archive.org/web/19961104070544/http://www.esri.com:80/base/data/i…

Title Navigation Bar

ANNOUNCING

ArcScene World Tour
Check out this educational tour of SPOT satellite images, new on CD-ROM from ESRI

Free data sets
Data sets you can download and try out right now

Data from ESRI
World, US, and educational data sets

The ArcData Catalog
A catalog of hundreds of data sets from leading data publishers, ready to use in your GIS

Selecting ArcData Sets
How to figure out what data sets you need

Understanding Data
The different kinds of data used by a GIS

MAPS

Data Publishing Services
We can help you publish your data

Data is the fuel for your GIS

Just as a car won't run without gasoline, a GIS without data has no information. You can get data for your mapping program or GIS from several sources.

Much useful data, such as US streets and demographic information, is included with BusinessMap, ArcView GIS, AtlasGIS, and other ESRI software, so you can start making maps right away. This data is also useful for mapping data that you may have gathered yourself, such as customer names, addresses, or sales figures.

You can add to your database by purchasing data from companies that specialize in collecting and publishing data or from government agencies. The ArcData Catalog lists hundreds of data sets that are ready to use with ESRI software, from ESRI and other companies that participate in ESRI's ArcData Publishing Program. ESRI and our business partners can help you choose your data sets so they work well together in your GIS.

Organizations in need of GIS data that is not readily available typically build their own map data using sources such as paper maps, surveys, aerial photographs, and other technologies such as Global Positioning Systems. ESRI's Database Services group can help you build your database.

More GIS data is in ARC/INFO format than any other data format in the world

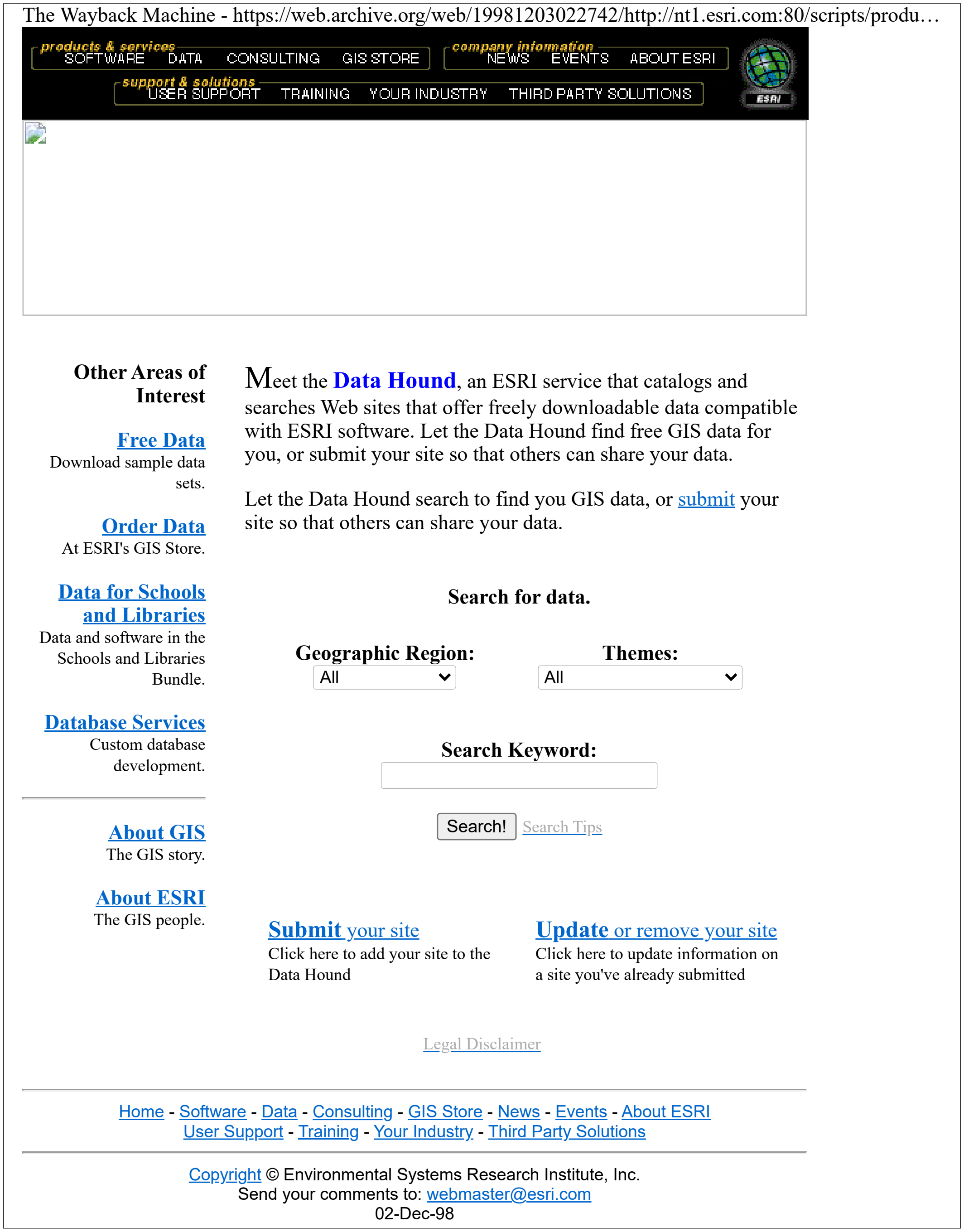
Because ARC/INFO software is the world's premier GIS, many organizations, especially government agencies, have developed their databases in ARC/INFO format. This format is readable by most ESRI software products (see specific software product pages for more information about data formats). Shapefile format, another ESRI-compatible software format, is also becoming widely available, especially for mapping and GIS on the desktop. In addition, ESRI software supports many other popular data formats. This means that when you choose ESRI software, you will be able to obtain the GIS data you need easily.

News | Products & Solutions | GIS at Work | Sales & Service | Training & Support | Community | The GIS Store

Top Copyright © 1996 Environmental Systems Research Institute. All rights reserved. *Home*
Send your comments to: webmaster@esri.com
October 24, 1996

FIG. 5.2 Archived webpage of Esri (captured on 4 November 1996) [Source: Esri (1996) from Wayback Machine by Internet Archive]

TOC


The Wayback Machine - https://web.archive.org/web/19981203022742/http://nt1.esri.com:80/scripts/produ…

products & services: SOFTWARE DATA CONSULTING GIS STORE
company information: NEWS EVENTS ABOUT ESRI
support & solutions: USER SUPPORT TRAINING YOUR INDUSTRY THIRD PARTY SOLUTIONS
ESRI

Other Areas of Interest

Free Data
Download sample data sets.

Order Data
At ESRI's GIS Store.

Data for Schools and Libraries
Data and software in the Schools and Libraries Bundle.

Database Services
Custom database development.

About GIS
The GIS story.

About ESRI
The GIS people.

Meet the **Data Hound**, an ESRI service that catalogs and searches Web sites that offer freely downloadable data compatible with ESRI software. Let the Data Hound find free GIS data for you, or submit your site so that others can share your data.

Let the Data Hound search to find you GIS data, or submit your site so that others can share your data.

Search for data.

Geographic Region: All
Themes: All

Search Keyword:

Search! Search Tips

Submit your site
Click here to add your site to the Data Hound

Update or remove your site
Click here to update information on a site you've already submitted

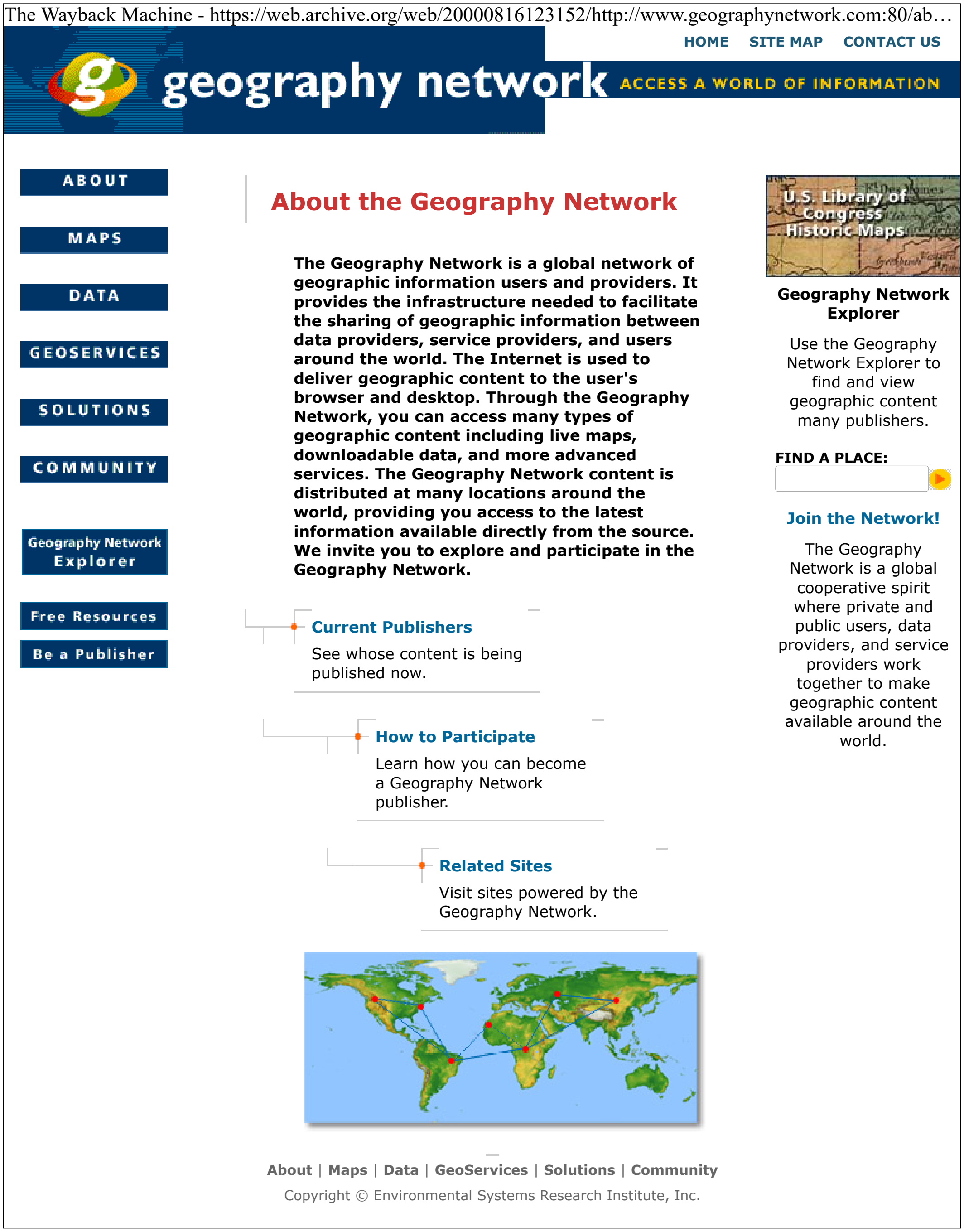
Legal Disclaimer

Home - Software - Data - Consulting - GIS Store - News - Events - About ESRI
User Support - Training - Your Industry - Third Party Solutions

---

Copyright © Environmental Systems Research Institute, Inc.
Send your comments to: webmaster@esri.com
02-Dec-98

FIG. 5.3 Archived webpage of Esri's Data Hound (captured on 3 December 1998) [Source: Esri (1998) from Wayback Machine by Internet Archive]



The Wayback Machine - https://web.archive.org/web/20000816123152/http://www.geographynetwork.com:80/ab…

HOME SITE MAP CONTACT US

geography network ACCESS A WORLD OF INFORMATION

ABOUT
MAPS
DATA
GEOSERVICES
SOLUTIONS
COMMUNITY

Geography Network Explorer

Free Resources
Be a Publisher

About the Geography Network

The Geography Network is a global network of geographic information users and providers. It provides the infrastructure needed to facilitate the sharing of geographic information between data providers, service providers, and users around the world. The Internet is used to deliver geographic content to the user's browser and desktop. Through the Geography Network, you can access many types of geographic content including live maps, downloadable data, and more advanced services. The Geography Network content is distributed at many locations around the world, providing you access to the latest information available directly from the source. We invite you to explore and participate in the Geography Network.

Current Publishers

See whose content is being published now.

How to Participate

Learn how you can become a Geography Network publisher.

Related Sites

Visit sites powered by the Geography Network.

U.S. Library of Congress Historic Maps

Geography Network Explorer

Use the Geography Network Explorer to find and view geographic content many publishers.

FIND A PLACE:

Join the Network!

The Geography Network is a global cooperative spirit where private and public users, data providers, and service providers work together to make geographic content available around the world.

About | Maps | Data | GeoServices | Solutions | Community

Copyright © Environmental Systems Research Institute, Inc.

FIG. 5.4 Archived webpage of Geography Network (captured on 16 August 2000) [Source: Geography Network (2000a) from Wayback Machine by Internet Archive]

TOC

In February 2023, Esri joined the Overture Maps Foundation, a collaboration founded by Amazon Web Services (AWS), Meta, Microsoft, and TomTom. Overture aims to create reliable, easy-to-use, and interoperable open global map data. Overture compiles and enhances OSM data with other sources to produce new open map datasets to be used by mapping platforms and service developers (Overture Maps Foundation, 2023), including Esri customers (interviewee 09-Esri-O). Esri contributes to Overture in three ways: by sharing data gathered through its Community Maps program, contributing human resources to help define data schemas and build certain information products, and providing infrastructure support (09-Esri-O). In sum, Table 5.2 presents the evolution of Esri's (open) data intermediation over the years.

TABLE 5.2 Evolution of Esri's (open) data intermediation

Year	(Open) data products/services by Esri
1996	**ArcData** *(Retired by 2003)* Esri provided hundreds of ready-to-use datasets from Esri and other companies participating in its ArcData Publishing Program. Some datasets were downloadable free of charge.
1998	**Data Hound** *(Retired by 2001)* Esri introduced a search engine that brought users to external websites offering freely downloadable data compatible with Esri software.
2000	**Geography Network** *(Retired in December 2009)* Esri introduced a website to access, share, and download geographic content worldwide, consolidating ArcData and Data Hound with additional features. Most of the content was free. Geography Network was retired in 2009 as the services were incorporated into the ArcGIS Online web-based software package introduced in the same year.
2010	**Community Maps Program** *(Active)* Esri welcomed organisational users to share their local data to improve the suite of basemaps created and hosted by Esri and offered to ArcGIS users.
2014	**ArcGIS Open Data** *(Active, rebranded as* ArcGIS Hub*)* Esri introduced ArcGIS Open Data as part of the ArcGIS Online software package to facilitate open data dissemination.
2014	**Living Atlas of the World** *(Active)* Esri launched a platform where Esri Inc. and its distributors actively curate geographic information (maps, apps, data layers) beyond basemaps and imagery.
2023	**Overture Maps Foundation** *(Active)* Esri joined a collaboration founded by AWS, Meta, Microsoft, and TomTom to create reliable, easy-to-use, and interoperable open global map data.

5.5 Esri's open data intermediation business model

5.5.1 Esri's open data intermediation value proposition

Generally speaking, there are four value propositions of Esri's open data intermediation. *First*, Esri provides a software system that is in itself an (open) geospatial data platform (01-Esri-A, 03-Esri-C, 09-Esri-O, 25-User-C, 26-User-C). ArcGIS software comes with a suite of basemaps that users can use immediately, which Esri enhanced through its Community Maps programme. Moreover, ArcGIS users can re-use millions of other objects (e.g., data files, layers, maps, apps, services) provided by other ArcGIS users, hosted in ArcGIS Online (Esri, 2023b). A subset of these are those curated by Esri Inc. and its distributors as part of the Living Atlas. Some of the objects in the Living Atlas are also reusable by non-ArcGIS users. Conversely, ArcGIS users can contribute their data in the ArcGIS system how they prefer, including by deciding the type of license they would like to affix to their data, including an open license (e.g., Creative Commons, Open Database Licence, Public Domain). Nevertheless, ArcGIS users can change their data sharing decisions, making ArcGIS software unsuitable for guaranteeing data permanence.

Second, Esri offers a software system that organisational data providers can easily use to create and disseminate open data (06-Esri-D, 09-Esri-O, 10-StO-C, 11-Prov-B, 16-Prov-C). It includes hosting the data in a cloud environment and publishing it as a ready-to-use service accessible through open application programming interfaces (APIs). ArcGIS Hub is the main product that delivers this value proposition. As of early 2018, over 5,000 government organisations, academic institutions, and others have published open data through ArcGIS Hub (Lafia et al., 2018). An account is not required to search, access, and re-use open data published through ArcGIS Hub.

Third, Esri Inc. and its distributors occasionally take on special projects related to open data, either on their initiative or in collaboration with other organisations (01-Esri-A, 02-Esri-B, 04-Esri-C, 06-Esri-D, 09-Esri-O, 16-Prov-C). Most of these projects aim to deliver social value from open data by making it accessible to the broadest audience possible. An excellent example is Esri's support of the COVID-19 Dashboard by Johns Hopkins University (JHU) by helping the JHU team with data scraping, automating the process of importing data, and scaling the infrastructure to withstand tremendous volumes of traffic (Barone, 2021; Geraghty, 2023; Milner, 2020; Perkel, 2020). Another example is the collaboration between Esri Netherlands and the Cadaster, Land Registry & Mapping Agency (Kadaster) to develop a web application called Topotijdreis (topo time travel), which presents historical 200-year topographic maps of the Netherlands (Esri Nederland, 2023). Meanwhile, Esri UK developed free GIS-based teaching resources based on the UK Meteorological Office's (Met Office) open data to help school children learn about climate change (Davies-Holloway, 2021; Esri UK, 2021). On its own initiative, Esri Spain developed a portal analysing and visualising traffic accidents in Madrid based on open data from the city council (Esri España, n.d.).

Fourth, Esri offers consultancy services to open data providers and re-users (01-Esri-A, 04-Esri-C, 07-Esri-E, 08-Esri-E). As described by interviewee 01-Esri-A,

'If I see the value of Esri beyond just the software, it's that we speak to everybody. We know the big problems. We know the little problems. And if we can fix them because we can see the shape of the jigsaw piece in the middle, then we can do that'.

A pertinent example of Esri providing consultancy services to an open data provider is Esri UK's work with the Met Office to improve how the latter publishes climate data. However, not all Esri distributors provide consultancy services, as at least one (05-Esri-D) explicitly said they do not do so due to a lack of resources.

5.5.2 Esri's open data intermediation value creation

Most of the resources and activities Esri deploys to offer its open data intermediation value propositions are also those required to maintain and develop its software offering (09-Esri-O). The enabling technology, in terms of desktop software for creating data and server software for disseminating data, has been around and continuously developed before Community Maps, Living Atlas, and ArcGIS Open Data (ArcGIS Hub) were even introduced (09-Esri-O). Esri claims to invest approximately 30% of its annual revenue into research and development (Esri, n.d.-a). Once customers subscribe to ArcGIS, they are supported by Esri's customer service, which ensures that the software, including open data-related products, is suited to their needs (11-Prov-B, 16-Prov-C).

Esri invested in an extensive physical and network infrastructure to host and make available petabytes of data to users in a highly scalable, reliable, and performant way (09-Esri-O). The company began this process in the early 2000s by purchasing physical hardware, establishing data centres, and developing technology to manage the infrastructure (09-Esri-O). Nevertheless, it was still not sufficiently reliable and scalable (09-Esri-O). Hence, in the last five years, the company has been migrating all of its infrastructure to the cloud, and almost all of Esri's content is now hosted in commercial cloud infrastructure. This, for example, made it possible for Esri to support JHU's COVID-19 Dashboard, which reached a peak of a billion hits per day (09-Esri-O).

To provide data in ArcGIS, such as the basemaps and other data in the Living Atlas, Esri performs the tasks of searching for, processing, and curating data (01-Esri-A, 02-Esri-B, 03-Esri-C, 05-Esri-D, 07-Esri-E). Esri also regularly updates datasets as they are made available by data providers, often through customised data routines that automatically download data from their sources on a scheduled basis, integrates them into the data model, and publishes them (02-Esri-B, 04-Esri-C, 05-Esri-D). In addition, Esri develops governance tools to inspect data nominated into the Living Atlas, communicate with data contributors, and accept it into the Living Atlas (09-Esri-O).

Esri also monitors the use of content in ArcGIS Online. This is to ensure that data request times are still within seconds and to keep abreast of local or global news that may result in traffic hikes (04-Esri-C and 05-Esri-D). For instance, Esri Netherlands has a large, self-developed monitoring environment including push notifications on mobile phones. This helped it react swiftly to events that require scaling up its infrastructure, such as the large flood in the south of the Netherlands in 2021, during which elevation maps were in high demand.

Some distributors also develop their own Esri national open data portal or service, separate from the Living Atlas. For example, Esri Germany Open Data Portal compiles geospatial data on Germany in various open formats, accessible to anyone, including non-ArcGIS users (Esri Deutschland, n.d.). The portal includes data on various topics, such as public safety, traffic, and the environment. This is in the context of a heterogeneous open data landscape in Germany, where every federal state decides on the open data it provides and how the data is provided. Meanwhile, Geoinfo Denmark offers Geoinfo DataLeverance (Geoinfo, n.d.), a free open data delivery service for ArcGIS users who sign up for the service. Data is delivered in file geodatabases via Secure File Transfer Protocol (SFTP).

Esri has been organising the Esri User Conference every year since 1981 (Esri, 2010b). It is the largest GIS practitioners' gathering in the world, where Esri showcases the new development of its technology and the applications of its software, including those involving open data. The conference also features exhibitions from over 200 organisations, including large companies (e.g., Airbus, AWS, and Maxar), government organisations, and academic institutions. In 2023, the event gathered nearly 18,000 in-person attendees from around the world, and over 20,000 online viewers (Esri, 2023a). Additionally, Esri distributors hold local conferences annually (Esri Deutschland, 2024; Esri España, 2024; Esri Nederland, 2024; Esri UK, 2024; Geoinfo, 2024), providing local organisations the opportunity to update themselves with GIS advancements and connect with other organisations in the country. Besides annual conferences, Esri ensures continuous customer engagement through other ways. For example, Esri Netherlands offers a content hub containing information on the open data it curates, including data changes log and recent data updates (Esri Nederland, n.d.). The content hub also provides inspiration on open data applications and tutorials on dealing with certain open data.

In some instances, Esri mediates feedback about open data from users to providers. Interviewee 07-Esri-E noted,

'A lot of end-users, especially ArcGIS users, will contact us if they have problems with gathering or accessing the data. After a few calls from different customers, we will start to look into it. […] So we are having some dialogue with the provider based on users' feedback and our own experience as well'.

Additionally, the Head of Climate Services of the UK's Met Office was offered centre stage during the 2022 Esri UK Annual Conference to solicit input on the agency's new data portal (Esri UK, 2022).

5.5.3 Esri's open data intermediation value capture

Esri captures value as an open data intermediary in *five* main ways. The *first* is through cross-subsidies. Intermediating open data creates a greater appeal for ArcGIS (11-Prov-B, 16-Prov-C, 17-Prov-B, 25-User-C). It is not only the software's capabilities that may attract customers but also its ready-to-use data, which reduces customers' burden of compiling and pre-processing data. The top-level executive from Esri Inc. interviewed noted that the company has seen exponential growth in its software adoption since around 2010. While this could be attributed to many factors, he stated that 'at the top of the list is the ability of users to access ready-to-use content, including open data content,' and inversely, 'the ability of them to share information'. Some Esri distributors (02-Esri-B, 04-Esri-C, 05-Esri-D, 07-Esri-E) also associated open data intermediation offerings with more software sales, or at least mentioned this as one of their goals.

Second is through nonmonetary marketing. By undertaking special projects related to open data, Esri aims to increase its visibility to a broader audience. These projects can demonstrate what ArcGIS can do. In the case of Esri UK working with the Met Office to offer free GIS-based climate change teaching resources, the interviewee from Esri UK shared,

'We wanted them [Met Office] to co-brand our resources into schools and universities, so we asked them for help on the climate science and the validation. They asked for help transferring knowledge to their GIS team. So we got something out of it. They got something out of it. No money changed hands. It was just time'.

Similarly, interviewee 04-Esri-C remarked,

'The more public good promotional apps are out there, the more people use them and see that they are powered by Esri. So that's more like marketing and brand awareness. That's a value for us'.

Third is through the freemium model. Esri also purchases datasets from commercial data providers, such as street data from Maxar and micro-boundaries data from Michael Bauer Research, since these datasets are not available as open data (09-Esri-O). While some of these purchased datasets are then provided for free to ArcGIS users, some as paid datasets (called premium data). Additionally, some datasets are initially obtained as open data but offered as premium data due to the intensive work involved in pre-processing and making them available in ArcGIS (09-Esri-O). Customers may find it convenient to purchase the premium data due to it already being contained in the ArcGIS system, where a vast quantity of other datasets, including basemaps, are already free to use.

Fourth is through consultancy fees. Some organisations require additional assistance in using ArcGIS, including related to their (open) data management. They typically have a service-level agreement that includes a certain number of days per year for consultation, but some may require more. Consultancy services are not only sources of revenue in and of themselves, but, more importantly, they are how Esri wants to retain its customers by supporting them with Esri products (01-Esri-A, 04-Esri-C, 08-Esri-E). Besides, by offering best practice advice to open data providers, Esri aims to help open data users among its customers obtain more usable data (01-Esri-A, 07-Esri-E).

Fifth, by intermediating open data, Esri benefits from self-learning. Through obtaining open data from various sources, pre-processing, and publishing it using ArcGIS, Esri and its distributors use the software 'as if [they] were the customers [themselves]', citing interviewee 04-Esri-C. Consequently, they do and learn a lot. '[We] eat [our] own dog food', said interviewee 06-Esri-D. The explicit and tacit knowledge they gathered from learning by doing may then be shared with their customers and used to improve their software and services.

TABLE 5.3 Links between the value proposition and value capture of Esri's open data intermediation

Value proposition	Value capture				
	Cross-subsidies	Nonmonetary marketing	Freemium model	Consultancy fees	Self-learning
A software system that is in itself an (open) geodata platform	X		X		X
A software system for data providers to create and disseminate open data	X				
Special projects related to open data		X			X
Consultancy services to open data providers and re-users				X	X

5.6 Current strengths, current weaknesses, potential opportunities, and potential threats of Esri's open data intermediation business model to the ODE

5.6.1 Current strengths

Driving values to open data providers and users

Esri's venture into open data has far more to do with improving its software's experience and sales than altruism. Arguably, this makes Esri's role as an open data intermediary rather enduring, as its cessation would hinder Esri's competitive advantage, at least until a (new) player with a new business model renders Esri's role irrelevant. The applicability of value drivers identified by Amit & Zott (2001) and Visnjic et al. (2017) was analysed to the Esri case.

Since introducing (open) data intermediation services as early as the 1990s, Esri sets the bar of what to expect from GIS software. Beyond just offering a set of GIS tools, Esri has long facilitated (open) data publishing and re-use by its customers. Hence, at least to a certain period, Esri drives (or drove) novelty value through its open data intermediation business model. By including open data in its software, Esri drives the efficiency value since, for many use cases, ArcGIS users could skip searching for and pre-processing datasets as they are already integrated into the software (16-Prov-C, 17-Prov-B, 24-User-C, 27-User-D). ArcGIS users can also publish their open data easily with ArcGIS software (particularly with ArcGIS Hub).

Open data providers occasionally institute structural changes to their data, which can impact users who implement automated data routines for data retrieval and pre-processing. However, as Esri takes care of fixing the data routines on its side, customers may not have to deal with this kind of disruption (04-Esri-C). This not only contributes to the efficiency value but also the accountability value where Esri manages risks on behalf of its customers.

Complementarity is also a value that Esri offers its customers, as its software comes with petabytes of data. Beyond that, Esri offers an integrated platform for GIS services, including enterprise, desktop, mobile, and cloud-based solutions, making it easy for open data to travel across different solutions (01-Esri-A, 02-Esri-B, 07-Esri-E, 11-Prov-B). In addition, Esri provides consultancy services and support to its customers, including among open data providers and users – indeed, some spoke very positively about the support they received (11-Prov-B, 16-Prov-C). Interviewee 16-Prov-C also related such support to the accountability value: 'With Esri, we can make an agreement, we want to use your software for our jobs, and we want proper support if needed', and according to the interviewee, this differentiates ArcGIS from its open source equivalent, QGIS.

Another value driver discovered from the Esri case but is less emphasised in the current literature on value drivers is the adaptability value. Notably, Esri customers appreciated the locally customised data, services, and projects that Esri's local distributors provided. Conversely, Esri customers expressed dissatisfaction when local customisation is lacking.

Leveraging network and technological capabilities

Consistent with the RBV (Barney, 1991), this study showed that Esri leverages its VRIN resources to intermediate open data, particularly its network and technological capabilities, which it has established for decades as the pioneer in GIS software. At this point, these resources are hardly imitable by other companies. In terms of network, for example, the interviewee from Esri Spain noted that in Spain, many public agencies are already Esri customers; hence, Esri Spain encouraged them to publish open data on the ArcGIS platform (aside from other platforms). Additionally, some of Esri's existing private-sector customers contribute open data for philanthropic or 'public good' reasons (01-Esri-A).

In terms of technological capabilities, along with the development of its software for the market, Esri has, at the same time, made it easier for Esri distributors themselves to process, curate, and host data on the Living Atlas. A couple of Esri distributors (01-Esri-A, 04-Esri-C) described that the availability of off-the-shelf Esri tools simplified their open data intermediation tasks. The cloud-based software of ArcGIS Online is leveraged for data hosting, reducing the need for Esri distributors to have large onsite servers (03-Esri-B, 04-Esri-C).

TOC

Promoting open data

Several interviewees credited Esri for bringing awareness about open data to the broader geospatial audience and indirectly promoting its provision and re-use (17-Prov-B, 19-OSG-O, 22-User-A, 24-User-C). An interviewee characterised Esri as a 'communicator of the open data' (17-Prov-B). The interviewee from OSM Foundation shared, '[Esri] have done a lot of good work to raise awareness of OpenStreetMap among their one of core customer bases, the public sector'. The interviewee from OSGeo observed, 'Esri's interaction with open data benefits Esri, but also benefits everyone else because it shines a stronger light onto open data; it shows the possibility'.

In addition, Esri also creates new types of data that were previously non-existent but had been deemed important. For example, Esri Inc. collaborated with Impact Observatory (an artificial intelligence company) and Microsoft to build the first high-resolution (10-meter) global land cover map based on the European Space Agency's (ESA) Sentinel-2 satellite imagery. The map was released as open data and is updated annually in the Living Atlas, and can also be used by non-ArcGIS users (Esri, 2021). By promoting open data and creating new types of open data, Esri contributes to the user-drivenness and circularity of the ODE.

5.6.2 Current weaknesses

Resting upon proprietary software

The main weakness of Esri's open data intermediation business model is the fact that it rests upon proprietary software (22-User-A, 27-User-D, 26-User-C). Interviewee 12-Prov-E strongly suggested that Esri's business model is 'very anti-open data ecosystem'. Others characterised it as potentially gatekeeping (15-Prov-D) and exclusionary (20-User-E, 29-User-C). On the other hand, interviewee 27-User-D viewed that, as a commercial software company, it is unsurprising that Esri offers (some) of its open data intermediation services only through its proprietary software. Other interviewees (24-User-C, 27-User-D) perceived that, despite Esri's proprietary software, the company does not deny the coexistence of other open data platforms and open source software. Furthermore, an interviewee among data providers (16-Prov-C) believed that if their organisation were to use open source software instead of ArcGIS, the organisation might have to hire companies specialising in open source to handle their data management, which would still incur costs.

Esri interviewees seemed familiar with the debate of using proprietary versus open source software with open data, as well as the conflation of open data and open source software. They generally emphasised the coexistence of open source and proprietary software, for instance:

‘We are different technologies that live in the same space, and we try to do our best to solve the problems of the users. It’s that simple. […]. There shouldn’t be, let’s say, dogmatic views about the way you approach these kinds of problems’ (02-Esri-B).

This echoes Sui’s (2014, p. 20) assertion that ‘an artful combination of both open and proprietary practices’ is likely ‘the most realistic option’. In any case, since most of Esri’s open data value propositions rest upon proprietary software, Esri’s contribution to the circularity of the ODE is limited, as not all actors can benefit from the value the company offers.

Local versus global tensions

Esri local distributors play a crucial role in the company’s open data intermediation. They provide data for the development of basemaps, curate local open data in the Living Atlas, perform local open data-based projects, and engage with local stakeholders. This decentralised approach seems apt since every country has different open data policies and landscapes (01-Esri-A, 04-Esri-C, 09-Esri-O). However, different priorities concerning open data intermediation were observed across the distributors. For a start, the number of items (e.g., maps and layers) curated by each distributor in the Living Atlas are starkly different: the Netherlands (1,401), Spain (1,047), Germany (638), the UK (439), and Denmark (77) (recorded in August 2024) (Esri, n.d.-b). The number of items curated by local distributors in the Living Atlas does not reflect the availability of open geospatial data in those countries. For example, even though Geoinfo Denmark curated the least number of items among the four distributors, a large number of Danish geospatial datasets were already open data as early as 2013, including topographic data, place names, elevation products (including LIDAR point cloud, a terrain model, and a surface model), administrative units, cadastral information and parcels, location-based addresses, and orthophotos (geometrically corrected aerial photography) (Copernicus In Situ, 2018; The Centre for Public Data, 2024).

Beyond those numbers, the interviews corroborated the diverging priorities. For instance, an interviewee from Geoinfo Denmark explicitly mentioned their preference to share data through the distributor's own service, Geoinfo DataLeverance, instead of the Living Atlas, due to 'some administrative, technical issues'. Meanwhile, the interviewee from Esri UK shared that their approach has shifted from curating data to mainly engaging their customers to publish their data on their own through ArcGIS Online or ArcGIS Hub. There may be organisational identity at play because when asked the reason for the shift, they said, 'We're a software company. […]. We make money from selling software and solutions. We're not a data company'. Furthermore, based on the interviews, the numbers of staff in charge of content-related products (including open data products) differed significantly across distributors. For instance, Esri Netherlands had eight full-time employees working on content, comprising developers, data engineers, cloud engineers, cartographers, and product managers, all in one team. On the other hand, Geoinfo Denmark had fewer than five employees working on content, all in different teams.

Therefore, while Esri Inc. may have a particular business model outlook to open data intermediation (09-Esri-O), it is not shared uniformly across Esri distributors. This may then impact the experience of ArcGIS users. For example, an ArcGIS user from Denmark remarked, '[in ArcGIS], it was not open data that was useful for me because it was not Danish open data; it would be some American open data'. Another interviewee reported that, despite being based in one country, they occasionally engaged in projects or research on another country (or across multiple countries), for which they noticed different data availability and quality levels in ArcGIS (29-User-C). This insight shows that while decentralisation strategies may help deliver the adaptability value, such strategies may also compromise other value drivers (particularly efficiency) if there are diverging views on organisational identity (e.g., software company versus more-than-software company).

That said, some interviewees highlighted that open data inconsistencies across geographical areas are not only an issue within ArcGIS but also reflect a persistent problem of the broader open data ecosystem (22-User-A, 26-User-C, 28-User-C, 29-User-C). Even across subnational boundaries, fragmentation exists (01-Esri-A, 02-Esri-B, 06-Esri-D, 15-Prov-D, 23-User-D). While problems and phenomena that need to be studied with (open) geospatial data may be transboundary, (open) geospatial data administration is not; an issue that has long been recognised (McLaughlin & Nichols, 1994).

Limited to geospatial data

Esri's open data intermediation is limited to geospatial data. Several interviewees highlighted the need to consider cross-domain open data business models, including those which integrate open geospatial data with other types of open data (10-StO-C, 12-Prov-E, 16-Prov-C, 20-User-E). Some interviewees also observed the general aversion of some non-geospatial data users towards geospatial data, which is likely due to the lack of awareness and skills (11-Prov-B, 20-User-E). As Masser et al. (2008, pp. 5-6) argued, most people are not 'spatially aware professional'. This predicament could also partly be attributed to different commonly used standards; for instance, while ISO standards are popular in the geospatial domain, DCAT standards are more prevalent in other fields (10-StO-C) (Ivánová et al., 2020). Hence, there is still a gap to bridge in reducing the barrier for users without geospatial backgrounds to use and integrate open geospatial data with other types of open data. Addressing this gap could improve the user-drivenness and skills-based aspects of the ODE.

5.6.3 Potential opportunities

Learning from the road taken

Esri has shown possible ways (and potential shortcomings) in how open data intermediation could be performed (16-Prov-C, 29-User-C). These insights can be used by others to undertake open data intermediation, including by the open source software community. The OSGeo interviewee said, 'what Esri has done in some way is analogous to how Windows helped Linux and some open source products'. Interviewees from Esri also observed how the open source software community often mimics or makes references to functionalities and services offered by Esri, which they perceived as positive (02-Esri-B, 04-Esri-C, 08-Esri-E).

However, Esri's approaches are not the only possible ways. The OSGeo interviewee encouraged schools, colleges, and universities to raise awareness about open source software as an alternative to proprietary software, consistent with calls in the literature (Brunsdon & Comber, 2021; Kedron et al., 2021; Singleton et al., 2016). This is a crucial point especially in the context of Esri being heavily involved in GIS education (Curran & Bowlick, 2022). This new generation could constitute a critical force to accelerate the development and adoption of open source software, encompassing new open data intermediation solutions. This is necessary for the inclusivity aspect of the ODE.

Further collaborations

There are opportunities for further collaboration between Esri and others to achieve shared goals. Interviewee 13-Prov-E wished for formal collaborations where Esri could systematically provide input and updated data back to open data providers. Interviewee 12-Prov-E proposed co-creation or sandbox collaboration where Esri could work with others in developing initiatives with shared benefits.

One significant collaboration that Esri has recently become involved in is the Overture Maps Foundation. While the foundation's claim to offer high quality open geospatial data based on OSM is desirable, and the interviewee from the OSM Foundation also described Overture as 'very much a good thing', it is still too early to assess its impacts. It is worth pointing out that Overture is led by four large tech companies as steering members, with Esri as the general member. The contributor members are almost entirely comprised of for-profit companies (as of August 2024) (Overture Maps Foundation, 2024b). Furthermore, Ballantyne & Berragan (2024) noted that while Overture's data offers great potential, 'accessing the data relies on computational resources beyond the skillset and capacity of the average researcher' (p. 1). Therefore, active involvement or new collaborative ventures in open data intermediation by public or CSOs at the global scale may still be necessary to seriously account for public interests and non-expert users.

Advocating for more open data

With Esri's network, resources, and market position, the company could significantly advocate for better open data availability and quality. Esri has already made such efforts to some extent. For example, the CEO of Esri Netherlands led the Breakthrough Project Open Geodata from 2013 to 2017, an initiative of the country's Ministry of Economic Affairs to identify and address obstacles around open geospatial data in the Netherlands. The project's outcomes include the release of the actual elevation map of the Netherlands (AHN) and satellite data from the Dutch Space Office (Blankena, 2016; Doorbraakproject Open Geodata, 2015). Nevertheless, gaps remain in terms of open data availability and quality in other countries, for which Esri could take a more active role in bridging. Besides, there is now limited open data from the private sector compared to the public sector (05-Esri-D, 06-Esri-D, 17-Prov-B). Therefore, Esri may contribute to the inclusivity of the ODE by incentivising companies to share more open data.

5.6.4 Potential threats

Dominant position

Although Esri may not have explicitly denied the space for other (potential) open data intermediaries to engage in similar activities, its dominant market position, along with its VRIN resources, renders its influence and power hard to match (19-OSG-O). This translates into, for instance, its perceived reluctance to adopt open standards, at least in the early days (10-StO-C, 15-Prov-D, 17-Prov-B, 23-User-D, 24-User-C, 25-User-C), despite having been a member of the Open Geospatial Consortium (OGC)[7] as early as 1996 (Open GIS Consortium, 1996).

Moreover, the company has also been perceived as steering open standards development to give itself an unfair advantage in the market (Dasgupta, 2013; OSGeo, 2013). However, Esri would argue that its proximity to users means that it is aware of 'real' user needs regarding standards beyond what is theoretically possible (Esri, 2018; Henriksen et al., 1994). Echoing Dasgupta (2013), the onus is on the OGC as a consortium to provide leadership for everyone's interests. This scenario highlights the larger question of what kind of governance is required for a self-organising ODE – beyond the issue of standards. This question is as necessary concerning Esri now as it concerns other current and future open data actors, especially for-profit companies in dominant positions (Johnson et al., 2017; Mahmoudi et al., 2024).

Lock-in system

The convenience that ArcGIS offers as an integrated platform may also result in high costs of switching away from the platform (10-StO-C), which could well be a double-edged sword. These costs may be due to technical restructuring of the open data infrastructure and staff re-skilling. There are various possible reasons why customers may wish to switch away from Esri products. For example, one interviewee shared that their organisation was considering moving to an open source infrastructure due to the high costs of ArcGIS subscription, as the required capacity

7 Open Geospatial Consortium (OGC), formerly Open GIS Consortium, is an international consortium aiming to make geospatial data FAIR (Findable, Accessible, Interoperable, and Reusable), including by developing open standards (OGC, n.d.).

for their infrastructure was reaching the limit set in their contract (11-Prov-B). Even though lock-in is one of the value drivers identified by (Amit & Zott, 2001), it is not a positive value driver to the circularity and inclusivity of the ODE.

Interviewee 10-StO-C suggested that Esri should consider compartmentalising its functionalities further than they currently do so, in order to allow open data providers and users to choose the specific services they want. As it stands, several users reported that they use ArcGIS alongside other software, such as Python and FME, due to their perceived strengths, despite the fact that the activities they conduct with other software can theoretically be performed in ArcGIS (20-User-E, 24-User-C, 26-User-C). Hence, not all ArcGIS users require all the functionalities that come in existing ArcGIS packages. However, another interviewee considered that current Esri services are already too compartmentalised (28-User-C). Notwithstanding, it would be worth further exploring the potential adverse impacts of Esri's lock-in system, as well as how these could be mitigated.

Sharing versus shifting responsibilities conundrum

One interviewee alluded to whether Esri is now conducting the work that ought to be addressed at the data provision stage by original open data providers, especially among public agencies (26-User-C). Others pointed out that even though Esri has seemingly reduced the open data accessibility and reusability barriers for ArcGIS users, these barriers are still left unaddressed for non-ArcGIS users (23-User-D, 29-User-C). One interviewee among data providers stressed that there should be an active role for open data intermediation within the public sector, and such a role should not be left only to the private sector (12-Prov-E). This resonates with the call by Johnson & Scassa (2023) for governments to critically consider their role within geospatial data collection, use, and dissemination. Sieber & Johnson (2015) highlighted various roles that public organisations could perform apart from publishing open data. They could also actively encourage the re-use of open data (e.g., by organising contests), accept contributions of citizen-generated data, and involve citizens in the decision-making related to open data.

Meanwhile, other interviewees among public organisations (13-Prov-E, 16-Prov-C) mentioned their lack of resources to regularly engage with open data users and address their needs; hence, they were appreciative of the role played by open data intermediaries outside the public sector, such as Esri. Additionally, some interviewees highlighted the blurred demarcation between the role of the public sector and the market from the eyes of competition laws (04-Esri-C, 13-Prov-E, 16-Prov-C).

Therefore, while the ODE implies that all actors share the responsibility for sustainable value creation from open data, there is no quick and easy answer regarding who should play what role, as one would expect from a purely top-down hierarchical system. It would be even more challenging to answer this question at the multinational level, since different countries have different open data approaches and preferences (Hossain et al., 2021), and issues such as competition law involve more factors beyond open data (as highlighted in Chapter 3).

5.7 Lessons from Esri: Developing an open data intermediation business model that supports the sustainability of the ODE

Based on the analysis of Esri's open data intermediation business model, its current strengths and weaknesses, and potential opportunities and threats, several aspects to consider in developing an open data intermediation business model that supports the sustainability of the ODE were recommended. These aspects especially apply to open data intermediation business models that share the same archetype as Esri, i.e., one-stop package around an (augmented) open data platform/ repository (archetype A6 in Chapter 4).

For the one-stop package archetype specifically, the core products or services are at the centre of the business model (in the Esri case, it is the ArcGIS software). Open data intermediation services should be developed around the core products/ services. This means offering open data intermediation services should not require radically different resources and activities from the core products/services. It also means that the core products/services and the open data intermediation services are mutually interdependent: the attractiveness of the open data intermediation services should ideally drive customers to engage with the core products/services. This tight integration between core products/services and open data intermediation services is consistent with theories of the RBV (where organisations gain competitive advantage by leveraging their VRIN resources) and the organisational identity theory (where the notion of 'who we are' drives strategic decisions).

Table 5.4 outlines 16 aspects to consider for developing an open data intermediation business model based on insights from the Esri case. In the table, the asterisk (*) next to several 'aspects to consider' indicates that the aspect can also be linked to a potential contribution(s) of open data intermediaries identified in Chapter 3, and the ID (e.g., P05) refers to the assigned ID of the contribution (see Table 3.4). Additionally, some aspects may fall under multiple categories in the first column; tabulating them according to those categories is simply meant to support the ideation process.

TABLE 5.4 Aspects to consider in developing an open data intermediation business model that supports the sustainability of the ODE (insights from the Esri case)

Categories	No.	Aspects to consider	Example insights from the Esri case
Resource-based view (RBV)	**1.**	Ensure that open data intermediation services offered are fittingly integrated with the existing core products or services.	Esri's open data intermediation is firmly linked to its business interests of improving its software's appeal and visibility, rather than to a philanthropic or ad hoc endeavour. This situation of 'having skin in the game' arguably renders its open data intermediation business model enduring. Additionally, Esri plays to its strengths by leveraging its existing network and technological capabilities in its open data intermediation instead of starting from scratch.
Organisational identity	**2.**	Offer open data intermediation services that are consistent with the organisational identity (i.e., do not involve a significant shift from the core business).	Esri's open data intermediation does not involve a significant shift from its core business as a software company.
	3.	Ensure consistency in how members understand the organisational identity, especially as the organisation becomes more decentralised.	Esri Inc. and Esri distributors have diverging approaches toward open data intermediation, partly contributed by different views of the company's organisational identity. This impacts the experience of its customers in different countries who engage in its open data intermediation services.
Value driver: novelty	**4.**	Offer new or demand-creating innovation instead of being entrenched in traditional paradigms.	Esri has set the expectation for a GIS software package that comes with ready-to-use data. While this may now be taken for granted, Esri's (open) data intermediation evolution since the 1990s shows that what Esri offers now results from conscious strategic decisions over decades.
Value driver: Complemen-tarity	**5.**	Offer diverse complementary products or services that leverage open data.	Besides the ready-to-use data within ArcGIS, Esri also offers consultancy services and conducts special projects related to open data.
Value driver: Efficiency	**6.**	Offer products or services that simplify, as much as possible, the process of supplying and using open data (e.g., some automation).* (P05)	Esri simplifies the process of reusing and supplying open data within ArcGIS.

>>>

TABLE 5.4 Aspects to consider in developing an open data intermediation business model that supports the sustainability of the ODE (insights from the Esri case)

Categories	No.	Aspects to consider	Example insights from the Esri case
Value driver: Accountability	7.	Offer services that minimise open data-associated risks that customers have to deal with.	Esri customers may not need to deal with occasional open data structural changes implemented by data providers, as Esri manages such disruptions. Esri, thus, manages the risks of data inconsistencies due to changes made by open data providers.
Value driver: Adaptability	8.	Offer customised data, services, or projects catering to local needs.* (P04)	Esri customers appreciate the locally customised data, services, and projects that Esri's local distributors provide.
Sustainable ODE feature: User-driven	9.	Offer open-source software, at least partially as part of the larger product suite (e.g., through a freemium model).* (P06)	Esri offers various value propositions that mainly address open data users' challenges. However, Esri's open data intermediation rests on proprietary software, thereby restricting non-Esri customers from benefiting from its services. Hence, there are gaps for open data intermediation integrated solutions based on open source software.
	10.	Stimulate potential multistakeholder collaborations, e.g., through projects or events.* (P11)	Esri Inc. has been organising the Esri User Conference every year since 1981. Additionally, Esri distributors hold local conferences annually, providing local organisations the opportunity to connect with other organisations.
	11.	Facilitate feedback on open data through a structured mechanism.* (P19)	To a limited extent, Esri mediates feedback about open data from users to providers. A more formal or standardised procedure for channelling feedback from open data users to providers could be explored.
Sustainable ODE feature: Circular	12.	Ensure that the unique resources or position are not leveraged in ways that unfairly stifle the growth of other actors (the winner-takes-it-all situation), e.g., by committing to the development of broadly adoptable open standards and technical interoperability.	Due to Esri's dominant position, it may directly or indirectly hinder other open data actors from also capturing value from the ODE (e.g., Esri being seen as influencing open standards to give itself an unfair advantage in the market).
	13.	Showcase the value of open data.* (P14)	Esri showed the value of open data to a broad audience, e.g., through its support towards the COVID-19 Dashboard by Johns Hopkins University (JHU). In turn, Esri concurrently promoted its open data intermediation services.

>>>

TABLE 5.4 Aspects to consider in developing an open data intermediation business model that supports the sustainability of the ODE (insights from the Esri case)

Categories	No.	Aspects to consider	Example insights from the Esri case
Sustainable ODE feature: Inclusive	**14.**	Advocate for the release of open data from non-public sectors.* (P12)	Esri has already made efforts to advocate for open data, e.g., the CEO of Esri Netherlands led the Breakthrough Project Open Geodata from 2013 to 2017 to identify and address obstacles around open geospatial data in the Netherlands. However, there is room for Esri to advocate for more and better open data, especially from non-public sectors.
	15.	Invest in open data-based collaborations.* (P13)	Esri invests in the development of Overture, which aims to create reliable, easy-to-use, and interoperable open global map data. Esri seeks to support its customers in obtaining open data through Overture.
Sustainable ODE feature: Skills-based	**16.**	Offer consultancy or training services.* (P07 & P16)	Esri provides consultancy services to its customers and publicly accessible tutorials and examples on using open data.

Note: The asterisk (*) next to several 'aspects to consider' indicates that the aspect can also be linked to a potential contribution(s) of open data intermediaries identified in Chapter 3.

5.8 Further discussion

Beyond identifying aspects to consider in developing an open data intermediation business model, this chapter also revealed several interrelated aspects that call for further research. First, this study calls into question how various responsibilities in the ODE should be allocated to ensure that diverse open data needs are addressed and public interests are protected. As shown, most of Esri's open data intermediation value propositions are only enjoyed by its customers, who can afford to subscribe to its proprietary software. Hence, challenges that Esri addresses may remain unaddressed for non-Esri customers. The situation is further complicated by existing competition laws that limit open data providers among public organisations from offering products and services similar to private sector open data intermediaries.

Second, this study reasserts the critical role of governing institutions (not only governments but also standards bodies such as OGC) in ensuring that the dominant position of any actors in the ODE does not jeopardise or overlook other actors' growth and mutual interests (Johnson et al., 2017). Furthermore, as highlighted by several interviewees from Esri (04-Esri-C, 06-Esri-D, 09-Esri-O), open data intermediation requires the availability of open data, first and foremost. Although open data intermediaries can also play a role in advocating for open data availability, their contribution could only go so far in the absence of legislation, policies, or organisational/political commitment on open data. In sum, without a forward-looking and robust governance system, the design of an open data intermediation business model alone cannot guarantee the sustainability of the ODE. In fact, the business model may perform well for open data intermediaries, but at the expense of other ODE actors.

Third, this study reaffirms the importance of the entire technology stack to the sustainability of the ODE. It is not only the characteristics of the open data (coverage, quality, format, license, etc.) that are important to the sustainability of the ODE, but also how it is disseminated and re-used, including the software used. Open source software that provides open data intermediation solutions is more desirable than proprietary software from the point of view of some actors in the ODE, since actors that may especially require open data intermediation services, such as grassroots groups (Elwood, 2008a), have limited financial resources. However, currently, while there is a wide range of open source tools available, understanding which tools and combination of tools are appropriate for specific purposes takes time (Lovelace, 2021), especially when documentation for open source software is often not as clear and detailed as proprietary software (Yap et al., 2022).

Several limitations in this study deserve further research. First, this study is based on a single-case study of a company that has long been a leader in GIS software. There are advantages to the study's methodological approach. As argued by several methodological scholars (Mariotto et al., 2014; Siggelkow, 2007), precisely because a case is exceptional or remarkable, a single-case study stands to offer revealing insights that may not be captured through a multiple-case study, since the latter focuses on common or comparable features. Having said that, further research is necessary to investigate the transferability of insights from this study to other cases. For instance, would insights from this study be transferable to open data intermediaries outside the geospatial data field?

Second, open data intermediaries can exist in various shapes and forms, employing different archetypes of business models. Certain insights from this study may not apply to all open data intermediation business model archetypes. For example, this study suggested that open data intermediation services be integrated with the core products/services of the organisation. However, certain business models do not differentiate the core products/services from open data intermediation services, as the business model entirely rests on providing open data intermediation services. Therefore, further research is necessary to investigate aspects to consider in developing open data intermediation business models of other archetypes.

Third, Esri has been intermediating (open) data for decades. Therefore, it might enjoy certain privileges unavailable to nascent open data intermediaries, such as its large customer base. The evolution of Esri's (open) data intermediation services presented in this chapter showed that they did not develop overnight but instead were gradually improved and refined over the years. Esri has been offering these services since its customer base was a lot smaller. Hence, one potential opportunity that other open data intermediaries can exploit is learning from the road taken by Esri. Having said that, it is still worth investigating unique challenges that nascent open data intermediaries may face nowadays, considering the maturity of (geospatial) data domains and related technologies compared to three decades ago.

5.9 Conclusion

This chapter has addressed the RQ4: What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem? The question was tackled through the case study of Esri, a significant player in the geospatial domain that has long been an open data intermediary. This chapter recommends 16 aspects to consider in developing an open data intermediation business model that supports the sustainability of the ODE (Table 5.4), which are especially applicable to one-stop package archetype. Among others, the aspects recommended revolve around ensuring that the open data intermediation services are fittingly integrated with the core products/services, offering products/services that simplify as much as possible the process of supplying and using open data and that minimise open data-associated risks, and investing in multistakeholder collaborations, such as through projects and events. The Esri case also showed the importance of ensuring different branches or franchises (distributors in the Esri case) of an organisation share a consistent outlook and approach on open data as the organisation grows and becomes more decentralised. Furthermore, one of the biggest weaknesses of Esri is its proprietary software package. Hence, offering similar (or better) open data intermediation value propositions based on open source software is an opportunity for other (potential) open data intermediaries to consider.

6 The case of OpenStreetMap (OSM)

6.1 Introduction

This chapter addresses the RQ4: What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem? Since there are nine open data intermediation business model archetypes identified in Chapter 4, this present chapter answers the research question specifically for the archetype of collaborative open data platform (archetype A1 in Chapter 4).

Towards that end, the case of OpenStreetMap (OSM) is selected. OSM is an open data platform where anyone can contribute and use geospatial data. The data is licensed under the Open Database Licence (ODbL). OSM was founded in 2004 to counter the high costs and restrictive licensing in accessing and using data from Ordnance Survey, the UK's national mapping agency (Coast, 2015). OSM has over five thousand daily active data contributors (Neis, 2025b), and the data is widely used, including by large companies, such as Apple and Meta (Anderson et al., 2019), and international organisations, such as the United Nations (UN) (UN Maps, n.d.-c). Thus, OSM is well-positioned to offer meaningful insights into the research question.

This chapter tackles three specific research objectives: (1) to detail OSM's open data intermediation business model, (2) to consider the current strengths and weaknesses, and potential opportunities and threats, of OSM's open data intermediation business model to the ODE, and (3) to formulate aspects to consider in developing an open data intermediation business model that can contribute to the sustainability of the ODE, specifically for the collaborative open data platform archetype.

The remainder of this chapter is organised as follows. Section 6.2 presents the methodology of this chapter. Section 6.3 briefly describes the historical background of OSM and its growth. Section 6.4 presents OSM's open data intermediation business model (Objective 1). Section 6.5 presents the analysis of the current strengths, current weaknesses, potential threats, and potential opportunities of OSM's business model (Objective 2). Section 6.6 formulates the aspects to consider in developing an open data intermediation business model that supports the sustainability of the ODE (Objective 3). Finally, Section 6.8 concludes the chapter.

6.2 Research methodology

The methodology of this chapter mirrors the previous chapter[8]. This chapter employs the single-case study methodology (R. K. Yin, 2018). OSM is a global project. Thus, to understand the business model of OSM, data from 29 semi-structured interviews (representing diverse communities from around the world) and publicly accessible sources, such as materials on websites (e.g., OSM Wiki), videos, and audio podcasts (cited accordingly), were gathered. The interviews were conducted between June 2023 and October 2024. The interviewees include former board members of OSM Foundation (OSMF) and representatives from Humanitarian OpenStreetMap Team (HOT), OSM local chapters or communities, YouthMappers, companies, a national mapping agency, a UN agency, and a research institution. Table 6.1 presents the identification (ID) of the interviewees to facilitate the presentation of the findings. The interviewees' roles were omitted to avoid the re-identification of the interviewees.

The data collected was analysed based on the abductive approach (Dubois & Gadde, 2002). The strengths, weaknesses, opportunities, and threats of OSM's open data intermediation business model were considered from the overall ODE point of view. These were then used to synthesise aspects to consider in developing open data intermediation business models. The theories and frameworks discussed in the previous chapter (i.e., on the Esri case) as well as the (provisional) features of a sustainable ODE (Section 1.1.3) and the potential contributions of open data intermediaries identified in Chapter 3 are used to guide the identification of the business model development aspects to consider.

8 Thus, refer to the Research methodology in Chapter 5 for justification of the methodology of the present chapter.

TABLE 6.1 Identification of interviewees

ID	Organisation	Type of organisation	Country/Region
COM-01	Development Seed	Company	US
COM-02	Gojek	Company	Indonesia
COM-03	Grab	Company	Southeast Asia
COM-04	TomTom	Company	Global
COM-05	TomTom	Company	Global
HOT-01	HOT Latin America	Humanitarian OpenStreetMap Team	Latin America
HOT-02	HOT Asia Pacific	Humanitarian OpenStreetMap Team	Asia Pacific
HOT-03	HOT Asia Pacific	Humanitarian OpenStreetMap Team	Asia Pacific
INS-01	UN Maps	Public institution	Global
INS-02	National Geographic Institute (NGI)	Public institution	Belgium
OSGE-01	OSGeo Ireland	OSGeo chapter	Republic of Ireland
OSMC-01	OSM Japan	OSM chapter/community	Japan
OSMC-02	FLOSS Kosovo (FLOSSK)	OSM chapter/community	Kosovo
OSMC-03	OSM UK	OSM chapter/community	UK
OSMC-04	OSM Uganda	OSM chapter/community	Uganda
OSMC-05	OSGeo Oceania	OSM chapter/community	Oceania
OSMC-06	OSM Indonesia*	OSM chapter/community	Indonesia
OSMC-07	OSM Ethiopia	OSM chapter/community	Ethiopia
OSMC-08	OSM community Nepal	OSM chapter/community	Nepal
OSMC-09	OSM community India	OSM chapter/community	India
OSMC-10	OSM Belgium*	OSM chapter/community	Belgium
OSMC-11	Wikimedia Italia	OSM chapter/community	Italy
OSMC-12	FOSSGIS Deutschland	OSM chapter/community	Germany
OSMC-13	OSM Korea**	OSM chapter/community	South Korea
OSMC-14	OSM Ghana	OSM chapter/community	Ghana
OSMF-01	OSM Foundation	OSM Foundation	Global
OSMF-02	OSM Foundation	OSM Foundation	Global
RES-01	TU Delft	Research organisation	Netherlands
YM-01	YouthMappers (Asia Pacific)	YouthMappers	Asia Pacific

Note: The asterisk (*) indicates that two interviewees were involved in the same interview. The double asterisk (**) indicates a written interview.

Similar to Chapter 5, while this dissertation is inspired by the SWOT analysis, the analysis is based on the temporal dimension (current strengths and weaknesses, and potential opportunities and threats) rather than internal versus external dimensions (i.e., the manner in which the SWOT analysis is conventionally used). Hence, the analysis in this chapter is not a SWOT analysis. Furthermore, open data intermediation business models exist in various archetypes (Chapter 4), and OSM's business model represents only a specific archetype (i.e., collaborative open data platform). Hence, not all insights from the Esri case may be directly transferable to all archetypes of open data intermediation business models. However, they may still set the groundwork for theorising other archetypes.

6.3 Background: OSM

OSM describes itself as a 'free, editable map of the whole world made by people like you' (OSM Foundation, n.d.-a). Those involved with OSM often refer to it as a community or a 'community of communities' (Solís et al., 2021) (interviewee OSMF-02). In the technical sense, OSM is both a database (the data can be downloaded and used in various applications) and a web-based map (the data is rendered for viewing and editing) (Biljecki et al., 2023). Solís et al. (2021) considered OSM as 'essentially serv[ing] as a data platform' (p. 600). OSM data is provided under the ODbL, which allows the data to be shared, used, and modified as long as attribution is provided and any adapted public version of the database is offered under the same licence (OKF, n.d.-a). OSM is not the same as the OSM Foundation (OSMF), which is the not-for-profit organisation 'supporting, but not controlling the OpenStreetMap Project' (OSM Foundation, 2024h). OSM and OSMF are run by a community of volunteers who help sustain the technical and social infrastructure and contribute data to OSM (OSM Foundation, n.d.-a).

OSM is often discussed in conjunction with several concepts. Goodchild (2007) cited OSM as an example of VGI, a term he coined to refer to 'the widespread engagement of large numbers of private citizens, often with little in the way of formal qualifications, in the creation of geographic information, a function that for centuries has been reserved to official agencies' (Goodchild, 2007, p. 212). OSM is also associated with the concept of neogeography (Glasze & Perkins, 2015; Haklay et al., 2008), which Turner (2006, p. 3) described as 'people using and creating their own maps, on their own terms, and by combining elements of an existing toolset'. Another concept typically associated with OSM is crowdsourcing (Huang et al., 2023), a term derived from outsourcing. While outsourcing refers to the hiring of third-party providers for specific business tasks (Howe, 2006), the providers in crowdsourcing (i.e., the crowd) were not directly selected (Howe, 2006). In any case, OSM is said to blur the boundaries between geospatial data producers and consumers, as reflected in neologisms such as *produser* (Budhathoki et al., 2008; Coleman et al., 2009; Ho & Rajabifard, 2010), which Bertolotto et al. (2020) regarded as transformative to the geospatial domain.

6.3.1 Raison d'être and technological enablers

OSM was founded in 2004 in the UK by Steve Coast and supported by other like-minded developers and map enthusiasts, primarily to counter the prohibitive costs and licensing in accessing and reusing data from Ordnance Survey, the UK's national mapping agency (Coast, 2015). The project quickly acquired appeal in other European countries where geospatial data was inaccessible, especially to individuals, small businesses, and community organisations (Haklay & Weber, 2008). Meanwhile, in the United States (US), while basic road data was available through the US Census Bureau's TIGER (Topologically Integrated Geographic Encoding and Referencing) programme, its coverage was limited to streets and roads, excluding features like green spaces and landmarks. Additionally, because of the high cost of mapping, TIGER updates occurred infrequently and failed to reflect rapid changes (Haklay & Weber, 2008). While commercial geographic data providers, such as NAVTEQ (now part of HERE), offered more comprehensive data, their products were expensive (Haklay & Weber, 2008). The business model behind OSM was largely inspired by Wikipedia, which was launched in 2001 (Coast, 2015). Naturally, OSM was dubbed the *wikification of GIS/mapping* (Perkins, 2014; Sui, 2008).

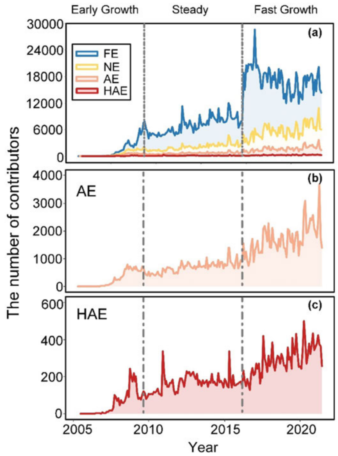
Historically, mapmaking has been the domain of professionals in governments and scientific institutions (Crampton, 2001). In contrast, OSM ostensibly offers an avenue for broader public participation in the generation of geospatial data (Chauhan et al., 2024; Schröder-Bergen et al., 2022). Following the footsteps of the open-source software movement, the early contributors developed accessible map editing tools, established mailing lists and wikis for coordination, and organised local mapping events (often called mapping parties or mapathons) along with annual conferences (Chauhan et al., 2024). Nevertheless, different from Wikipedia, OSM decided to allow only registered users to contribute data so that the data source can be traced in case of disputes, such as due to copyright infringement (Haklay & Weber, 2008).

Two major technological enablers ushered in the emergence and growth of OSM. First was the accessibility of GPS receivers. In May 2000, US President Clinton announced the removal of the Selective Availability of the Global Positioning System (GPS) signal, consequently allowing much-improved accuracy for simple, low-cost civilian GPS receivers. By mid-2001, a receiver unit could be purchased for approximately US$100 (Haklay & Weber, 2008). OSM contributors can thus use GPS receivers to collect and upload data into OSM. Second, the availability of satellite imagery. At the end of 2006, Yahoo granted the right to use its satellite imagery to trace roads and other features, essentially enabling remote mapping apart from field mapping (Haklay & Weber, 2008). Since then, many other organisations such as Microsoft, Esri, DigitalGlobe, and Mapbox have also provided their satellite imagery for OSM contributors to use (Mandourah & Hochmair, 2024; OSM Wiki, 2023g).

A note on terminology: Adding or deleting data on OSM is often called 'editing'. Thus, in this chapter, editing is also used interchangeably with contributing data or simply, mapping. The data contributors to OSM are also referred to as editors or mappers, and the tools used to contribute data are called mapping tools or (map) editing tools.

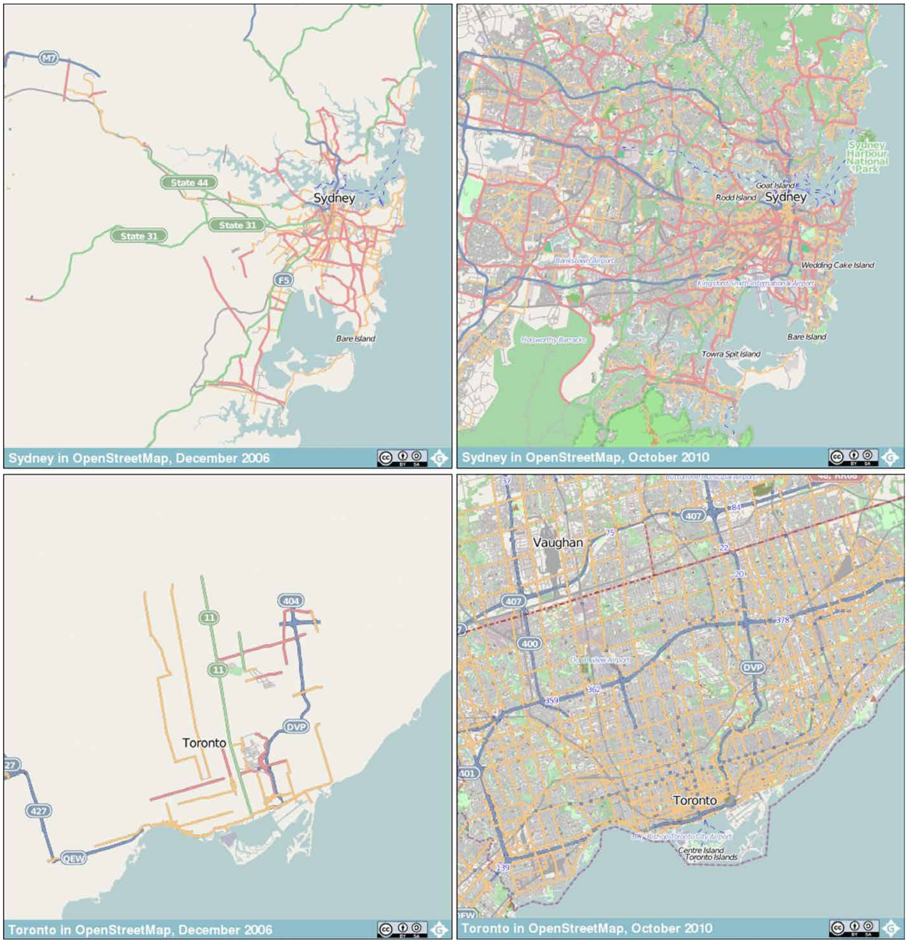
6.3.2 Growth of OSM

OSM has grown tremendously over the past two decades from a small group of developers to a large and heterogeneous global community (Chauhan et al., 2024; Choe et al., 2023a). OSM is, by far, the most successful and extensive mapping project (Schröder-Bergen et al., 2022), with more than 10 million registered editors, and over 5 thousand daily active editors, collectively making around 3 million changes every day (Neis, 2025b). Figure 6.1 shows the number of OSM editors from 2005 to 2021, analysed and visualised by Zhang et al. (2024). In the figure, FE stands for first-time editors (one edit uploaded), NE for new editors (> 100 edits in total), AE for active editors (> 1,000 edits/month), and HAE for highly active editors (> 10,000 edits/month). Zhang et al. (2024) divided the period into three phases. The pioneers of OSM contributed to the early growth phase (2005-2007). As the pioneers began to promote OSM widely, the number of editors increased steadily between 2007 and 2016. Figure 6.2 illustrates the growth of OSM map coverage in Sydney, Australia and Toronto, Canada, over four years, from 2006 to 2010. Starting in 2016, the number of first-time editors increased drastically, which can be attributed to the intensive involvement of corporate editors in OSM (Zhang et al., 2024).



◄ FIG. 6.1 Monthly numbers of OSM editors, analysed by Zhang et al. (2024) [Source: Zhang et al. (2024), provided under CC-BY 4.0]

▼ FIG. 6.2 Sydney, Australia (top) and Toronto, Canada (bottom) in OSM in December 2006 (left) and October 2010 (right) [Source: Geofabrik (2024a)]



TOC

6.4 Business model of OSM

Note that compared to Chapter 5 of the Esri case study, the value capture dimension of OSM's business model is intentionally presented before its value creation. This is because understanding the value capture aspects of OSM, i.e., why OSM is developed and maintained, is necessary to follow its value creation, i.e., the activities and resources involved in developing and maintaining OSM.

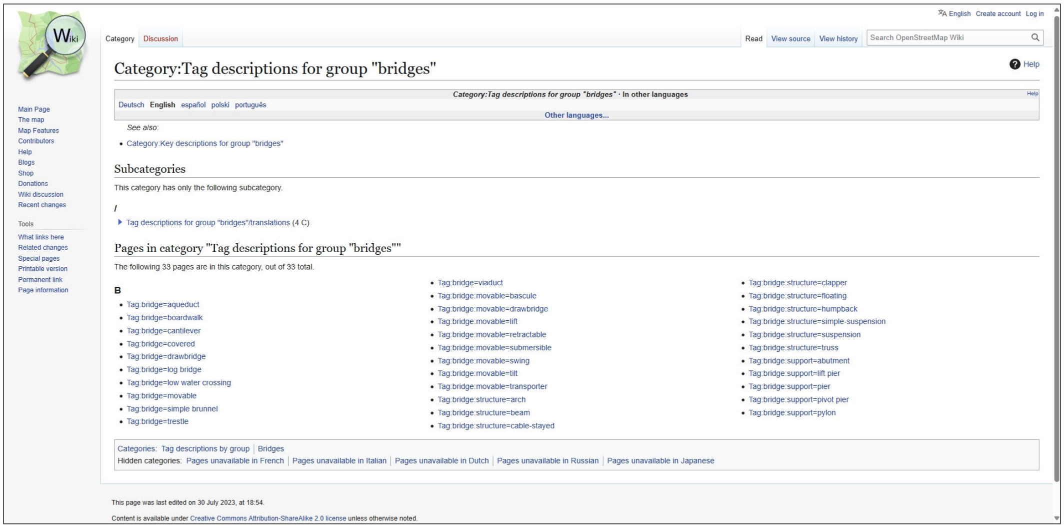
6.4.1 OSM's value proposition: What does OSM offer?

Database based on a simple data model

OSM offers a simple data model compared to traditional geospatial data models (Plennert et al., 2019). Based on the OSM data model, every object is defined by a geometric and a semantic component. The geometric component can be described by (i) a *node* (a point feature with a latitude and a longitude), (ii) a *way* (an ordered list of nodes to represent a linear feature, e.g., a highway and a river, and a polygon feature, e.g., a building and a forest), and (iii) a *relation* (to represent the relationships between existing nodes and ways, e.g., a turn restriction on roads). The semantic component is described by attributes, called *tags*, consisting of 'key=value' pairs, such as 'amenity=restaurant' (Minghini et al., 2022; OSM Wiki, 2024d).

There are at least two main differences between the OSM data model and traditional geospatial data models (Plennert et al., 2019). First, the OSM data model does not have a polygon element. A polygon feature (i.e., boundaries of an area) is represented by ways. This has an implication, for example, in how addresses are represented. The address tags can be attributed to single nodes outside, inside, or on the perimeter of a building footprint. These varied mapping practices are typically established by local, regional, or national OSM communities (Sarretta & Minghini, 2021).

Second, while most governmental and commercial geospatial data providers employed a comprehensive attribution system (e.g., TIGER definitions and ISO standards), OSM utilises a folksonomy, a tagging system that allows OSM editors to freely create semantic annotations to classify data. Consequently, the ontology of OSM data evolves organically rather than being pre-defined (Plennert et al., 2019).



Category Discussion | Read View source View history | Search OpenStreetMap Wiki | English Create account Log in

Main Page
The map
Map Features
Contributors
Help
Blogs
Shop
Donations
Wiki discussion
Recent changes

Tools
What links here
Related changes
Special pages
Printable version
Permanent link
Page information

Category:Tag descriptions for group "bridges"

Help

Category:Tag descriptions for group "bridges" · In other languages Help
Deutsch **English** español polski português
Other languages...

See also:

- Category:Key descriptions for group "bridges"

Subcategories

This category has only the following subcategory.

/

- Tag descriptions for group "bridges"/translations (4 C)

Pages in category "Tag descriptions for group "bridges""

The following 33 pages are in this category, out of 33 total.

B

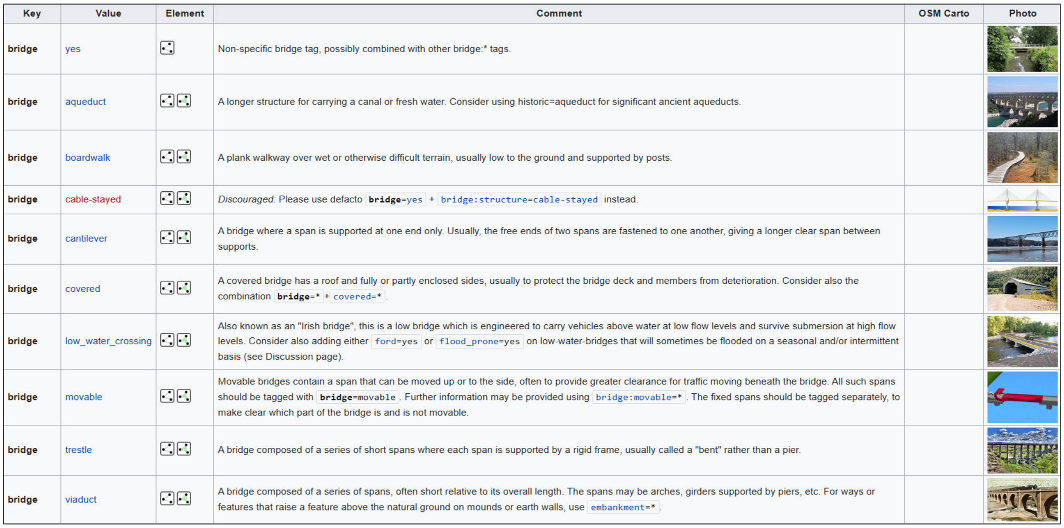
- Tag:bridge=aqueduct
- Tag:bridge=boardwalk
- Tag:bridge=cantilever
- Tag:bridge=covered
- Tag:bridge=drawbridge
- Tag:bridge=log bridge
- Tag:bridge=low water crossing
- Tag:bridge=movable
- Tag:bridge=simple brunnel
- Tag:bridge=trestle
- Tag:bridge=viaduct
- Tag:bridge:movable=bascule
- Tag:bridge:movable=drawbridge
- Tag:bridge:movable=lift
- Tag:bridge:movable=retractable
- Tag:bridge:movable=submersible
- Tag:bridge:movable=swing
- Tag:bridge:movable=tilt
- Tag:bridge:movable=transporter
- Tag:bridge:structure=arch
- Tag:bridge:structure=beam
- Tag:bridge:structure=cable-stayed
- Tag:bridge:structure=clapper
- Tag:bridge:structure=floating
- Tag:bridge:structure=humpback
- Tag:bridge:structure=simple-suspension
- Tag:bridge:structure=suspension
- Tag:bridge:structure=truss
- Tag:bridge:support=abutment
- Tag:bridge:support=lift pier
- Tag:bridge:support=pier
- Tag:bridge:support=pivot pier
- Tag:bridge:support=pylon

Categories: Tag descriptions by group | Bridges
Hidden categories: Pages unavailable in French | Pages unavailable in Italian | Pages unavailable in Dutch | Pages unavailable in Russian | Pages unavailable in Japanese

This page was last edited on 30 July 2023, at 18:54.

Content is available under Creative Commons Attribution-ShareAlike 2.0 license unless otherwise noted.

FIG. 6.3 Common tags associated with bridges documented in the OSM Wiki [Source: OSM Wiki (2023), provided under CC-BY SA 2.0]



Key	Value	Element	Comment	OSM Carto	Photo
bridge	yes		Non-specific bridge tag, possibly combined with other bridge:* tags.		
bridge	aqueduct		A longer structure for carrying a canal or fresh water. Consider using historic=aqueduct for significant ancient aqueducts.		
bridge	boardwalk		A plank walkway over wet or otherwise difficult terrain, usually low to the ground and supported by posts.		
bridge	cable-stayed		*Discouraged*: Please use defacto `bridge=yes` + `bridge:structure=cable-stayed` instead.		
bridge	cantilever		A bridge where a span is supported at one end only. Usually, the free ends of two spans are fastened to one another, giving a longer clear span between supports.		
bridge	covered		A covered bridge has a roof and fully or partly enclosed sides, usually to protect the bridge deck and members from deterioration. Consider also the combination `bridge=*` + `covered=*`.		
bridge	low_water_crossing		Also known as an "Irish bridge", this is a low bridge which is engineered to carry vehicles above water at low flow levels and survive submersion at high flow levels. Consider also adding either `ford=yes` or `flood_prone=yes` on low-water-bridges that will sometimes be flooded on a seasonal and/or intermittent basis (see Discussion page).		
bridge	movable		Movable bridges contain a span that can be moved up or to the side, often to provide greater clearance for traffic moving beneath the bridge. All such spans should be tagged with `bridge=movable`. Further information may be provided using `bridge:movable=*`. The fixed spans should be tagged separately, to make clear which part of the bridge is and is not movable.		
bridge	trestle		A bridge composed of a series of short spans where each span is supported by a rigid frame, usually called a "bent" rather than a pier.		
bridge	viaduct		A bridge composed of a series of spans, often short relative to its overall length. The spans may be arches, girders supported by piers, etc. For ways or features that raise a feature above the natural ground on mounds or earth walls, use `embankment=*`.		

FIG. 6.4 Further descriptions of bridge types in the OSM Wiki [Source: OSM Wiki (2024b), provided under CC-BY SA 2.0]

The adoption of folksonomies was made to avoid hierarchical and ‘over-engineered’ standards and provide an easy way to contribute data, including by non-geospatial professionals (Plennert et al., 2019). In 2008, OSM editors began documenting frequently and widely used tags in the OSM Wiki to foster a more consistent view among editors. Before the tags were included in OSM Wiki, they were discussed, for instance, via mailing lists (Mocnik et al., 2017). As an example, Figure 6.3 shows documented tags associated with bridges in the OSM Wiki (OSM Wiki, 2023a). Figure 6.4 shows descriptions, including illustrative images, of different types of bridges documented in the OSM Wiki (OSM Wiki, 2024j).

Infrastructure to contribute data

Besides field mapping (using GPS devices) and remote mapping (using aerial or street-level imagery) to collect and contribute data to OSM, importing external datasets is the third way. These external datasets range from small-scale (e.g., tree datasets in a city) to large-scale, which is ‘more than a few hundred nodes or for a larger area like a whole country’ (OSM Wiki, 2025i). External data importing is conducted for various reasons, such as to generate a baseline dataset for stimulating further contributions by OSM editors, especially in countries or regions with small or less active OSM communities (Witt et al., 2021). In general, external data importing is cautioned as it is non-trivial and potentially disruptive to the existing data in OSM (Minghini et al., 2022; Witt et al., 2021). The importing must be planned, documented, and discussed with the relevant OSM communities and on the dedicated OSM imports mailing list in advance (OSM Wiki, 2025aa).

The fourth and much less common way of contributing data in OSM is through scripts and bots. Scripts are ‘executed manually and consequently run at an irregular schedule or are executed only once’ (van Berkel & Pohl, 2024, p. 38:8). In other words, scripts are non-routine automated edits (i.e., edits with little human oversight), conducted, for example, to fix incorrect data import (van Berkel & Pohl, 2024). Bots are routine-based automated edits where the bots ‘crawl through the data of contributors to identify predefined aspects within the data and subsequently alter this data in accordance with the bot’s instructions’ (van Berkel & Pohl, 2024, p.38:8). The objective of scripts and bots is typically to reduce repetitive tasks, such as standardising certain tags (van Berkel & Pohl, 2024). In addition to scripts and bots, some types of importing (e.g., bulk import) are also considered automated edits (sometimes called mechanical edits) (OSM Wiki, 2024b). Like external data importing, other automated edits must be discussed with the relevant OSM communities and documented (OSM Wiki, 2025ab). Any unresolved disputes among OSM members concerning automated edits shall be mediated and addressed by the OSMF Data Working Group (DWG) (OSM Wiki, 2025ab).

In general, OSM editors do not interact with the OSM database directly (van Berkel & Pohl, 2024). Instead, they use software editing tools developed by the communities, including companies. Some of the popular editing tools for desktops are iD (found by default on the OpenStreetMap homepage), JOSM (a standalone Java desktop application with more advanced features), and Rapid (developed by Meta and includes features to verify AI-generated data and integrate it into the OSM database) (OSM Wiki, 2025a). In addition, although it is way less popular, ArcGIS, GIS desktop software by Esri, can also be used to contribute data to OSM. This is not directly possible with QGIS, the open-source equivalent of ArcGIS (OSM Wiki, 2025a)

Meanwhile, some popular mobile editing tools include StreetComplete (Android app with simple functionalities), Vespucci (Android app), Go Map!! (iOS app), and Every Door (Android and iOS app) (OSM Wiki, 2025a). Different editing tools offer different functionalities with different limitations. Table 6.2 describes selected editing tools, including their usage shares based on 2024 statistics (OSM Wiki, 2025b). Based on the table, while only 7% of the total OSM editors used JOSM in 2024 (as it is a relatively advanced tool), the share of edits made through JOSM constituted almost 60% of the total edits. Meanwhile, 78% of the overall OSM editors used iD in 2024. This is expected because, apart from being the default editing tool on the OSM homepage, it is also the default editing tool in the Humanitarian OpenStreetMap Team (HOT) Tasking Manager (HOT, 2019). Furthermore, edits via desktop editing tools (JOSM, iD, Rapid) contributed more than 96% of the total.

TABLE 6.2 Common editing tools to contribute OSM data

Editing tool	Platform	Functionalities			Share of users (%)	Share of edits (%)
		Add, edit, or delete POIs	Edit tags of existing OSM objects	Edit geometries		
JOSM	Windows, Linux, macOS	Yes	Yes	Yes	6.9	59.1
iD	Web-based	Yes	Yes	Yes	78.0	33.1
Rapid	Web-based	Yes	Yes	Yes	1.4	4.0
StreetComplete	Android	Limited	Limited	Limited	10.3	1.0
Vespucci	Android	Yes	Yes	Yes	2.0	0.8
Go Map!!	iOS, macOS	Yes	Yes	Yes	1.6	0.3
Every Door	Android, iOS	Yes	Limited	No	1.6	0.07

Note: POIs stand for points of interest. Shares of users and edits are based on 2024 statistics. In calculating the share of users, some users may be counted multiple times, as they might have used multiple editing tools within a year; consequently, the sum of the shares of users across editing tools exceeds 100%.
Source: OSM Wiki (2025a, 2025b).

When one registers to become a member of OSM, they are presented with a contributor agreement that clarifies the intellectual property rights of the data contributed. Among other items, the agreement states that the contributor grants to OSMF ‘a worldwide, royalty-free, non-exclusive, perpetual, irrevocable licence to do any act that is restricted by copyright, database right or any related right’ on the data contributed, including for commercial use (OSM Foundation, 2022a).

Infrastructure to use data

There are three common ways to retrieve data from the OSM database (Sarretta & Minghini, 2021). First is through the Application Programming Interfaces (APIs), namely OSM API and Overpass API (OSM Wiki, 2022a). While OSM API is optimised for editing, Overpass API is a read-only API, optimised for using OSM data by selecting parts of the database (OSM Wiki, 2025r). The former is maintained by the OSMF Operations Working Group (OWG) (OSM OWG, 2025), while the latter by an OSM volunteer, Roland Olbricht (OSM Wiki, 2025r). Second is by downloading the entire OSM database through Planet OSM (Planet OSM, 2025). Third is by downloading pre-defined extracts (selected regions or countries) through third-party providers such as Geofabrik (Geofabrik, 2024b; OSM Wiki, 2025z). Besides providing OSM data in its raw PBF format, Geofabrik also gives it in the shapefile format, enabling its direct use with almost any GIS software (Geofabrik, 2024b).

Apart from the three ways described above, OSM data can be directly displayed in an interactive map on a website or application called slippy maps. This is one of the most common ways of using OSM data (OSM Wiki, 2025y). To display an interactive OSM map on a website, several options of tile servers can be used at the backend, including OSM Carto (maintained by OSMF), MapTiler, and Stadia Maps (OSM Wiki, 2025d, 2025an). At the frontend, several options of JavaScript map libraries can be used, including OpenLayers, Leaflet, MapLibre GL JS, Tangram, and Mapbox GL JS (OSM Wiki, 2025y).

Figure 6.5 visualises the value proposition of OSM based on the key technical infrastructure to contribute and use OSM data. It is clear from the figure that most of the tooling is provided and maintained by a third-party instead of by the OSMF.

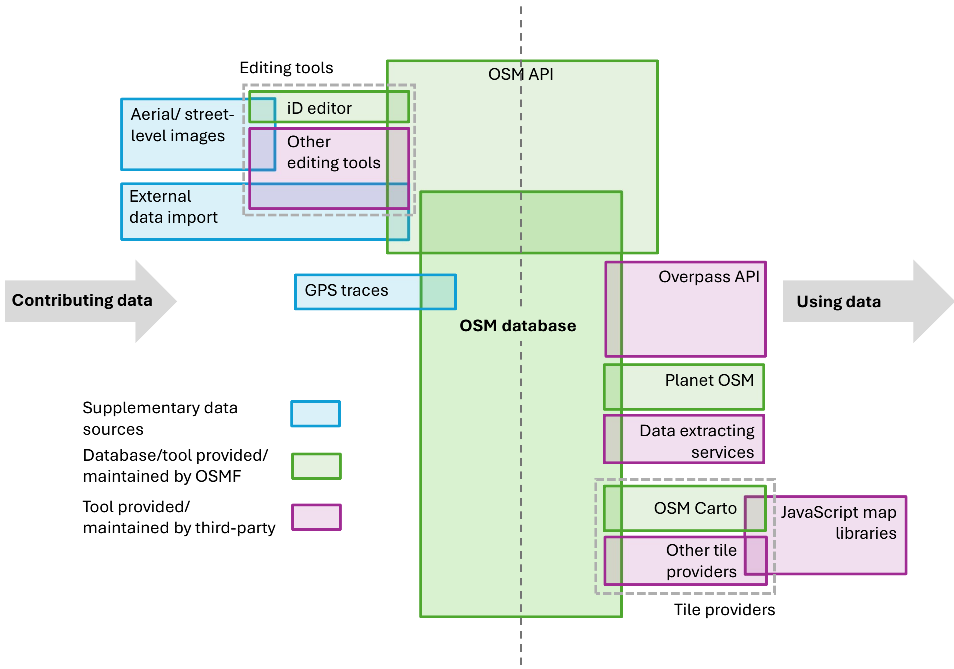


FIG. 6.5 OSM value proposition based on the key technical infrastructure to contribute and use data

6.4.2 OSM's value capture: Why is OSM developed and maintained?

OSM captures value mainly through the contributions of labour from volunteers (e.g., developing and maintaining OSM technical infrastructure and providing data) and funds from sponsors (e.g., voluntary donations and corporate membership fees). Unlike organisations that offer products and services with specific expected benefits (e.g., profits) to be obtained from others outside their organisations, OSM is developed and maintained by those who expect benefits to be gained from OSM itself (e.g., companies that contribute and use OSM data simultaneously). Therefore, the value capture dimension of OSM can be studied through the motivations of the contributors (not only data contributions but also other resources such as funding and labour).

Ideological drive

Many OSM contributors are driven by the ideology for open data (Coast, 2015) (OSMC-09, OSMF-01). Nagaraj & Piezunka (2024) conducted an econometric analysis to compare data contributions to OSM before and after the entry of Google Maps in 87 countries, spanning five waves of entry between 2009 and 2011. They found that contributors who had previously contributed data to OSM intensified their contributions after the introduction of Google Maps. Upon further analysis, their study highlighted the role of ideology-based attachment that motivated these contributors to intensify their efforts. As one of them said:

‘The point is that when you use any (traditional) map provider, you are handing them the controls—letting them determine what features get emphasised, or what features may not be displayed at all’ (Nagaraj & Piezunka, 2024).

This sentiment is also illustrated in the opinion piece titled ‘Why the world needs OpenStreetMap’ by Wroclawski (2014) in The Guardian, who wrote:

‘Every time I tell someone about OpenStreetMap, they inevitably ask “Why not use Google Maps?” From a practical standpoint, it’s a reasonable question, but ultimately this is not just a matter of practicality, but of what kind of society we want to live in’.

Furthermore, many OSM advocates are also open-source software advocates (OSMC-04, OSMC-05), which lends further promotion of OSM through the intermingling of the two communities. Notably, several OSM local chapters are also a chapter of OSGeo, such as OSGeo Oceania, FOSSGIS Germany, FLOSSK in Kosovo, and GeoLibres in Argentina (OSM Wiki, 2025ac).

Mapping passions

Some OSM contributors are motivated by their interests in mapping and the power of creating their own data, as illustrated by the remark by an interviewee (YMAP-01):

‘We have the power to create our own data, and we’re not just the consumers of the data that we want to use’.

Similarly, another interviewee (OSMC-03) said that

‘Contributing to OpenStreetMap is strangely very selfish because you map what interests you. And so if you find something deficient, you can make it right’.

The interviewee from OSM Japan shared a personal story of how OSM allowed him to pursue his interest in mapping, which was inspired by a Japanese icon, Inō Tadataka, who was responsible for the first map of Japan based on scientific surveying methods (Steele, 2019).

Community feeling

The community feeling in OSM is also one of the primary motivations for OSM contributors, anchored by shared interests and goals (Budhathoki & Haythornthwaite, 2013). Interviewee OSMC-07 shared:

‘Many contributors value the collaborative and inclusive nature of the community we created, which is the OSM community, where individuals from diverse backgrounds come together and create a shared resource’.

Another interviewee (OSMC-08) shared:

‘It’s also this feeling of community. It’s not one thing that we are doing, but we are building this larger thing from just a single thing that we are doing’

In addition, several interviewees (OSMC-01, OSMC-13) alluded to how the community feeling is not only confined to within OSM but transcends across other open data and open source communities. Such community feeling is also associated with enjoyment from social interactions (OSMC-02), especially through OSM mapping events (Coast, 2015; Schott et al., 2021).

Humanitarian and civic causes

Humanitarian crises, such as natural disasters, often attract a surge of OSM contributors who produce large amounts of data in a short time (Anderson et al., 2018; Madubedube et al., 2021). This resonates with the ethos of altruism as one of the motivations to participate in OSM (Budhathoki & Haythornthwaite, 2013). For example, during the 2010 Haiti earthquake, in just over 3 weeks, around 600 remote mappers managed to build a basemap for the Port-au-Prince region and its immediate surroundings nearly from scratch. This map then served as the default map for rescue and relief efforts (Soden & Palen, 2014; Zook et al., 2010).

Similarly, Seto (2022) discovered that the first wave of OSM editors mapping in Japan occurred during the Great East Japan Earthquake in March 2011. Another surge later happened in April 2016, which can be attributed to the Kumamoto earthquake (Seto, 2022). During the 2013 Typhoon Haiyan in the Philippines, around 2,000 OSM editors from over 80 countries contributed to mapping the affected regions (C. H. Park et al., 2020). Comparing the edits in Nepal between before and after the 2015 Gorkha earthquake within one month, Poiani et al. (2016) found that 99% of the edited nodes and ways were made after the earthquake. There were also 4,287 new OSM editors registered during the earthquake (Poiani et al., 2016). In Mozambique, Madubedube et al. (2021) discovered that more new editors were mapping the country during Cyclones Idai and Kenneth.

Apart from natural disasters, OSM editors also contributed to mapping conflict areas to facilitate humanitarian responses. In fact, the first mobilisation of OSM editors was in response to the conflict in Gaza in January 2009 (Klapper et al., 2020), where fundraising was done to buy satellite imagery from a Digital Globe reseller to be used for the mapping (Coast, 2015). More recently, following the escalation of conflict in Sudan, there was a significant increase in the creation of new nodes in the country starting in the second half of 2023 (Scholz et al., 2024). Overall, since 2009, the global OSM community has responded to over 90 disasters or crises (C. H. Park et al., 2020).

Beyond post-crisis responses, OSM contributors are also motivated by the potential of OSM for proactive humanitarian or civic mapping (Herfort et al., 2021). A notable example is a project initiated in 2009 to map one of the largest informal settlements, Kibera, in Nairobi, Kenya. The project received funding from UNICEF, and four themes were given attention: water and sanitation, education, health, and security (Coast, 2015; Haklay, 2013; Map Kibera, n.d.). Further examples: The OSM community in Bangladesh gathered field data on arsenic-contaminated wells and included it in OSM to improve water quality in the country (Hunt, 2017; C. H. Park et al., 2020). Tanzania Development Trust initiated a project called Crowd2Map to map rural Tanzania using OSM, intending to tackle female genital mutilation and support community development (Crowd2Map, n.d.; C. H. Park et al., 2020).

Business purposes

Business motives have been one of the motivations behind the development and maintenance of OSM since its inception. Steve Coast co-founded CloudMade in 2007, just a few years after launching OSM (Coast, 2015; OSM Wiki, 2024c). Fast forward to 2020, Accenture roughly estimated the value of OSM at more than US$1.67 billion based on the business ecosystem built around it (Accenture, 2020). A significant trigger that could be linked to the gradual interest of businesses, including large corporations, towards using and contributing OSM data was the changes made by Google to its Maps API licensing model in 2011 (Hardy, 2012; Lardinois, 2012; Zielstra et al., 2013). Over the next five years, companies such as Apple (Arthur, 2012), Foursquare (Lardinois, 2012), and Strava (De Neef, 2015) transitioned from using Google Maps to OSM. During that time, Microsoft had already supported OSM and hired Steve Coast for its map service (Callaham, 2012). A more recent display of companies' interest in OSM can be seen with the establishment of the Overture Maps Foundation in December 2022 by AWS, Meta, Microsoft and TomTom, which leverages OSM data to produce new open datasets for mapping platforms and service developers (Overture Maps Foundation, 2023).

Ochoa-Ortiz & Re (2025) explored the motivations of companies to use OSM. Interviewees in their study ascertained that the high fees of the Google Maps API, in contrast to the free OSM data, were one of the reasons they chose OSM, especially since the latter's quality is, in many cases, comparable to that of the former. Moreover, Singapore-based ride-sharing company, Grab, noted that many commercial geospatial data providers do not capture the specific road typologies in Southeast Asia; hence, it decided to move from relying on companies like Google Maps and HERE to leveraging OSM (Anindya, 2024; Schröder-Bergen, 2024). Indonesia-based ride-sharing company, Gojek, also uses OSM, especially for its routing engine (OSM Wiki, 2022b). The interviewee from Gojek (COM-02) noted that specific information, such as whether a road is accessible by four-wheel or only two-wheel vehicles, is vital in the context of Indonesia. Such information can be added to OSM data.

The global nature of OSM also motivated its use by companies (Ochoa-Ortiz & Re, 2025). This prevents them from having to deal with individual countries' government data laws and bureaucracies. To illustrate: in South Korea, data from the national mapping agency is legally restricted from being exported to anyone accessing from abroad, resulting in long-standing negotiations between the country's government and multinational companies such as Google and Apple (KH Digital, 2016; I.-K. Kim & Seo, 2023). However, since OSM in the country is based on contributions from individual editors and not government data, it is relatively unaffected by such restrictions (OSM Community Forum, 2024).

Another attraction of OSM to companies is the vibrant community that maintains the data and tools around it, allowing companies to direct resources to other parts of their business operations (Ochoa-Ortiz & Re, 2025). As an interviewee from TomTom remarked:

'Consider TomTom as a canvas and we were missing the community part. Now, with OpenStreetMap, we have the community as a very important stakeholder, as part of this canvas. That's why we want to leverage OpenStreetMap because [we] cannot keep spending millions and trillions on maps. Yet, [we're] willing to spend money to empower, to enable, and to make sure that [we] sponsor local communities to make the map better'.

Some companies also recognised the better interoperability potential of open data, including OSM, compared to proprietary data. As one interviewee (COM-04) noted:

'The problem with having a proprietary map is that […] it's not very flexible to change the data models because you own the data model. So, if someone else comes in with new data, you cannot enable them to provide that data. It is costly'.

Companies typically use OSM data as the basemap, layering it with their proprietary data or using it for specific applications. This explains how companies in the same sector can use OSM data simultaneously. For example, ride-sharing companies such as Uber, Lyft, and Grab use OSM for back-end scenarios, such as calculating estimated time of arrival (ETA) and cost, and optimising driver and rider matching (Accenture, 2020). India-based ride-sharing company Ola also uses OSM for the road network, traffic signs and signals, buildings, natural features, land use, and some POIs (OSM Wiki, 2025j). Esri offers an extensive collection of 2D and 3D basemaps on its software, developed with OSM and its proprietary data (Esri, 2024). TomTom leverages OSM and layers it with the company's proprietary data to offer Orbis Maps (TomTom, 2025).

Accordingly, since the quality and coverage of OSM data affect the products and services of companies that use it, these companies are also incentivised to contribute to its development and maintenance (Ochoa-Ortiz & Re, 2025). One way companies contribute to OSM is by being a corporate member of OSMF, with fees starting at €750 annually for the lowest tier and up to €30,000 for the highest (i.e., platinum) tier (OSM Foundation, 2023b). There are 51 corporate members in 2025, including four platinum members: TomTom, Microsoft, Esri, and Meta (OSM Foundation, 2025b).

Another way is through organised editing[9], which involves coordinating a group of OSM editors (for companies, often comprising their staff) to edit OSM with specific and targeted objectives. For example, Uber hired a team to add and modify turn restrictions, directionalities, and road geometries (OSM Community Forum, 2018). Amazon Logistics team carried out similar operations, using driver feedback as a primary source, as well as GPS traces of delivery partners (Accenture, 2020; OSM Wiki, 2025p).

Companies also contribute to OSM by providing aerial or street-level imagery for OSM communities to utilise. For example, to date, Grab has donated almost 2 million street-level images captured by its drivers to KartaView (an open platform for street-level imagery) with the intention for its use to improve OSM (COM-03) (KartaView, 2025). Microsoft released building footprints in the US by employing AI on its imagery, and encouraged their use by the OSM community (Bing Maps, 2018). In addition, companies also develop tools around OSM (Ochoa-Ortiz & Re, 2025). For example, Meta developed and maintained Rapid, one of the OSM editing tools (OSM Wiki, 2025e).

Public institutions' tasks

Some governments leverage OSM to improve their own data or carry out their duties. For instance, in collaboration with Mapbox and the OSM community, New York City has imported over 1 million new buildings, including height information, and more than 900,000 new addresses into OSM. One goal of this data import was to enable the New York City GIS department to monitor changes in OSM, which may indicate that updates to its original data are necessary (Mapbox, 2014). The government department of Natural Resources Canada (NRCan) released the topographic map of Canada (CanVec), which was subsequently imported into OSM. One goal of this activity was also for NRCan to track changes in OSM data, verify those changes in the field, and update its own database accordingly (Haklay et al., 2014).

9 OSMF Organised Editing Guidelines defined organised editing as 'any edits that involve more than one person and can be grouped under one or more sizeable, substantial, coordinated editing initiatives. While primarily aimed at map editing, they can also be applied to other aspects of the project, e.g., the Wiki' (OSM Foundation, 2020b).

Apart from an integrative feedback-channelling strategy of improving government agencies' own data based on community updates in OSM, some governments leverage OSM for specific use cases. For example, the Jakarta Disaster Management Agency needed better data for flood planning and reporting. The agency, in collaboration with international organisations and a university in the city, worked with 267 urban villages to include the critical infrastructure of those villages in OSM (Haklay et al., 2014).

Besides government organisations, international organisations such as the UN also use and contribute OSM data for their tasks. UN Maps is tasked to provide precise, up-to-date geographical information and services to UN Peacekeeping Missions, agencies, and others in conflict-affected areas (UN Maps, n.d.-a). The unit has been using and editing OSM data to overcome the limited availability of data in the field mission areas (OSM Wiki, 2025t). Apart from editing topographic data, such as river networks, roads, bridges, and land cover and use in OSM, UN Maps also occasionally edits POIs, such as communication towers and places of worship (INS-01). OSM is chosen over commercial map providers because it is an open database that provides up-to-date data to personnel on the ground, and it is less affected by some commercial and government perspectives (INS-01). UN Maps has also contributed to editing OSM for disaster relief efforts, such as during the earthquakes in Syria and Morocco and the flood in Derna, Libya, all in 2023 (INS-01). In addition to having paid OSM editors, UN Maps built a community of OSM volunteers called UN Mappers to help respond to the needs of UN field operations (UN Maps, n.d.-c).

Employment and learning opportunities

Some individuals are motivated by employment and learning opportunities from being involved in OSM (Budhathoki & Haythornthwaite, 2013). For example, UN Mappers are often given certificates of recognition for their contributions. Some volunteers appreciate such extracurricular certificates, especially among students, as they can be used during job searching to showcase their skills and interests (INS-01). Similarly, Grab offers the Grab*Star programme, which can be integrated into the core curriculum of university courses or run as an extracurricular activity. The programme offers OSM editing workshops to students (Grab, 2019). Students may join the programme for educational or career goals while indirectly contributing to OSM (OSMC-09). Even without specific programmes such as UN Mappers and Grab*Star, OSM and many open-source tools surrounding it offer opportunities for learning geospatial skills at a low cost (OSMC-08, OSMC-09).

6.4.3 OSM's value creation: What activities and resources are involved in developing and maintaining OSM?

OSMF as the legal representation

OSMF was established in 2006, about two years after OSM was founded (OSM Foundation, n.d.-b). The foundation is registered in the UK as a company limited by guarantee (OSM Blog, n.d.). Key functions of OSMF include acting as a legal entity for the OSM project, being the custodian for the servers and services hosting OSM, providing a vehicle for fundraising activities, organising the annual conference (called State of the Map), and supporting and communicating with the working groups (OSM Blog, n.d.). To contribute to OSM, one must register for an OSM account, but is not required to be an OSMF member. However, only individual members of OSMF have the exclusive right to vote in the foundation's affairs. The voting rights also do not extend to corporate members that comprise sponsoring organisations (i.e., not 'natural persons') (OSM Foundation, 2023d).

There are five bodies within OSMF governing and supporting the foundation and OSM (OSM Foundation, 2024e) (see Figure 6.6). First is the board. The board members are volunteers who are elected by the OSMF members. Within the board, there are three officer roles: chairperson, secretary, and treasurer. Board members with or without officer roles typically have non-officer board-related roles, such as being a liaison to working groups (OSM Foundation, 2024g). Second is the advisory board, a group of individuals who may raise issues or suggestions to the OSMF board or be consulted on important decisions. Currently, the advisory board consists of representatives from Bronze or higher-level corporate members and official local chapters of the foundation (OSM Foundation, 2025a). Third are the local chapters, which are OSMF-recognised country-level or region-level not-for-profit legal entities representing the country/region's OSM editor and OSM data when dealing with the government, business, and media (OSM Foundation, 2024c).

Fourth are the working groups, consisting of volunteers supporting OSM in specific areas. Currently, there are eight working groups: licensing, data, operations, engineering, communication, State of the Map organising, membership, and local chapters and communities (OSM Foundation, 2024c). The fifth is committees and other groups. Committees must be chaired by a board member, and there are two running committees: fundraising and personnel. As necessary, the board may establish special committees to address specific tasks. Other groups are not committees (hence, they do not have to be chaired by a board member) nor working groups (as their tasks are narrower and more operational than those of working groups). Currently, there are three other groups which are responsible for microgrants, software dispute resolution, and mailing list moderation (OSM Foundation, 2024a).

Before 2022, the OSMF had no paid employees, only individual contractors (OSM Foundation, 2025e). Currently, OSMF has an employee as a senior site reliability engineer and three long-term contractors consisting of an accountant, an administrative assistant, and an iD editing tool developer (OSM Foundation, 2025e). The OSMF board, working groups, and committees and other groups are all filled by volunteers. The advisory board and local chapters may be represented by paid employees of other organisations outside of OSMF.

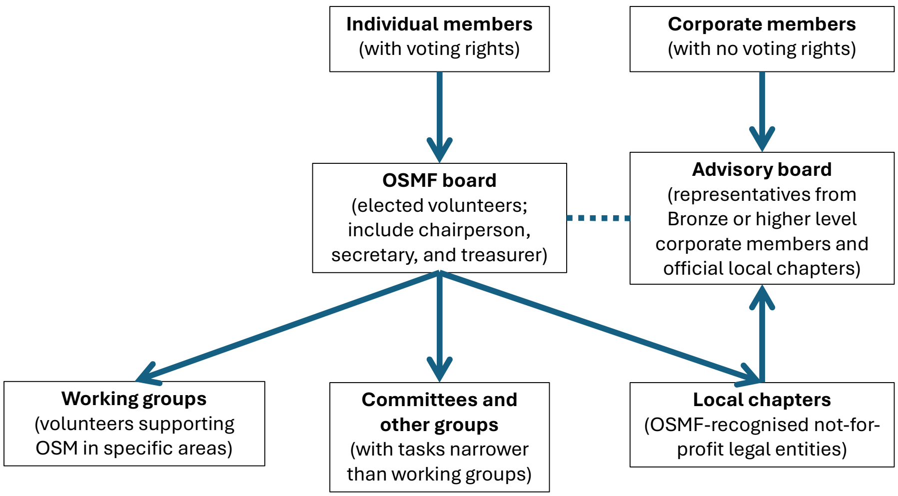


FIG. 6.6 Organisational chart of OSMF

Local chapters and communities

Local chapters recognised by the OSMF scheme started in 2015, with OSM Iceland being the first local chapter, followed by OSM Italy in the following year (OSM Foundation, 2024d). OSM local chapters are separate legal entities from the OSMF. This means, legally, changes to bylaws can be done independently, and OSM local chapters cannot be sued for something on the openstreetmap.org website because it is owned by the OSMF (OSM Foundation, 2024d). Financially, it implies that OSM local chapters manage their own income and expenditure (including membership

TOC

fees and donations, as relevant), and there is no flow of funds from the OSMF to local chapters and vice versa (OSM Foundation, 2024d). Consequently, members of local chapters do not automatically become members of OSMF and vice versa.

Not all local OSM communities establish organisations with legal status, and not all those with legal status are part of the OSMF-recognised local chapters. In other words, there can be (i) informal local OSM communities without legal status, (ii) local OSM communities with legal status but are not part of OSMF-recognised local chapters, and (iii) local OSM communities with legal status that are part of OSMF-recognised local chapters. The existence of OSM local communities predates the formalisation of local chapters. A clear example is that the OSM community in the UK began with the launch of OSM in 2004. Nevertheless, it formed a Community Interest Company (CIC) under UK law in 2016 and became the OSMF-recognised local chapter in 2017 (Companies House, 2024). The OSM community in Japan formed a general corporate association under Japanese law in 2010 (OSM Wiki, 2020d), which became the OSMF-recognised local chapter only in 2021 (OSM Foundation, 2024d). There are 18 OSMF-recognised local chapters as of mid-March 2025, of which 13 are in Europe, and one each in North America, South America, Asia, Africa, and Oceania (OSM Foundation, 2024d).

Local OSM communities establish legal entities, among other reasons, to facilitate financial transactions through a corporate bank account rather than personal accounts. This can help them raise and distribute funds since companies or government organisations often require a corporate bank account to transfer money to (OSMC-09, OSMC-10) (Coast, 2015). Once they establish legal entities, becoming OSMF-registered local chapters is a step further, aiming to increase their visibility (OSM Foundation, 2024d). In addition, they may also sit on the OSMF advisory board, a privilege that is not granted to non-OSMF-registered local communities (OSM Foundation, 2025a). However, both OSMF-registered local chapters and non-registered local communities can be part of the Local Chapters and Communities Working Group (LCCWG) that discusses ideas and issues of OSM communities (OSM Foundation, 2025c).

While OSM local chapters and communities generally aim to promote the contribution and use of OSM data in their region, some have broader objectives, such as simultaneously promoting the adoption of open-source geospatial software (e.g., FLOSS Kosovo and OSGeo Oceania). OSM Belgium was previously part of Open Knowledge Belgium, which has goals extending beyond the geospatial domain, until it became a separate organisation in 2023 (OSM Foundation, 2024d). Meanwhile, OSM Italy is part of Wikimedia Italy. Hence, in some events conducted by Wikimedia Italy, activities targeting OSM and Wikimedia were combined; for example, participants walked around collecting data for OSM, while others took notes and photos for Wikipedia content (OSMC-11).

Local chapters and communities can be viewed as another layer of intermediaries within OSM, as they facilitate local engagement with OSM that the global OSM(F) may not otherwise be able to reach or assist. They speak the local language and are familiar with the local context. Some provide training and consultations on contributing to and using OSM data, carry out OSM (or broadly, open data) advocacy, organise mapping events, and act as a point of contact for local organisations interested in OSM (based on various interviews).

Humanitarian OpenStreetMap Team (HOT)

The potential of OSM for humanitarian purposes has been recognised and explored since the formation of OSM (Klapper et al., 2020). Such promises are one of the factors that attract some of the OSM pioneers to get involved in OSM (Coast, 2015). In August 2010, after the valuable contribution of OSM in responding to the 2010 Haiti earthquake was demonstrated, the Humanitarian OpenStreetMap Team (HOT) was registered in the US as a non-profit organisation (HOT, 2012; OSM Wiki, 2024f), separate from the OSMF.

Over the years, HOT projects have been funded by many organisations, including USAID, the World Bank, the Australian government, UNDP, Bill & Melinda Gates Foundation, Knight Foundation, JP Morgan Chase, Accenture, and AWS (various annual reports in (HOT, 2025a)). In 2020, HOT received a 5-year core funding from The Audacious Project (an initiative of TED, known for TED Talks) (HOT, 2020). Through this funding, HOT decided to become more decentralised and introduced four regional hubs: East Africa, West Africa, Asia and the Pacific, and Latin America and the Caribbean (Radford & Firth, 2020). Each hub defines its own strategy, projects, and work approach, including those related to its funding. With the Audacious funding, HOT expanded from approximately 15 staff to nearly 100 in five years (HOT-01).

The primary and original activity of HOT is coordinating organised OSM editing in response to humanitarian crises (HOT, 2018c). The maps were created together and used by various organisations, including local government agencies, Médecins Sans Frontières (MSF), the International Federation of the Red Cross/Red Crescent (IFRC), the World Central Kitchen (WCK), and the UN World Food Programme (WFP) (HOT, 2024c; Masters & de los Reyes, 2023; Pechmann & de los Reyes, 2023). In addition to crisis responses, HOT also undertakes proactive or participatory mapping projects in line with its five impact areas: disaster and climate resilience, sustainable cities and communities, public health, displacement and safe migration, and gender equality. HOT, together with MSF, the American Red Cross, and the British Red Cross, founded the Missing Maps project in 2014 to join the effort of building a mapping community that proactively maps areas prone to disasters and crises in OSM (Missing Maps, n.d.; OSM Wiki, 2025g).

HOT has been involved in over 60 countries on over 100 projects (Klapper et al., 2020), with a focus on developing countries (Yang et al., 2024). HOT contributed around half of the total OSM changesets[10] in Africa in the past decade (Boateng et al., 2023). This demonstrates the impact of HOT towards OSM data coverage. HOT also exported its datasets into HDX (Humanitarian Data Exchange), a platform by the UN Office for the Coordination of Humanitarian Affairs (OCHA) (HDX, 2025), promoting further re-use of OSM data.

HOT has evolved from coordinating organised OSM editing to also developing open-source mapping tools (HOT-01). Most notably, HOT developed HOT Tasking Manager, which is an open-source software tool to coordinate the organised editing on OSM (OSM Wiki, 2025u) (Figure 6.7). The tool itself has been used by many organisations around the world, of which some do not necessarily engage directly with HOT, including OSM local chapters and communities, MSF, UN agencies, national Red Cross societies, and companies such as Kaart and Grab (OSM Wiki, 2025s). HOT also developed Field Mapping Tasking Manager (FMTM, an Android and web application that facilitates field data collection) and Drone Tasking Manager (DTM, a drone application to facilitate imagery acquisition, processing, and dissemination) (HOT, 2018b). More recently, HOT released an open AI-assisted mapping service called fAIr, which employs computer vision techniques to detect objects such as buildings, roads, waterways, and trees from aerial imagery (HOT, 2018a; Najjar, 2024).

10 Changeset in OSM is 'a group of edits to the database by a single user over a short period of time' (OSM Wiki, 2025ad)

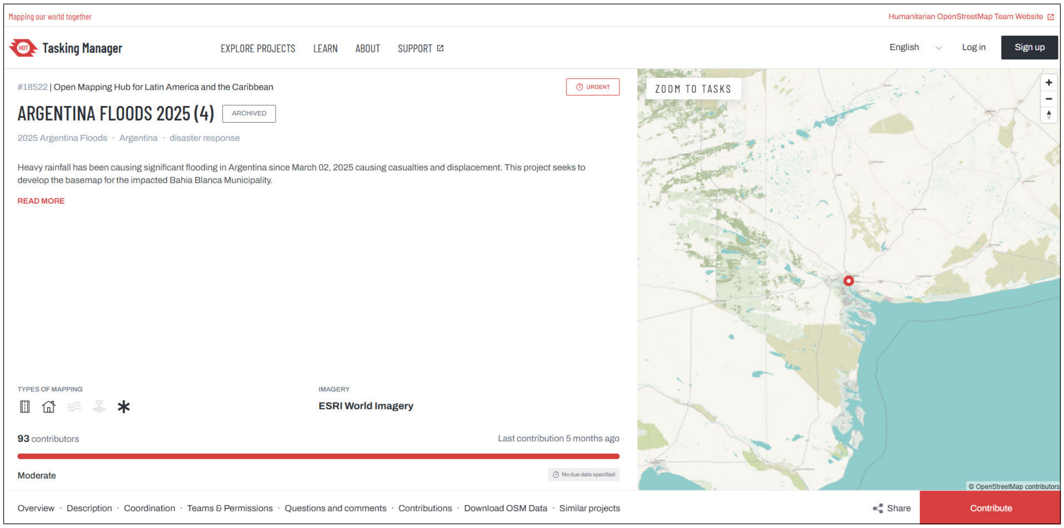


FIG. 6.7 Example of HOT Tasking Manager [Source: HOT (2025), provided under CC-BY SA 4.0]

Furthermore, HOT also aims to install capacities and technical knowledge around mapping in local communities. For example, HOT Asia Pacific (AP) Hub launched She Leads & She Inspires, which is a training and development programme designed by, with, and for women to equip them with open mapping knowledge and leadership skills for addressing social issues at the community level (HOT, 2023a). HOT AP also introduced the Open Mapping Guru project, which aims to empower a network of experienced OSM contributors to lead training, produce learning materials, and support mapping activities (OSM Wiki, 2024g).

Corporate editing

Although companies have been involved in OSM in terms of contributing funds and satellite imagery since the early days, large-scale organised editing by paid editors hired by companies (termed corporate editing) started no earlier than 2015 (Patel et al., 2023; Sarkar & Anderson, 2022; Yang et al., 2024). While organised editing had been carried out for humanitarian purposes long before corporate editing emerged, the latter represents a new form of engagement, as the contributions of the editors are compensated and meant for specific business purposes (Patel et al., 2023; Sarkar & Anderson, 2022).

Based on the recent publicly available figures, Amazon Logistics hired almost 500 people to edit on OSM, not including reviewers (OSM Wiki, 2025q), while Apple had around 250 (OSM Wiki, 2024l), Uber around 90 (Uber, 2020), Meta around 50 (OSM Wiki, 2024m), and Microsoft 14 (OSM Wiki, 2025af). Yang et al. (2024) identified the top ten countries where different companies have made OSM edits between 2015 and 2020 (Table 6.3).

TABLE 6.3 Top ten countries with the largest OSM edits by different companies, 2015 – 2020 [Source: Yang et al. (2024)]

Amazon	Apple	Uber	Meta	Microsoft
US	Brazil	New Zealand	India	Australia
UK	Mexico	India	Thailand	Serbia
Germany	Indonesia	Philippines	Indonesia	Peru
Italy	Philippines	Japan	Vietnam	Venezuela
India	Russia	Kenya	Malaysia	US
Canada	South Africa	Colombia	Tanzania	Indonesia
France	Malaysia	Mexico	South Africa	New Zealand
UAE	China	Brazil	US	Myanmar
Spain	Chile	United States	Laos	Fiji
Kenya	Ukraine	Egypt	Myanmar	Nigeria

The top ten countries are significantly different across companies, reflecting each company's business priorities and markets (Yang et al., 2024). For example, nine out of ten countries where Amazon focused its OSM editing are also among the top ten countries in its market (Hjorth, 2021). India, being the top country for Meta, aligns with the number of Facebook and Instagram users in that country, which is by far the highest (Dixon, 2024a, 2024b). Furthermore, Yang et al. (2024) also discovered that some companies are involved in many countries, but each with a small number of edits, while some are the other way around. For instance, Meta has edited in 159 countries but with a median number of edits of only 291, whereas Amazon has only edited in 30 countries, but the maximum number of edits made in a single country was more than 61 million (Yang et al., 2024). This may reflect the different purposes for which OSM data is used by Meta (for its social media) and Amazon (primarily for logistics).

Sarkar & Anderson (2022) found that compared to other OSM editors, corporate editors tend to edit the works of other editors and have their work edited by others. Nevertheless, corporate editors are more likely to edit other corporate editors' works than non-corporate editors (Sarkar & Anderson, 2022). Besides, corporate editors mainly focused on road networks, whereas non-corporate editors also gave attention to other features such as buildings and POIs (Sarkar & Anderson, 2022). As the number of corporate editors increases over time, these patterns may create isolated corporate editing clusters within OSM, even though such a scenario is not yet happening (Sarkar & Anderson, 2022).

YouthMappers

Another important group within the OSM community is YouthMappers, which has more than 200 university chapters in over 40 countries (Solís et al., 2021). YouthMappers was founded in 2014 by academics from Texas Tech University, The George Washington University, and West Virginia University with support from USAID (Solís et al., 2021). The consortium is now administered at Arizona State University (Solís et al., 2021). The goal of YouthMappers is 'to fulfil the demand for open geospatial data access in all parts of the world' by establishing a community of students, researchers, educators, and scholars (YouthMappers, n.d.-c). Students participated in YouthMappers for their careers (by acquiring and demonstrating geospatial skills) and global citizenry motivations (real-world issues and social changes) (Hite et al., 2018; Solís et al., 2021).

The main activity of YouthMappers is organising mapping activities and adding data to OSM (YouthMappers, n.d.-b). YouthMappers have edited over 9 million buildings and 145,000 km of roads on OSM (Solís et al., 2021), and almost all of the features edited are in Africa and Asia (Boateng et al., 2023). In line with its tagline, 'We don't just build maps. We build mappers.' (YouthMappers, n.d.-c), YouthMappers also conducts training and leadership programmes. For instance, it organised YouthMappers Leadership Fellowship, where selected students were provided travel support to attend an on-site workshop to receive training and mentorship (YouthMappers, 2024). The consortium also developed an online learning platform called YouthMappers Academy, where its members can follow courses related to OSM (including mapping with the iD editing tool, interpreting satellite imagery, understanding the OSM data model), as well as related to conducting field surveys and managing a YouthMappers chapter (YouthMappers, n.d.-d). Additionally, YouthMappers launched the 'Everywhere She Maps' campaign to improve women's participation and women-focused data inclusion in mapping (YouthMappers, n.d.-a).

Mapping events

Mapping events, colloquially called mapping parties or mapathons, where OSM editors gather physically (or virtually at the same time) to edit OSM data of a particular area (either via field or remote mapping), have been an integral part of the OSM community (Coetzee et al., 2018). One of the first OSM mapping events was organised in the Isle of Wight, UK, in May 2006, where OSM contributors attempted to map the entire island using GPS receivers (Haklay & Weber, 2008; OSM Wiki, 2020c). Mapping events offer opportunities for OSM editors to socialise and for new editors to learn from the more experienced ones. They may also attract the involvement of people who may not be initially aware of or interested in OSM but are drawn by the social aspect of the events (Schott et al., 2021; Solís et al., 2021). This is in line with the findings from a large-scale study on technology gatherings (not particularly on OSM mapping events) that found that they serve as a social forum for sponsorship, social learning, knowledge exchange, and social coordination (Fang et al., 2021).

Studies have shown that retaining new users in the long run after mapping events is still a challenge (Juhász & Hochmair, 2018; Schott et al., 2021). However, this indicates the diverse motivations of people contributing to OSM; for example, some new editors may become involved in OSM editing in response to a humanitarian crisis and are motivated by the desire to help rather than the OSM project itself.

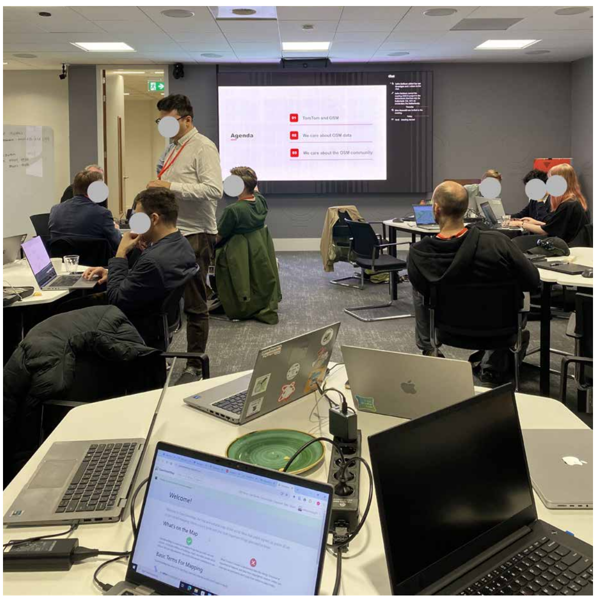


FIG. 6.8 OSM mapping event in Amsterdam, the Netherlands, organised by TomTom in December 2024 [Source: Author.]

Note: In the photo, experienced OSM editors helping new participants.

Established rules and norms

Over time, the OSM community has established editing standards and conventions (OSM Wiki, 2025f, 2025l). Notably, OSM follows the 'on the ground principle', which means that 'street names and other proper names are generally entered as they appear on signs, even if those names deviate from the general spelling rules' (OSM Wiki, 2025f). In cases where the on-the-ground signs do not match the official names, the latter can be added as an additional tag of 'official_name=*' (OSM Wiki, 2025f). OSM editors are not supposed to map historic events or features because 'such features cannot be verified'[11] (OSM Wiki, 2025l). For disputed names, such as in contested areas, the default rule is 'whatever name [that is] used by the people on the ground at that location' (OSM Wiki, 2025l). Additionally, although in theory anyone can use any tag on objects in OSM, in practice, OSM editors often follow the established OSM folksonomy or introduce new tags through open voting processes on the OSM Wiki (Minghini et al., 2022; Quinn & Bull, 2019).

Triggered by the rise of corporate editing, the OSMF approved the Organised Editing Guidelines (OEG) in 2018 (OSM Foundation, 2020b; OSM Wiki, 2025v). The guidelines lay out the process that must be followed by parties that intend to conduct organised editing based on a 'best-effort approach' (OSM Foundation, 2020b). Problematic edits may be reverted, and severe or repeated issues may lead to a ban. The power of banning (generally, blocking, which includes temporary blocking) lies with the DWG (OSM Foundation, 2023a).

DWG publicly documented OSM editors that have been blocked, and the reason for each block (OSM Foundation, 2025g). Quinn & Bull (2019) categorised the blocking reasons into four themes: nefariousness, obstinance, ignorance, and mechanical problems. Nefariousness refers to edits that demonstrated no interest in the overall goals of OSM, including vandalism (actions that intentionally harm OSM data accuracy or credibility), politically-motivated edits, sock-puppetry (new OSM accounts created by the same individuals/organisation that were previously blocked), and spam (using OSM tags for marketing). Obstinance refers to the behaviour of editors who refuse to engage constructively with other editors, and attempts at correction or collaboration are met with stubbornness or hostility. Ignorance refers to well-meaning edits based on a misunderstanding of OSM practices or software, copyright violations, or incorrect data. This type of problem often involves only temporary blocking. Mechanical problems refer to problematic edits done via automated editing (Quinn & Bull, 2019).

11 However, the OSM US chapter runs OpenHistoricalMap (OHM), which uses OSM technology to map objects that existed in the past (thus, do not belong in OSM database). OHM data is provided under the CC0 license, and not the ODbL like the OSM data (OHM, n.d.; OSM Wiki, 2025n).

Financial support

One of the earliest significant donations to OSMF was the US$575,000 grant from the Knight Foundation in 2012 to improve the OSM infrastructure, particularly the iD editing tool (Barth, 2013; Zielstra et al., 2013). Currently, corporate membership is a primary source of revenue for OSMF. The first corporate membership scheme was introduced in 2014 (OSM Blog, 2014), and in 2016, a new tiered membership scheme was implemented (OSM Blog, 2016). About half of the corporate members are headquartered in the US and Germany (Ochoa-Ortiz & Re, 2025) (Table 6.5). Apart from corporate membership, another main source of revenue for OSMF is through donations, which can be categorised into regular donations received throughout the year[12] and hardware donations (OSM Foundation, 2025d). Other sources of revenue include individual memberships and the sponsorships of the annual State of the Map conference. Table 6.4 shows that OSMF's revenue increased from £260,624 in 2019 to £445,794 in 2023 (OSM Foundation, 2025d). The majority of the expenditure went to wages, administrative fees, and website and computer costs (OSM Foundation, 2025d).

At the same time, OSM local chapters and communities obtain funds in various ways. One of them is through specific projects and grants, mainly from companies and international organisations, such as Global Basel Infrastructure (GBI), UN Development Programme (UNDP), World Bank, and USAID (KLL, 2023; POI, 2023) (OSMC-02, OSMC-06, OSMC-07, OSMC-08). Occasionally, they also conduct projects with government agencies, such as OSM Indonesia with the Special District Capital of Jakarta, to develop an urban farming platform called SiPetani (Sulistioningrum, 2023). Some OSM local chapters and communities, such as OSGeo Oceania, FOSSGIS Germany, and OSMF Japan, generate funds mainly through conference sponsorship (OSMC-05). Others, such as OSM UK, rely on membership fees (OSMC-03). Some of them, such as OSMF Japan, also receive in-kind donations in terms of server services (OSMC-01).

12 https://supporting.openstreetmap.org/donate/

TABLE 6.4 OSMF revenue, 2019 – 2023 [Source: Various OSMF financial statements (OSM Foundation, 2025d)]

Revenue (£) / Year	2019	2020	2021	2022	2023
Donations	27,815	84,187	254,520	133,759	175,307
Miscellaneous income	6	67	0	0	2,848
Conference registrations	45,873	0	0	30,550	0
Merchandising and commission	42	3,749	65	82	59
Conference sponsorship	100,671	40,524	10,263	91,216	9,654
Fiscal sponsor income	5,326	0	0	0	33,779
Membership - corporate	65,881	81,527	74,390	104,083	205,383
Membership - individual	15,010	16,219	15,323	14,734	18,764
Total	260,624	226,273	354,561	374,424	445,794

TABLE 6.5 OSM corporate members as of April 2025, according to different tiers [Source: OSM Foundation (2023b)]

Platinum (€30,000 annually)	Corporate (€15,000 annually)	Silver (€6,000 annually)	Bronze (€2,250 annually)	Supporter (€750 annually)
TomTom Microsoft Esri Meta	Mapbox Grab Gojek Komoot	Cesium Radar OpenCage Geofabrik NextBillion.ai Elastic Graphhopper Bolt HOT Regrid QGIS Calimoto Mapy.cz Niantic	Geotab Omniscale Kaart Krick.com YellowMap NextGIS VK Maps inDrive INIT RINKAI Stadia Maps SUSE Verso Interline E-Smart LandClan	Maptoolkit LocationIQ Passenger Intevation Skyhost PeakFinder Safe Sky Industries Kendall County, Illinois MobilityLabel con.sens mobilitätsdesign farmeye.ie BTC Map Infrageometics

Hardware infrastructure

OSM started at the University College London (UCL) when Steve Coast was working there (Coast, 2015), and UCL has been supporting and hosting a server for OSM since the beginning (Haklay & Weber, 2008; OSM Foundation, 2025f). Currently, servers owned and managed by OSMF are also located in various locations, with the main ones in Amsterdam and Dublin (OSM Foundation, 2025f), storing not only the OSM database but also other services, including tiles[13], aerial imagery, mailing lists, and OSM Wiki. The Operations Working Group (OWG), which currently has seven members, is responsible for proposing the budget for and running the servers owned by the OSMF (OSM Wiki, 2023d).

Some local chapters or communities offer their own map tile servers, which are different from those hosted by OSMF. One of the reasons for providing local tile servers is to comply with national mapping laws, such as those regarding the display of international borders or the removal of military installations (OSM Wiki, 2019, 2023c) (OSMC-01, OSMC-13). Another reason is that the use of tiles provided by OSMF is limited by its usage policy due to the server's capacity constraints (OSM Operations Working Group, n.d.); thus, local tile servers offer an alternative option to users (OSM Belgium, 2024). While OSM users can also engage with third-party providers or host their own tile servers (OSM Wiki, 2025a, 2025e, 2025s), by providing local tile servers, local OSM chapters or communities aim to make OSM data easier to use, especially by domestic users. Examples of OSM local chapters or communities offering local tile servers include OSM Belgium[14], OSMF Japan[15], FOSSGIS Germany[16], OSM communities in Korea[17], and OSM communities in India[18].

13 Map tiles are smaller parts of a larger map that help efficiently render and display a map (instead of rendering the entire map each time the use zooms in and out) (Forrest, 2023; OSM Foundation, 2024b).

14 https://tile.osm.be/

15 https://tile.openstreetmap.jp/

16 https://tile.openstreetmap.de/

17 https://tiles.osm.kr/

18 https://www.openstreetmap.in/

Aerial and street-level imagery

The availability of aerial (including satellite and drone) imagery and street-level imagery has catalysed the growth of OSM as it has allowed the emergence of remote mappers, also known as armchair mappers, who do not have to travel to and wander around a particular location to map it (Mandourah & Hochmair, 2024; OSM Wiki, 2024a). Consequently, not only has such availability facilitated the mapping of unreachable areas (e.g., conflict or disaster areas), but it also has expedited the mapping activity (OSM Wiki, 2023f). Furthermore, it has also allowed those who do not have prior mapping experience, such as using GPS receivers, to be involved in editing OSM (Juhász & Hochmair, 2018; Mandourah & Hochmair, 2024).

A decade ago, with the exceptions of satellite imagery collected through NASA's Landsat and the ESA's Copernicus programmes that was already made open in 2008 and 2013, respectively (Commission Delegated Regulation (EU) No 1159/2013, 2013; Miller et al., 2013), organisations mostly offered their imagery for OSM editing on an ad-hoc basis in response to humanitarian crises. For instance, DigitalGlobe donated its imagery to OSM in response to the 2011 Van Earthquake in Turkey and the 2011 Tōhoku Earthquake in Japan, and several companies, including Airbus Defence and Space, DigitalGlobe, and Mapbox, donated their imagery following the 2014 Ebola outbreak (C. H. Park et al., 2020). The US Department of State's Humanitarian Information Unit initiated the Imagery to the Crowd (IttC) programme, where the organisation published high-resolution commercial satellite imagery purchased by the US government in a web-based format that allowed humanitarian mapping on OSM (Haklay et al., 2014; C. H. Park et al., 2020; Verhulst, 2013).

Nowadays, more organisations have granted continuous special permission for their aerial imagery to be used for OSM, such as Microsoft via Bing Maps and DigitalGlobe through a partnership with MapBox (OSM Wiki, 2023f). In addition, HOT relaunched[19] OpenAerialMap (OAM) in 2015 with a grant from Elrha's Humanitarian Innovation Fund (HIF), providing an open-source platform for searching, sharing, and utilising openly licensed aerial imagery. Imagery provided by OAM is licensed under the CC BY 4.0 (OAM, n.d.; OSM Wiki, 2018). OAM allows OSM editors to select the OSM editing tool, such as iD or JOSM, that they want to use with the selected imagery (HOT, 2024a). By 2021, OAM hosted around 15,000 images shared by 1,300 users worldwide, of which around 40% were taken by drones, 37% by aircraft, and 23% by satellite (Mandourah & Hochmair, 2024).

19 OAM was previously provided by HOT between November 2007 and December 2008.

Besides aerial imagery, street-level imagery has been increasingly capitalised in OSM as it facilitates the refinement of OSM data by including features that cannot be seen from overhead imagery, such as address numbers, storefronts, and street signage (Alvarez Leon & Quinn, 2019; Biljecki & Ito, 2021). OSM editors are not allowed to use Google Street View, even though it is the market leader in street-level imagery, due to its copyrights (Alvarez Leon & Quinn, 2019; OSM Wiki, 2025c). Nevertheless, several (crowdsourced) street-level imagery platforms provide openly licensed imagery that can be used for OSM editing, including KartaView (formerly OpenStreetCam), Mapillary, Mapilio, and Panoramax (OSM Wiki, 2025c). KartaView was developed by Telenav and then transferred to Grab in 2019, whereas Mapillary was created by a Swedish startup and sold to Facebook (now Meta) in 2020 (Ilisei, 2019; Paul, 2020). Imagery from these platforms is (or can be) integrated into OSM editing tools; for example, Mapillary is already in Rapid, and Mapillary and KartaView plug-ins are available for JOSM (Mapillary, 2025; OSM Wiki, 2024i).

External datasets

Many external datasets have been (or are being) integrated into OSM. Those datasets could be under an open licence compatible with ODbL, or the providers gave explicit permission for the data to be imported into OSM (OSM Wiki, 2025i). Although importing external datasets to OSM is not easy and requires careful planning, community engagement, and documentation (OSM Wiki, 2024h; Quinn & Bull, 2019), to date, more than 100 one-time bulk imports and more than 50 community imports have been completed, with many more planned or in progress (OSM Wiki, 2025i). The former is a one-off automated (i.e., via script) import, whereas the latter is a manual import by OSM community members based on external open data. In addition, there are about 15 semi-automated and six fully scripted ongoing imports, which are imports that are regularly updated (OSM Wiki, 2025i).

Some of the earliest one-time bulk imports were street data donated by AND (now GeoJunxion) in India and the Netherlands in 2007, TIGER data in the US in 2007, house numbers data from the Danish National Survey and Cadastre in 2009, and Spanish administrative borders data from National Geographic Institute (IGN) in 2009 (OSM Wiki, 2025i). Meanwhile, notable community imports include Paris building heights from the City of Paris, geospatial data of Antarctica from the Scientific Committee on Antarctic Research, basic registration of addresses and buildings (BAG) data from the Netherlands' Cadastre, Land Registry and Mapping Agency (Kadaster), and building footprints in the US data from Microsoft (OSM Wiki, 2025i).

The effects of external data import are mixed. Witt et al. (2021) found that the number of active OSM editors increased significantly after the introduction of AND data imports in India and the Netherlands in 2007. It could be due to the publicity generated from the data donation (e.g., through blog posts and radio interviews), especially since the imports happened in the early years of OSM (Witt et al., 2021). This finding aligns with the results of Yang et al. (2016), who discovered that in countries without significant imports, the distribution of OSM editing became more unequal over time, with the share of passive OSM editors increasing, whereas the opposite occurred in countries with substantial imports. On the other hand, Nagaraj (2021) found that the TIGER data import in the US led to fewer follow-up edits, resulting in a deterioration of data quality. This might be because the import limited the OSM community's ability to create objects from scratch and develop a sense of ownership (Nagaraj, 2021).

Figure 6.9 shows an overview of the value creation of OSM, through the diverse and interweaving contributions of different actors. Indeed, (some of) these actors are also conversely benefited from OSM data, but those flows (i.e., from OSM to the actors) are not shown in the figure because OSM data is open data and by definition, anyone can benefit from it. Table 6.6 summarises OSM's value proposition, value capture, and value creation.

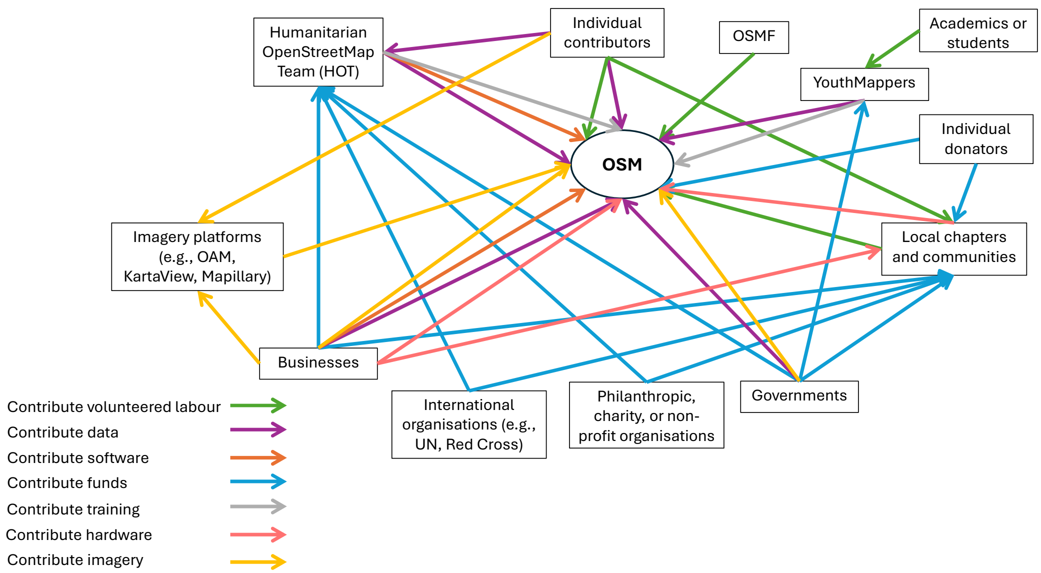


FIG. 6.9 Contributions of various actors to OSM

TABLE 6.6 Summary of the value proposition, value capture, and value creation of OSM's business model

Value proposition (What does OSM offer?)	Value capture (Why is OSM developed and maintained?)	Value creation (What activities and resources are involved in developing and maintaining OSM?)
• Database based on a simple data model • Infrastructure to contribute data • Infrastructure to use data	• Ideological drive • Mapping passions • Community feeling • Humanitarian and civic causes • Business purposes • Public institutions' tasks • Employment and learning opportunities	• OSMF as the legal representation • Local chapters and communities • Humanitarian OpenStreetMap Team (HOT) • Corporate editing • YouthMappers • Mapping events • Established rules and norms • Financial support • Hardware infrastructure • Aerial and street-level imagery • External datasets

6.5 Current strengths, current weaknesses, potential opportunities, and potential threats of OSM's business model to the ODE

6.5.1 Current strengths

Relatively low barrier for contributing data

From the start, OSM was designed to reduce barriers to contributing geospatial data (Coast, 2015; Lenormand, 2022). This is demonstrated, among other things, through its simple data model, based on a free tagging system rather than complex or traditional geospatial data standards, and the absence of an approval process for contributing data, which allows contributors to view the objects they have mapped on OSM immediately. Such design decisions triggered creativity and instant gratification of participants, including those without a geospatial data background, to be involved in OSM (OSMF-01). In turn, OSM attracts a diverse group of contributors with various interests (OSMF-01). This diversity is arguably a key recipe to OSM's growth and sustainability. As interviewee OSMF-01 said,

‘If you see there are shared interests among many different entities and find something that would attract them to be a part of it, then it is much more sustainable […]. If you need to keep something going in the future, and it could not just be money, it could just be interest in the project, you need to have an ecosystem; you need to have very different species. […] If you just have a monoculture, then a disease could come in and totally wipe out that area. So what's good for ecological and agricultural practices is also good for developing open data communities'.

Diverse use cases

With the free tagging system in OSM, the data can be much richer than traditional geospatial data since objects in OSM can theoretically contain indefinite attributes, as a result of the bottom-up instead of the top-down ontology (Arnold & Hukal, 2024; Biljecki et al., 2023). This translates into a wide variety of OSM use cases. For instance, OSM has been used for the development of land use/land cover datasets (Arsanjani & Vaz, 2015; Fan et al., 2021; Schultz et al., 2017) and to estimate urban land value (Carranza et al., 2022). Furthermore, OSM data can also be used as training data to develop a high-resolution classification of urban surfaces based on a deep learning method (Fan et al., 2021).

In the built environment, OSM has been used to develop 3D city models (Ma et al., 2024; Over et al., 2010) and global building morphology indicators (Biljecki & Chow, 2022). OSM can also be used to estimate population distribution (Bakillah et al., 2014) and as an alternative to official cadastral surveys (Basiouka et al., 2015). In addition, it can also be leveraged to analyse the distribution of public or urban green spaces (Teeuwen et al., 2024; Weigand et al., 2023), urban infrastructure for emergency responses (Peixoto et al., 2023), and access to food (Quinn & Yapa, 2016). In the field of mobility, OSM has been used to develop a national cycling infrastructure dataset in Canada (Ferster et al., 2023), a tool for citizens' mobility assessment (Truden et al., 2022), and the street network of India (Tripathy et al., 2021) and the African continent (Prieto-Curiel et al., 2022).

OSM has also been used in socioeconomic studies, including to analyse ethnic divisions and the provision of public goods (e.g., schools, hospitals, and libraries) nationally (Seidel, 2023) and interethnic group relations (Dementeva et al., 2024). A non-profit organisation, Digital Democracy, built an offline mapping tool based on OSM to facilitate indigenous communities in the Amazon to map their neighbourhood (MacLennan, 2016). OSM also found its place in cultural and historical studies, such as in studying the history of cities based on street names (Carmona-Derqui et al., 2023) and in documenting endangered historical monuments (Heidelberg Academy of Humanities and Sciences, 2018).

In the field of public health, OSM has been utilised to develop a decision-making tool for community health programmes (Randriamihaja et al., 2024) and the strategy for the radiology outreach programme for underserved populations (Daniels et al., 2021). OSM has also been leveraged to evaluate the distribution of health services and access to health facilities (Hu et al., 2023; Khazi-Syed et al., 2023), including in a complex humanitarian emergency setting (Garber et al., 2020). Besides, OSM has played a role in responding to and studying epidemic diseases such as Ebola and COVID-19 (Minghini et al., 2022; C. H. Park et al., 2020) (OSMC-04).

OSM has also been utilised for climate or natural disaster risk modelling and assessment (Cerri et al., 2021; Mühlhofer et al., 2024; Scholz et al., 2024). It has also been leveraged to develop datasets of the global inventory of electricity infrastructures (Kalt et al., 2021) and power grid modelling (Medjroubi et al., 2017), which can help support the energy transition. In addition, OSM has also been used to simulate heating energy demand at the urban scale (Schiefelbein et al., 2019).

Better coverage than other datasets

In general, the notion of data completeness is flimsy in OSM, as in theory, an endless number of tags and objects can be created and refined in OSM (Neis & Zielstra, 2014). One can argue that this ever-evolving nature is the case for any map data, as maps are 'of-the-moment' (Kitchin & Dodge, 2007) and 'always-in-the-making' (Gerlach, 2015). However, the fluidity of the producer and user roles in OSM means that OSM data is not an activity exclusive to specific (authoritative) organisations, but of anyone. In other words, OSM data can be considered user-driven by nature; what a data user can conceive, they can create.

Consequently, OSM has proven its potential to capture data that would otherwise be neglected by governmental or commercial data providers, such as informal settlements and slums (Panek & Sobotova, 2015; Soman et al., 2020). Scholz et al. (2024) found that in conflict areas with weak institutions, some citizens map critical infrastructure in OSM (e.g., schools and hospitals) to call for accountability in case those infrastructures were 'accidentally' targeted. Furthermore, in some countries, OSM data may complement the coverage of government data (Brovelli & Zamboni, 2018; de Arruda et al., 2024) or be the only available open data for specific use cases (Yen et al., 2021). In addition, the global nature of OSM data means that international organisations, such as multinational companies, can rely on a single (primary) source of data (Ochoa-Ortiz & Re, 2025).

Self-correcting mechanism

OSM benefits from Linus's law, 'Given enough eyeballs, all bugs are shallow', a term coined by Raymond (1999), in tribute to Linus Torvalds, the inventor of the open-source operating system, Linux. Specifically, OSM data is collectively created and edited by many contributors and errors that a contributor makes will likely be detected and corrected by others (Almendros-Jiménez et al., 2021; Haklay, 2010). Co-editing (iterative editing of the same object by multiple contributors) was empirically found in OSM (Mooney & Corcoran, 2014; Sarkar & Anderson, 2022), which gives evidence to its self-correcting mechanism. Such a mechanism thus supports the credibility and trustworthiness of OSM data quality (Sarkar & Anderson, 2022).

The widespread use of OSM data by various large multinational companies indicates that the quality of OSM data is well-regarded. Comparing the Microsoft Building Footprint that is generated using machine learning and the OSM building footprint in Sudan, Scholz et al. (2024) found that while the coverage of the former was more extensive, the latter's accuracy was better. This means that, currently, even AI cannot supersede the accuracy of community-generated data in OSM. The self-correcting mechanism also leads to OSM data being deemed as up-to-date data, often more than other open data sources (Huang et al., 2023). For instance, Sarretta & Minghini (2021) found that while the address data from the National Land Survey of Finland was generally more complete, the OSM data was more up-to-date and detailed.

Advantages of the ODbL licence

On 12 September 2012, OSM changed its licence from Creative Commons Attribution Share-Alike 2.0 (CC BY-SA 2.0) to ODbL (OSM Wiki, 2022c). Several arguments were presented in support of this change (Amos et al., 2009; OSMF, 2016). First, it was argued that OSM data is likely not protected by the US (and some other jurisdictions') copyright laws. This is in accordance with the landmark case of Feist v. Rural (*Feist Publications, Inc. V. Rural Tel. Serv. Co., 499 U.S. 340 (1991)*, 1991), which established that a collection of facts (in the case of OSM, including names, reference codes, and house numbers) is not copyrightable[20]. Since CC BY-SA 2.0 relies on copyright in the data, the licence thus does not protect OSM data in the US and other jurisdictions with similar laws. While Creative Commons effectively recommended OSM data to be public domain (CC0 licence), a substantial portion of the OSM community was in favour of a reciprocal licence that has the same spirit as CC BY-SA 2.0 (Amos et al., 2009; OSMF, 2016).

20 This is different than in the UK and the EU that reward the efforts of collecting information, applying the principle called 'Sweat of the brow' doctrine. For the UK, see item 17 (R. Griggs Group Ltd & Ors v Evans & Ors [2003] EWHC 2914 (Ch), 2003). For the EU, see article 7 of the EU Directive on the legal protection of databases (Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases, 1996).

Second, some companies were advised by their legal team against using the OSM data due to the unclear boundary between collective and derived works according to the CC BY-SA 2.0 licence, which was illustrated by the hypothetical question posed by OSMF (2016), 'Does the derived work extend only to the basemap, the basemap plus any overlays, or to the work around it, such as a news show or a book?' In practice, the wide use of OSM data depends on the possibility to combine it with proprietary or sensitive data to generate (commercial) products such as thematic maps (e.g., in mobility apps) or news articles (Haklay & Weber, 2008). The uncertainty over the extent of the derived work and whether proprietary data can even be used with CC BY-SA 2.0 data was deemed to discourage the use of OSM data (OSMF, 2016).

To tackle these issues, the ODbL, which was developed by OKF in close cooperation with the OSM community (OKF, 2017), was considered the most viable option. It maintains the reciprocity spirit of CC BY-SA 2.0 while not only using copyright law, but also contract and database rights. Unlike CC BY-SA 2.0, ODbL distinguishes 'database' (e.g., OSM database) and 'produced work' (e.g., a paper map). This allows OSM to protect the core database under a reciprocal licence (i.e., share-alike) while freeing the end-use work (Amos et al., 2009). Evidently, beginning less than five years after the licence change, not only has the use of OSM data soared, but data contributions to OSM have also increased drastically.

Collection of tools

OSM is a federative project, which means that the software for contributing and using OSM data is developed mainly by third-party providers (OSM Wiki, 2025w). This modularity was intentional from the start in order to relieve the burden of a single party from having to develop all the technology needed (and possible) around OSM (Lenormand, 2022). Apart from the editing tools already mentioned in Section 6.4.1, such as Rapid (developed by Meta), HOT Tasking Manager, and ArcGIS Editor for OSM (by Esri), there is also MapRoulette, which is a gamified microtask OSM editing app developed by the OSM US (OSM Wiki, 2025m). The app provides the exciting feeling of completing challenges while contributing to OSM, even with limited knowledge about it (van Berkel & Pohl, 2024) (OSMC-09).

Meanwhile, MapSwipe, developed by the Missing Maps project, is a gamified app that facilitates data contribution to OSM, but the app users are not the OSM editors themselves. Instead, app users help humanitarian OSM editors from having to sift through thousands of satellite images to identify those relevant to their mapping objectives. The app has three tasks options: swipe through satellite images and select those that contain the requested features (e.g., roads, buildings, and waterways); review before and after satellite images to identify changes in the environment and help inform damage assessment and data accuracy verification; and validate the accuracy of the mapped building footprints (MapSwipe, 2025; OSM Wiki, 2025aj).

For accessing and using OSM data, besides the tools mentioned in Section 6.4.1, there is also OSMnx, which is a Python package that interacts with OSM APIs to download, model, analyse and visualise OSM data, developed by Geoff Boeing, an academic at the University of Southern California (Boeing, 2025). The team at Heidelberg Institute for Geoinformation Technology (HeiGIT) developed the Ohsome API that enables OSM data history analysis (HeiGIT, 2018).

Furthermore, various quality assurance tools help detect, assess, and validate abnormalities in OSM data (Almendros-Jiménez et al., 2021; OSM Wiki, 2025ag; Quinn & Bull, 2019). For example, OSMCha, initially supported by Mapbox and now by OSM US, helps editors to analyse and review data changes of others, so that potentially problematic edits can be identified more efficiently (OSMCha, n.d.). OSMCha has been used by organisations such as Apple, Meta, and HOT, and the larger OSM community (Marcel, 2024). Another quality assurance tool is Osmose, developed primarily by OSM France members, which detects a wide range of issue types, and users can indicate if the issues are false positives or have been corrected (OSM Wiki, 2025h; Osmose, 2025).

6.5.2 Current weaknesses

Persistent inequalities in coverage and participation

Several researchers have cautioned against the overoptimism of the democratisation potential that VGI could offer (Elwood, 2008b; Haklay, 2013; Solomon et al., 2024). As OSM grows, it has become clear that there are persistent inequalities within OSM. Lin (2015) conceptualised three forms of interaction in OSM: user-to-documents (admittedly a somewhat confusing term; it refers to the interactions between OSM contributors and OSM data), user-to-system (interactions between OSM contributors and OSM software), and user-to-user (interactions among OSM contributors). Inequalities can be observed in all three forms of interactions, which, for better clarity, are rephrased and discussed in terms of (i) inequalities in OSM data and data contributors, (ii) inequalities in OSM's software development, and (iii) inequalities in OSM's social participation.

First, inequalities in OSM data and data contributors: Even though OSM data coverage has improved globally, inequality patterns persist. The coverage of OSM data is much better in developed countries, especially in Europe, compared to developing countries in Africa, Asia, and South America (Yang et al., 2024; Zhou et al., 2022). Furthermore, within countries, OSM data is richer in more densely populated (i.e., urban) and higher-income areas (Klinkhardt et al., 2023; Moradi et al., 2021; Sarretta & Minghini, 2021).

Based on the OSM data trajectory analysis, Yang et al. (2024) found that maintaining business as usual in OSM data contribution means that data equality will never be achieved. This is in line with the findings by Thebault-Spieker et al. (2018) who found that the behaviour of OSM editors are 'born, not made' (i.e., OSM editors are relatively consistent in the places and types of data they edit in their lifespans) and active OSM editors, unfortunately, tend to overlook rural and socioeconomically disadvantaged areas from the moment they started to contribute to OSM. Such biases may be attributed to the fact that OSM editors tend to edit what interests them (Bégin et al., 2013), a phenomenon known as the self-focus bias (Das et al., 2019; Hecht & Gergle, 2009). This means that user-driven and supplier-driven notions in the OSM are, in fact, two sides of the same coin.

Due to the self-focus bias, it follows that in order to improve the equality of OSM data, the diversity of OSM editors is likely have to be enhanced. However, the backgrounds of OSM editors are highly biased. While identifying the country-based of OSM editors is not straightforward, as this information is not captured in the OSM membership registry, by assuming the first or the most edited country as the country-based of the editors, several analyses have shown that OSM editors are disproportionately based in Europe and North America, when adjusted for country population or land area (Anderson, 2021; Neis, 2025a; Poole, 2017; Shin & Basiri, 2022). On the use of OSM data in academic studies, Grinberger et al. (2022) found that 63% and 21% of authors who used OSM data are based in Europe and North America, respectively. Likewise, 45% and 16% of the study areas of the research that used OSM data are in Europe and North America, respectively (Grinberger et al., 2022). Hence, one may infer that the background of OSM editors influences what is included in the OSM data, which in turn influences its use.

Furthermore, the gender dimension of OSM editors is highly unbalanced. Based on a survey, close to 90% of the OSM editors identified as male (Gardner et al., 2020). Besides, on average, male editors have statistically more active days (158.94 to 66.86) and a higher number of changesets (156.26 to 84.33) in a year than their female counterparts (Gardner et al., 2020). In addition, male editors use more bots to contribute to OSM (with-bots datasets 9.51 times higher than no-bots datasets) compared to female editors (only 3.55 times higher) (Das et al., 2019). This may be related to an observation where female contributors tend to generally be less involved in the more technical or expertise-driven types of crowdsourcing projects (Steinmann et al., 2013).

Consequently, OSM is inscribed with predominantly male worldviews. A case in point: Stephens (2013) discussed that while tags for venues primarily attended by men were classified in detail in OSM Wiki, for example, distinguishing between ‘bar’, ‘pub’, and ‘biergarten’ and between ‘night club’, ‘swinger club’, ‘strip club’, and ‘brothel’, there were only two tags described for childcare facilities, namely ‘kindergarten’ and ‘baby hatch’. A proposal to include ‘childcare’ as a documented tag in OSM Wiki was rejected with the opponents arguing that it is the same as ‘kindergarten’, even though in the proposal, ‘childcare’ was explicitly described as ‘place for children to do homework, play and spend time otherwise *after school or kindergarten* [emphasis added]’ (OSM Wiki, 2023e; Stephens, 2013).

Meanwhile, a share of female editors involved in HOT (around 30%) is higher than in OSM as a whole (HOT, 2018a), indicating diverging motivations between male and female editors in OSM, which subsequently determines where and what is mapped in OSM. In the same vein, Solomon et al. (2024) found that female editors tend to map in more diverse countries than their male counterparts. Therefore, addressing female participation in OSM editing may also help tackle coverage inequalities in OSM data.

Second, inequalities in OSM’s software development: Although OSM is built on open-source software infrastructure, and in theory, any member can shape or at least have a say in the OSM software development, only a very few have the (social) power and (technical) skills to do so (Perkins, 2011). Rather than being egalitarian, Perkins (2011) argued that there are clear bureaucratised hierarchies in OSM. Plennert (2018) conceptualised three software layers in OSM (): the front end/presentation tier (where most OSM editors and users reside), the application tier (where expert geospatial professionals typically reside), and the back end/data tier (where software developers are; some are also active in other open-source projects). Each layer depends on the one more complex below it to function (Plennert, 2018).

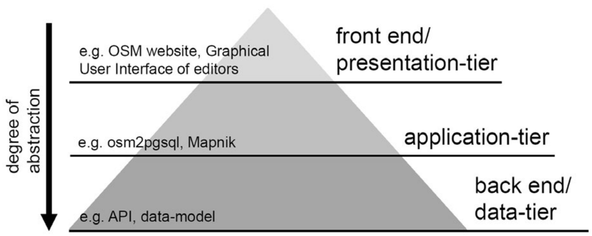


FIG. 6.10 OSM software tiers [Source: Plennert (2018)]

Based on 15 semi-structured interviews, Hall et al. (2017) showed that OSM members with higher technical ability (e.g., the ability to modify the OSM editing tool) have greater power to influence the nature of OSM data. Although OSM has established communication media such as the mailing list to discuss OSM's technology, Plennert et al. (2019) argued that non-expert members can only allegedly participate in OSM's technological orientation. In practice, OSM upholds *do-ocracy*, one of the core values expressed by OSMF and widely embraced by the OSM community (OSM Wiki, 2020a), where those who do the coding work determine how the technology is, regardless of what is discussed within the community (Plennert, 2018). Do-ocracy essentially creates a paradox where OSM is both bottom-up (anyone can seemingly contribute changes in OSM) and top-down (those who lack technical power can only accept what is determined by those who do). Not only does do-ocracy potentially exclude the ideas and values of non-expert OSM members from serious consideration, but it has also led to conflicts. For instance, some members expressed dissatisfaction with the perceived privilege held by bot writers whose works sometimes caused issues in the data, which then had to be corrected manually by OSM editors (van Berkel & Pohl, 2024).

Third, inequalities in OSM's social participation: The community is the lifeblood of OSM, but involvement in OSM's decision-making processes and social interactions is still dominated by the privileged few. To begin with, of the 18 OSMF-recognised local chapters, 13 are in Europe (OSM Wiki, 2025ac). The representatives of these local chapters serve on the OSMF advisory board. Around 64% and 20% of the OSMF members are from Europe and North America, respectively (Hormann, 2024; OSM Foundation, 2024f). Out of 16 OSM annual State of the Map conferences held until 2024, nine were organised in Europe, twice online, twice in Asia (Japan), and once each in North America (US), South America (Argentina), and Africa (Kenya) (OSM Wiki, 2025ai). Furthermore, interviewees involved in the OSMF board and working groups shared that OSMF meetings were almost always organised following the time zone in Europe or North America, limiting participation of members from other regions. By studying OSM mailing list archives, Chauhan et al. (2024) found that debates within OSM sometimes created an unwelcoming environment, leading many active OSM members to opt out of the conversation, often along gendered and geographical lines.

Schröder-Bergen et al. (2022) observed that the development of OSM data in many previously underrepresented regions in the Global South was not driven by local OSM communities, but by humanitarian organisations and commercial companies. While this scenario indicates the contribution of institutional actors in narrowing the data inequalities in OSM, Herfort et al. (2021) argued that empowering and developing sustainable local communities is crucial to support local perspectives and benefits from OSM. Meanwhile, there were also instances where the contributions of local OSM communities were unacknowledged or dismissed. In 2020, the OSM community in the Philippines released a public statement calling out the lack of recognition of the local OSM community's efforts in responding to Typhoon Yolanda (Haiyan) in 2013 in Amazon Prime's documentary series, Now Go Build, which only acknowledged the role of HOT (Vicario et al., 2020). The documentary was also perceived to relegate Filipinos to mere workers or beneficiaries (Vicario et al., 2020). Having said that, in some cases, external factors limit the active participation of individuals from certain countries in OSM. For instance, mapping OSM in China is effectively illegal (W. Lin, 2018; OSM Wiki, 2025ah), even though, ironically, OSM data in China has grown exponentially in the past years (Zhao et al., 2015).

Furthermore, the voice of the indirect impact recipients (sometimes called the non-users) of OSM is often rendered invisible (Chauhan et al., 2024). These are people who do not contribute or use OSM directly but are (positively or negatively) affected by it; for instance, disaster victims that were rescued by responders who used OSM or citizens of an occupied area according to the internation law of which the occupation is, to some extent, legitimised in OSM following the 'on the ground' principle (Bittner & Glasze, 2021; Chauhan et al., 2024). Even concerning the seemingly benign humanitarian or civic mapping, there were potential disconnects between HOT projects and the contributions needed by local or field NGOs (Y. Yin et al., 2024).

Inconsistent or limited semantic attributes

On the one hand, OSM's free tagging system encouraged data contribution without being bogged down by complicated taxonomies, but on the other, it resulted in inconsistent tagging practices (Biljecki et al., 2023; Mayer et al., 2020). Even though the OSM community has documented tags folksonomy in the OSM Wiki, different tags would have different interpretations in local contexts. For instance, the tags 'highway=primary' and 'highway=secondary' have different meanings in Europe and Australia (Wilmott, 2019). Some tags emerged from a specific context, such as 'highway=living_street', which initially referred to *Erf/Woonerf* in the Netherlands and Flanders, or *Spielstraße* in Germany. A general feature of these streets is that pedestrians are legally granted a higher or equal right of way over other road users, such as cars (different from exclusively pedestrian streets). Such legislation, however, varies across jurisdictions and, in many cases, is absent, resulting in inconsistent interpretations of what constitutes a living street (OSM Wiki, 2025ak; Wilmott, 2019).

Moreover, studies suggested that OSM editors pay more attention to the geometric aspects of the features in OSM than the semantic aspects (Biljecki et al., 2023; Davidovic et al., 2016). This is illustrated by the study of Sarretta & Minghini (2021) who discovered that the inconsistencies between the names of streets or cities in

OSM and government data are partly due to misspellings (or different spellings) in OSM. Additionally, Yamashita et al. (2023) found that in Japan, discrepancies in names can largely be attributed to the use of Kanji characters, where names can be written differently, albeit with the same pronunciation and meaning.

Meanwhile, Mandourah & Hochmair (2024) pointed out that the predominant use of imagery to contribute OSM data restricted the semantic attributes that can be captured by editors (Biljecki et al., 2023; Mandourah & Hochmair, 2024). Biljecki et al. (2023) highlighted that such deficiencies, for example, in terms of the building material and height, limited the potential use of OSM data in built environment studies. Additionally, Klinkhardt et al. (2023) found that more visible POIs, such as shops with notable signs or display windows, are better mapped in OSM compared to less visible POIs, such as doctors and small agencies with less notable signs.

Difficulties in integrating other databases into OSM and vice versa

Integrating external datasets into the OSM database is technically and legally challenging. In terms of technical aspects, importing external datasets requires careful consideration of how to convert the data to OSM XML, handle conflation, and map original data attributes to OSM tags, among others (OSM Wiki, 2025aa). Besides, editors who want to import datasets must document and discuss their plans in the OSM Wiki and obtain the community's buy-in (OSM Wiki, 2025aa).

Regarding the legal aspect, complexities arose primarily due to the requirements of the licences for the external datasets. For example, many open government datasets are published under the CC-BY 4.0 International licence, which is not exactly compatible with the ODbL on two fronts (OSM License Working Group, 2017). First, even though CC-BY 4.0 gives more flexibility in terms of the content attribution requirement compared to the previous versions, it remains a risk if the licensors would be satisfied with the indirect attribution practice of OSM (i.e., listing all the external data sources on OSM Wiki instead of on individual data on OSM map) (OSM Wiki, 2025al). Second, CC-BY 4.0 has a strict requirement against imposing any additional conditions or Effective Technological Measures (ETMs)[21] on the downstream use of the content (Section 2(a)(5)(B) of the licence). This requirement conflicts with

21 In the EU, Effective Technological Measures (ETMs) are defined as follows: 'any technology, device or component that, in the normal course of its operation, is designed to prevent or restrict acts, in respect of works or other subject-matter, which are not authorised by the rightholder of any copyright or any right related to copyright as provided for by law or the sui generis right provided for in Chapter III of Directive 96/9/EC' (Directive 2001/29/EC, 2019).

the term under the ODbL licence (Section 4.7(b) of the licence) that allows parallel distribution where, for example, a company can provide a proprietary map service with advanced features based on OSM data behind a paywall and offer the raw OSM data for free under the same ODbL terms. Hence, due to these two points of (potential) incompatibility between the CC-BY licence and the ODbL, OSM's standard practice is to request special permission or a waiver to include CC-BY licensed data into OSM from the licensor (OSM Foundation, 2023c; OSM License Working Group, 2017; OSM Wiki, 2025x). The role of local OSM chapters and communities in interacting with and negotiating with local data providers in such cases is highly valuable.

On the other end, some organisations deemed integrating the OSM database into their databases risky, as they need to trace the derivative products they built using OSM data to comply with the attribution and reciprocity requirements of the ODbL (OSMC-01, OSMC-03, OSMC-05). Having said that, many other organisations, especially companies such as TomTom with Orbis Maps (TomTom, 2025), have taken advantage of section 4.5(a) of the ODbL, which releases the requirement to apply the ODbL to the collective database (i.e., a collection of independent databases, including the ODbL-licensed database, that form a collective whole).

6.5.3 Potential threats

Tensions between companies and the OSM community at large

The heterogeneous interests in OSM are a double-edged sword: On the one hand, they increase the support towards OSM, including attracting new editors and financial sponsors, and on the other, they fuel tensions between seemingly competing interests (Chauhan et al., 2024). In recent years, tensions between for-profit companies and the OSM community at large have been growing (Schröder-Bergen et al., 2022). Beyond ideological clashes that have existed since the early days of OSM, where some members believe that the spirit of OSM is incompatible with commercial interests (Budhathoki & Haythornthwaite, 2013; Perkins, 2011), more recently, frustrations from the OSM community at large towards for-profit companies were due to the material actions of the latter.

For example, a controversy ensued when there was an unusual membership registration of more than 100 individuals from the same IP address within a few hours, just as the window for eligibility for voting in the 2018 OSM board election was closing. The IP address was traced to GlobalLogic, which is an outsourcing firm in India that is known for providing OSM editing services to Grab and Apple (OSM Wiki, 2020b). A subsequent investigation by the OSMF Membership Working Group strongly suggested that the registration en masse was indeed an attempt to influence the 2018 OSMF board election, with the explanation from the company found disingenuous (Friedl & Rischard, 2018). The same company also caused outrage among OSM members in Thailand when it overwrote dozens of edits created by local Thai mappers when the company was working for Grab. Grab later acknowledged the mistakes and said that the issues were rectified as soon as it was alerted to them (Russell, 2018).

Even though OSM editing is always selective in nature, determined by the interests and motivations of the editors, the scale of the corporate editing raised a new concern about its impact on OSM data (Chauhan et al., 2024). Thus, some OSM community members were unhappy with the intransparency of companies in disclosing their corporate editing activities (Chauhan et al., 2024), prompting OSMF to publish the OEG that requires public documentation in the OSM Wiki (OSM Wiki, 2025v). Some pointed to the increasing burden on DWG in resolving or moderating conflicts or issues because of large-scale corporate editing practices (Chauhan et al., 2024). Some also characterised the potential exploitative nature of paid editing, allegedly done by low-paid workers in Asia or South America on behalf of large multinational corporations (Chauhan et al., 2024; Schröder-Bergen et al., 2022). The scepticism towards commercial companies was also extended to the software tools they developed. For example, when Facebook released the AI-assisted editing tool, RapiD, in 2019 (now spelt as Rapid), the reactions from OSM community members were mixed, with some raising concerns around data quality and others supporting the use of AI for the long-term maintenance and growth of OSM data (Chauhan et al., 2024).

Some companies have created community manager roles to specifically handle communication and engagement affairs between the companies and the OSM community at large in order to build better relationships (OSMC-09) (TomTom Blog, 2023). Some of the individuals who were hired for these roles were already involved in OSM before joining those companies (Sarkar & Anderson, 2022; Schröder-Bergen et al., 2022), which is consistent with the 2017 OSM Foundation survey that found 55% of the respondents who were associated with organisations engaged in paid editing had been contributing to OSM at least 3 years prior to joining those organisations (OSM Foundation, 2017). Thus, Schröder-Bergen et al. (2022) highlighted the potential risk where the relation between OSM members hired by companies and the OSM community at large is merely one of ‘housekeeping’.

In 2021, the OSMF board formed the Special Committee on Takeover Protection, whose remits include, but are not limited to, giving specific attention to paid voting, triggered by the 2018 board election incident (OSM Foundation, 2021b). The committee identified other potential mechanisms of takeover, including through Working Group participation (i.e., slow infiltration) or significant financial contributions (OSM Foundation, 2021a). Sieber & Brandusescu (2025) linked the involvement of for-profit companies in crowdmapping projects (not limited to OSM) to the concept of *philanthrocapitalism*, which could lead to the prioritisation of market interests over the overall community needs.

OSM vulnerabilities amplified by the Overture Maps Foundation

The use of OSM data by Overture proves the value of OSM data. However, Overture may pose certain risks to OSM, not necessarily as direct outcomes of Overture per se, but the worsening of OSM vulnerabilities facilitated by the emergence of initiatives like Overture. First, Overture may diminish the visibility of OSM (OSMF-01). Even though OSM is one of the primary sources of Overture, the consortium also leverages data from its members and other sources to create new open datasets (Overture Maps Foundation, 2025a). Additionally, Overture offers a range of open datasets (OSM Wiki, 2024e)that are more readily usable for specific applications (van Rees, 2024). For example, certain characteristics of the OSM data (e.g., the transportation network is split whenever road attributes, such as road type, speed limit, or lane count, change) complicate its use for routing applications. Overture, thus, preprocesses OSM data for such use cases and creates Overture's Transportation network data model (Shakeri, 2023).

In addition, the preprocessing of OSM data by Overture also involves improving its technical interoperability (e.g., standardising inconsistent tags) and validating the data (Overture Maps Foundation, 2024a), with one interviewee characterising Overture as 'a gold copy of OSM' (COM-01). Even though Overture, to some extent, is powered by OSM data, the better usability of Overture data may render OSM more invisible over time. Not to mention, Overture datasets (apart from those based on OSM data) are provided under the Community Database License Agreement (CDLA)-Permissive-2.0, which is more flexible than ODbL, especially in terms of reciprocity requirements (Overture Maps Foundation, 2024a). Overture is also a project supported by the Linux Foundation (Linux Foundation, 2022), a de facto umbrella of the open source developers community. OSM's invisibility may lead to a decrease of new (unpaid or corporate) OSM editors. This concern is not unfounded when looking at a past event where newly recruited OSM editors were 27% short after the introduction of Google Maps, in comparison to the projected trend of OSM editors had Google Maps not emerged (Nagaraj & Piezunka, 2024). This is compounded by the fact that in many countries, especially those without legally registered or OSMF-recognised local chapters, the use and impacts of OSM data have always been invisible (OSMC-03).

Second, arguably more concerning than the visibility aspect is that OSM risks of technological stagnation, or worse, decay. While there are many software tools developed around OSM, its core technology, namely the database and the API, has not changed much since 2007 (Plennert, 2018; Wroclawski, 2018). In fact, Steve Coast expressed his disappointment with OSM's technological stagnation in his interview in 2015 (Coast, 2015). Various reasons could explain OSM's technological stagnation, such as the practice of do-ocracy that excludes (fresh) ideas from a significant part of the community, suspicion towards the external for-profit actors and monetisation, the fear of disrupting the stability of the community, and the implicit ideology of safeguarding hobby-like values of OSM (Plennert, 2018; Wroclawski, 2018).

On the other hand, Overture seems to be a workaround established by for-profit companies to circumvent the bureaucracy of OSM while still leveraging the community asset of OSM (Holovin, 2022a, 2022b; Ochoa-Ortiz & Re, 2025). As a result, there could be less involvement of technological leaders in the development of OSM's technology in the future (Vichot, 2025). Instead, more resources from the same companies that are now OSM contributors may be channelled to Overture, which is already ahead of OSM in certain aspects. For instance, Overture adopts a cloud-native strategy of building, deploying, and managing applications in cloud computing environments (Bullock, 2024). This allows anyone to use Overture data with cloud-based tools, which is increasingly becoming the norm (Dong et al., 2024). Another example, OSM lacks permanent IDs of objects, resulting in interoperability issues for various applications such as navigation (Wroclawski, 2018). Overture overcomes this problem by introducing the Global Entity Reference System (GERS) that assigns a permanent unique ID to objects that exist in the real world, such as office buildings, grocery stores, and roads (Breunig, 2024; Overture Maps Foundation, 2025b). Meta recently announced that it has transitioned its suite of global basemaps across its apps, such as Facebook and Instagram, to Overture's base data layers (Overture Maps Foundation, 2025c).

Some members from Overture and the OSM community view the two as complementary and not in competition, especially since OSM has a large community that Overture does not have (Lenormand, 2023; OSM Wiki, 2024e; Overture Maps Foundation, 2024a; van Rees, 2024). However, one can argue that the sustainability of the OSM community relies on its technological, social, and market relevance.

Perhaps unimaginable now, but if organisations had moved from using Google Maps to OSM at one point, it is not impossible for them to move to other data sources that offer better value propositions than OSM in the future. Therefore, Overture serves as a call for action to OSM to stay relevant.

Sustainability of contributors and contributions

Over three-quarters of OSM members are the silent majority who have not contributed data to OSM at all (Zhang et al., 2024). Based on one estimate, 98.5% of the OSM changes were made by only 2.3% of all registered members (Zhang et al., 2024). According to another estimate, the top 1% of the OSM editors contributed 85.4% of the edits, not including bots and data imports (Sarkar & Anderson, 2022). This distribution pattern is not unique to OSM but is common in online communities, consistent with Nielsen's 90-9-1 rule, where '90% of users are lurkers who never contribute, 9% of users contribute a little, and 1% of users account for almost all the action' (Nielsen, 2006).

However, several aspects raise concerns. Active OSM editors tend to be 'born, not made', whereby if an editor does not become active early in their lifespan, the probability of them transitioning to be an active editor is low (Zhang et al., 2024). In other words, editors are unlikely to become active gradually over time. In the past decade, a significant portion of 'born, not made' editors are in fact corporate editors (Zhang et al., 2024). Thus, it calls into question what would happen if companies no longer saw the value of contributing to OSM. This relates to the recognition that not all contributions in OSM are voluntary (Sarkar & Anderson, 2022), which is increasingly apparent with the growth of paid editors, challenging the association of OSM with the VGI concept.

Moreover, the retention rates of OSM editors are low. Even for highly active editors (>10,000 edits per month), around 75% stayed active for only less than 2 years (Zhang et al., 2024). While humanitarian organised editing events were found to recruit many new editors quickly, their retention rates are lower than the overall OSM editor base (Mahmud et al., 2022). This scenario is also evident from the bottleneck that exists in the validation phase of HOT Tasking Manager, which is typically done by more experienced editors (Herrera-Murillo et al., 2024). A similar pattern of high recruitment yet low retention rates is also seen during other kinds of mapping events (Bégin et al., 2018; Hristova et al., 2013; Khanal et al., 2019). There is also a concern that corporate editing is squeezing out existing (non-paid) local editors (Anderson et al., 2019), further accelerating the editor attrition rate.

Conflicts among OSM members

Notwithstanding the tension between companies and the OSM community at large, conflicts also exist among other OSM members. According to a survey conducted by (Choe et al., 2023b), organised editing is the most important topic of conflict in OSM, followed by tagging, diversity and inclusion, and code of conduct. Meanwhile, factors leading to conflict include interest, priority, value and secondary goals, followed by 'Us vs. Them' (i.e., subgroup animosities), different cultural background, and uncooperative behaviour (Choe et al., 2023b). The most significant effect of conflicts is discouraging contributions to OSM, followed by a toxic atmosphere (Choe et al., 2023b). One notable instance of OSM community members ceasing their contributions occurred after the fallout from the license change (DiverCTH, 2024; TimSC, 2011). Thus, conflicts within OSM are not limited to a simple disagreement on an edit, but are interlinked with the social dimension of the editors, such as their backgrounds, ideologies, and interests (Ballatore & Mooney, 2015; Choe et al., 2023b; Grinberger et al., 2021).

In 2020, OSMF published a Diversity Statement (OSM Foundation, 2020a), and in 2021, it revised its etiquette guidelines that apply to OSM communication channels (Table 6.7) (OSM Wiki, 2023b). The latter was in response to an open letter endorsed by around 30 organisations and over 300 OSM individual members titled 'A Call to Take Action and Confront Systemic Offensive Behaviour in the OSM Community' (OSM members, 2020).

TABLE 6.7 Snippet of OSM Etiquette Guidelines [Source: OSM Wiki (2023b)]

OpenStreetMap community members should do the following: • **Act in good faith.** It is surprisingly easy to misunderstand each other, whether online or in person, particularly in such a culturally and linguistically diverse setting as OpenStreetMap. Misunderstandings can easily arise when we are debating topics or when we are in a rush or distracted. Please ask the other person to explain before assuming that a communication was inappropriate or not made in good faith • **Be respectful.** Communicate with the same level of respect as you would use in person. Enthusiastic discussions are vital in a successful project and there are bound to be disagreements. However, we should keep discussions and disagreements appropriate and calm. • **Be welcoming.** OpenStreetMap aims to be a community that welcomes and supports people of all backgrounds, cultures and identities. Some examples of behavior that can help create a positive environment include using welcoming and inclusive language, respecting different viewpoints and experiences, showing empathy towards other community members and communicating with a global audience in mind. Remember that people may be new to OpenStreetMap. We should be tolerant and supportive towards new members. • **When we disagree, try to understand why.** Disagreements, both social and technical, happen easily and often. It is important to try to understand each other and work to settle any disagreements and differing views constructively. If someone contradicts your own views, try to understand where the other person is coming from. Try to ask questions that will allow the other person to explain the situation rather than make the disagreement worse, or consider private messaging for one-to-one conversations.

To improve the communication situation among OSM members, OSMF introduced a new discussion forum (OSM Foundation, 2022b), encouraging OSM sub-communities, including local chapters and communities, to adopt the platform in order to streamline discussions and allow moderation to take place (OSM Community Forum, 2025). Nevertheless, different OSM communities are more familiar with different communication media. Anecdotally, OSM communities in Asia and Africa were said to commonly use Facebook, Telegram, and WhatsApp, whereas their European and North American counterparts typically use mailing lists and Slack (OSMF-02, OSMC-09). Hence, streamlining communication media and ensuring a welcoming and constructive environment is not straightforward for OSMF.

Besides, disagreements were also triggered by policies initially introduced to address other related conflicts. For example, OEG, which was introduced to address conflicts around corporate editing, was deemed to disproportionately impact local OSM editors because of the potential extra work required to organise local mapping events (Chauhan et al., 2024). Learning from Wikipedia, whose quality control policy was found to be a key cause of the decrease in newcomer retention (Halfaker et al., 2013), implementing policies in OSM is thus a delicate act of balancing the interests and expectations of heterogeneous community groups.

Representations and ethical responsibilities

Y. Lin (2011) broadly categorised four *social worlds* involved in OSM: business, government, NGO/third sector, and individuals and less organised communities. OSM artefacts (data, software, policies, etc.) are the outcomes of the social processes by these different actors (Bittner, 2017; Carraro, 2021; Y. Lin, 2011; Mayer et al., 2020; Quinn & Tucker, 2017; Sarkar & Anderson, 2022). Scholz et al. (2024) gave an example where a certain local government requested water points to be mapped in OSM so that they could collect taxes for those water points, and on the other hand, the citizens wanted waste dumps to be mapped to hold the government accountable for waste disposal. Even the categorisation by Y. Lin (2011) grossly generalises the diversity in each social world. For instance, the interests, worldviews, mindsets, and practices of governments in countries in Europe versus in Africa and at the local versus national level can differ significantly. When certain disagreements have become invisible online (e.g., when edit wars seemed to stop), it does not necessarily mean that they have been resolved offline. Instead, some participants have decided to stop engaging with them due to the perceived futility of such negotiations or discussions in OSM (Carraro, 2021; Carraro & Wissink, 2018). OSM, and other forms of data for that matter, not only selectively represent the reality but also affect the reality that it (mis-/un-)represents. Therefore, not only is diverse participation necessary to ensure more diverse representations, but the governance and social environment in OSM should also facilitate challenging yet pertinent interactions among OSM members and their indirect impact recipients, i.e., the non-users (Chauhan et al., 2024).

Relatedly is the question of who bears the ethical[22] responsibilities of OSM data? The majority of interviewees in Scholz et al. (2024) regarded that the primary responsibility of taking ethical considerations in conflict-affected areas is on the OSM editors because once the data is in the OSM database, users would simply use it. At the same time, some of the interviewees are under the impression that while ethical considerations are a valid concern, the armed forces in the conflict-affected areas already have good quality data without having to use OSM data (Scholz et al., 2024). Meanwhile, Gerlach (2010) questioned the ethical responsibilities of remote mappers who are physically and contextually detached from the areas they map. In a similar vein, So & Duarte (2020) questioned who benefited from OSM data in North Korea, where 30 cities in the country had been mapped by the end of 2018, when it is unlikely that the citizens can access the data.

With the introduction of the EU General Data Protection Regulation (GDPR) that aims to protect personal data, OSMF has taken several measures to comply with it, including removing certain metadata from general public distribution and revising the privacy policy and terms of use (OSM License Working Group, 2018; OSM Wiki, 2025k). In addition to personal data, there is informal understanding on certain data that should not be captured in OSM, for example, when it involves safety concerns (e.g., safe houses for victims of domestic violence and places of worship for a religion persecuted in a given region), protection of endangered species (e.g., rare plants and an eagle's nest), and indigenous sacred sites of which the traditional owners or representatives have asked for the sites to be kept private (OSM Wiki, 2025o). Nevertheless, there are no written rules on the aforementioned aspects; thus, they remain at the discretion of the OSM contributors. Furthermore, for a similar type of data, contexts play an important role in whether it should be included in OSM. For instance, while mapping LGBTQ+ spaces such as bars and healthcare facilities in OSM is well-intentioned (OSM Wiki, 2024k, 2025ae), it may endanger the LGBTQ+ community in regions where members of the group are still discriminated against or persecuted (Scholz et al., 2024).

---

22 In this section, the question of ethics is discussed in a broader term beyond axiomatic (good vs. bad) and normative (do's vs. don'ts) standpoints to include the instrumentalist standpoint (value-sensitive approach that accounts for the sociotechnical entanglements between people and technology throughout the development processes of the technology), following (Calzati & Ploeger, 2024)in this article we make the case for an ecosystemic understanding of data ethics (for the city.

TOC

6.5.4 Potential opportunities

Steady stream of new editors

Even though OSM retention rates are low, as described in the previous section, OSM is still attracting a steady stream of new editors (Zhang et al., 2024). Furthermore, the socialisation process of OSM editors (measured by the transitions of new editors to become active editors and then highly active editors) is becoming faster (Zhang et al., 2024). Nevertheless, it is unclear if this accelerated socialisation is mainly attributed to the increasing involvement of corporate editors who are ‘born, not made’ to become highly active editors (Zhang et al., 2024).

The education sector, for example, through YouthMappers, also plays a role in promoting OSM to the new generation (Solís et al., 2021). Collaborations between the OSM community and the education sector can also take shape in other forms. For example, as part of the Miami-Dade County’s building footprints data import process, students enrolled in two courses of the Geomatics programme at the University of Florida were introduced to the OSM import task as part of their assignment (Juhász & Hochmair, 2018). Even though none of the students seemed to be actively contributing to OSM after the academic deadline (Juhász & Hochmair, 2018). This kind of initiative is worth exploring and experimenting with more. In Nepal, a not-for-profit company, Kathmandu Living Labs (KLL), recruited recent high school graduates and undergraduate students for its 2-month Digital Internship and Leadership (DIAL) internship programme that combined leadership skills training with OSM mapping activities for remote and vulnerable areas (Khanal et al., 2019). The programme attracted students from diverse backgrounds such as business administration, crisis management, architecture, public health, computer science, and geomatics engineering (Khanal et al., 2019).

Organised editing improves data coverage and quality

By comparing regions with active corporate editing with a baseline region with limited corporate editing, Patel et al. (2023) found that corporate editing improves the intrinsic completeness of OSM data (measured by examining changes over time). This finding is consistent with the scenario analysis of Yang et al. (2024) that showed contributions from organisational editors led to better equality of data coverage compared to contributions only from non-organisational editors. Notably, most organisations, especially humanitarian organisations such as HOT and Missing Maps, concentrated their effort in regions that initially lacked data in OSM (Yang et al., 2024).

Furthermore, the involvement of organisational editors likely resulted in spillover participation of non-organisational editors through the publicity on OSM that those organisations created. For example, a blog post revealing that Tesla potentially used OSM for detailed maps of parking aisles within parking lots led to an immediate spike in editors, presumably Tesla owners or fans, adding parking aisles in OSM (Anderson & Sarkar, 2020). Another example is that after Niantic, the developer of the mobile augmented reality game Pokémon Go, switched its base map from Google Maps to OSM, many Pokémon Go players started to edit OSM data in an attempt to improve their gameplay (OSM Wiki, 2024n). In general, Sarkar & Anderson (2022) found that the growth rates of new editors are higher in places with more corporate editors. Furthermore, Patel et al. (2023) argued that while corporate editors have been mostly editing road networks, such data can have a map seeding effect, where more features around the road network will be added by other editors. Likewise, the media coverage of HOT projects may also attract participation of new editors who are driven by the desire to help others (Y. Yin et al., 2024).

In contrast to the scepticism towards the quality of data contributed by organisational editors (Chauhan et al., 2024), the repeated co-editing between organisational editors in a relatively short duration reflects the internal editing workflows (where more experienced editors validate data edited by other editors), which likely increases the data quality (Sarkar & Anderson, 2022). HOT Latin America and the Caribbean (LAC) hub runs the Humanitarian Mapping Brigade programme, where it pays a select group of technically skilled mappers to help complete and validate time-bound projects (HOT, 2023b, 2024b). Similarly, UN Maps also hired skilled consultants to improve and validate data on OSM (INS-01).

Moreover, several organisations that are involved in OSM, such as UN Maps and Grab, also offer educational materials or training on OSM editing, further enhancing the enrolment of new OSM editors (Grab, 2019; UN Maps, n.d.-b). The interviewee from Gojek shared the company's vision of being the OSM ambassador in Indonesia to facilitate the onboarding of new editors.

Tensions create progress

Even though conflicts have adverse effects, such as discouraging contributions, they are also an opportunity for progress. For example, tensions between companies and OSM community at large have triggered productive debates in the community over the risks of organised editing towards data quality, the use of novel technologies such as AI for automated editing, and the transparency of organisational contributions, which have culminated in, among others, the publication of OEG and Automated Edits code of conduct (Chauhan et al., 2024). Another example was when OSM released the API version 0.5 (v0.5) in 2006 and introduced relations (in addition to nodes and ways) in the OSM data model (OSM Wiki, 2025am). Nevertheless, there was a lack of formal standardisation of relations in v0.5, especially in terms of relation members ordering that is crucial, for example, for routing directions and multipolygon, resulting in debates within the OSM community whether standardisation is necessary (Arnold & Hukal, 2024). Consequently, in 2009, API v0.6 was released (which to date is still the current version of OSM API) with better standardisation of relations (OSM Wiki, 2025am). The updates were found to increase the innovative potential of OSM to generate new content, even though initially it led to a decrease in new objects (Arnold & Hukal, 2024).

Similarly, some interviewees (e.g., OSMF-01, OSMF-02, OSMC-14) have considered that the emergence of Overture is an opportunity for OSMF and OSM community to recalibrate its current landscape (e.g., who uses OSM data, for what purposes, how do they contribute back to OSM, and what are the alternatives/competitors of OSM) and re-strategise its priorities (e.g., increasing the OSM visibility and creating a welcoming environment for new contributors). To some extent, the establishment of Overture reflects the changing landscape of geospatial data that warrants adaptation from OSM to remain relevant.

6.6 Lessons from OSM: Developing an open data intermediation business model that supports the sustainability of the ODE

Based on the analysis of OSM's business model, its current strengths and weaknesses, and potential opportunities and threats, several aspects to consider in developing an open data intermediation business model that supports a sustainable ODE are recommended (Table 6.8). These aspects especially apply to open data intermediation business models that share the same archetype as OSM, i.e., the collaborative open data platform (archetype A1 in Chapter 4).

Table 6.8 outlines 21 aspects to consider based on the insights from the OSM case. In the table, the asterisk (*) next to several 'aspects to consider' indicates that the aspect can also be linked to a potential contribution(s) of open data intermediaries identified in Chapter 3, and the ID (e.g., P05) refers to the assigned ID of the contribution (see Table 3.4). Additionally, some aspects may fall under multiple categories (first column); tabulating them according to those categories is meant to support the ideation process.

TABLE 6.8 Aspects to consider in developing an open data intermediation business model that supports a sustainable ODE (insights from the OSM case)

Categories	No.	Aspects to consider	Example insights from the OSM case
Resource-based view (RBV)	**1.**	Foster a healthy and constructive community of contributors.	The largest asset of OSM is its community of volunteer contributors. Thus, ensuring that the community members work collaboratively and constructively and that their issues are addressed are important to OSMF.
	2.	Protect the overall interests of the contributors by implementing an open communication policy and putting in place a mechanism that prevents potential takeover or hijacking.	While the growing involvement of for-profit companies (especially large corporations) has contributed to the increasing data coverage and quality, it has also led to conflicts and undesirable behaviours (e.g., OSM board election interference attempt).
	3.	Invest in enhancing the visibility and reach of the organisation or community.	One of the factors that led to the growth of OSM in the first few years since its launch was the promotional work done by its pioneers. In order to stay relevant and recruit more new contributors, the value of OSM needs to be recognised more prominently, especially now with the emergence of Overture, which can potentially overshadow OSM.
	4.	Ensure a transparent, efficient, civility-focused communication mechanism.	OSM's large and diverse community has led to some members communicating offensively. Besides, some policies introduced by OSMF were not received well by some OSM members, which may be attributed to shortcomings in communication.
Organisational identity	**5.**	Ensure consistency in how members view the (potentially multifaceted) organisational identity, especially as the community expands.	Since OSM is a heterogeneous global community, there are diverging views on what OSM should be, which sometimes lead to conflicts. OSM has, in fact, benefited from diverse participation driven by various interests, and this should be perceived as an asset instead of a source of division.
Value driver: Novelty	**6.**	Consider non-traditional solutions to addressing traditional problems.	OSM offers a radical solution to addressing the limited availability of open data. Instead of (continuously) asking governments or companies to release their data as open data, OSM harnesses the collaborative power of the lay people to collect data and make it open.
	7.	Consider the application of AI to provide recommendations for certain metadata (e.g., semantic attributes).	OSM's free tagging system resulted in inconsistent tagging practices. Besides, OSM data contributors tend to pay more attention to the geometric aspects than semantic attributes.
Value driver: Complemen-tarity	**8.**	Facilitate and encourage the development of (especially open-source) tools around the platform (i.e., federated architecture).* (P01 & P06)	OSM is a federative project where the software for contributing and using the data is developed mainly by third-party providers. This modularity was intentional to relieve the burden of a single party to develop all the technology needed and to give freedom for complementary software innovation.

>>>

TOC

TABLE 6.8 Aspects to consider in developing an open data intermediation business model that supports a sustainable ODE (insights from the OSM case)

Categories	No.	Aspects to consider	Example insights from the OSM case
Value driver: Efficiency	**9.**	Leverage a self-correcting mechanism by cultivating a ‘team-minded’ culture and having a clear process for dispute resolution.* (P19)	OSM benefits from co-editing (iterative editing of the same object by multiple contributors) that improves the accuracy and credibility of the data quality. Additionally, a dedicated working group within OSMF mediates unresolved data editing disputes.
Value driver: Adaptability	**10.**	Support local chapters or communities in engaging with local organisations (e.g., local governments or NGOs).* (P12)	OSM local chapters or communities play an important role in the development of OSM, e.g., by requesting special permission or a waiver to include open government data in OSM and helping local NGOs to use or contribute OSM data.
	11.	Invest in the adoption of new and emerging technology for the software infrastructure.	The core technology of OSM (i.e., its database and API) have not changed much for almost two decades. In contrast, Overture is already ahead of OSM in certain technological aspects (e.g., cloud-native implementation). This exposes OSM to the risk of becoming technologically irrelevant.
	12.	Transform disagreements and conflicts into opportunities for enhancing governance mechanisms and technical development.	Conflicts in OSM have prompted some improvements in its governance mechanism and technical development. For example, tensions between companies and the OSM community at large have led OSMF to release the Organised Editing Guidelines and Automated Edits code of conduct.
Sustainable ODE feature: User-driven	**13.**	Take initiatives to include the perspectives of indirect impact recipients (i.e., the non-users), e.g., by working with on-the-ground NGOs and prioritising their needs.	The perspectives of the indirect impact recipients (sometimes called the non-users) of OSM are often rendered invisible, especially with the prevalence of remote mappers. To ensure that the use of OSM data is translated into tangible and well-received outcomes, non-users’ needs should not be overlooked.
Sustainable ODE feature: Circular	**14.**	Choose a reciprocal licence that does not impose (many) restrictions on the end-use (i.e., derivative) work.	In less than five years since the adoption of the ODbL by OSM, not only has the use of OSM data soared, but data contributions to OSM have also increased drastically.
	15.	Organise events or conferences to facilitate collaboration, recruit new contributors, and showcase the value of open data.* (P11)	OSMF and OSM local chapters and communities often organise mapping events and conferences that help introduce OSM to new contributors and stimulate networking.

>>>

TABLE 6.8 Aspects to consider in developing an open data intermediation business model that supports a sustainable ODE (insights from the OSM case)

Categories	No.	Aspects to consider	Example insights from the OSM case
Sustainable ODE feature: Inclusive	**16.**	Minimise barriers for data contribution.* (P15)	The relatively low barrier for contributing data to OSM, even by non-geospatial professionals, through the simple data model, facilitates diverse actors (from companies to individuals) to contribute open data (instead of typically only governments).
	17.	Allow the contribution of diverse types of data to accommodate broad and diverse use cases.	The diversity of OSM data use cases has attracted diverse groups of users (e.g., businesses, researchers, governments, humanitarian organisations) in various domains from various parts of the world. They thus have a stake in ensuring the continuous development and maintenance of OSM.
	18.	Proactively recruit new data contributors among typically marginalised or disadvantaged groups, and nurture a welcoming environment to retain their participation.	OSM data remains unequal, which can be attributed to the lack of participation from certain groups (e.g., citizens in less developed countries, rural inhabitants, and female participants).
	19.	Meaningfully consider feedback related to the technological development from non-technical expert members.	While the principle of do-ocracy in OSM seemingly opens the processes of OSM technological development, it sidelines those who lack technical skills, thus potentially overlooking new innovative ideas.
	20.	Cultivate a culture of deliberating ethical responsibilities in the contribution and use of data, beyond providing general ethics guidelines.	In OSM, there are no formal rules on data that should not be captured apart from personal data, leaving the collection of data, such as that representing communities that are persecuted/ discriminated against in certain regions, at the discretion of data contributors. While having a formal guideline may not be possible for all complex ethical dilemmas, cultivating a culture where they are deliberated is necessary.
Sustainable ODE feature: Skills-based	**21.**	Leverage the education sector to recruit new data contributors and users and to nurture data literacy and skills.* (P20)	YouthMappers, which has more than 200 university chapters in over 40 countries, plays a role in promoting OSM and recruiting new contributors. Additionally, initiatives by some local OSM communities, such as Kathmandu Living Labs in Nepal, complement OSM promotion with broader data literacy and skills training.

Note: The asterisk (*) next to several ‘aspects to consider’ indicates that the aspect can also be linked to a potential contribution(s) of open data intermediaries identified in Chapter 3.

6.7 Further discussion

Beyond identifying aspects to consider in developing an open data intermediation business model, this study also reveals several insights that deserve attention. First, the case of OSM shows that the supplier-driven and user-driven sustainable ODE features, as suggested by van Loenen et al. (2021), may be two sides of the same coin. Just like governments collecting data initially for their internal objectives and eventually making it open data to be re-used by others (see e.g., Sexton et al. (2017) and Shaharudin, Reyes, et al. (2023)), the rapid growth of companies editing OSM data following their commercial objectives is a stark display of the same pattern (Sarkar & Anderson, 2022). While some (Ho & Rajabifard, 2010; Huang et al., 2023) considered the existence of *produsers* as a relatively new or unique phenomenon, it has, in fact, always been the case even for governments. Thus, moving from supplier-driven towards user-driven ODE, as envisioned by van Loenen et al. (2021), likely requires the alignment of the data suppliers and users' interests. Specifically in the case of OSM, such a condition can be achieved by having data contributors from diverse backgrounds who would be aware of the needs of different users. In the words of Gardner et al. (2020, p. 1606), 'who contributes the data, matters'.

Second, Verhulst et al., 2020 (p. 9) argued that a new 'wave' of open data, dubbed 'the third wave', should represent 'a much more purpose-directed approach to data provision than prior waves' for more impactful data re-use. However, similar to the previous point, purposes have always been ingrained in what open data becomes available. In OSM, the purpose-directed approach has translated into the limitations of its data usability, for example, in terms of the data coverage and semantic attributes. This is because all the current and future purposes of open data cannot be imagined now or by a group of certain actors. The key philosophy behind open data is about (re)generating its value through its re-use for different purposes. Quoting Tim Berners-Lee, the inventor of the World Wide Web, and his co-founder of the Open Data Institute, Nigel Shadbolt:

> 'It's re-use of data in new - and often unexpected - ways that creates both social value and opportunities for economic growth. *It's not our job to say where data might be useful* [emphasis added]; it's our job to unleash it and allow businesses and independent developers to build innovative services which they can then deliver to users. That's the story of technology through the years - and the way the World Wide Web itself has grown over the last twenty years.' (Berners-Lee & Shadbolt, 2010).

This does not mean that a purpose-driven approach is not valuable, but a sustainable ODE should focus more on inclusive participation of data contributors and users than on being bogged down predefining the purposes of open data.

Third, the OSM case shows that an ODE where government data is not the only (or main) open data in circulation, described by the inclusive feature by van Loenen et al. (2021), already exists. Interestingly, the contributions of data to OSM by various actors were not compelled through any form of legislation. Instead, they voluntarily contributed data because they appreciated the tangible value of such data either to them (e.g., their own commercial interests) or to others (e.g., to disaster victims) and/or due to the reciprocal requirement of the ODbL licence. As an open data intermediary, OSM effectively addressed the need for technical and governance infrastructure for open data contributions from various actors. Nevertheless, and ironically, some government organisations seem to neglect the value of collectively contributing to OSM or have a blanket scepticism towards its quality (INS-02, OSMC-05, OSMC-06, OSMC-07) (Quinn & Bull, 2019). Such a predicament may hinder the integration of open government data with open non-government data.

Fourth, insights from the OSM case question whether the skills-based feature, as described by van Loenen et al. (2021), is necessarily a required or realistic feature for a sustainable ODE. The do-ocracy principle in OSM privileges those with high-level skills in determining the technological direction of OSM. It potentially overlooks ideas and needs from non-expert community members. At the same time, certain high-level skills are indeed required to carry out certain tasks, such as maintaining the OSM database. However, it is unrealistic to expect everyone to have such a high-level technical proficiency, even by providing various skills training.

Several limitations in this study deserve further research. First, this study is based on a single-case study of a well-established open data platform. As argued in the previous chapter, there are advantages to this methodological approach. Having said that, further research is necessary to investigate the transferability of insights from this study to other cases. Second, certain insights from this study may not apply to all open data intermediation business model archetypes. For example, this study suggested the adoption of a reciprocal licence that does not impose many restrictions on the end-use (i.e., derivative) work. However, this recommendation may not apply to open data intermediaries that adopt business models other than the collaborative open data platform archetype, since they may not necessarily offer open data as their products or services. Third, being one of the most well-established collaborative open data platforms, OSM might enjoy certain privileges that are unavailable to nascent open data intermediaries. Nevertheless, OSM's story tells us that allowing learning to happen is key to the longevity of a project or an organisation.

6.8 Conclusion

This chapter has addressed the RQ4: What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem? The question was tackled through OSM's case study.
This chapter recommends 21 aspects to consider in developing an open data intermediation business model that supports the sustainability of the ODE (Table 6.8), which are especially applicable to the collaborative open data platform archetype. These aspects revolve around fostering a healthy and constructive community of contributors, ensuring a transparent and civility-focused communication mechanism, facilitating and encouraging the development of (open-source) tools around the platform, and choosing a reciprocal licence that does not impose (many) restrictions on the end-use work (e.g., ODbL). Furthermore, shortcomings of OSM taught us the importance of proactively recruiting contributors from typically marginalised or disadvantaged groups, meaningfully considering feedback from non-technical expert members, and investing in the adoption of new and emerging technologies.

PART D

(Re)framing the big picture

From the globe to the forest and trees, and back

Kia whakatōmuri te haere whakamua
I walk backwards into the future with my eyes fixed on my past
Māori's proverb

This last part, that constitutes Chapter 7, concludes this dissertation by summarising key findings for each of the research questions. This dissertation has proposed a common definition of open data intermediaries, identified their potential contributions in the open data ecosystem and business model archetypes, and suggested aspects to consider in developing open data intermediation business models that support a sustainable open data ecosystem. Additionally, this chapter offers several theoretical reflections on the concept of (sustainable) open data ecosystem and highlights the contributions of the dissertation and future research direction.

TOC

7 Discussion and conclusions

7.1 Introduction

Within the open data ecosystem (ODE), open data intermediaries are regarded as playing a pivotal role. They can enhance the access to and the (re-)use of open data and connect other open data actors in the ODE.

The overarching objective of this dissertation is to understand how open data intermediation business models can support a sustainable ODE. Towards that end, four sub-research questions were addressed:

- **RQ1:** What are open data intermediaries?
- **RQ2:** What are potential contributions of open data intermediaries in addressing challenges in an open data ecosystem?
- **RQ3:** What are archetypes of open data intermediation business models?
- **RQ4 (also the overarching RQ):** What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem?

The organisation of this chapter is as follows. Section 7.2 summarises the results of each research question of this dissertation. Section 7.3 reflects on and refines the (sustainable) ODE concept based on insights gathered through this dissertation. Section 7.4 highlights the contributions of this dissertation. Section 7.5 elaborates on the future research agenda. Section 7.6 offers the final conclusion.

7.2 Answering the research questions

7.2.1 RQ1: What are open data intermediaries?

Various definitions and conceptualisations of open data intermediaries were described in the academic literature. Some contradict each other, while some are narrower than others. At the same time, there is no established definition of open data intermediaries in policy documents that can be relied upon. For example, the definition of 'data intermediaries' (without 'open' at the front) in the EU Digital Governance Act is too narrow to apply to open data intermediaries. This is the first stumbling block before further research and development on open data intermediaries can proceed.

Therefore, RQ1 clarified the obscurity of open data intermediaries by proposing a common definition through a systematic literature review (SLR) that compiled existing definitions of open data intermediaries and identified the wide range of actor types, tasks, and objectives of open data intermediaries. Chapter 2 addressed this research question. Eventually, this dissertation proposed a definition of open data intermediaries as:

third-party actors who provide specialised resources and capabilities to (i) enhance the supply, flow, and/or use of open data and/or (ii) strengthen the relationships among various open data actors.

7.2.2 RQ2: What are potential contributions of open data intermediaries in addressing challenges in an open data ecosystem?

To better situate open data intermediaries within the ODE, it is necessary to look into what they (can) contribute to other actors. This can help refine the understanding of their characteristics and relationships with other actors in the ODE. Most studies have focused solely on examining the current activities of open data intermediaries. While this baseline understanding is valuable, by exploring what they can potentially contribute, RQ2 also aimed to identify gaps in the ODE that open data intermediaries may be able to close or narrow. Chapter 3 addressed this research question.

Moving from only surveying the literature, RQ2 involved interacting with 19 practitioners from 15 organisations to explore the connections between challenges in the ODE and potential contributions by open data intermediaries to address them. A two-stage methodology was employed. In Stage 1, data was gathered through semi-structured interviews. These interviews derived challenges in the ODE and the potential contributions of open data intermediaries. In Stage 2, the links between the individual potential contributions of open data intermediaries and specific challenges in the ODE that they can address were explored. These links were validated by organisations that were interviewed and additional practitioners.

The study showed that open data intermediaries can help overcome ODE challenges through various technical, non-technical, and combination contributions (Table 7.1). Not only can open data intermediaries help the ecosystem strengthen the four features of a sustainable ODE as proposed by van Loenen et al. (2021), but they can also help address foundational issues around data management systems and mitigate the broad political factors impacting the ODE.

TABLE 7.1 Potential contributions of open data intermediaries

Technical contributions
• Implement federated architecture • Integrate data (e.g., across sectors/administrations) • Transform data into open standards (esp. web standards) • Customise data (based on use cases) • Offer process automation • Develop open-source tooling • Provide direct technical services • Offer a freemium data platform
Non-technical contributions
• Foster public-private collaboration • Foster public-civic collaboration • Implement multistakeholder collaboration • Perform open data advocacy • Invest in open data-based civic tech • Showcase open data value • Promote open non-governmental data
Combination contributions
• Provide consultancy • Streamline cross-administrative processes • Facilitate internal re-use of open data • Facilitate feedback on open data • Provide education on data literacy and skills

The study also suggested that an open data provider or user could benefit from the contributions of multiple open data intermediaries simultaneously, in parallel and/or sequentially. This also means there could be various orderings of open data providers, intermediaries, and users in the ODE. Thus, apart from provider-intermediary-user relationships, they could also take the form of provider-intermediary-intermediary-user or provider-parallel intermediaries-user relationships, among others. Thus, open data intermediaries are not merely a 'bridge' between open data providers and users.

7.2.3 RQ3: What are archetypes of open data intermediation business models?

Since there can be an endless number of business models, each with its unique elements and arrangements of value proposition, value creation, and value capture, RQ3 identified common archetypes of open data intermediation business models that exist in practice. Such knowledge is necessary before any business model recommendations can be prescribed, as different business model archetypes may need to consider different aspects. Chapter 4 addressed this research question.

RQ3 was answered through a four-stage methodology. First, an initial codebook consisting of categories and elements of open data intermediation business models was developed through an SLR. Second, relevant qualitative data was collected from 190 samples of existing open data intermediaries facilitated by the initial codebook developed. The codebook was iteratively modified based on the learning throughout the data-gathering process. Third, K-means clustering was employed to group the business models of the sample cases. Fourth, the K-means clustering results were interpreted, and the archetypes were identified.

Nine open data intermediation business model archetypes were determined in Chapter 4. Each archetype was described based on its value proposition (i.e., what the organisation offers?), value creation (i.e., what are the resources and activities deployed?), and value capture (i.e., how the organisation is compensated?) dimensions (Table 7.2).

TABLE 7.2 Archetypes of open data intermediation business models

ID	Name	Salient characteristics based on value dimensions
A1	**Collaborative open data platform**	**Value proposition:** Open data platform freely available for both open data providers and users. **Value creation:** The critical open data stage is preparation, and the consumer relationship is collaborative (co-creation or community-based). **Value capture:** Funded by external contribution (crowdfunding or sponsorship)
A2	**Paid self-service data delivery**	**Value proposition:** Augmented open data (i.e., in combination with non-open data) delivered via various types of products to data users. **Value creation:** The critical open data stage is preparation, and the consumer relationship is self-service. **Value capture:** Revenue generated from augmented open data through freemium or subscription models.
A3	**Personalised open data service**	**Value proposition:** Multiple service units based on augmented open data, providing personalised services to open data providers and users. **Value creation:** The consumer relationship is personal assistance. **Value capture:** Revenue typically generated through service delivery.
A4	**Interactive app with other complementary products**	**Value proposition:** Interactive app with other complementary products. **Value creation:** The critical open data stage is re-use, and the consumer relationship is self-service. **Value capture:** Revenue generated mainly from (augmented) open data via various means such as subscription fees, app sales, and sponsorship. Complementary products may enhance the benefit, visibility, or appeal of the interactive app.
A5	**Open data repository funded by sponsorship**	**Value proposition:** Open data repository mainly targeted at generic open data re-users and is free. **Value creation:** The critical open data stage is the preparation, and the consumer relationship is self-service. **Value capture:** Funded by public or private sponsorship.
A6	**One-stop package around an (augmented) open data platform/ repository**	**Value proposition:** Multiple product units with complementary products, centred around a restricted data platform/ repository based on augmented open data. The target consumers are typically (but not necessarily) highly skilled data users and providers. **Value creation:** The critical open data stage is the preparation, and the consumer relationship is self-service. **Value capture:** Revenue generated through subscription fees or software sales.

>>>

TOC

TABLE 7.2 Archetypes of open data intermediation business models

ID	Name	Salient characteristics based on value dimensions
A7	**Single-purpose app**	**Value proposition:** Single-purpose app based on open data, targeting generic data users. **Value creation:** The critical open data stage is re-use, and the consumer relationship is self-service. **Value capture:** Various means of revenue generation, such as cross-subsidy and sponsorship.
A8	**Interactive app without complementary products**	**Value proposition:** Interactive app without other complementary products. **Value creation:** The critical open data stage is re-use, and the consumer relationship is self-service or personal assistance. **Value capture:** Various means of revenue generation, such as subscription fees, brokerage, or app sales.
A9	**Open data advocacy**	**Value proposition:** Multiple units of open data advocacy, campaigning, or lobbying services. **Value creation:** Various critical stages of the open data lifecycle and various forms of consumer relationships. **Value capture:** Mainly funded via external contributions (sponsorship or crowdfunding), but in some cases, through service delivery.

Notably, findings from RQ3 also showed that open data intermediation business models do not have to rely on generating revenue solely or mainly from open data products. Many open data intermediaries also offer other open data-based or non-open data-based products. For certain archetypes, these other products are complementary to the open data products. For one archetype in particular (i.e., the one-stop package around an (augmented) open data platform/ repository), the main source of revenue is in fact from the non-open data-based products instead of the open data-based products.

Furthermore, it was also confirmed that contributions of open data intermediaries can happen at various stages of the open data lifecycle (i.e., identification, preparation, publication, re-use, and evaluation). This again reaffirmed that open data intermediaries are not merely the ‘bridge’ between open data providers and users.

Additionally, the business model archetypes of open data intermediaries discovered are starkly different from those of data intermediaries that mainly deal with non-open data, as identified by Schweihoff et al. (2024). Except for one, all of the other patterns of data intermediation services they found offer services related to data control, consent management, or identity management. Neither of these aspects is particularly focal in the open data intermediation business models. The more crucial aspect for open data intermediaries is gaining benefits from intermediating data that is already freely reusable by everyone under an open license (i.e., not requiring registration, consent, or identity verification). This ascertains the peculiarity of open data intermediaries compared to generic or non-open data intermediaries.

7.2.4 RQ4 (also the overarching RQ): What aspects should be considered in developing open data intermediation business models that support a sustainable open data ecosystem?

RQ4, which is also the overarching research question of this dissertation, aimed to recommend aspects to consider in developing open data intermediation business models that support sustainable ODE. Since there are multiple open data intermediation business model archetypes (identified in RQ3), RQ4 was addressed specifically for two archetypes, each through an in-depth single-case study. They are Esri in Chapter 5 (representing the one-stop package around an (augmented) open data platform/repository archetype) and OpenStreetMap (OSM) in Chapter 6 (representing the collaborative open data platform archetype). The decision to study two archetypes was guided by feasibility considerations for a doctoral dissertation. Additionally, this dissertation focused on the geospatial data domain to address RQ4. Esri and OSM are two prominent open data intermediaries in the geospatial domain. By focusing on a specific domain through the selection of the two case studies, the role of open data intermediaries can be understood more deeply by having a close familiarity with the context in which they operate.

Based on the Esri and OSM cases, 6 recommendations that apply to both one-stop package and collaborative open data platform archetypes, 11 specifically to the one-stop package archetype, and 14 specifically to the collaborative open data platform archetype were identified (Table 7.3). In total, this dissertation offers 31 aspects to consider. The 6 recommendations that apply to both archetypes revolve around organisational identity, non-traditional solutions, adopting new and emerging technology, simplifying the open data supply and use processes, stimulating multistakeholder collaborations through projects and events, and offering consultancy or training. In the table, the asterisk (*) indicates that the aspect can also be linked to a potential contribution(s) of open data intermediaries identified in Chapter 3.

TABLE 7.3 Factors to consider in developing open data intermediation business models that support a sustainable ODE

Applicable to both the one-stop package around an (augmented) open data platform/repository archetype (based on the Esri case) and the collaborative open data platform archetype (based on the OSM case)	
1.	Ensure consistency in how members view the (potentially multifaceted) organisational identity, especially as the community or organisation expands or becomes more decentralised.
2.	Consider offering new or non-traditional solutions instead of being entrenched in traditional paradigms.
3.	Invest in the adoption of new and emerging technology for the software infrastructure (e.g. the application of AI for metadata recommendations).
4.	Offer products or services that simplify, as much as possible, the process of supplying and using open data.*
5.	Stimulate potential multistakeholder collaborations, e.g., through projects or events.*
6.	Offer consultancy or training services, including through the formal education sector (e.g., schools or universities).*
Applicable specifically to the one-stop package around an (augmented) open data platform/repository archetype (based on the Esri case)	
1.	Ensure that open data intermediation services offered are fittingly integrated with the existing core products or services.
2.	Offer open data intermediation services that are consistent with the organisational identity (i.e., does not involve a significant shift from the core business).
3.	Offer diverse complementary products or services that leverage open data.
4.	Offer services that minimise open data-associated risks that customers have to deal with.
5.	Offer customised data, services, or projects catering to local needs.*
6.	Offer open-source software, at least partially as part of the larger product suite (e.g., through a freemium model).*
7.	Facilitate feedback on open data through a structured mechanism.*
8.	Ensure that the unique resources or position are not leveraged in ways that unfairly stifle the growth of other actors (the winner-takes-it-all situation), e.g., by committing to the development of broadly adoptable open standards and technical interoperability.
9.	Showcase the value of open data.*
10.	Advocate for the release of open data from non-public sectors.*
11.	Invest in open data-based collaborations.*
Applicable specifically to the collaborative open data platform archetype (based on the OSM case)	
1.	Foster a healthy and constructive community of contributors.
2.	Protect the overall interests of the contributors by implementing a transparency policy and putting in place a mechanism to prevent potential takeover or hijacking.
3.	Invest in enhancing the visibility and reach of the organisation or community.
4.	Ensure a clear, efficient, civility-focused communication mechanism.
5.	Facilitate and encourage the development of (especially open-source) tools around the platform (i.e., federated architecture).*
6.	Leverage a self-correcting mechanism by cultivating a 'team-minded' culture and having a clear process for dispute resolution.*
7.	Support local chapters or communities in engaging with local organisations (e.g., local governments or NGOs).*
8.	Transform disagreements and conflicts into opportunities for enhancing governance mechanisms and technical development.
9.	Allow diverse types of data to be contributed to accommodate broad and diverse use cases.
10.	Take initiatives to include the perspectives of indirect impact recipients (i.e., the non-users), e.g., by working with on-the-ground NGOs and prioritising their needs.
11.	Choose a reciprocal licence that does not impose (many) restrictions on the end-use (i.e., derivative) work.
12.	Proactively recruit new data contributors among typically marginalised or disadvantaged groups and nurture a welcoming environment to retain their participation.
13.	Meaningfully consider feedback related to the technological development from non-technical expert members.
14.	Cultivate a culture of deliberating ethical responsibilities in the contribution and use of data, beyond providing general ethics guidelines.

Note: Some rephrasing was done to merge recommendations that apply to both one-stop package and collaborative open data platform archetypes. The asterisk (*) indicates that the aspects to consider can also be associated with the potential contribution(s) of open data intermediaries found in Chapter 3.

7.3 Reflection on (sustainable) open data ecosystem

This dissertation began with a particular theoretical understanding of what is meant by a (sustainable) open data ecosystem (ODE). The ODE is understood as an analytical lens, inspired by the ANT, that foregrounds the complex interrelations of open data actors. The definition of the ODE by Csáki (2019) was adopted: ‘way of looking at how participating actors and groups create shared meaning and generate value around open data and how the structural properties of their interactions shape this process, which in turn enables or constrains the growth and health of the ecosystem itself’. Furthermore, likewise inspired by the ANT, a sustainable ODE is understood as a condition where stable networks of open data actors’ aligned interests are created and maintained. Four features of a sustainable ODE proposed by van Loenen et al. (2021) were adopted in this dissertation, namely, user-driven (open data supply matches the demands of users of different types and domains), circular (all actors mutually create and capture value), inclusive (all actors, not only government organisations, are incentivised to contribute open data and participate in decision-making processes), and skills-based (appropriate data skills and competencies are applied).

By the end of this dissertation, several reflections are offered. First, this dissertation reaffirmed the need to distinguish a *role* from an *actor* in the ODE, as elucidated by Oliveira & Lóscio (2018). Diverse ODE actors can undertake open data intermediation, including public organisations, for-profit companies, and CSOs. Notably, open data intermediaries do not only exist outside of the public sector, as some have implied (Balvert & van Maanen, 2019; Schrock & Shaffer, 2017). Having said that, what a government organisation can do as an open data intermediary may differ from what a for-profit company or a CSOs can do in that role due to the different legal obligations, societal expectations, resources, and other factors.

Second, some of the challenges of the ODE identified in Chapter 3 are foundational issues around open data management systems (e.g., heterogenous data administrations, technical difficulties in establishing open data systems, and poor data quality) or related to broader political factors (e.g., inflexible/unclear government-market boundary according to competition laws, reliance on political agenda, and inflexible law), and do not necessarily fit into the four features of a sustainable ODE suggested by van Loenen et al. (2021). This implies that the four features may be inadequate to determine the sustainability of the ODE. Additional layers of criteria may be necessary and deserve future attention. At the very least, the four features noted by van Loenen et al. (2021) may have to be clarified or refined to readily incorporate those foundational and broader political issues around open data. Alternatively, those features may only become relevant when open data provision and use in a particular jurisdiction or domain reached a certain level of maturity and broad sociopolitical support. Further research on this aspect may offer useful insights.

Third, the case of OSM in Chapter 6 showed that the supplier-driven and user-driven features as discussed by van Loenen et al. (2021) may be two sides of the same coin. The rapid growth of companies editing OSM data following their commercial objectives validated that the open data that is currently available may not cater to the needs of diverse re-users precisely because it was initially created following the interests of the open data producers, who then made the data open for re-use. Thus, moving from supplier-driven towards user-driven ODE indeed requires the alignment of the data suppliers and users' interests. Specifically, in the case of OSM, such a condition can be achieved by having data contributors from diverse backgrounds, representing or addressing the needs of diverse users. For the case of open *government* data, it may involve continuously convincing government organisations on how publishing certain open data in certain ways may also benefit them (e.g., through economic contributions or public sector innovation).

Fourth, in the same vein as the previous point, Verhulst et al., 2020 (p. 9) argued that a new 'wave' of open data, dubbed 'the third wave', should represent 'a much more purpose-directed approach to data provision than prior waves' for more impactful data re-use. However, purposes have always been tacitly embedded in what open data becomes available. Paradoxically, the key philosophy behind open data is about (re)generating its value through its re-use for different purposes, and all the current and future purposes of open data cannot be imagined now or by a group of certain actors. This is not to suggest that 'purpose-directed approach' is not a way to go, but rather, a sustainable ODE should place greater emphasis on the inclusive and equitable participation of data contributors and users (i.e., the inclusive and circular features of van Loenen et al. (2021)) instead of being occupied pre-defining the purposes of open data.

Fifth, the OSM case in Chapter 6 also showed that an ODE where government data is not the only (or main) open data in circulation, described by the inclusive feature by van Loenen et al. (2021), already exists. Interestingly, the contributions of data to OSM by various actors were not compelled through any form of legislation (such as the EU Open Data Directive). Instead, they voluntarily contributed data because they appreciated the tangible value of such data either to them (e.g., their own commercial interests) or to others (e.g., to disaster victims) and/or due to the reciprocal requirement of the ODbL licence. However, some government organisations are still reluctant to be involved in contributing to and using OSM data. Such a situation highlights that the inclusive feature of a sustainable ODE by van Loenen et al. (2021) should also emphasise not only the availability of the open non-government data but also the integration of the open government data and open non-government data, instead of the two circulating in silos.

Sixth, insights from the OSM case also questioned whether the skills-based feature, as described by van Loenen et al. (2021), is necessarily a required or realistic feature for a sustainable ODE. The do-ocracy principle in OSM privileges those with high-level skills in determining the technological direction of OSM. It potentially overlooked ideas and needs from non-expert community members. At the same time, certain high-level skills are indeed required to carry out certain tasks, such as maintaining the OSM database. However, even through various skills training, it is unrealistic to expect everyone to have such a high-level technical proficiency. The question is thus, what is the so-called optimal balance between the inclusive and skills-based features of a sustainable ODE?

Seventh, the challenges of the ODE and the potential contributions of open data intermediaries found in Chapter 3 called the boundaries of the ODE into question. Notably, the ODE challenges identified include those associated with external factors beyond just open data, and the contributions of open data intermediaries may involve many more activities than directly handling open data. Nevertheless, drawing from ANT, the boundaries of the ODE do not have to (or cannot) be defined beforehand. However, for a specific assessment, research inquiry, or intervention, one should identify relevant actors, trace their associations, and analyse their interactions, which would involve making and remaking boundaries. Striking a balance between being pragmatic and reductionist is indeed a constant and iterative process.

Eighth, Chapters 5 and 6 reaffirmed the importance of the entire technology stack to the sustainability of the ODE. It is not only the characteristics of the open data (coverage, quality, format, license, etc.) that are important to the sustainability of the ODE, but also how it is disseminated and re-used, including the software used. This insight is less emphasised in any of the sustainable ODE's features by van Loenen et al. (2021), which can be explained by the fact that those features are heavily centred on the open data per se instead of the open data actors and their (complex) relationships. One potential remedy is to highlight, in the circular feature, the critical role of open data intermediaries in providing the technical and social enablers for effective value creation and capture by various ODE actors.

Lastly, Chapters 5 and 6 also called into question how various responsibilities in the ODE should be allocated to ensure that diverse open data needs are addressed and public interests are protected. For example, most of Esri's open data intermediation value propositions are only enjoyed by its customers, who can afford to subscribe to its proprietary software. Hence, challenges that Esri addresses may remain unaddressed for non-Esri customers. With regard to public interests, while the rise of corporate involvement in OSM may have long-term benefits, it may also expose OSM to some forms of co-optation if not managed carefully. Thus, this dissertation reasserted the critical role of governance and governing institutions, not only governments but also organisations such as OGC and OSMF.

7.4 Contributions of the dissertation

Each research question of this dissertation responded to a past knowledge gap that also has practical relevance. The common definition proposed in Chapter 2 can be used by researchers and practitioners to mutually identify open data intermediaries and build knowledge about them on top of a mutual understanding of what constitutes open data intermediaries. In the past, various definitions and conceptualisations of open data intermediaries were described in the academic literature, with some contradicting each other and some being narrower than others. Additionally, in the process of formulating a common definition, a range of actors, tasks, and objectives of open data intermediaries was identified. This has, in itself, enhanced the understanding of their roles. The resulting overview highlights the diversity among open data intermediaries, an important consideration when developing related policies or business models. Moreover, it can assist actors within the ODE in reflecting on questions such as 'Who is currently doing what?' and, by extension, 'What gaps remain to be addressed?'.

The findings from Chapter 3 contributed to the body of knowledge by exploring the links between the potential contributions of open data intermediaries and the challenges in the ODE. Most past studies have focused solely on examining the current activities of open data intermediaries. Findings from this chapter reaffirmed the critical role of open data intermediaries in the ODE, including in addressing challenges related to foundational issues around data management systems and broad political factors. For practice, those contributions can be transformed by open data intermediaries into opportunities to address the pain points encountered by various ODE actors.

The findings from Chapter 4 offered a broad overview (i.e., nine archetypes) and more detailed account (i.e., across the value proposition, value creation, and value capture dimensions) of existing open data intermediation business model archetypes. This knowledge was missing from the past literature. Besides, the study was based on a large number of cases across many countries, whereas previous studies were based on only a handful of cases and/or a single country. The study offered insights for existing and potential open data intermediaries on the business model they can adopt. Such knowledge may be particularly illuminating to public organisations and NGOs/NPOs since, currently, they only employ a limited number of business model archetypes. Furthermore, the codebook developed during the identification of the archetypes is in and of itself useful to practitioners as a morphological box for designing and experimenting with new open data intermediation business models, by mixing and matching elements across different categories. In this sense, the codebook has the potential to be turned into a more detailed and tailored equivalent of Osterwalder & Pigneur's (2010) business model canvas, specifically for open data intermediaries.

Finally, findings from Chapters 5 and 6, which addressed the overarching research question of this dissertation, offered practical recommendations on aspects to consider in developing open data intermediation business models that support a sustainable ODE. Such recommendations are currently absent in the academic literature. They can be utilised by existing or potential open data intermediaries for their business model development and innovation. The aspects recommended also highlighted the importance of the social dimension (e.g., organisational identity, engagement with the users, and community management) and the legal dimension (e.g., the choice of licence), apart from the technological dimension, in the development of open data intermediation business models.

7.5 Future research

While the focus of this dissertation has been on the business models of open data intermediaries, future research should look into the policy aspects that could steer open data intermediaries to act in ways that support the sustainability of the ODE, for example, the viability and benefits of extending the open data legislation to non-government open data intermediaries. Without a forward-looking and robust governance system, for example in terms of anticipating the evolution of emerging technology and its associated risks, and proactively facilitating multistakeholder engagement in open data decision-making processes, the design of open data intermediation business models alone cannot ensure the sustainability of the ODE.

Additionally, open data intermediaries hidden and potentially undesirable impacts on other ODE actors deserve attention, taking into account that these impacts may evolve over the course of their business model development (i.e., nascent versus mature and potentially dominant stage). While the business model design of open data intermediaries may perform well for themselves, it may, on the other hand, endanger the sustainability of the ODE. Towards that end, various concepts offered by ANT, such as translation, immutable mobile, inscription, and black box, can be leveraged.

Moreover, this dissertation has only focused on two case studies representing two open data intermediation business model archetypes. Extending the research to studying other archetypes is thus necessary. Besides, the two case studies chosen, namely Esri and OSM, are both well-established open data intermediaries from the same domain, i.e., geospatial data. Even within the geospatial domain, there are significantly different types of open data. In particular, the emergence of open data intermediaries dealing with aerial and street-level images, such as Mapillary and KartaView, and historical archival maps, such as Allmaps, presents opportunities for essential and exciting future research. Furthermore, exploring the transferability of the recommendations derived from the two cases to other cases, including new, public sector-based, and/or non-geospatial open data intermediaries is needed.

Stepping out of the scope of open data intermediaries, further research is necessary to refine the understanding of a sustainable ODE, including in terms of the adequateness (e.g., where do broad political factors and the entire technology stack fall into) and practical relevance (e.g., how realistic is the skills expectations) of the features proposed by van Loenen et al. (2021). Furthermore, the mechanisms to build and maintain the alignment of open data providers' and users' interests as a potential means to achieve a more user-driven ODE are worth investigating. Additionally, given the already extensive and widespread provision and re-use of open non-government data, as demonstrated through the case of OSM, further research is necessary to understand how open data from government and non-government sectors interact and how can it be better integrated to avoid siloed development.

7.6 Final conclusion

In conclusion, this dissertation advances the understanding of open data intermediaries, their position in the ODE, and their business models. In particular, this dissertation theoretically contributed to the definition of open data intermediaries, potential contributions of open data intermediaries, and the archetypes of open data intermediation business models. It also identified practical aspects to consider in developing open data intermediation business models that contribute to a sustainable ODE, ultimately enhancing the generation of open data value. This value, in turn, can be leveraged to foster innovation, promote economic well-being, and address pressing social and environmental challenges.

References

Accenture. (2020). *Navigating OpenStreetMap*. https://wiki.openstreetmap.org/w/images/8/80/Accenture-Navigating-OpenStreetMap.pdf

Acciarini, C., Cappa, F., Boccardelli, P., & Oriani, R. (2023). How can organizations leverage big data to innovate their business models? A systematic literature review. *Technovation*, *123*, 102713. https://doi.org/10.1016/j.technovation.2023.102713

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, *16*(1), 3–9.

Adaktylou, N., Stratoulias, D., & Landenberger, R. (2020). Wildfire Risk Assessment Based on Geospatial Open Data: Application on Chios, Greece. *ISPRS International Journal of Geo-Information*, *9*(9), Article 9. https://doi.org/10.3390/ijgi9090516

Afuah, A. (2018). *Business Model Innovation: Concepts, Analysis, and Cases*. Routledge. https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1907925&site=ehost-live

Ahmadi Zeleti, F., Ojo, A., & Curry, E. (2016). Exploring the economic value of open government data. *Government Information Quarterly*, *33*(3), 535–551. https://doi.org/10.1016/j.giq.2016.01.008

Albert, S., & Whetten, D. A. (2003). Organizational identity. In J. M. T. Balmer & S. A. Greyser (Eds.), *Revealing the corporation: Perspectives on identity, image, reputation, corporate branding, and corporate-level marketing: An anthology* (pp. 77–105). Routledge.

Al-Debei, M. M., & Avison, D. (2010). Developing a unified framework of the business model concept. *European Journal of Information Systems*, *19*(3), 359–376. https://doi.org/10.1057/ejis.2010.21

Almendros-Jiménez, J. M., Becerra-Terón, A., Merayo, M. G., & Núñez, M. (2021). Metamorphic testing of OpenStreetMap. *Information and Software Technology*, *138*, 106631. https://doi.org/10.1016/j.infsof.2021.106631

Alvarez Leon, L. F., & Quinn, S. (2019). The value of crowdsourced street-level imagery: Examining the shifting property regimes of OpenStreetCam and Mapillary. *GeoJournal*, *84*(2), 395–414. https://doi.org/10.1007/s10708-018-9865-4

Amit, R., & Zott, C. (2001). Value creation in E-business. *Strategic Management Journal*, *22*(6–7), 493–520. https://doi.org/10.1002/smj.187

Amos, M., Coast, S., Collinson, M., Hoff, H., Möller, U., & Slater, G. (2009, October 20). *New License Proposal by The OpenStreetMap Foundation Licensing Working Group October 20th, 2009*. OSM Foundation. https://osmfoundation.org/w/images/archive/3/3c/20091023173924!License_Proposal.pdf

Anderson, J. (2021, February 25). *Editing Activity by Country Since 2020 and the 2021 OSMF Survey*. OpenStreetMap. https://www.openstreetmap.org/user/Jennings%20Anderson/diary/395950

Anderson, J., & Sarkar, D. (2020, July 4). Curious Cases of Corporations in OpenStreetMap. *Proceedings of the Academic Track at the State of the Map 2020*. https://doi.org/10.5281/zenodo.3923050

Anderson, J., Sarkar, D., & Palen, L. (2019). Corporate Editors in the Evolving Landscape of OpenStreetMap. *ISPRS International Journal of Geo-Information*, *8*(5), Article 5. https://doi.org/10.3390/ijgi8050232

Anderson, J., Soden, R., Keegan, B., Palen, L., & Anderson, K. M. (2018). The Crowd is the Territory: Assessing Quality in Peer-Produced Spatial Data During Disasters. *International Journal of Human–Computer Interaction*, *34*(4), 295–310. https://doi.org/10.1080/10447318.2018.1427828

Andrason, A., & van Schalkwyk, F. (2017). *Opportune Niches in Data Ecosystems: Open Data Intermediaries in the Agriculture Sector in Ghana* (SSRN Scholarly Paper No. 2949722). Social Science Research Network. https://doi.org/10.2139/ssrn.2949722

Andreini, D., & Bettinelli, C. (2017). *Business Model Innovation: From Systematic Literature Review to Future Research Directions*. Springer International Publishing. https://doi.org/10.1007/978-3-319-53351-3

Anindya, M. (2024, November 13). *Grab built its own map in Southeast Asia, and is now going after Google*. Rest of World. https://restofworld.org/2024/grab-google-maps-southeast-asia/

Arnold, L., & Hukal, P. (2024). The varying effects of standardisation on digital platform innovation: Evidence from OpenStreetmap. *Innovation*, 1–26. https://doi.org/10.1080/14479338.2024.2345319

Arsanjani, J. J., & Vaz, E. (2015). An assessment of a collaborative mapping approach for exploring land use patterns for several European metropolises. *International Journal of Applied Earth Observation and Geoinformation*, *35*, 329–337. https://doi.org/10.1016/j.jag.2014.09.009

Attard, J., Orlandi, F., & Auer, S. (2016). Value Creation on Open Government Data. *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 2605–2614. https://doi.org/10.1109/HICSS.2016.326

Aydinoglu, A. C., & Bilgin, M. S. (2015). Developing an open geographic data model and analysis tools for disaster management: Landslide case. *Natural Hazards and Earth System Sciences*, *15*(2), 335–347. https://doi.org/10.5194/nhess-15-335-2015

Baack, S. (2015). Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism. *Big Data & Society*, *2*(2), 2053951715594634. https://doi.org/10.1177/2053951715594634

Bakillah, M., Liang, S., Mobasheri, A., Jokar Arsanjani, J., & Zipf, A. (2014). Fine-resolution population mapping using OpenStreetMap points-of-interest. *International Journal of Geographical Information Science*, *28*(9), 1940–1963. https://doi.org/10.1080/13658816.2014.909045

Ballantyne, P., & Berragan, C. (2024). Overture Point of Interest data for the United Kingdom: A comprehensive, queryable open data product, validated against Geolytix supermarket data. *Environment and Planning B: Urban Analytics and City Science*, 23998083241263124. https://doi.org/10.1177/23998083241263124

Ballatore, A., & Mooney, P. (2015). Conceptualising the geographic world: The dimensions of negotiation in crowdsourced cartography. *International Journal of Geographical Information Science*, *29*(12), 2310–2327. https://doi.org/10.1080/13658816.2015.1076825

Balvert, A., & van Maanen, G. (2019). Open for whom? The role of intermediaries in data publication. *Revista Publicum*, *5*(2), Article 2. https://doi.org/10.12957/publicum.2019.47205

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Barone, E. (2021, September 29). What Happens When the World's Most Popular COVID-19 Dashboard Can't Get Data? *Time*. https://time.com/6101967/covid-19-data-gaps/

Barry, E., & Bannister, F. (2014). Barriers to open data release: A view from the top. *Information Polity*, *19*(1–2), 129–152. https://doi.org/10.3233/IP-140327

Barth, A. (2013, May 30). *Collaborating to improve OpenStreetMap infrastructure*. Knight Foundation. https://knightfoundation.org/articles/collaborating-improve-openstreetmap-infrastructure/

Basiouka, S., Potsiou, C., & Bakogiannis, E. (2015). OpenStreetMap for cadastral purposes: An application using VGI for official processes in urban areas. *Survey Review*, *47*(344), 333–341. https://doi.org/10.1179/1752270615Y.0000000011

Beck, L. W. (1956). Kant's Theory of Definition. *The Philosophical Review*, *65*(2), 179–191. https://doi.org/10.2307/2182830

Bégin, D., Devillers, R., & Roche, S. (2018). The life cycle of contributors in collaborative online communities -the case of OpenStreetMap. *International Journal of Geographical Information Science*, *32*(8), 1611–1630. https://doi.org/10.1080/13658816.2018.1458312

Benitez-Paez, F., Degbelo, A., Trilles, S., & Huerta, J. (2018). Roadblocks Hindering the Reuse of Open Geodata in Colombia and Spain: A Data User's Perspective. *ISPRS International Journal of Geo-Information*, *7*(1), Article 1. https://doi.org/10.3390/ijgi7010006

Bergman, R., Abbas, A. E., Jung, S., Werker, C., & de Reuver, M. (2022). Business model archetypes for data marketplaces in the automotive industry. *Electronic Markets*, *32*(2), 747–765. https://doi.org/10.1007/s12525-022-00547-x

Berners-Lee, T., & Shadbolt, N. (2010, January 21). Our manifesto for government data. *The Guardian*. https://www.theguardian.com/news/datablog/2010/jan/21/timbernerslee-government-data

Berry, L. (2024, April 23). Your Living Atlas Questions Answered. *ArcGIS Blog*. https://www.esri.com/arcgis-blog/products/arcgis-living-atlas/sharing-collaboration/your-living-atlas-questions-answered/

Bertolotto, M., McArdle, G., & Schoen-Phelan, B. (2020). Volunteered and crowdsourced geographic information: The OpenStreetMap project. *Journal of Spatial Information Science*, *20*, Article 20. https://josis.org/index.php/josis/article/view/115

Biljecki, F., Chew, L. Z. X., Milojevic-Dupont, N., & Creutzig, F. (2021). *Open government geospatial data on buildings for planning sustainable and resilient cities* (No. arXiv:2107.04023). arXiv. http://arxiv.org/abs/2107.04023

Biljecki, F., & Chow, Y. S. (2022). Global Building Morphology Indicators. *Computers, Environment and Urban Systems*, *95*, 101809. https://doi.org/10.1016/j.compenvurbsys.2022.101809

Biljecki, F., Chow, Y. S., & Lee, K. (2023). Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes. *Building and Environment*, *237*, 110295. https://doi.org/10.1016/j.buildenv.2023.110295

Biljecki, F., & Ito, K. (2021). Street view imagery in urban analytics and GIS: A review. *Landscape and Urban Planning*, *215*, 104217. https://doi.org/10.1016/j.landurbplan.2021.104217

Bing Maps. (2018, June 28). *Microsoft Releases Building Footprints*. Bing Maps API. https://www.microsoft.com/en-us/maps/news/microsoft-releases-130-million-building-footprints-in-the-usa-as-open-data

Bittner, C. (2017). Diversity in volunteered geographic information: Comparing OpenStreetMap and Wikimapia in Jerusalem. *GeoJournal*, *82*(5), 887–906. https://doi.org/10.1007/s10708-016-9721-3

Bittner, C., & Glasze, G. (2021). Who Maps Middle Eastern Geographies in the Digital Age? Inequalities in Web 2.0 Cartographies in Israel/Palestine. In A. Strohmaier & A. Krewani (Eds.), *Media and Mapping Practices in the Middle East and North Africa: Producing Space*. Amsterdam University Press.

Blankena, F. (2016, July 1). Doorbraakproject Open Geodata kan zonder doorbraak. *iBestuur*. https://ibestuur.nl/artikel/doorbraakproject-open-geodata-kan-zonder-doorbraak/

Boateng, E. N. K., Manyungwa, Z., & Anderson, J. (2023). Assessing YouthMappers Contributions to the Generation of Open Geospatial Data in Africa. In P. Solís & M. Zeballos (Eds.), *Open Mapping towards Sustainable Development Goals: Voices of YouthMappers on Community Engaged Scholarship* (pp. 171–179). Springer International Publishing. https://doi.org/10.1007/978-3-031-05182-1_15

Bocquet, R., Cotterlaz-Rannard, G., & Ferrary, M. (2020). How Do NPOs Get Funding? A Business Model Perspective Based on the Conversion of Symbolic Capital. *Nonprofit and Voluntary Sector Quarterly*, *49*(6), 1233–1258. https://doi.org/10.1177/0899764020925912

Boeing, G. (2025). *Getting Started*. OSMnx 2.0.2 Documentation. https://osmnx.readthedocs.io/en/stable/getting-started.html

Bojovic, N., Sabatier, V., & Coblence, E. (2020). Becoming through doing: How experimental spaces enable organizational identity work. *Strategic Organization*, *18*(1), 20–49. https://doi.org/10.1177/1476127019864673

Breunig, D. (2024, July 31). Towards Standardizing Place. *Drew Breunig*. https://www.dbreunig.com/2024/07/31/towards-standardizing-place.html

Brovelli, M. A., & Coetzee, S. (2021). Open geospatial data for responding to the COVID-19 challenge. In A. Rajabifard, D. Paez, & G. Foliente (Eds.), *COVID-19 Pandemic, Geospatial Information, and Community Resilience* (pp. 31–54). CRC Press.

Brovelli, M. A., & Zamboni, G. (2018). A New Method for the Assessment of Spatial Accuracy and Completeness of OpenStreetMap Building Footprints. *ISPRS International Journal of Geo-Information*, *7*(8), Article 8. https://doi.org/10.3390/ijgi7080289

Brugger, J., Fraefel, M., Riedl, R., Fehr, H., Schöeneck, D., & Weissbrod, C. S. (2016). Current Barriers to Open Government Data Use and Visualization by Political Intermediaries. *2016 Conference for E-Democracy and Open Government (CeDEM)*, 219–229. https://doi.org/10.1109/CeDEM.2016.18

Brunsdon, C., & Comber, A. (2021). Opening practice: Supporting reproducibility and critical spatial data science. *Journal of Geographical Systems*, *23*(4), 477–496. https://doi.org/10.1007/s10109-020-00334-2

Budhathoki, N. R., Bruce, B. (Chip), & Nedovic-Budic, Z. (2008). Reconceptualizing the role of the user of spatial data infrastructure. *GeoJournal*, *72*(3), 149–160. https://doi.org/10.1007/s10708-008-9189-x

Budhathoki, N. R., & Haythornthwaite, C. (2013). Motivation for Open Collaboration: Crowd and Community Models and the Case of OpenStreetMap. *American Behavioral Scientist*, *57*(5), 548–575. https://doi.org/10.1177/0002764212469364

Bullock, K. (2024, April 22). *Overture Maps and the Cloud Native Global Map*. Development Seed. https://developmentseed.org/blog/2024-04-22-overture-summit/

Cahlikova, T., & Mabillard, V. (2020). Open Data and Transparency: Opportunities and Challenges in the Swiss Context. *Public Performance & Management Review*. https://www.tandfonline.com/doi/full/10.1080/15309576.2019.1657914

Callaham, J. (2012, March 27). Microsoft supporting competitor to Google Maps. *Neowin*. https://www.neowin.net/news/microsoft-supporting-competitor-to-google-maps/

Callon, M., & Law, J. (1997). After the Individual in Society: Lessons on Collectivity from Science, Technology and Society. *Canadian Journal of Sociology / Cahiers Canadiens de Sociologie*, *22*(2), 165. https://doi.org/10.2307/3341747

Calzati, S., & Ploeger, H. (2024). Problem-solving? No, problem-opening! A method to reframe and teach data ethics as a transdisciplinary endeavour. *Big Data & Society*, *11*(3), 20539517241270687. https://doi.org/10.1177/20539517241270687

Cambridge Dictionary. (n.d.). *Archetype*. Retrieved 10 April 2024, from https://dictionary.cambridge.org/dictionary/english/archetype

Camero, J. (2015, March 26). *House okays National Mangrove Forests Conservation and Rehabilitation Act* [Republic of the Philipines House of Representatives Press Releases]. https://www.congress.gov.ph/press/details.php?pressid=8604

Cañares, M. P. (2014). Opening the local: Full disclosure policy and its impact on local governments in the Philippines. *Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance*, 89–98. https://doi.org/10.1145/2691195.2691214

Carmona-Derqui, D., Gutiérrez-Mora, D., & Oto-Peralías, D. (2023). Tabulating and visualizing street-name data in the US and Europe. *Environment and Planning B: Urban Analytics and City Science*, *50*(7), 1981–1987. https://doi.org/10.1177/23998083231190711

Carolan, L. (2016). *Open data, transparency and accountability: Topic guide*. GSDRC, University of Birmingham.

Carranza, J. P., Piumetto, M. A., Lucca, C. M., & Da Silva, E. (2022). Mass appraisal as affordable public policy: Open data and machine learning for mapping urban land values. *Land Use Policy*, *119*, 106211. https://doi.org/10.1016/j.landusepol.2022.106211

Carraro, V. (2021). Naming Jerusalem on OpenStreetMap. In *Jerusalem Online: Critical Cartography for the Digital Age* (pp. 87–109). Springer. https://doi.org/10.1007/978-981-16-3314-0_5

Carraro, V., & Wissink, B. (2018). Participation and marginality on the geoweb: The politics of non-mapping on OpenStreetMap Jerusalem. *Geoforum*, *90*, 64–73. https://doi.org/10.1016/j.geoforum.2018.02.001

Cerri, M., Steinhausen, M., Kreibich, H., & Schröter, K. (2021). Are OpenStreetMap building data useful for flood vulnerability modelling? *Natural Hazards and Earth System Sciences*, *21*(2), 643–662. https://doi.org/10.5194/nhess-21-643-2021

Chan, M., Johnson, P. A., & Shookner, M. (2016). Assessing the Use of Government Open Data and the Role of Data Infomediaries The Case of Nova Scotia's Community Counts Program. *JeDEM - eJournal of eDemocracy and Open Government*, *8*(1), Article 1. https://doi.org/10.29379/jedem.v8i1.370

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Höchtl, J., & Ferro, E. (2018). *The World of Open Data: Concepts, Methods, Tools and Experiences*. Springer International Publishing AG.

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., & Ferro, E. (2018). Open Data Value and Business Models. In Y. Charalabidis, A. Zuiderwijk, C. Alexopoulos, M. Janssen, T. Lampoltshammer, & E. Ferro (Eds.), *The World of Open Data: Concepts, Methods, Tools and Experiences* (pp. 115–136). Springer International Publishing. https://doi.org/10.1007/978-3-319-90850-2_7

Chattapadhyay, S. (2014). Access and use of government data by research and advocacy organisations in India: A survey of (potential) open data ecosystem. *Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance*, 361–364. https://doi.org/10.1145/2691195.2691262

Chauhan, A., Sarkar, D., Agrawaal, T. S., & Soden, R. (2024). Value Tensions in OpenStreetMap: Openness, Membership, and Policy in Online Communities. *Proc. ACM Hum.-Comput. Interact.*, *8*(CSCW2), 380:1-380:25. https://doi.org/10.1145/3686919

Chesbrough, H. (2010). Business Model Innovation: Opportunities and Barriers. *Long Range Planning*, *43*(2), 354–363. https://doi.org/10.1016/j.lrp.2009.07.010

Chirumalla, K., Kulkov, I., Parida, V., Dahlquist, E., Johansson, G., & Stefan, I. (2024). Enabling battery circularity: Unlocking circular business model archetypes and collaboration forms in the electric vehicle battery ecosystem. *Technological Forecasting and Social Change*, *199*, 123044. https://doi.org/10.1016/j.techfore.2023.123044

Choe, Y., Tomko, M., & Kalantari, M. (2023a). Assessing Mapper Conflict in OpenStreetMap Using the Delphi Survey Method. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3544548.3580758

Choe, Y., Tomko, M., & Kalantari, M. (2023b). Assessing Mapper Conflict in OpenStreetMap Using the Delphi Survey Method. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3544548.3580758

Citymapper. (2025). *Data Sources*. Citymapper. https://content.citymapper.com/news/1435/data-sources

Claessens, C. (2016, June 23). *ArcGIS Open Data—New Beta Sites!* [Online post]. Esri Community. https://community.esri.com/t5/arcgis-hub-blog/arcgis-open-data-new-beta-sites/ba-p/885429

Coast, S. (2015). *The Book of OSM*. CreateSpace Independent Publishing Platform. https://books.google.nl/books?id=awW9jgEACAAJ

Coetzee, S., Ivánová, I., Mitasova, H., & Brovelli, M. A. (2020). Open Geospatial Software and Data: A Review of the Current State and A Perspective into the Future. *ISPRS International Journal of Geo-Information*, *9*(2), Article 2. https://doi.org/10.3390/ijgi9020090

Coetzee, S., Minghini, M., Solis, P., Rautenbach, V., & Green, C. (2018). TOWARDS UNDERSTANDING THE IMPACT OF MAPATHONS – REFLECTING ON YOUTHMAPPERS EXPERIENCES. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-4-W8*, 35–42. FOSS4G 2018 – Academic Track (Volume XLII-4/W8) - 29–31 August 2018, Dar es Salaam, Tanzania. https://doi.org/10.5194/isprs-archives-XLII-4-W8-35-2018

Coleman, D., Georgiadou, Y., & Labonte, J. (2009). Volunteered geographic information: The nature and motivation of produsers. *International Journal of Spatial Data Infrastructures Research*, *4*(4), 332–358.

Commission Delegated Regulation (EU) No 1159/2013 of 12 July 2013 Supplementing Regulation (EU) No 911/2010 of the European Parliament and of the Council on the European Earth Monitoring Programme (GMES) by Establishing Registration and Licensing Conditions for GMES Users and Defining Criteria for Restricting Access to GMES Dedicated Data and GMES Service Information Text with EEA Relevance, No. 911/2010, European Commission (2013). https://eur-lex.europa.eu/eli/reg_del/2013/1159/oj/eng

Companies House. (2024). *OpenStreetMap United Kingdom Community Interest Company overview*. Find and Update Company Information - GOV.UK. https://find-and-update.company-information.service.gov.uk/company/10527963

Conradie, P., & Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly*, *31*, S10–S17. https://doi.org/10.1016/j.giq.2014.01.003

Cook, R. (2009). *Dictionary of Philosophical Logic*. Edinburgh University Press.

Copernicus In Situ. (2018, February 22). Open Data in Denmark: Interview with Olav Eggers, SDFE. *Copernicus In Situ Newsletter #4*. https://insitu.copernicus.eu/news/open-data-in-denmark-interview-with-olav-eggers-sdfe

Corbett, J., Templier, M., & Takeda, H. (2018, January 3). The Making of a 'Top' Open Data City: A Case Study of Edmonton's Open Data Initiative. *Open Data, Information Processing, and Datification in Government.* Hawaii International Conference on System Sciences 2018. https://doi.org/10.24251/HICSS.2018.308

Corbett, J., Templier, M., Townsend, H., & Takeda, H. (2020). Integrating Across Sustainability, Political, and Administrative Spheres: A Longitudinal Study of Actors' Engagement in Open Data Ecosystems in Three Canadian Cities. *Communications of the Association for Information Systems*, *47*(1). https://doi.org/10.17705/1CAIS.04728

Crampton, J. W. (2001). Maps as social constructions: Power, communication and visualization. *Progress in Human Geography*, *25*(2), 235–252. https://doi.org/10.1191/030913201678580494

Crowd2Map. (n.d.). *About Us*. Crowd2Map Tanzania. Retrieved 11 March 2025, from http://About%20Us

Crusoe, J., Simonofski, A., Clarinval, A., & Gebka, E. (2019). The Impact of Impediments on Open Government Data Use: Insights from Users. *2019 13th International Conference on Research Challenges in Information Science (RCIS)*, 1–12. https://doi.org/10.1109/RCIS.2019.8877055

Csáki, C. (2019). Open Data Ecosystems: A Comparison of Visual Models. In A. Kő, E. Francesconi, G. Anderst-Kotsis, A. M. Tjoa, & I. Khalil (Eds.), *Electronic Government and the Information Systems Perspective* (pp. 16–30). Springer International Publishing. https://doi.org/10.1007/978-3-030-27523-5_2

Curran, E. E., & Bowlick, F. J. (2022). Geographic Information Science education at Esri Development Center institutions. *Transactions in GIS*, *26*(1), 341–361. https://doi.org/10.1111/tgis.12858

da Silva Craveiro, G., & Albano, C. (2017). Open data intermediaries: Coproduction in budget transparency. *Transforming Government: People, Process and Policy*, *11*(1), 119–131. https://doi.org/10.1108/TG-12-2015-0057 da Silva Craveiro, G., & Albano, C. S. (2015). Budgetary Data (in an Open Format) Benefits, Advantages, Obstacles and Inhibitory Factors in the View of the Intermediaries of this System: A Study in Latin American Countries. In M. Janssen, M. Mäntymäki, J. Hidders, B. Klievink, W. Lamersdorf, B. van Loenen, & A. Zuiderwijk (Eds.), *Open and Big Data Management and Innovation* (pp. 223–235). Springer International Publishing. https://doi.org/10.1007/978-3-319-25013-7_18

Daniel, Z. (2018, September 11). It's taken 50 years and $225m to save this wild coastline. *ABC News - Foreign Correspondent*. https://www.abc.net.au/news/2018-09-11/esri-founders-spend-225-million-on-california-nature-preserve/10221170

Daniels, M. J., Game, A., Mollura, D. J., & England, R. W. (2021). Strategic Radiology Outreach Planning for Underserved Populations Using Geographic Information Systems. *Journal of the American College of Radiology*, *18*(4), 537–544. https://doi.org/10.1016/j.jacr.2020.08.022

Das, M., Hecht, B., & Gergle, D. (2019). The Gendered Geography of Contributions to OpenStreetMap: Complexities in Self-Focus Bias. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3290605.3300793

Dasgupta, A. (2013, June 25). Has OGC Lost its Way? *Geospatial World*. https://www.geospatialworld.net/article/has-ogc-lost-its-way/

DaSilva, C. M., & Trkman, P. (2014). Business Model: What It Is and What It Is Not. *Long Range Planning*, *47*(6), 379–389. https://doi.org/10.1016/j.lrp.2013.08.004

Davidovic, N., Mooney, P., Stoimenov, L., & Minghini, M. (2016). Tagging in Volunteered Geographic Information: An Analysis of Tagging Practices for Cities and Urban Regions in OpenStreetMap. *ISPRS International Journal of Geo-Information*, *5*(12), Article 12. https://doi.org/10.3390/ijgi5120232

Davies, T. (2011). *Open Data: Infrastructures and ecosystems*. ACM Web Science Conference, Evanston, Illinois. https://www.emergentworks.net/sites/default/files/ikmemergent_archive/Social_Life_of_Data_-_Infrastructure_and_Ecosystem_Paper.pdf

Davies, T., & Edwards, D. (2012). Emerging Implications of Open and Linked Data for Knowledge Sharing in Development. *IDS Bulletin*, *43*. https://doi.org/10.1111/j.1759-5436.2012.00372.x

Davies, T., & Perini, F. (2016). Researching the emerging impacts of open data: Revisiting the ODDC conceptual framework. *The Journal of Community Informatics*, *12*(2), Article 2. https://doi.org/10.15353/joci.v12i2.3246

Davies-Holloway, C. (2021, October 21). Esri UK and Met Office join forces to create climate change teaching resources. *Association For Geographic Information*. https://www.agi.org.uk/esri-uk-and-met-office-join-forces-to-create-climate-change-teaching-resources/

de Arruda, H. F., Reia, S. M., Ruan, S., Atwal, K. S., Kavak, H., Anderson, T., & Pfoser, D. (2024). An OpenStreetMap derived building classification dataset for the United States. *Scientific Data*, *11*(1), 1210. https://doi.org/10.1038/s41597-024-04046-w

de Reuver, M., Bouwman, H., & MacInnes, I. (2009). Business models dynamics for start-ups and innovating e-businesses. *International Journal of Electronic Business*, *7*(3), 269. https://doi.org/10.1504/IJEB.2009.026530

Degbelo, A. (2022). FAIR geovisualizations: Definitions, challenges, and the road ahead. *International Journal of Geographical Information Science*, *36*(6), 1059–1099. https://doi.org/10.1080/13658816.2021.1983579

Dementeva, D., Meeusen, C., & Meuleman, B. (2024). Using OpenStreetMap, Census, and Survey Data to Predict Interethnic Group Relations in Belgium: A Machine Learning Approach. *Social Science Computer Review*, 08944393241269098. https://doi.org/10.1177/08944393241269098

Dempsey, C. (2000, July 5). The Geography Network. *GIS Lounge*. https://www.gislounge.com/the-geography-network/

den Haan, D. (2018). *Intermediation as a dissolver of barriers: How intermediaries can help overcome barriers in open data use.* [Master's thesis, Wageningen University]. https://studenttheses.uu.nl/handle/20.500.12932/31148

Di Valentin, C., Burkhart, T., Vanderhaeghen, D., Werth, D., & Loos, P. (2012). Towards a Framework for Transforming Business Models into Business Processes. *AMCIS 2012 Proceedings*. https://aisel.aisnet.org/amcis2012/proceedings/EnterpriseSystems/10

Díaz Andrade, A., & Urquhart, C. (2010). The affordances of actor network theory in ICT for development research. *Information Technology & People*, *23*(4), 352–374. https://doi.org/10.1108/09593841011087806

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases (1996). https://eur-lex.europa.eu/eli/dir/1996/9/oj/eng

Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society (2019). http://data.europa.eu/eli/dir/2001/29/2019-06-06/eng

Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the Re-Use of Public Sector Information (2003). http://data.europa.eu/eli/dir/2003/98/oj/eng

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) (2007). http://data.europa.eu/eli/dir/2007/2/oj/eng

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-Use of Public Sector Information, No. 2019/1024 (2019). http://data.europa.eu/eli/dir/2019/1024/oj/eng

DiverCTH. (2024). *User:DiverCTH*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/User:DiverCTH

Dixon, S. J. (2024a, February 15). *Countries with most Instagram users 2024*. Statista. https://www.statista.com/statistics/578364/countries-with-most-instagram-users/

Dixon, S. J. (2024b, May 22). *Facebook users by country 2024*. Statista. https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/

Dong, H., Zhang, C., Li, G., & Zhang, H. (2024). Cloud-Native Databases: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, *36*(12), 7772–7791. IEEE Transactions on Knowledge and Data Engineering. https://doi.org/10.1109/TKDE.2024.3397508

Doolin, B., & Lowe, A. (2002). To Reveal is to Critique: Actor–Network Theory and Critical Information Systems Research. *Journal of Information Technology*, *17*(2), 69–78. https://doi.org/10.1080/02683960210145986

Doorbraakproject Open Geodata. (2015). *Doorbraakproject Open Geodata als grondstof voor groei en innovatie. Belemmeringen rondom Open Geodata wegnemen*. https://adoc.pub/doorbraakproject-open-geodata-als-grondstof-voor-groei-en-inc207c91981124eda8cc49d9e21c39fbc80957.html

Dove, G., Shanley, J., Matuk, C., & Nov, O. (2023). Open Data Intermediaries: Motivations, Barriers and Facilitators to Engagement. *Proc. ACM Hum.-Comput. Interact.*, *7*(CSCW1), 78:1-78:22. https://doi.org/10.1145/3579511

Dubois, A., & Gadde, L.-E. (2002). Systematic combining: An abductive approach to case research. *Journal of Business Research*, *55*(7), 553–560. https://doi.org/10.1016/S0148-2963(00)00195-8

Dubosson-Torbay, M., Osterwalder, A., & Pigneur, Y. (2002). E-business model design, classification, and measurements. *Thunderbird International Business Review*, *44*(1), 5–23. https://doi.org/10.1002/tie.1036

Duguay, R., Rauter, T., & Samuels, D. (2023). The Impact of Open Data on Public Procurement. *Journal of Accounting Research*, *61*(4), 1159–1224. https://doi.org/10.1111/1475-679X.12479

Dumpawar, S. (2015). *Open government data intermediaries: Mediating data to drive changes in the built environment* [Master's thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/97994

Duparc, E., Möller, F., Jussen, I., Stachon, M., Algac, S., & Otto, B. (2022). Archetypes of open-source business models. *Electronic Markets*, *32*(2), 727–745. https://doi.org/10.1007/s12525-022-00557-9

Dyer, W. G., & Wilkins, A. L. (1991). Better Stories, Not Better Constructs, to Generate Better Theory: A Rejoinder to Eisenhardt. *The Academy of Management Review*, *16*(3), 613–619. https://doi.org/10.2307/258920

Edwards, P. (1967). *The encyclopedia of philosophy*. Macmillan.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, *50*(1), 25–32.

Eisenhardt, K. M., & Schoonhoven, C. B. (1996). Resource-based View of Strategic Alliance Formation: Strategic and Social Effects in Entrepreneurial Firms. *Organization Science*, *7*(2), 136–150. https://doi.org/10.1287/orsc.7.2.136

Elwood, S. (2008a). Grassroots groups as stakeholders in spatial data infrastructures: Challenges and opportunities for local data development and sharing. *International Journal of Geographical Information Science*, *22*(1), 71–90. https://doi.org/10.1080/13658810701348971

Elwood, S. (2008b). Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal*, *72*(3), 173–183. https://doi.org/10.1007/s10708-008-9186-0

Enaholo, P. (2017). Beyond mere advocacy: CSOs and the role of intermediaries in Nigeria's open data ecosystem. In F. van Schalkwyk, S. Verhulst, G. Magalhaes, J. Pane, & J. Walker (Eds.), *The Social Dynamics of Open Data* (pp. 89–108). African Minds. http://www.africanminds.co.za/the-social-dynamics-of-open-data/

Enaholo, P., & Dina, D. (2020). Journalists and the intermediation of open data: A Nigerian perspective. In *Situating Open Data: Global Trends in Local Contexts* (pp. 31–49). African Minds. https://www.africanminds.co.za/situation-open-data-global-trends-in-local-contexts/#:~:text=Open%20data%20and%20its%20effects,focus%20on%20its%20situated%20use.

Esri. (n.d.-a). *About Esri*. Esri. Retrieved 28 August 2023, from https://www.esri.com/en-us/about/about-esri/company

Esri. (n.d.-b). *ArcGIS Living Atlas of the World*. Esri. Retrieved 30 August 2023, from https://livingatlas.arcgis.com/

Esri. (n.d.-c). *ArcGIS Open Data*. Esri. Retrieved 30 August 2023, from https://www.esri.com/en-us/arcgis/products/arcgis-open-data

Esri. (n.d.-d). *Community Maps Program*. Esri. Retrieved 30 August 2023, from https://communitymaps.arcgis.com/home/

Esri. (n.d.-e). *Esri Products*. Esri. Retrieved 30 August 2023, from https://www.esri.com/en-us/arcgis/products/index

Esri. (1996, November 4). *Products & Solutions: Data*. https://web.archive.org/web/19961104070544/http://www.esri.com/base/data/index.html

Esri. (1998, December 3). *Data Hound*. https://web.archive.org/web/19981203022742/http://nt1.esri.com/scripts/production/esri/marketing/datahound/main.cfm

Esri. (2000). *An Overview of the Geography Network*. ArcNews Online. https://www.esri.com/news/arcnews/summer00articles/anoverview.html

Esri. (2001, August 15). *ESRI Data for Your GIS*. https://web.archive.org/web/20010815111640/http://www.esri.com/data/index.html

Esri. (2003, February 7). *ArcData Online No Longer Available*. https://web.archive.org/web/20030207073427/http://www.esri.com/data/online/index.html

Esri. (2010a, June 28). *ArcGIS Online | Community Maps Program*. https://web.archive.org/web/20100628164605/http://www.esri.com/software/arcgis/arcgisonline/comm-maps-contribute.html

Esri. (2010b, July 17). *(Archive) Esri History*. Esri. https://web.archive.org/web/20100717005732/http://www.esri.com/about-esri/about/history.html

Esri. (2014, December 26). *ArcGIS Online | Key Features*. https://web.archive.org/web/20141226094123/http://www.esri.com/software/arcgis/arcgisonline/features

Esri (Director). (2018, April 25). *Esri's Open Vision | A Conversation with Jack Dangermond and Satish Sankaran* [Video recording]. https://www.youtube.com/watch?v=NwPesTVNeg8

Esri. (2021, June 24). *Esri Releases New 2020 Global Land Cover Map*. Esri. https://www.esri.com/about/newsroom/announcements/esri-releases-new-2020-global-land-cover-map/

Esri. (2023a, July 13). 2023 Esri User Conference Highlights. *Esri*. https://www.esri.com/about/newsroom/announcements/2023-esri-user-conference-highlights/

Esri. (2023b, August 3). *ArcGIS Online: Best Practices for Content Management, Item Discoverability and Search*. Esri Videos: GIS, Events, ArcGIS Products & Industries. https://mediaspace.esri.com/media/t/1_pgm6twv9

Esri. (2024). The Emergence of 3D Basemaps | Fall 2024 | ArcNews. *Esri*. https://www.esri.com/about/newsroom/arcnews/the-emergence-of-3d-basemaps/

Esri Deutschland. (n.d.). *Esri Deutschland Open Data Portal*. Esri Deutschland Open Data Portal. Retrieved 4 September 2023, from https://opendata-esri-de.opendata.arcgis.com/

Esri Deutschland. (2024). *Esri Konferenz 2024*. https://www.esri.de/de-de/esri-konferenz/2024/uebersicht

Esri España. (n.d.). *Análisis Geográfico de la Siniestralidad en Madrid (2010-2020)*. Esri España. Retrieved 5 October 2023, from https://experience.arcgis.com/experience/10bc2a493a974d5caf024ff13d7891b3/?data_id=dataSource_7-accidentes_por_barrios_2019_2020_600%3A23&views=view_1
Esri España. (2024). *Conferencia Esri España 2024*. https://www.esri.es/es-es/acerca-de/eventos/cesri24/conferencia-esri24
Esri Events (Director). (2016, August 4). *Web GIS: Server and Online: ArcGIS Online: Getting started with Open Data* [Video recording]. https://www.youtube.com/watch?v=6ou15TIw1h4
Esri Nederland. (n.d.). *Esri Nederland Content*. Esri Nederland Content. Retrieved 11 October 2023, from https://esri-nederland-content-esrinl-content.hub.arcgis.com/
Esri Nederland. (2023). *Hoe Topotijdreis de Nederlandse geschiedenis voor iedereen inzichtelijk maakt. 2*. https://magazine.esri.nl/esri-magazine-2-2023/hoe-topotijdreis-de-nederlandse-geschiedenis-voor-iedereen-zichtbaar-maakt
Esri Nederland. (2024). *Esri Connect 2025*. https://www.esri.nl/nl-nl/evenementen/esri-connect/home
Esri UK. (2021). *Climate Crisis*. Teach with GIS UK. https://teach-with-gis-uk-esriukeducation.hub.arcgis.com/pages/climate
Esri UK (Director). (2022, November 24). *Using data to address the climate change (Opening Plenary)—Met Office—AC22* [Video recording]. https://www.youtube.com/watch?v=o6_ecDBBUiA
Esri UK. (2024). *Esri UK Annual Conference 2024*. https://www.esriuk.com/en-gb/about/events/ac/overview
European Commission. (2023, February 15). *Press Release: 4 Member States referred to the Court of Justice of the EU*. European Commission. https://ec.europa.eu/commission/presscorner/detail/en/IP_23_706
Fan, Y., Ding, X., Wu, J., Ge, J., & Li, Y. (2021). High spatial-resolution classification of urban surfaces using a deep learning method. *Building and Environment*, *200*, 107949. https://doi.org/10.1016/j.buildenv.2021.107949
Fang, T. P., Wu, A., & Clough, D. R. (2021). Platform diffusion at temporary gatherings: Social coordination and ecosystem emergence. *Strategic Management Journal*, *42*(2), 233–272. https://doi.org/10.1002/smj.3230
Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991) (U.S. Supreme Court 27 March 1991). https://supreme.justia.com/cases/federal/us/499/340/
Ferro, E., & Pizzamiglio, A. (2023). *New Business Models for Data-Driven Services* [Discussion paper]. Publications Office of the European Union. https://data.europa.eu/sites/default/files/report/New%20Business%20Models%20for%20Data-Driven%20Services_ENG.pdf
Ferster, C., Nelson, T., Manaugh, K., Beairsto, J., Laberee, K., & Winters, M. (2023). Developing a national dataset of bicycle infrastructure for Canada using open data sources. *Environment and Planning B: Urban Analytics and City Science*, 23998083231159905. https://doi.org/10.1177/23998083231159905
Flowerdew, J. (1992). Definitions in Science Lectures. *Applied Linguistics*, *13*(2), 202–221. https://doi.org/10.1093/applin/13.2.202
Forrest, M. (2023, January 23). Map Tiles: Everything You Need To Know. *Carto*. https://carto.com/blog/map-tiles-guide
Frank, M., & Waddell, P. (2014). *Solving the Democratic Deficit: The Role of Open Data and Intermediaries*. Conference for E-Democracy and Open Governement, Krems an der Donau, Austria.
Friedl, S., & Rischard, G. (2018). *Investigation into the Unusual Signups: Prepared by the OSMF Membership Working Group*.
Gao, Y., & Janssen, M. (2022). The Open Data Canvas—Analyzing Value Creation from Open Data. *Digital Government: Research and Practice*, *3*(1), 5:1-5:15. https://doi.org/10.1145/3511102
Garatzogianni, A., Orlandi, F., Fawcett, J., & Chauvet, L. (2017). *ODINE D6.3 Business models, lessons learned and success stories* (Open Data Incubator for Europe). https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5b42e545f&appId=PPGMS
Garber, K., Fox, C., Abdalla, M., Tatem, A., Qirbi, N., Lloyd-Braff, L., Al-Shabi, K., Ongwae, K., Dyson, M., & Hassen, K. (2020). Estimating access to health care in Yemen, a complex humanitarian emergency setting: A descriptive applied geospatial analysis. *The Lancet Global Health*, *8*(11), e1435–e1443. https://doi.org/10.1016/S2214-109X(20)30359-4
Gardner, Z., Mooney, P., De Sabbata, S., & Dowthwaite, L. (2020). Quantifying gendered participation in OpenStreetMap: Responding to theories of female (under) representation in crowdsourced mapping. *GeoJournal*, *85*(6), 1603–1620. https://doi.org/10.1007/s10708-019-10035-z

Gebka, E., Crusoe, J., & Ahlin, K. (2021). Methods to Capture User Information Needs: Design Principles for Open Data Intermediaries and Data Providers. *Data Science Journal*, *20*(1), 37. https://doi.org/10.5334/dsj-2021-037

Geissdoerfer, M., Pieroni, M. P. P., Pigosso, D. C. A., & Soufani, K. (2020). Circular business models: A review. *Journal of Cleaner Production*, *277*, 123741. https://doi.org/10.1016/j.jclepro.2020.123741

Geofabrik. (2024a). *OpenStreetMap History*. Geofabrik. https://www.geofabrik.de/gallery/history/

Geofabrik. (2024b). *Our Download Server*. Geofabrik. https://www.geofabrik.de/data/download.html

Geography Network. (2000a, August 16). *About the Geography Network*. https://web.archive.org/web/20000816123152/http://www.geographynetwork.com/about/index.html

Geography Network. (2000b, August 16). *Geography Network—Maps*. https://web.archive.org/web/20000816123323/http://www.geographynetwork.com/maps/index.html

Geography Network. (2009, December 17). *Geography Network*. https://web.archive.org/web/20091217235022/http://www.geographynetwork.com/

Geoinfo. (n.d.). *Geoinfo Dataleverance | Skab værdi med frie geografiske data*. Retrieved 15 August 2024, from https://www.geoinfo.dk/Vi-tilbyder/L%C3%B8sningskatalog/geoinfo-dataleverance

Geoinfo. (2024). *Dansk Esri Konference 2024*. https://www.geoinfo.dk/Events/dansk-esri-konference-2024/alle-indl%C3%A6g-2024

Geospatial Media and Communications. (2018). *geoBuiz Geospatial Industry Outlook & Readiness Index*. https://geobuiz.com/pdf/geoBuiz-2018-report.pdf

Geraghty, E. (2023, March 11). A Thank You to Johns Hopkins University. *Industry Blogs*. https://www.esri.com/en-us/industries/blog/articles/a-thank-you-to-johns-hopkins-university/

Gerlach, J. (2010). Vernacular mapping, and the ethics of what comes next. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *45*(3), 165–168.

Gerlach, J. (2015). Editing worlds: Participatory mapping and a minor geopolitics. *Transactions of the Institute of British Geographers*, *40*(2), 273–286. https://doi.org/10.1111/tran.12075

Germano, E. C., Souza, C. A. de, & Sun, V. (2016). Business models adopted by intermediaries in the use of open government data. *REBRAE*, *9*(1), Article 1. https://doi.org/10.7213/rebrae.09.001.AO05

Gioia, D. A., Patvardhan, S. D., Hamilton, A. L., & Corley, K. G. (2013). Organizational Identity Formation and Change. *The Academy of Management Annals*, *7*(1), 123–193. https://doi.org/10.1080/19416520.2013.762225

Glassey, O. (2017). Proactive Transparency and Open Data: A Tentative Analysis. In M. Janssen, K. Axelsson, O. Glassey, B. Klievink, R. Krimmer, I. Lindgren, P. Parycek, H. J. Scholl, & D. Trutnev (Eds.), *Electronic Government* (pp. 314–323). Springer International Publishing. https://doi.org/10.1007/978-3-319-64677-0_26

Glasze, G., & Perkins, C. (2015). Social and Political Dimensions of the OpenStreetMap Project: Towards a Critical Geographical Research Agenda. In J. Jokar Arsanjani, A. Zipf, P. Mooney, & M. Helbich (Eds.), *OpenStreetMap in GIScience: Experiences, Research, and Applications* (pp. 143–166). Springer International Publishing. https://doi.org/10.1007/978-3-319-14280-7_8

González-Zapata, F., & Heeks, R. (2015, May 20). *Understanding Multiple Roles of Intermediaries in Open Government Data*. 13th International Conference on Social Implications of Computers in Developing Countries, Negombo, Sri Lanka.

Goodchild, B., & Ferrari, E. (2024). Intermediaries and mediators: An actor-network understanding of online property platforms. *Housing Studies*, *39*(1), 102–123. https://doi.org/10.1080/02673037.2021.2015297

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, *69*(4), 211–221. https://doi.org/10.1007/s10708-007-9111-y

Grab. (2019, July 26). *Join us as a Grab Geo*Star!* Grab. https://www.grab.com/my/blog/join-us-as-a-grab-geostar-2/

Gray, J. (2014, September 3). *Towards a Genealogy of Open Data*. General Conference of the European Consortium for Political Research, Glasgow, UK. https://doi.org/10.2139/ssrn.2605828

Grinberger, A. Y., Minghini, M., Yeboah, G., Juhász, L., & Mooney, P. (2022). Bridges and Barriers: An Exploration of Engagements of the Research Community with the OpenStreetMap Community. *ISPRS International Journal of Geo-Information*, *11*(1), Article 1. https://doi.org/10.3390/ijgi11010054

Grinberger, A. Y., Schott, M., Raifer, M., & Zipf, A. (2021). An analysis of the spatial and temporal distribution of large-scale data production events in OpenStreetMap. *Transactions in GIS*, *25*(2), 622–641. https://doi.org/10.1111/tgis.12746

Gulati, R., Nohria, N., & Zaheer, A. (2000). Strategic networks. *Strategic Management Journal*, *21*(3), 203–215. https://doi.org/10.1002/(SICI)1097-0266(200003)21:3<203::AID-SMJ102>3.0.CO;2-K

Gurin, J. (2014, October 6). *Driving Innovation With Open Data* [Text]. U.S. Chamber of Commerce Foundation. https://www.uschamberfoundation.org/driving-innovation-open-data

Gutierrez, M., & Landa, M. (2021). From Available to Actionable Data: An Exploration of Expert and Re-users Views of Open Data. *Journal of Urban Technology*, *0*(0), 1–28. https://doi.org/10.1080/10630732.2021.1939626

Haarsma, D. (2012, May). The Digital Agenda in the Light of Geoinformation—GIM International Interviews European Commissioner Neelie Kroes. *GIM International*, 15–19.

Hablé, J. (2019). *Value Drivers and Inhibitors in Municipal Open Government Data Ecosystems: An in-depth analysis of stakeholder perceptions on values, barriers and success factors in open government data initiatives in Dutch municipalities* [Master's thesis, Delft University of Technology]. https://repository.tudelft.nl/islandora/object/uuid%3A1347feeb-dce7-44c7-b8b7-3f58f2dd254b

Haklay, M. (2010). How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703. https://doi.org/10.1068/b35097

Haklay, M. (2013). Neogeography and the Delusion of Democratisation. *Environment and Planning A: Economy and Space*, *45*(1), 55–69. https://doi.org/10.1068/a45184

Haklay, M. (muki), Antoniou, Vyron, Basiouka, Sofia, Soden, Robert, & Mooney, Peter. (2014). *Crowdsourced Geographic Information Use in Government*. World Bank. https://doi.org/10.1596/978-1-4648-0484-7

Haklay, M., Singleton, A., & Parker, C. (2008). Web Mapping 2.0: The Neogeography of the GeoWeb. *Geography Compass*, *2*(6), 2011–2039. https://doi.org/10.1111/j.1749-8198.2008.00167.x

Haklay, M., & Weber, P. (2008). OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Computing*, *7*(4), 12–18. IEEE Pervasive Computing. https://doi.org/10.1109/MPRV.2008.80

Halfaker, A., Geiger, R. S., Morgan, J. T., & Riedl, J. (2013). The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist*, *57*(5), 664–688. https://doi.org/10.1177/0002764212469365

Hall, A., McRoberts, S., Thebault-Spieker, J., Lin, Y., Sen, S., Hecht, B., & Terveen, L. (2017). Freedom versus Standardization: Structured Data Generation in a Peer Production Community. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6352–6362. https://doi.org/10.1145/3025453.3025940

Hardy, Q. (2012, March 20). Facing Fees, Some Sites Are Bypassing Google Maps. *The New York Times*. https://query.nytimes.com/gst/fullpage.html?res=9904E4D61E3AF933A15750C0A9649D8B63

Harrison, T. M., Pardo, T. A., & Cook, M. (2012). Creating Open Government Ecosystems: A Research and Development Agenda. *Future Internet*, *4*(4), Article 4. https://doi.org/10.3390/fi4040900

HDX. (2025). *Humanitarian OpenStreetMap Team (HOT) Humanitarian Data*. HDX. https://data.humdata.org/organization/hot?sort=metadata_modified+desc

Hecht, B., & Gergle, D. (2009). Measuring self-focus bias in community-maintained knowledge repositories. *Proceedings of the Fourth International Conference on Communities and Technologies*, 11–20. https://doi.org/10.1145/1556460.1556463

Hedman, J., & Kalling, T. (2003). The business model concept: Theoretical underpinnings and empirical illustrations. *European Journal of Information Systems*, *12*(1), 49–59. https://doi.org/10.1057/palgrave.ejis.3000446

Heidelberg Academy of Humanities and Sciences. (2018). *Nepal Heritage Documentation Project (NHDP)*. https://www.hadw-bw.de/en/research/research-center/nepal-heritage-documentation-project-nhdp

HeiGIT. (2018, April 26). The Ohsome API: Dynamic OSM statistics for real-world applications. *Heidelberg Institute for Geoinformation Technology*. https://heigit.org/the-ohsome-api-dynamic-osm-statistics-for-real-world-applications/

Helms, M. M., & Nixon, J. (2010). Exploring SWOT analysis – where are we now? A review of academic research from the last decade. *Journal of Strategy and Management*, *3*(3), 215–251. https://doi.org/10.1108/17554251011064837

Henriksen, C., Lauzon, J. P., & Morehouse, S. (1994). Open geodata access through standards. *StandardView*, *2*(3), 169–174. https://doi.org/10.1145/202749.202763

Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., & Zipf, A. (2021). The evolution of humanitarian mapping within the OpenStreetMap community. *Scientific Reports*, *11*(1), 3037. https://doi.org/10.1038/s41598-021-82404-z

Herrera-Murillo, D. J., Ochoa-Ortiz, H., Ahmed, U., López-Pellicer, F. J., Re, B., Polini, A., & Nogueras-Iso, J. (2024). Process Analysis in Humanitarian Voluntary Geographic Information: The case of the HOT Tasking Manager. *AGILE: GIScience Series*, *5*, 1–12. https://doi.org/10.5194/agile-giss-5-5-2024

Hielkema, H., & Hongisto, P. (2013). Developing the Helsinki Smart City: The Role of Competitions for Open Data Applications. *Journal of the Knowledge Economy*, *4*(2), 190–204. https://doi.org/10.1007/s13132-012-0087-6

Hite, R., Solís, P., Wargo, L., & Larsen, T. B. (2018). Exploring Affective Dimensions of Authentic Geographic Education Using a Qualitative Document Analysis of Students' YouthMappers Blogs. *Education Sciences*, *8*(4), Article 4. https://doi.org/10.3390/educsci8040173

Hjorth, A. (2021, April 2). Amazon Market Watch: The Top 10 Amazon Markets [2022 update]. *Innovell - Digital Marketing Insights*. https://www.innovell.com/amazon-market-watch-the-top-10-amazon-markets/

Ho, S., & Rajabifard, A. (2010). Learning from the crowd: The role of volunteered geographic information in realising a spatially enabled society. *Proceedings of GSDI*, *12*.

Hoffman, A. (Director). (2021, June 10). *ESRI: The Global Leader in GIS and Location Intelligence with over $1 billion in Revenue* [Video recording]. https://www.youtube.com/watch?v=QL6BgXUcb_0

Holovin, A. (2022a, December 24). *What does Overture Map mean for the future of OpenStreetMap?* Andrii Holovin – Blog. https://blog.andygol.co.ua/en/2022/12/24/creating-overture-map-and-future-of-osm/

Holovin, A. (2022b, December 27). What does Overture Map mean for the future of OpenStreetMap? (Part 2). *Andrii Holovin – Blog*. https://blog.andygol.co.ua/en/2022/12/27/creating-overture-map-and-future-of-osm-2/

Holzmann, P., Breitenecker, R. J., & Schwarz, E. J. (2019). Business model patterns for 3D printer manufacturers. *Journal of Manufacturing Technology Management*, *31*(6), 1281–1300. https://doi.org/10.1108/JMTM-09-2018-0313

Hormann, C. (2024, February 2). Representation of mappers in OSMF membership | Imagico.de. *IMAGICO.DE*. https://imagico.de/blog/en/representation-of-mappers-in-osmf-membership/

Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. *Journal of Organizational Computing and Electronic Commerce*, *26*(1–2), 14–40. https://doi.org/10.1080/10919392.2015.1124007

Hossain, M. A., Rahman, S., Quaddus, M., Hooi, E., & Olanrewaju, A.-S. (2021). Factors Affecting Performance of Open Government Data Initiatives: A Multi-Method Approach Using Sem and FSQCA. *Journal of Organizational Computing and Electronic Commerce*, *31*(4), 300–319. https://doi.org/10.1080/10919392.2021.2018258

HOT. (2012). *Form 990: Return of Organization Exempt from Income Tax 2012*. Github. https://github.com/hotosm/hotosm-website/tree/gh-pages/downloads

HOT. (2018a, February 6). *fAIr*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/tech-suite/fair/

HOT. (2018b, February 6). *HOT Tech Product Suite*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/tech-suite.html

HOT. (2018c, February 6). *What We Do*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/what-we-do.html

HOT. (2019). *Starting with a Tasking Manager—iD editor*. LearnOSM. https://learnosm.org/en/hot-tips/getting-started/

HOT. (2020, June 18). *The Audacious Project*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/audacious

HOT. (2023a, June 30). *She Leads and She Inspires*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/projects/she-leads-and-she-inspires/

HOT. (2023b, December 26). *Hub LAC Calls for Participation in the Humanitarian Mapping Brigade*. Humanitarian OpenStreetMap Team. https://hotosm-staging-new.fly.dev/en/news/hub-lac-calls-for-participation-in-the-humanitarian-mapping-brigade/

HOT. (2024a). *User Guide: How to use?* OpenAerialMap Docs. https://docs.openaerialmap.org/browser/user-guide/

HOT. (2024b, July 10). *Humanitarian Mapping Brigade*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/projects/humanitarian-mapping-brigade/

HOT. (2024c, November 14). *Morocco Earthquake: OpenStreetMap data in support of response efforts*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/projects/morocco-earthquake-open-streetmap-data-in-support-of-response-efforts/

HOT. (2025a). *Annual Reports*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/annual-reports/

HOT. (2025b). *Argentina Floods 2025 (4)*. HOT Tasking Manager. https://tasks.hotosm.org/projects/18522

Howe, J. (2006, June 1). The Rise of Crowdsourcing. *Wired*. https://www.wired.com/2006/06/crowds/

Hristova, D., Quattrone, G., Mashhadi, A., & Capra, L. (2013). The Life of the Party: Impact of Social Mapping in OpenStreetMap. *Proceedings of the International AAAI Conference on Web and Social Media*, *7*(1), Article 1. https://doi.org/10.1609/icwsm.v7i1.14416

Hu, S., Zhao, R., Cui, Y., Zhang, D., & Ge, Y. (2023). Identifying the uneven distribution of health and education services in China using open geospatial data. *Geography and Sustainability*, *4*(2), 91–99. https://doi.org/10.1016/j.geosus.2023.01.002

Huang, X., Li, X., Yang, D., & Zou, L. (2023). Chapter 6 - Crowdsourced geospatial data in human and Earth observations: Opportunities and challenges. In N. Stathopoulos, A. Tsatsaris, & K. Kalogeropoulos (Eds.), *Geoinformatics for Geosciences* (pp. 109–129). Elsevier. https://doi.org/10.1016/B978-0-323-98983-1.00007-7

Huber, F., Ponce ,Alan, Rentocchini ,Francesco, & and Wainwright, T. (2022). The wealth of (Open Data) nations? Open government data, country-level institutions and entrepreneurial activity. *Industry and Innovation*, *29*(8), 992–1023. https://doi.org/10.1080/13662716.2022.2109455

Huijboom, N., & Van den Broek, T. (2011). Open data: An international comparison of strategies. *European Journal of ePractice*, *12*(1), 4–16.

Hunt, A. (2017, December 6). Mapping Arsenic Contamination—Youth Mappers at AUW Bangladesh. *Humanitarian OpenStreetMap Team*. https://www.hotosm.org/updates/2017-12-06_mapping_arsenic_contamination_-_youth_mappers_at_auw_bangladesh

Hurley, P. J. (2014). *A concise introduction to logic* (12th ed.). Cengage Learning.

Ilisei, A. (2019, December 12). *OpenStreetCam and ImproveOSM are moving to Grab – OpenStreetMap @ Grab*. OpenStreetMap @ Grab. https://web.archive.org/web/20210124030406/http://blog.improveosm.org/en/2019/12/openstreetcam-and-improveosm-are-moving-to-grab/

IODC. (n.d.). Principles. *International Open Data Charter*. https://opendatacharter.net/principles/

Ivánová, I., Siao Him Fa, J., McMeekin, D. A., Arnold, L. M., Deakin, R., & Wilson, M. (2020). From spatial data to spatial knowledge infrastructure: A proposed architecture. *Transactions in GIS*, *24*(6), 1526–1558. https://doi.org/10.1111/tgis.12656

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, *29*(4), 258–268. https://doi.org/10.1080/10580530.2012.716740

Janssen, M., & Zuiderwijk, A. (2014). Infomediary Business Models for Connecting Open Data Providers and Users. *Social Science Computer Review*, *32*(5), 694–711. https://doi.org/10.1177/0894439314525902

Jarvenpaa, S. L., & Essén, A. (2023). Data sustainability: Data governance in data infrastructures across technological and human generations. *Information and Organization*, *33*(1), 100449. https://doi.org/10.1016/j.infoandorg.2023.100449

Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2019). The Sustainable Value of Open Government Data. *Journal of the Association for Information Systems*, *20*(6). https://doi.org/10.17705/1jais.00549

Johannesson, P. (2007). The Role of Business Models in Enterprise Modelling. In J. Krogstie, A. L. Opdahl, & S. Brinkkemper (Eds.), *Conceptual Modelling in Information Systems Engineering* (pp. 123–140). Springer. https://doi.org/10.1007/978-3-540-72677-7_8

Johnson, P. A., & Greene, S. (2017). Who Are Government OpenData Infomediaries? A Preliminary Scan and Classification of Open Data Users and Products. *Journal of the Urban and Regional Information Systems Association*, *28*, 10.

Johnson, P. A., & Scassa, T. (2023). Who owns the map? Data sovereignty and government spatial data collection, use, and dissemination. *Transactions in GIS*, *27*(1), 275–289. https://doi.org/10.1111/tgis.13024

Johnson, P. A., Sieber, R., Scassa, T., Stephens, M., & Robinson, P. (2017). The Cost(s) of Geospatial Open Data. *Transactions in GIS*, *21*(3), 434–445. https://doi.org/10.1111/tgis.12283

Johnson, P. A., & Varga, C. (2022). Challenges to the Access of Government Open Data by Private Sector Companies. In P. Robinson & T. Scassa (Eds.), *The Future of Open Data* (pp. 103–119). University of Ottawa Press. https://ruor.uottawa.ca/items/be42caf9-1d66-495c-b792-2ebaba2fb534

Juell-Skielse, G., Hjalmarsson, A., Juell-Skielse, E., Johannesson, P., & Rudmark, D. (2014). Contests as innovation intermediaries in open data markets. *Information Polity*, *19*(3–4), 247–262. https://doi.org/10.3233/IP-140346

Juhász, L., & Hochmair, H. H. (2018). OSM Data Import as an Outreach Tool to Trigger Community Growth? A Case Study in Miami. *ISPRS International Journal of Geo-Information*, *7*(3), Article 3. https://doi.org/10.3390/ijgi7030113

Kalt, G., Thunshirn, P., & Haberl, H. (2021). A global inventory of electricity infrastructures from 1980 to 2017: Country-level data on power plants, grids and transformers. *Data in Brief*, *38*, 107351. https://doi.org/10.1016/j.dib.2021.107351

Kamariotou, M., & Kitsios, F. (2022). Bringing Digital Innovation Strategies and Entrepreneurship: The Business Model Canvas in Open Data Ecosystem and Startups. *Future Internet*, *14*(5), Article 5. https://doi.org/10.3390/fi14050127

KartaView. (2025, March 12). *Grabsg*. KartaView. https://kartaview.org/user/grabsg

Kassen, M. (2018). Open data and its intermediaries: A cross-country perspective on participatory movement among independent developers. *Knowledge Management Research & Practice*, *16*(3), 327–342. https://doi.org/10.1080/14778238.2018.1471377

Kedron, P., Li, W., Fotheringham, S., & Goodchild, M. (2021). Reproducibility and replicability: Opportunities and challenges for geospatial research. *International Journal of Geographical Information Science*, *35*(3), 427–445. https://doi.org/10.1080/13658816.2020.1802032

Kensok, D. (2020a, April 1). Live OpenStreetMap data in ArcGIS. *ArcGIS Blog*. https://www.esri.com/arcgis-blog/products/arcgis-living-atlas/mapping/live-openstreetmap-data-in-arcgis/

Kensok, D. (2020b, June 30). Community Maps Data Sharing. *ArcGIS Blog*. https://www.esri.com/arcgis-blog/products/arcgis-living-atlas/mapping/community-maps-data-sharing/

Kesting, P., & Günzel-Jensen, F. (2015). SMEs and new ventures need business model sophistication. *Business Horizons*, *58*(3), 285–293. https://doi.org/10.1016/j.bushor.2015.01.002

KH Digital. (2016, August 10). *Controversy persists in Korea over restrictions on Google Maps*. The Korea Herald. https://www.koreaherald.com/article/1049236

Khanal, K., Budhathoki, N. R., & Erbstein, N. (2019). Filling OpenStreetMap data gaps in rural Nepal: A digital youth internship and leadership Programme. *Open Geospatial Data, Software and Standards*, *4*(1), 12. https://doi.org/10.1186/s40965-019-0071-1

Khazi-Syed, A., Pecherskiy, M., & Krambeck, H. (2023). Analyzing Access to Health Facilities by Road Using Unconventional Data Sources. *Transportation Research Record*, *2677*(1), 169–177. https://doi.org/10.1177/03611981221098693

Kim, D. (2018). *Data + disasters: Rethinking the role of Metro Boston's data intermediaries in disaster management* [Master's thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/117828

Kim, I.-K., & Seo, J.-E. (2023, March 6). Apple request for detailed Korean mapping data rejected. *Korea JoongAng Daily*. https://koreajoongangdaily.joins.com/2023/03/06/business/tech/Korea-Apple-map/20230306174727032.html

Kitchin, R., Collins, S., & Frost, D. (2015). Funding models for Open Access digital data repositories. *Online Information Review*, *39*(5), 664–681. https://doi.org/10.1108/OIR-01-2015-0031

Kitchin, R., & Dodge, M. (2007). Rethinking maps. *Progress in Human Geography*, *31*(3), 331–344. https://doi.org/10.1177/0309132507077082

Klapper, C., Kleve, L., O'Hearn, C., & Dorger, R. (2020, November 17). *A brief introduction of HOT as well as their history with mapathons*. Humanitarian OpenStreetmap Team and Mapathons. https://storymaps.arcgis.com/stories/c4936c3ecbab4a6cab84338fa8f685e4

Klecuń, E. (2004). Conducting Critical Research in Information Systems: Can Actor-Network Theory Help? In B. Kaplan, D. P. Truex, D. Wastell, A. T. Wood-Harper, & J. I. DeGross (Eds.), *Information Systems Research: Relevant Theory and Informed Practice* (pp. 259–274). Springer US. https://doi.org/10.1007/1-4020-8095-6_15

Klinkhardt, C., Kühnel, F., Heilig, M., Lautenbach, S., Wörle, T., Vortisch, P., & Kuhnimhof, T. (2023). Quality Assessment of OpenStreetMap's Points of Interest with Large-Scale Real Data. *Transportation Research Record*, 03611981231169280. https://doi.org/10.1177/03611981231169280

KLL. (2023). *Kathmandu Living Labs*. Kathmandu Living Labs. https://kathmandulivinglabs.org/

Kohtamäki, M., Parida, V., Oghazi, P., Gebauer, H., & Baines, T. (2019). Digital servitization business models in ecosystems: A theory of the firm. *Journal of Business Research*, *104*, 380–392. https://doi.org/10.1016/j.jbusres.2019.06.027

Koski, H. (2011). *Does marginal cost pricing of public sector information spur firm growth?* (Working Paper No. 1260). ETLA Discussion Papers. https://www.econstor.eu/handle/10419/87764

Krivý, M. (2023). Digital ecosystem: The journey of a metaphor. *Digital Geography and Society*, *5*, 100057. https://doi.org/10.1016/j.diggeo.2023.100057

Lafia, S., Turner, A., & Kuhn, W. (2018). *Improving Discovery of Open Civic Data*. https://doi.org/10.4230/LIPIcs.GISCIENCE.2018.9

Lambert, S. C., & Davidson, R. A. (2013). Applications of the business model in studies of enterprise success, innovation and classification: An analysis of empirical research from 1996 to 2010. *European Management Journal*, *31*(6), 668–681. https://doi.org/10.1016/j.emj.2012.07.007

Lardinois, F. (2012, June 22). Google Maps API Gets Massive Price Cut In The Wake Of Developer Defections. *TechCrunch*. https://techcrunch.com/2012/06/22/google-maps-api-gets-massive-price-cut-in-the-wake-of-developer-defections/

Latour, B. (2007). *Reassembling the social: An introduction to actor-network-theory*. Oup Oxford.

Lavie, D. (2006). The Competitive Advantage of Interconnected Firms: An Extension of the Resource-Based View. *Academy of Management Review*, *31*(3), 638–658. https://doi.org/10.5465/amr.2006.21318922

Law, J. (2008). Actor Network Theory and Material Semiotics. In B. S. Turner (Ed.), *The new Blackwell companion to social theory* (pp. 141–158). Wiley-Blackwell.

Lee, N., & Hassard, J. (1999). Organization Unbound: Actor-Network Theory, Research Strategy and Institutional Flexibility. *Organization*, *6*(3), 391–404. https://doi.org/10.1177/135050849963002

LeMay, S., Helms, M. M., Kimball, B., & McMahon, D. (2017). Supply chain management: The elusive concept and definition. *The International Journal of Logistics Management*, *28*(4), 1425–1453. https://doi.org/10.1108/IJLM-10-2016-0232

Lenormand, M. (Director). (2022, October 1). *Steve Coast: Discussing the Future of Maps & Tech with the Founder of OpenStreetMap (& why he thinks maps will disappear)* [Video recording]. https://youtu.be/uB9NxbnE1i0?si=96Nf1dxc84nYRp7P

Lenormand, M. (Director). (2023, May 16). *The Overture Maps Foundation: Marc Prioleau - MBM#43* [Video recording]. https://www.youtube.com/watch?v=OSK4DlFePzk

Leppänen, P., George, G., & Alexy, O. (2023). When Do Novel Business Models Lead to High Performance? A Configurational Approach to Value Drivers, Competitive Strategy, and Firm Environment. *Academy of Management Journal*, *66*(1), 164–194. https://doi.org/10.5465/amj.2020.0969

Lin, W. (2015). Revealing the making of OpenStreetMap: A limited account. *Canadian Geographies / Géographies Canadiennes*, *59*(1), 69–81. https://doi.org/10.1111/cag.12137

Lin, W. (2018). Volunteered Geographic Information constructions in a contested terrain: A case of OpenStreetMap in China. *Geoforum*, *89*, 73–82. https://doi.org/10.1016/j.geoforum.2018.01.005

Lin, Y. (2011). A qualitative enquiry into OpenStreetMap making. *New Review of Hypermedia and Multimedia*, *17*(1), 53–71. https://doi.org/10.1080/13614568.2011.552647

Linders, D. (2013). Towards open development: Leveraging open data to improve the planning and coordination of international aid. *Government Information Quarterly*, *30*(4), 426–434. https://doi.org/10.1016/j.giq.2013.04.001

Lindman, J., Kinnari, T., & Rossi, M. (2014). Industrial Open Data: Case Studies of Early Open Data Entrepreneurs. *2014 47th Hawaii International Conference on System Sciences*, 739–748. https://doi.org/10.1109/HICSS.2014.99

Lindman, J., Kinnari, T., & Rossi, M. (2016). Business Roles in the Emerging Open-Data Ecosystem. *IEEE Software*, *33*(5), 54–59. IEEE Software. https://doi.org/10.1109/MS.2015.25

Linneberg, M. S., & Korsgaard, S. (2019). Coding qualitative data: A synthesis guiding the novice. *Qualitative Research Journal*, *19*(3), 259–270.

Linux Foundation. (2022, December 15). *Linux Foundation Newsletter: December 2022*. The Linux Foundation Blog. https://www.linuxfoundation.org/blog/linux-foundation-newsletter-decemberr-2022

Lovelace, R. (2021). Open source tools for geographic analysis in transport planning. *Journal of Geographical Systems*, *23*(4), 547–578. https://doi.org/10.1007/s10109-020-00342-2

Lucivero, F. (2020). Big Data, Big Waste? A Reflection on the Environmental Sustainability of Big Data Initiatives. *Science and Engineering Ethics*, *26*(2), 1009–1030. https://doi.org/10.1007/s11948-019-00171-7

Lüdeke-Freund, F., Gold, S., & Bocken, N. M. P. (2019). A Review and Typology of Circular Economy Business Model Patterns. *Journal of Industrial Ecology*, *23*(1), 36–61. https://doi.org/10.1111/jiec.12763

Ma, R., Chen, J., Yang, C., & Li, X. (2024). OSMsc: A framework for semantic 3D city modeling using OpenStreetMap. *International Journal of Geographical Information Science*, *38*(1), 1–26. https://doi.org/10.1080/13658816.2023.2266824

Ma, R., & Lam, P. T. I. (2019). Investigating the barriers faced by stakeholders in open data development: A study on Hong Kong as a "smart city". *Cities*, *92*, 36–46. https://doi.org/10.1016/j.cities.2019.03.009

Maail, A. G. (2017). The relational impact of open data intermediation: Experience from Indonesia and the Philippines. In F. van Schalkwyk, S. Verhulst, G. Magalhaes, J. Pane, & J. Walker (Eds.), *The Social Dynamics of Open Data* (pp. 153–166). African Minds. http://www.africanminds.co.za/the-social-dynamics-of-open-data/

MacLennan, G. (2016, July 22). Technology Preview: Participatory mapping in the Amazon with Mapeo. *Digital Democracy*. https://www.digital-democracy.org/blog/technology-preview-participatory-mapping-in-the-amazon-with-mapeo

Madubedube, A., Coetzee, S., & Rautenbach, V. (2021). A Contributor-Focused Intrinsic Quality Assessment of OpenStreetMap in Mozambique Using Unsupervised Machine Learning. *ISPRS International Journal of Geo-Information*, *10*(3), Article 3. https://doi.org/10.3390/ijgi10030156

Magalhaes, G., & Roseira, C. (2020). Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly*, *37*(3), 1–10. https://doi.org/10.1016/j.giq.2017.08.004

Magalhaes, G., Roseira, C., & Manley, L. (2014). Business models for open government data. *Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance*, 365–370. https://doi.org/10.1145/2691195.2691273

Magnusson, E., & Marecek, J. (2015). *Doing interview-based qualitative research: A learner's guide*. Cambridge University Press.

Magretta, J. (2002, May 1). Why Business Models Matter. *Harvard Business Review*. https://hbr.org/2002/05/why-business-models-matter

Maguire, M. (2009). The Nonprofit Business Model: Empirical Evidence From the Magazine Industry. *Journal of Media Economics*, *22*(3), 119–133. https://doi.org/10.1080/08997760903129333

Mahmoudi, D., Thatcher, J., Imaoka, L. B., & O'Sullivan, D. (2024). From FOSS to profit: Digital spatial technologies and the mode of production. *Digital Geography and Society*, *7*, 100101. https://doi.org/10.1016/j.diggeo.2024.100101

Mahmud, Z., Chauhan, A., Sarkar, D., & Soden, R. (2022). Revisiting Engagement in Humanitarian Mapping: An Updated Analysis of Contributor Retention in OpenStreetMap. *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3491101.3519728

Mandourah, A., & Hochmair, H. H. (2024). Analysing the use of OpenAerialMap images for OpenStreetMap edits. *Geo-Spatial Information Science*, *0*(0), 1–16. https://doi.org/10.1080/10095020.2024.2341747

Map Kibera. (n.d.). *Map Kibera*. Retrieved 11 March 2025, from https://www.mapkibera.org/

Mapbox. (2014, August 22). Over 1 million New York City buildings and addresses imported to OpenStreetMap [Medium]. *Mapbox - Maps for Developers*. https://blog.mapbox.com/over-1-million-new-york-city-buildings-and-addresses-imported-to-openstreetmap-924ef0d7e82d

Mapillary. (2025). *Images and Data for OpenStreetMap*. Mapillary. https://www.mapillary.com/osm

MapSwipe. (2025). *MapSwipe*. https://mapswipe.org/en/

Marcel, W. (2024, June 4). OSMCha: A Decade of Validating OpenStreetMap. *Developement Seed*. https://developmentseed.org/blog/2024-06-04-osmcha/

Mariotto, F. L., Zanni, P., & Moraes, G. (2014). *What is the use of a single-case study in management research?* https://doi.org/10.1590/S0034-759020140402

Masser, I., Rajabifard, A., & Williamson, I. (2008). Spatially enabling governments through SDI implementation. *International Journal of Geographical Information Science*, *22*(1), 5–20. https://doi.org/10.1080/13658810601177751

Masters, P., & de los Reyes, C. (2023, March 9). Turkey And Syria Response: A Month After The Earthquake. *Humanitarian OpenStreetMap Team*. https://www.hotosm.org/updates/turkey-and-syria-response-a-month-after-the-earthquake/

Mayer, M., Heck, D. W., & Mocnik, F.-B. (2020, January 27). *Shared Mental Models as a Psychological Explanation for Converging Mental Representations of Place – the Example of OpenStreetMap*. Second International Symposium on Platial Information Science (PLATIAL'19), Coventry, UK. https://doi.org/10.5281/zenodo.3628871

Mayer-Schönberger, V., & Zappia, Z. (2011). *Participation and power: Intermediaries of open data*. 1st Berlin Symposium on Internet and Society, Berlin, Germany. https://www.hiig.de/wp-content/uploads/2017/08/participation_and_power.pdf

McLaughlin, J., & Nichols, S. (1994). Developing a National Spatial Data Infrastructure. *Journal of Surveying Engineering*, *120*(2), 62–76. https://doi.org/10.1061/(ASCE)0733-9453(1994)120:2(62)

Medjroubi, W., Müller, U. P., Scharf, M., Matke, C., & Kleinhans, D. (2017). Open Data in Power Grid Modelling: New Approaches Towards Transparent Grid Models. *Energy Reports*, *3*, 14–21. https://doi.org/10.1016/j.egyr.2016.12.001

Meijer, A., de Hoog, J., van Twist, M., van der Steen, M., & Scherpenisse, J. (2014). Understanding the Dynamics of Open Data: From Sweeping Statements to Complex Contextual Interactions. In M. Gascó-Hernández (Ed.), *Open Government: Opportunities and Challenges for Public Governance* (pp. 101–114). Springer. https://doi.org/10.1007/978-1-4614-9563-5_7

Meijer, A., & Potjer, S. (2018). Citizen-generated open data: An explorative analysis of 25 cases. *Government Information Quarterly*, *35*(4), 613–621. https://doi.org/10.1016/j.giq.2018.10.004

Meng, A. (2016). *The social impact of open government data* [PhD thesis, Georgia Institute of Technology]. https://smartech.gatech.edu/handle/1853/56292

Meng, A., DiSalvo, C., Tsui, L., & Best, M. (2019). The social impact of open government data in Hong Kong: Umbrella Movement protests and adversarial politics. *The Information Society*, *35*(4), 216–228. https://doi.org/10.1080/01972243.2019.1613464

Micheli, M., Farrell, E., Carballa-Smichowski, B., Posada-Sánchez, M., Signorelli, S., & Vespe, M. (2023). *Mapping the landscape of data intermediaries—Emerging models for more inclusive data governance* (JRC Science for Policy Report). Publications Office of the European Union. https://dx.doi.org/10.2760/261724

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society*, *7*(2), 2053951720948087. https://doi.org/10.1177/2053951720948087

Miller, H. M., Richardson, L., Koontz, S. R., Loomis, J., & Koontz, L. (2013). *Users, Uses, and Value of Landsat Satellite Imagery—Results from the 2012 Survey of Users* (Nos. 2013–1269). USGS. http://dx.doi.org/10.3133/ofr20131269.

Milner, G. (2020, July 16). COVID-19: Inside Look at the Johns Hopkins Dashboard, Keeping Tabs on the Virus. *Esri Blog*. https://www.esri.com/about/newsroom/blog/how-researchers-built-johns-hopkins-dashboard/

Minghini, M., Cetl, V., Kotsev, A., Tomas, R., & Lutz, M. (2021). INSPIRE: The Entry Point to Europe's Big Geospatial Data Infrastructure. In M. Werner & Y.-Y. Chiang (Eds.), *Handbook of Big Geospatial Data* (pp. 619–641). Springer International Publishing. https://doi.org/10.1007/978-3-030-55462-0_24

Minghini, M., Sarretta, A., & Napolitano, M. (2022). OpenStreetMap Contribution to Local Data Ecosystems in COVID-19 Times: Experiences and Reflections from the Italian Case. *Data*, *7*(4), Article 4. https://doi.org/10.3390/data7040039

Missing Maps. (n.d.). *About*. Missing Maps. Retrieved 24 March 2025, from https://www.missingmaps.org/about/

Mocnik, F.-B., Zipf, A., & Raifer, M. (2017). The OpenStreetMap folksonomy and its evolution. *Geo-Spatial Information Science*, *20*(3), 219–230. https://doi.org/10.1080/10095020.2017.1368193

Moloney, A. (2021, July 12). Peru's indigenous tribes use tech to track Amazon deforestation. *Thomson Reuters Foundation News*. https://news.trust.org/item/20210712174822-u31bt/

Mooney, P., & Corcoran, P. (2014). Analysis of Interaction and Co-editing Patterns amongst OpenStreetMap Contributors. *Transactions in GIS*, *18*(5), 633–659. https://doi.org/10.1111/tgis.12051

Mooney, P., Grinberger, A. Y., Minghini, M., Coetzee, S., Juhasz, L., & Yeboah, G. (2021). OpenStreetMap data use cases during the early months of the COVID-19 pandemic. In A. Rajabifard, D. Paez, & G. Foliente (Eds.), *COVID-19 pandemic, geospatial information, and community resilience* (pp. 171–186). CRC Press.

Moradi, M., Roche, S., & Mostafavi, M. A. (2021). Exploring five indicators for the quality of OpenStreetMap road networks: A case study of Québec, Canada. *Geomatica*, *75*(4), 178–208. https://doi.org/10.1139/geomat-2021-0012

Mühlhofer, E., Kropf, C. M., Riedel, L., Bresch, D. N., & Koks, E. E. (2024). OpenStreetMap for multi-faceted climate risk assessments. *Environmental Research Communications*, *6*(1), 015005. https://doi.org/10.1088/2515-7620/ad15ab

Müller, M. (2015). Assemblages and Actor-networks: Rethinking Socio-material Power, Politics and Space. *Geography Compass*, *9*(1), 27–41. https://doi.org/10.1111/gec3.12192

Nagaraj, A. (2021). Information Seeding and Knowledge Production in Online Communities: Evidence from OpenStreetMap. *Management Science*, *67*(8), 4908–4934. https://doi.org/10.1287/mnsc.2020.3764

Nagaraj, A., & Piezunka, H. (2024). The Divergent Effect of Competition on Platforms: Deterring Recruits, Motivating Converts. *Strategy Science*, *9*(3), 277–296. https://doi.org/10.1287/stsc.2022.0125

Najjar, O. (2024, June 4). *fAIr in Production is Available for Everyone*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/updates/fair-in-production-is-available-for-everyone/

Navalkha, C. (2021). *Data for Housing Justice: Examining Activists' Use of Open Government Data for Housing Justice in Boston, MA and New York, NY* [Master's thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/138973

Nehemia-Maletzky, M., Iyamu, T., & Shaanika, I. (2018). The use of activity theory and actor network theory as lenses to underpin information systems studies. *Journal of Systems and Information Technology*, *20*(2), 191–206. https://doi.org/10.1108/JSIT-10-2017-0098

Neis, P. (2025a). *Edits per country for Mar, 27th 2025*. OSMstats - Statistics of the Free Wiki World Map. https://osmstats.neis-one.org/?item=countries&date=28-3-2025

Neis, P. (2025b, March 9). *OSMstats—Statistics of the free wiki world map*. OSMstats. https://osmstats.neis-one.org/

Neis, P., & Zielstra, D. (2014). Recent Developments and Future Trends in Volunteered Geographic Information Research: The Case of OpenStreetMap. *Future Internet*, *6*(1), Article 1. https://doi.org/10.3390/fi6010076

Neves, F. T., de Castro Neto, M., & Aparicio, M. (2020). The impacts of open data initiatives on smart cities: A framework for evaluation and monitoring. *Cities*, *106*, 1–15. https://doi.org/10.1016/j.cities.2020.102860

Nielsen, J. (2006, October 8). *Participation Inequality: The 90-9-1 Rule for Social Features*. Nielsen Norman Group. https://www.nngroup.com/articles/participation-inequality/

Nikiforova, A., & Zuiderwijk, A. (2022). Barriers to Openly Sharing Government Data: Towards an Open Data-adapted Innovation Resistance Theory. *15th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2022), October 04–07, 2022, Guimarães, Portugal*, 215–220. https://doi.org/10.1145/3560107.3560143

OAM. (n.d.). *About OpenAerialMap*. OpenAerialMap. Retrieved 23 March 2025, from http://openaerialmap.org/

Obama. (2013, May 9). *Executive Order—Making Open and Machine Readable the New Default for Government Information*. Whitehouse.Gov. https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-

Ochoa-Ortiz, H., & Re, B. (2025). Exploring why and how commercial organizations contribute to OpenStreetMap. *Annals of GIS*, 1–12. https://doi.org/10.1080/19475683.2025.2457396

OGC. (n.d.). *Home*. Open Geospatial Consortium. Retrieved 16 August 2024, from https://www.ogc.org/

OHM. (n.d.). *Copyright and License*. OpenHistoricalMap. Retrieved 23 March 2025, from https://www.openhistoricalmap.org/copyright

Okamoto, K. (2016). What is being done with open government data? An exploratory analysis of public uses of New York City open data. *Webology*, *Volume 13*(No 1), 1–12. http://www.webology.org/abstract.php?id=236

OKF. (n.d.-a). *Open Data Commons Open Database License (ODbL)*. Open Data Commons: Legal Tools for Open Data. Retrieved 7 March 2025, from https://opendatacommons.org/licenses/odbl/

OKF. (2017, January 16). *Maintenance and future of the Open Data Commons licences—Open Definition*. Open Knowledge Forums. https://discuss.okfn.org/t/maintenance-and-future-of-the-open-data-commons-licences/4460

OKF, O. K. F. (n.d.-b). *The Open Definition—Open Definition—Defining Open in Open Data, Open Content and Open Knowledge*. Retrieved 10 August 2022, from http://opendefinition.org/

Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1954824

Oliveira, M. I. S., & Lóscio, B. F. (2018). What is a data ecosystem? *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, 1–9. https://doi.org/10.1145/3209281.3209335

Open GIS Consortium. (1996, November 12). *Members Directory by Level*. Open GIS. https://web.archive.org/web/19961112220459/http://ogis.org/levelmem.html

OSGeo. (n.d.). *About OSGeo*. OSGeo. Retrieved 18 August 2024, from https://www.osgeo.org/about/

OSGeo. (2013, May). *Geoservices REST API - OSGeo*. https://wiki.osgeo.org/wiki/Geoservices_REST_API

OSM Belgium. (2024). *Free Belgium Baselayers*. OpenStreetMap Belgium. https://openstreetmap.be/en/projects/belgium-baselayer.html

OSM Blog. (n.d.). *About*. Retrieved 13 March 2025, from https://blog.openstreetmap.org/about/

OSM Blog. (2016, October 27). Introducing the new Corporate Membership Program. *OpenStreetMap Blog*. https://blog.openstreetmap.org/2016/10/27/introducing-the-new-corporate-membership-program/

OSM Community Forum. (2018, July 9). *Uber planning to explore and contribute to OpenStreetMap in Delhi—Communities / India* [Online post]. OpenStreetMap Community Forum. https://community.openstreetmap.org/t/uber-planning-to-explore-and-contribute-to-openstreetmap-in-delhi/83829

OSM Community Forum. (2024, May 10). *Use OSM in App in Korea—Communities / ☒☒/☒☒ (Korea)* [Online post]. OpenStreetMap Community Forum. https://community.openstreetmap.org/t/use-osm-in-app-in-korea/113007

OSM Community Forum. (2025). *FAQ*. OpenStreetMap Community Forum. https://community.openstreetmap.org/faq

OSM Foundation. (n.d.-a). *What is OpenStreetMap?* Welcome Mat. Retrieved 6 March 2025, from https://welcome.openstreetmap.org/what-is-openstreetmap/

OSM Foundation. (n.d.-b). *What is the history of OSM?* Welcome Mat. Retrieved 13 March 2025, from https://welcome.openstreetmap.org/about-osm-community/history-of-osm/

OSM Foundation. (2017, October 18). *Data Working Group/Results of Organised Editing Survey 2017*. https://osmfoundation.org/wiki/Data_Working_Group/Results_of_Organised_Editing_Survey_2017

OSM Foundation. (2020a). *Diversity Statement*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Diversity_Statement

OSM Foundation. (2020b). *Organised Editing Guidelines*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Organised_Editing_Guidelines

OSM Foundation. (2021a). *Board/Minutes/2021-02-S2S/Takeover protection*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Board/Minutes/2021-02-S2S/Takeover_protection#Mechanisms_of_takeover

OSM Foundation. (2021b). *Special Committee on Takeover Protection*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Special_Committee_on_Takeover_Protection

OSM Foundation. (2022a). *Licence/Contributor Terms*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Licence/Contributor_Terms

OSM Foundation. (2022b, March 21). A new discussion forum for OSM. *OpenStreetMap Blog*. https://blog.openstreetmap.org/2022/03/21/a-new-discussion-forum-for-osm/

OSM Foundation. (2023a). *Ban Policy*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Ban_Policy

OSM Foundation. (2023b). *Join as a corporate member*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Join_as_a_corporate_member

OSM Foundation. (2023c). *Licence/Waiver and Permission Templates/Cover letter and waiver template for CC BY 4.0*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Licence/Waiver_and_Permission_Templates/Cover_letter_and_waiver_template_for_CC_BY_4.0

OSM Foundation. (2023d). *Membership*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Membership

OSM Foundation. (2024a). *Committees*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Committees

OSM Foundation. (2024b). *FAQ*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/FAQ

OSM Foundation. (2024c). *Local Chapters*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Local_Chapters

OSM Foundation. (2024d). *Local Chapters*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Local_Chapters

OSM Foundation. (2024e). *Main page*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Main_Page

OSM Foundation. (2024f). *Membership/Statistics*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Membership/Statistics

OSM Foundation. (2024g). *Officers & Board*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Officers_%26_Board

OSM Foundation. (2024h, June 10). *OpenStreetMap Foundation*. https://osmfoundation.org/

OSM Foundation. (2025a, January 27). *Advisory Board*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Advisory_Board

OSM Foundation. (2025b, February 4). *Corporate Members*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Corporate_Members

OSM Foundation. (2025c, February 6). *Local Chapters and Communities Working Group*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Local_Chapters_and_Communities_Working_Group

OSM Foundation. (2025d, February 10). *Finances*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Finances

OSM Foundation. (2025e, March 6). *Contractors and employees*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Contractors_and_employees

OSM Foundation. (2025f, March 13). *OSMF Server Info*. https://hardware.openstreetmap.org/

OSM Foundation. (2025g, March 20). *List of user blocks*. OpenStreetMap Foundation. https://www.openstreetmap.org/user_blocks?before=17382

OSM License Working Group. (2017, March 17). Use of CC BY 4.0 licensed data in OpenStreetMap. *OpenStreetMap Blog*. https://blog.openstreetmap.org/2017/03/17/use-of-cc-by-data/

OSM License Working Group. (2018). *File:GDPR Position Paper.pdf*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/File:GDPR_Position_Paper.pdf

OSM members. (2020). *A Call to Take Action and Confront Systemic Offensive Behaviour in the OSM Community*. OSM Wiki. https://wiki.openstreetmap.org/w/images/5/56/A_Call_to_Take_Action_and_Confront_Systemic_Offensive_Behaviour_in_the_OSM_Community.pdf

OSM Operations Working Group. (n.d.). *Tile Usage Policy*. OSMF OWG. Retrieved 13 March 2025, from https://operations.osmfoundation.org/policies/tiles/

OSM OWG. (2025). *API Usage policy*. OSM OWG. https://operations.osmfoundation.org/policies/api/

OSM Wiki. (2018). *OpenAerialMap*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/OpenAerialMap

OSM Wiki. (2019). *India/Rendering server*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/India/Rendering_server

OSM Wiki. (2020a). *Foundation/Core Values*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Foundation/Core_Values

OSM Wiki. (2020b). *GlobalLogic*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/GlobalLogic

OSM Wiki. (2020c). *Isle of Wight workshop 2006*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Isle_of_Wight_workshop_2006

OSM Wiki. (2020d). *OSMFJ*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/OSMFJ

OSM Wiki. (2022a). *API*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/API

OSM Wiki. (2022b). *Gojek*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Gojek

OSM Wiki. (2022c). *Open Database License*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Open_Database_License

OSM Wiki. (2023a). *Category:Tag descriptions for group 'bridges'*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Category:Tag_descriptions_for_group_%22bridges%22
OSM Wiki. (2023b). *Etiquette/Etiquette Guidelines*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Etiquette/Etiquette_Guidelines
OSM Wiki. (2023c). *Japan/OSMFJ Tileserver*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Japan/OSMFJ_Tileserver
OSM Wiki. (2023d). *Operations Working Group*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Operations_Working_Group
OSM Wiki. (2023e). *Proposal:Childcare—OpenStreetMap Wiki*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Proposal:Childcare
OSM Wiki. (2023f). *Vertical Aerial Photographs*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Vertical_Aerial_Photographs#Outside_of_UK
OSM Wiki. (2023g). *Vertical Aerial Photographs- OpenStreetMap Wiki*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Vertical_Aerial_Photographs
OSM Wiki. (2024a). *Armchair mapping*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Armchair_mapping
OSM Wiki. (2024b). *Automated edits*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Automated_edits
OSM Wiki. (2024c). *CloudMade*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/CloudMade
OSM Wiki. (2024d). *Elements—OpenStreetMap Wiki*. https://wiki.openstreetmap.org/wiki/Elements
OSM Wiki. (2024e). *Foundation/AGM2024/Election to Board/Answers and manifestos/Q14 Perspective on Overture Maps*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Foundation/AGM2024/Election_to_Board/Answers_and_manifestos/Q14_Perspective_on_Overture_Maps
OSM Wiki. (2024f). *Humanitarian OSM Team*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Humanitarian_OSM_Team
OSM Wiki. (2024g). *Humanitarian OSM Team/Open Mapping Hub—Asia Pacific/Open Mapping Guru Project*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Humanitarian_OSM_Team/Open_Mapping_Hub_-_Asia_Pacific/Open_Mapping_Guru_Project#What_is_an_Open_Mapping_Guru_(Guru)
OSM Wiki. (2024h). *Import/Past Problems*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Import/Past_Problems
OSM Wiki. (2024i). *JOSM/Plugins/KartaView*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/JOSM/Plugins/KartaView
OSM Wiki. (2024j). *Key:bridge*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Key:bridge
OSM Wiki. (2024k). *MapBeks*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/MapBeks
OSM Wiki. (2024l). *Organised Editing/Activities/Apple/Data Team*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Organised_Editing/Activities/Apple/Data_Team
OSM Wiki. (2024m). *Organised Editing/Activities/Meta*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Organised_Editing/Activities/Meta#United_States
OSM Wiki. (2024n). *Pokémon Go*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Pok%C3%A9mon_Go
OSM Wiki. (2025a, January 4). *Editors*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Editors
OSM Wiki. (2025b, January 8). *Editor usage stats*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Editor_usage_stats
OSM Wiki. (2025c, January 9). *Street-level imagery services*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Street-level_imagery_services
OSM Wiki. (2025d, January 11). *Raster tile providers*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Raster_tile_providers
OSM Wiki. (2025e, January 12). *Rapid*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Rapid
OSM Wiki. (2025f, January 14). *Editing Standards and Conventions*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Editing_Standards_and_Conventions
OSM Wiki. (2025g, January 15). *Missing Maps Project*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Missing_Maps_Project
OSM Wiki. (2025h, January 16). *Osmose*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Osmose
OSM Wiki. (2025i, January 17). *Import/Catalogue*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Import/Catalogue

OSM Wiki. (2025j, January 26). *Ola—Organised Editing*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Ola_-_Organised_Editing

OSM Wiki. (2025k, January 28). *GDPR*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/GDPR

OSM Wiki. (2025l, January 28). *Good practice*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Good_practice#Map_what's_on_the_ground

OSM Wiki. (2025m, January 31). *MapRoulette*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/MapRoulette

OSM Wiki. (2025n, January 31). *OpenHistoricalMap/FAQ*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/OpenHistoricalMap/FAQ

OSM Wiki. (2025o, February 4). *Mapping private information*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Mapping_private_information

OSM Wiki. (2025p, February 7). *Organised Editing/Activities/Amazon*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Organised_Editing/Activities/Amazon

OSM Wiki. (2025q, February 7). *Organised Editing/Activities/Amazon*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Organised_Editing/Activities/Amazon

OSM Wiki. (2025r, February 7). *Overpass API*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Overpass_API

OSM Wiki. (2025s, February 11). *Humanitarian OSM Team/HOT Tasking Manager Organizations*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Humanitarian_OSM_Team/HOT_Tasking_Manager_Organizations

OSM Wiki. (2025t, February 11). *Organised Editing/Activities/UN Maps*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Organised_Editing/Activities/UN_Maps

OSM Wiki. (2025u, February 11). *Tasking Manager*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Tasking_Manager

OSM Wiki. (2025v, February 19). *Organised Editing/Guidelines*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Organised_Editing/Guidelines

OSM Wiki. (2025w, February 22). *Downloading data*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Downloading_data

OSM Wiki. (2025x, February 22). *Import/ODbL Compatibility*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Import/ODbL_Compatibility

OSM Wiki. (2025y, February 23). *Deploying your own Slippy Map*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Deploying_your_own_Slippy_Map

OSM Wiki. (2025z, March 1). *Planet.osm*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Planet.osm#Extracts

OSM Wiki. (2025aa, March 3). *Import/Guidelines*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Import/Guidelines#Step_1:_Prerequisites

OSM Wiki. (2025ab, March 5). *Automated Edits code of conduct*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Automated_Edits_code_of_conduct

OSM Wiki. (2025ac, March 6). *Foundation/Local Chapters*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Foundation/Local_Chapters

OSM Wiki. (2025ad, March 7). *Changeset*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Changeset

OSM Wiki. (2025ae, March 9). *Key:lgbtq*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Key:lgbtq

OSM Wiki. (2025af, March 10). *Organised Editing/Activities/Microsoft*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Organised_Editing/Activities/Microsoft

OSM Wiki. (2025ag, March 10). *Quality assurance*. OpenStreetMap Wiki.

OSM Wiki. (2025ah, March 19). *China*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/China

OSM Wiki. (2025ai, March 21). *State of the Map*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/State_of_the_Map

OSM Wiki. (2025aj, March 25). *MapSwipe*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/MapSwipe

OSM Wiki. (2025ak, April 2). *Tag:highway=living_street*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Tag:highway%3Dliving_street

OSM Wiki. (2025al, April 3). *Contributors*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Contributors

OSM Wiki. (2025am, April 10). *API v0.6*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/API_v0.5
OSM Wiki. (2025an, April 19). *Vector tiles*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Vector_tiles#Providers
OSMCha. (n.d.). *OSMCha Guide*. OSMCha. Retrieved 28 March 2025, from https://osmcha.org/about#introduction-to-osmcha
OSMF. (2016). *Licence and Legal FAQ/Why CC BY-SA is Unsuitable*. OpenStreetMap Foundation. https://osmfoundation.org/wiki/Licence_and_Legal_FAQ/Why_CC_BY-SA_is_Unsuitable
Osmose. (2025). *Osmose*. https://osmose.openstreetmap.fr
Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, game changers, and challengers* (Vol. 1). John Wiley & Sons.
Over, M., Schilling, A., Neubauer, S., & Zipf, A. (2010). Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany. *Computers, Environment and Urban Systems*, *34*(6), 496–507. https://doi.org/10.1016/j.compenvurbsys.2010.05.001
Overture Maps Foundation. (2023). *FAQ*. Overture Maps Foundation. https://overturemaps.org/resources/faq/
Overture Maps Foundation. (2024a). FAQs. *Overture Maps Foundation*. https://overturemaps.org/about/faq/
Overture Maps Foundation. (2024b, August 16). *Home*. Overture Maps Foundation. https://overturemaps.org/
Overture Maps Foundation. (2025a, March 24). *Attribution and Licensing*. Overture Maps Documentation. https://docs.overturemaps.org/attribution/
Overture Maps Foundation. (2025b, March 24). *Global Entity Reference System (GERS)*. Overture Maps Documentation. https://docs.overturemaps.org/gers/
Overture Maps Foundation. (2025c, March 25). Meta's Adoption of Overture Maps: A Case Study. *Overture Maps Foundation*. https://overturemaps.org/case-study/2025/metas-adoption-of-overture-maps-a-case-study/
Oxford. (n.d.). Sustainable. In *Oxford Advanced Learner's Dictionary*. Retrieved 12 October 2022, from https://www.oxfordlearnersdictionaries.com/definition/english/sustainable?q=sustainable
Oxford University Press. (n.d.). *Definition of third party noun*. Oxford Advanced Leaner's Dictionary. https://www.oxfordlearnersdictionaries.com/definition/english/third-party?q=third+party
Pandey, N., Dé, R., & Ravishankar, M. (2022). Improving the governance of information technology: Insights from the history of Internet governance. *Journal of Information Technology*, *37*(3), 266–287. https://doi.org/10.1177/02683962211054513
Panek, J., & Sobotova, L. (2015). Community Mapping in Urban Informal Settlements: Examples from Nairobi, Kenya. *THE ELECTRONIC JOURNAL OF INFORMATION SYSTEMS IN DEVELOPING COUNTRIES*, *68*(1), 1–13. https://doi.org/10.1002/j.1681-4835.2015.tb00487.x
Park, C. H., Longo, J., & Johnston, E. W. (2020). Exploring Non-State Stakeholder and Community-Led Open Governance: Beyond the Three Pillars of Open Government. *Public Performance & Management Review*. https://www.tandfonline.com/doi/abs/10.1080/15309576.2019.1677253
Park, S., & Gil-Garcia, J. R. (2017). Understanding Transparency and Accountability in Open Government Ecosystems: The Case of Health Data Visualizations in a State Government. *Proceedings of the 18th Annual International Conference on Digital Government Research*, 39–47. https://doi.org/10.1145/3085228.3085318
Patel, J., Dickinson, C., Sengupta, R., & Sarkar, D. (2023). *Intrinsic data quality of corporate contributions to OpenStreetMap: Assessing completeness*. Spatial Knowledge and Information Canada 2023, Banff, Alberta, Canada. https://doi.org/10.17605/OSF.IO/N6DU5
Paul, K. (2020, June 19). Facebook acquires crowdsourced mapping company Mapillary. *Reuters*. https://www.reuters.com/article/technology/facebook-acquires-crowdsourced-mapping-company-mapillary-idUSKBN23Q0FZ/
Pechmann, J., & de los Reyes, C. (2023, March 17). Using OSM Data in the Turkey and Syria Earthquake Response. *Humanitarian OpenStreetMap Team*. https://www.hotosm.org/updates/using-osm-data-for-the-turkey-and-syria-earthquake-response/
Peixoto, J. P. J., Costa, D. G., Portugal, P., & Vasques, F. (2023). A geospatial dataset of urban infrastructure for emergency response in Portugal. *Data in Brief*, *50*, 109593. https://doi.org/10.1016/j.dib.2023.109593

Peric, M., Durkin, J., & Vitezic, V. (2017). The Constructs of a Business Model Redefined: A Half-Century Journey. *Sage Open*, *7*(3), 1–13. https://doi.org/10.1177/2158244017733516

Perkel, J. (2020, April 7). Behind the Johns Hopkins University coronavirus dashboard. *Nature Index*. https://www.nature.com/nature-index/news/behind-the-johns-hopkins-university-coronavirus-dashboard

Perkins, C. (2011). Researching mapping: Methods, modes and moments in the (im)mutability of OpenStreetMap. *Global Media Journal - Australian Edition*.

Perkins, C. (2014). Plotting practices and politics: (Im)mutable narratives in OpenStreetMap. *Transactions of the Institute of British Geographers*, *39*(2), 304–317. https://doi.org/10.1111/tran.12022

Perkmann, M., & Schildt, H. (2015). Open data partnerships between firms and universities: The role of boundary organizations. *Research Policy*, *44*(5), 1133–1143. https://doi.org/10.1016/j.respol.2014.12.006

Planet OSM. (2025). *Planet OSM*. https://planet.openstreetmap.org/

Plennert, M. (2018). The social construction of technological stasis: The stagnating data structure in OpenStreetMap. *Big Data & Society*, *5*(2), 2053951718790591. https://doi.org/10.1177/2053951718790591

Plennert, M., Glasze, G., & Schlieder, C. (2019). The Socio-Technical Background of an Unconventional Software Architecture in OpenStreetMap: Understanding the Implementation of 'Folksonomy'. *Computational Culture*, *7*. http://computationalculture.net/the-socio-technical-background-of-an-unconventional-software-architecture-in-openstreetmap-understanding-the-implementation-of-folksonomy/

POI. (2023, July 14). Engaging the Crowd: Mapping Lakes in Jabodetabek-Punjur for Flood Mitigation. *Perkumpulan OpenStreetMap Indonesia*. https://openstreetmap.or.id/blog/engaging-the-crowd-mapping-lakes-in-jabodetabek-punjur-for-flood-mitigation/

Poiani, T. H., Rocha, R., Degrossi, L., & De Albuquerque, J. (2016, January). *Potential of Collaborative Mapping for Disaster Relief: A Case Study of OpenStreetMap in the Nepal Earthquake 2015*. https://doi.org/10.1109/HICSS.2016.31

Poikola, A., Kola, P., & Hintikka. (2011). *Public data: An introduction to opening information resources*. Ministry of Transport and Communications. https://core.ac.uk/download/pdf/198191501.pdf

Poole, S. (2017, January 15). *OpenStreetMap Community Statistics Revisited*. OpenStreetMap. https://www.openstreetmap.org/user/SimonPoole/diary/40246

Prieto-Curiel, R., Schumann, A., Heo, I., & Heinrigs, P. (2022). Detecting cities with high intermediacy in the African urban network. *Computers, Environment and Urban Systems*, *98*, 101869. https://doi.org/10.1016/j.compenvurbsys.2022.101869

Publications Office of the European Union. (2023). *Rethinking the impact of open data: A first step towards a European impact assessment for open data.* Publications Office. https://data.europa.eu/doi/10.2830/911822

Quarati, A., De Martino, M., & Rosim, S. (2021). Geospatial Open Data Usage and Metadata Quality. *ISPRS International Journal of Geo-Information*, *10*(1), Article 1. https://doi.org/10.3390/ijgi10010030

Quinn, S., & Bull, F. (2019). Understanding threats to crowdsourced geographic data quality through a study of OpenStreetMap contributor bans. In N. Valcik & D. Dean (Eds.), *Geospatial Information System Use in Public Organizations: How and Why GIS Should be Used in the Public Sector* (pp. 80–96). Routledge. https://doi.org/10.4324/9780429272851

Quinn, S., & Tucker, D. A. (2017). How geopolitical conflict shapes the mass-produced online map. *First Monday*. https://doi.org/10.5210/fm.v22i11.7922

Quinn, S., & Yapa, L. (2016). OpenStreetMap and Food Security: A Case Study in the City of Philadelphia. *The Professional Geographer*, *68*(2), 271–280. https://doi.org/10.1080/00330124.2015.1065547

R. Griggs Group Ltd & Ors v Evans & Ors [2003] EWHC 2914 (Ch) (EWHC (Ch) 2 December 2003). https://www.bailii.org/ew/cases/EWHC/Ch/2003/2914.html

Radford, T., & Firth, R. (2020, June 17). *Frequently Asked Questions: About The Audacious Project*. Humanitarian OpenStreetMap Team. https://www.hotosm.org/audacious-faq.html

Randriamihaja, M., Ihantamalala, F. A., Rafenoarimalala, F. H., Finnegan, K. E., Rakotonirina, L., Razafinjato, B., Bonds, M. H., Evans, M. V., & Garchitorena, A. (2024). Combining OpenStreetMap mapping and route optimization algorithms to inform the delivery of community health interventions at the last mile. *PLOS Digital Health*, *3*(11), e0000621. https://doi.org/10.1371/journal.pdig.0000621

Raymond, E. (1999). The cathedral and the bazaar. *Knowledge, Technology & Policy*, *12*(3), 23–49. https://doi.org/10.1007/s12130-999-1026-0

Reggi, L., & Dawes, S. S. (2022). Creating Open Government Data ecosystems: Network relations among governments, user communities, NGOs and the media. *Government Information Quarterly*, *39*(2). https://doi.org/10.1016/j.giq.2022.101675

Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European Data Governance and Amending Regulation (EU) 2018/1724 (Data Governance Act), 152 OJ L (2022). http://data.europa.eu/eli/reg/2022/868/oj/eng

Reinhardt, R., Christodoulou, I., García, B. A., & Gassó-Domingo, S. (2020). Sustainable business model archetypes for the electric vehicle battery second use industry: Towards a conceptual framework. *Journal of Cleaner Production*, *254*, 119994. https://doi.org/10.1016/j.jclepro.2020.119994

Robinson, P., & Mather, L. W. (2017). Open Data Community Maturity: Libraries as Civic Infomediaries. *URISA Journal*, *28*(1), 31–39. https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=10458077&v=2.1&it=r&id=GALE%7CA559830503&sid=googleScholar&linkaccess=abs

Robinson, P., & Mather, L. W. (2022). Reflections on the Future of Open Data. In P. Robinson & T. Scassa (Eds.), *The Future of Open Data*.

Rossi, F., Caloffi, A., Colovic, A., & Russo, M. (2022). New business models for public innovation intermediaries supporting emerging innovation systems: The case of the Internet of Things. *Technological Forecasting and Social Change*, *175*, 121357. https://doi.org/10.1016/j.techfore.2021.121357

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Rowley, J. (2007). The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of Information Science*, *33*(2), 163–180. https://doi.org/10.1177/0165551506070706

Russell, J. (2018, December 20). *Grab is messing up the world's largest mapping community's data in Southeast Asia*. TechCrunch. https://techcrunch.com/2018/12/19/grab-maps-osm-thailand-southeast-asia/

Sangiambut, S., & Sieber, R. (2017). The Civic Open Data and Crowdsourcing App Ecosystem: Actors, Materials, and Interventions. *Journal of the Urban & Regional Information Systems Association*, *28*(1).

Sarkar, D., & Anderson, J. T. (2022). Corporate editors in OpenStreetMap: Investigating co-editing patterns. *Transactions in GIS*, *26*(4), 1879–1897. https://doi.org/10.1111/tgis.12910

Sarretta, A., & Minghini, M. (2021). Towards The Integration of Authoritative and OpenStreetMap Geospatial Datasets in Support of The European Strategy for Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLVI-4-W2-2021*, 159–166. https://doi.org/10.5194/isprs-archives-XLVI-4-W2-2021-159-2021

Schiefelbein, J., Rudnick, J., Scholl, A., Remmen, P., Fuchs, M., & Müller, D. (2019). Automated urban energy system modeling and thermal building simulation based on OpenStreetMap data sets. *Building and Environment*, *149*, 630–639. https://doi.org/10.1016/j.buildenv.2018.12.025

Scholz, C., Jaime, C., Raju, E., Coughlan De Perez, E., & van Aalst, M. (2024). Off the grid: Utilizing OpenStreetMap for early warning and early action in conflict settings in Sudan. *Frontiers in Climate*, *6*. https://doi.org/10.3389/fclim.2024.1439940

Schott, M., Grinberger, A. Y., Lautenbach, S., & Zipf, A. (2021). The Impact of Community Happenings in OpenStreetMap—Establishing a Framework for Online Community Member Activity Analyses. *ISPRS International Journal of Geo-Information*, *10*(3), Article 3. https://doi.org/10.3390/ijgi10030164

Schrock, A., & Shaffer, G. (2017). Data ideologies of an interested public: A study of grassroots open government data intermediaries. *Big Data & Society*, *4*(1), 2053951717690750. https://doi.org/10.1177/2053951717690750

Schröder-Bergen, S. (2024, June 12). *Google Maps and OpenStreetMap: The coexistence of two unequal siblings*. Geoawesome. https://geoawesome.com/google-maps-and-openstreetmap-the-coexistence-of-two-unequal-siblings/

Schröder-Bergen, S., Glasze, G., Michel, B., & Dammann, F. (2022). De/colonizing OpenStreetMap? Local mappers, humanitarian and commercial actors and the changing modes of collaborative mapping. *GeoJournal*, *87*(6), 5051–5066. https://doi.org/10.1007/s10708-021-10547-7

Schroeder, R. (2016). Big data business models: Challenges and opportunities. *Cogent Social Sciences*, *2*(1), 1166924. https://doi.org/10.1080/23311886.2016.1166924

Schultz, M., Voss, J., Auer, M., Carter, S., & Zipf, A. (2017). Open land cover from OpenStreetMap and remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, *63*, 206–213. https://doi.org/10.1016/j.jag.2017.07.014

Schweihoff, J., Lipovetskaja, A., Jussen-Lengersdorf, I., & Möller, F. (2024). Stuck in the middle with you: Conceptualizing data intermediaries and data intermediation services. *Electronic Markets*, *34*(1), 48. https://doi.org/10.1007/s12525-024-00729-9

Seddon, P. B. (2014). Implications for strategic IS research of the resource-based theory of the firm: A reflection. *The Journal of Strategic Information Systems*, *23*(4), 257–269. https://doi.org/10.1016/j.jsis.2014.11.001

Seddon, P. B., Lewis, G. P., Freeman, P., & Shanks, G. (2004). The Case for Viewing Business Models as Abstractions of Strategy. *Communications of the Association for Information Systems*, *13*(1). https://doi.org/10.17705/1CAIS.01325

Seidel, A. (2023). A global map of amenities: Public goods, ethnic divisions and decentralization. *Journal of Development Economics*, *164*, 103113. https://doi.org/10.1016/j.jdeveco.2023.103113

Seto, T. (2022). Development of OpenStreetMap Data in Japan. In Y. Wakabayashi & T. Morita (Eds.), *Ubiquitous Mapping: Perspectives from Japan* (pp. 113–126). Springer Nature. https://doi.org/10.1007/978-981-19-1536-9_7

Sexton, A., Shepherd, E., Duke-Williams, O., & Eveleigh, A. (2017). A balance of trust in the use of government administrative data. *Archival Science*, *17*(4), 305–330. https://doi.org/10.1007/s10502-017-9281-4

Shaharudin, A., Reyes, M. E. L., Pantazatou, K., Storm, I., Larsen, B., van Loenen, B., & Mazzoli, U. K. (2023). *Workshop Report on Data Ecosystems and Spatial Data Infrastructure—Facilitators for Data Value Creation* [Official Workshop Report]. European Spatial Data Research (EuroSDR). https://www.eurosdr.net/publications/workshop-report-data-ecosystems-and-spatial-data-infrastructure-facilitators-data-value

Shaharudin, A., van Loenen, B., & Janssen, M. (2023). Towards a Common Definition of Open Data Intermediaries. *Digital Government: Research and Practice*, *4*(2), 6:1-6:21. https://doi.org/10.1145/3585537

Shakeri, S. (2023, November 8). *Mapping the Future: How Overture Transportation Network Data Model Enhances Geospatial Applications.* TomTom Engineering Blog. https://engineering.tomtom.com/overture-transportation-network-linear-referencing/

Shin, H., & Basiri, A. (2022, April). *Geographic Biases in OSM Contributions: How do the Geographic Extent of Contributions Differ among Demographic groups?* 30th GISRUK Conference 2022, Liverpool, UK. https://doi.org/10.5281/zenodo.6411602

Sieber, R., & Brandusescu, A. (2025). Disparity, Instability, and Power in the Crowdmapping Ecosystem. *The Journal of Community Informatics*, *21*(1), Article 1. https://doi.org/10.15353/joci.v21i1.6071

Siggelkow, N. (2007). Persuasion With Case Studies. *Academy of Management Journal*, *50*(1), 20–24. https://doi.org/10.5465/amj.2007.24160882

Şimşek, T., Öner, M. A., Kunday, Ö., & Olcay, G. A. (2022). A journey towards a digital platform business model: A case study in a global tech-company. *Technological Forecasting and Social Change*, *175*, 121372. https://doi.org/10.1016/j.techfore.2021.121372

Singleton, A. D., Spielman, S., & Brunsdon, C. (2016). Establishing a framework for Open Geographic Information science. *International Journal of Geographical Information Science*, *30*(8), 1507–1521. https://doi.org/10.1080/13658816.2015.1137579

Sloan, M. C. (2010). Aristotle's Nicomachean Ethics as the Original Locus for the Septem Circumstantiae. *Classical Philology*, *105*(3), 236–251. https://doi.org/10.1086/656196

Smith, G., Ofe, H. A., & Sandberg, J. (2016). Digital Service Innovation from Open Data: Exploring the Value Proposition of an Open Data Marketplace. *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 1277–1286. https://doi.org/10.1109/HICSS.2016.162

Snihur, Y. (2016). Developing optimal distinctiveness: Organizational identity processes in new ventures engaged in business model innovation. *Entrepreneurship & Regional Development*, *28*(3–4), 259–285. https://doi.org/10.1080/08985626.2016.1155745

Snihur, Y., & Markman, G. (2023). Business Model Research: Past, Present, and Future. *Journal of Management Studies*, *60*(8), e1–e14. https://doi.org/10.1111/joms.12928

So, W., & Duarte, F. (2020). Cartographers of North Korea: Who are they and what are the technical, political, and social issues involved in mapping North Korea. *Geoforum*, *110*, 147–156. https://doi.org/10.1016/j.geoforum.2020.02.008

Soden, R., & Palen, L. (2014). From Crowdsourced Mapping to Community Mapping: The Post-earthquake Work of OpenStreetMap Haiti. In C. Rossitto, L. Ciolfi, D. Martin, & B. Conein (Eds.), *COOP 2014—Proceedings of the 11th International Conference on the Design of Cooperative Systems, 27-30 May 2014, Nice (France)* (pp. 311–326). Springer International Publishing. https://doi.org/10.1007/978-3-319-06498-7_19

Solís, P., Anderson, J., & Rajagopalan, S. (2021). Open geospatial tools for humanitarian data creation, analysis, and learning through the global lens of YouthMappers. *Journal of Geographical Systems*, *23*(4), 599–625. https://doi.org/10.1007/s10109-020-00339-x

Solomon, G., Sutton, D., Polat Kayali, M., Yuan, X., Gardner, Z., & Basiri, A. (2024). Evaluating geotemporal behaviours of OpenStreetMap contributors. *AGILE: GIScience Series*, *5*, 1–10. https://doi.org/10.5194/agile-giss-5-14-2024

Soman, S., Beukes, A., Nederhood, C., Marchio, N., & Bettencourt, L. M. A. (2020). Worldwide Detection of Informal Settlements via Topological Analysis of Crowdsourced Digital Maps. *ISPRS International Journal of Geo-Information*, *9*(11), Article 11. https://doi.org/10.3390/ijgi9110685

Spencer, B. (2013). *Business Model Design and Learning: A Strategic Guide*. Business Expert Press. http://ebookcentral.proquest.com/lib/delft/detail.action?docID=1048406

Spieth, P., Schneider, S., Clauß, T., & Eichenberg, D. (2019). Value drivers of social businesses: A business model perspective. *Long Range Planning*, *52*(3), 427–444. https://doi.org/10.1016/j.lrp.2018.04.004

Steele, M. W. (2019). The Man who Walked 40 Million Steps: Ino Tadataka and the Modern Mapping of Japan. アジア文化研究, *45*, 93–101.

Steinley, Douglas. (2006). K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology*, *59*(1), 1–34. https://doi.org/10.1348/000711005X48266

Steinmann, R., Häusler, E., Klettner, S., Schmidt, M., & Lin, Y. (2013). Gender Dimensions in UGC and VGI: A Desk-Based Study. *GI_Forum 2013: Creating the GISociety*, *Volume 1*, 355–364. https://doi.org/10.1553/giscience2013s355

Stephens, M. (2013). Gender and the GeoWeb: Divisions in the production of user-generated cartographic information. *GeoJournal*, *78*(6), 981–996. https://doi.org/10.1007/s10708-013-9492-z

Sterk, F., Stocker, A., Heinz, D., & Weinhardt, C. (2024). Unlocking the value from car data: A taxonomy and archetypes of connected car business models. *Electronic Markets*, *34*(1), 13. https://doi.org/10.1007/s12525-024-00692-5

Sugg, Z. (2022). Social barriers to open (water) data. *WIREs Water*, *9*(1), e1564. https://doi.org/10.1002/wat2.1564

Sui, D. (2008). The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Computers, Environment and Urban Systems*, *32*(1), 1–5. https://doi.org/10.1016/j.compenvurbsys.2007.12.001

Sui, D. (2014). Opportunities and Impediments for Open GIS. *Transactions in GIS*, *18*(1), 1–24. https://doi.org/10.1111/tgis.12075

Sulistioningrum, D. (2023, September 22). *SiPetani DKI: Empowering Agriculture in Jakarta Through Digital Innovation*. Perkumpulan OpenStreetMap Indonesia. https://openstreetmap.or.id/blog/sipetani-dki-empowering-agriculture-in-jakarta-through-digital-innovation/

Susha, I., Flipsen, M., Agahari, W., & de Reuver, M. (2020). Towards Generic Business Models of Intermediaries in Data Collaboratives: From Gatekeeping to Data Control. In G. Viale Pereira, M. Janssen, H. Lee, I. Lindgren, M. P. Rodríguez Bolívar, H. J. Scholl, & A. Zuiderwijk (Eds.), *Electronic Government* (pp. 304–315). Springer International Publishing. https://doi.org/10.1007/978-3-030-57599-1_23

Susha, I., Janssen, M., & Verhulst, S. (2017). Data Collaboratives as a New Frontier of Cross-Sector Partnerships in the Age of Open Data: Taxonomy Development. *Hawaii International Conference on System Sciences 2017 (HICSS-50)*. https://aisel.aisnet.org/hicss-50/eg/open_data_in_government/4

Sutherland, M. K., & Cook, M. E. (2017). Data-Driven Smart Cities: A Closer Look at Organizational, Technical and Data Complexities. *Proceedings of the 18th Annual International Conference on Digital Government Research*, 471–476. https://doi.org/10.1145/3085228.3085239

Täuscher, K., & Laudien, S. M. (2018). Understanding platform business models: A mixed methods study of marketplaces. *European Management Journal*, *36*(3), 319–329. https://doi.org/10.1016/j.emj.2017.06.005

Teece, D. J. (2010). Business Models, Business Strategy and Innovation. *Long Range Planning*, *43*(2), 172–194. https://doi.org/10.1016/j.lrp.2009.07.003

Teeuwen, R., Milias, V., Bozzon, A., & Psyllidis, A. (2024). How well do NDVI and OpenStreetMap data capture people's visual perceptions of urban greenspace? *Landscape and Urban Planning*, *245*, 105009. https://doi.org/10.1016/j.landurbplan.2024.105009

Temiz, S., Holgersson, M., Björkdahl, J., & Wallin, M. W. (2022). Open data: Lost opportunity or unrealized potential? *Technovation*, *114*, 102535. https://doi.org/10.1016/j.technovation.2022.102535

Thakuriah, P. (Vonu), Dirks, L., & Keita, Y. M. (2017). Digital Infomediaries and Civic Hacking in Emerging Urban Data Initiatives. In P. (Vonu) Thakuriah, N. Tilahun, & M. Zellner (Eds.), *Seeing Cities Through Big Data: Research, Methods and Applications in Urban Informatics* (pp. 189–207). Springer International Publishing. https://doi.org/10.1007/978-3-319-40902-3_11

The Centre for Public Data. (2024, May). *Geospatial data institutions around the world*. https://static1.squarespace.com/static/5ee7a7d964aeed7e5c507900/t/664b58b94817b-4004854c71d/1716213967727/Geospatial+data+institutions+around+the+world.pdf

Thebault-Spieker, J., Hecht, B., & Terveen, L. (2018). Geographic Biases are 'Born, not Made': Exploring Contributors' Spatiotemporal Behavior in OpenStreetMap. *Proceedings of the 2018 ACM International Conference on Supporting Group Work*, 71–82. https://doi.org/10.1145/3148330.3148350

Thomas, D. R. (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*, *27*(2), 237–246. https://doi.org/10.1177/1098214005283748

Thomé, A. M. T., Scavarda, L. F., & Scavarda, A. J. (2016). Conducting systematic literature review in operations management. *Production Planning & Control*, *27*(5), 408–420. https://doi.org/10.1080/09537287.2015.1129464

Timmers, P. (1998). Business models for electronic markets. *Electronic Markets*, *8*(2), 3–8.

TimSC. (2011). *User:TimSC/Quit*. OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/User:TimSC/Quit

TomTom. (2025). *TomTom Orbis Maps—Digital Mapping Solutions*. TomTom. https://www.tomtom.com/tomtom-orbis-maps/

TomTom Blog. (2023, March 8). *Q&A: TomTom OSM community manager*. TomTom. https://www.tomtom.com/newsroom/explainers-and-insights/q-a-tomtom-osm-community-manager/

Trapp, C. T. C., & Kanbach, D. K. (2021). Green entrepreneurship and business models: Deriving green technology business model archetypes. *Journal of Cleaner Production*, *297*, 126694. https://doi.org/10.1016/j.jclepro.2021.126694

Tripathy, P., Rao, P., Balakrishnan, K., & Malladi, T. (2021). An open-source tool to extract natural continuity and hierarchy of urban street networks. *Environment and Planning B: Urban Analytics and City Science*, *48*(8), 2188–2205. https://doi.org/10.1177/2399808320967680

Truden, C., Kollingbaum, M. J., Reiter, C., & Schasché, S. E. (2022). A GIS-based analysis of reachability aspects in rural public transportation. *Case Studies on Transport Policy*, *10*(3), 1827–1840. https://doi.org/10.1016/j.cstp.2022.07.012

Tummons, J. (2021). Ontological Pluralism, Modes of Existence, and Actor-Network Theory: Upgrading Latour with Latour. *Social Epistemology*, *35*(1), 1–11. https://doi.org/10.1080/02691728.2020.1774815

Turner, A. (2006). *Introduction to neogeography*. O'Reilly Media, Inc.

Uber. (2020). *Uber-OSM/DataTeam: Captures the usernames of our team*. GitHub. https://github.com/Uber-OSM/DataTeam

UN Maps. (n.d.-a). *Homepage*. UN Maps. Retrieved 12 March 2025, from https://maps.un.org/

UN Maps. (n.d.-b). *Learn with UN Maps*. UN Maps. Retrieved 15 April 2025, from https://maps.un.org/learn-un-maps

UN Maps. (n.d.-c). *The Community*. UN Maps. Retrieved 12 March 2025, from https://maps.un.org/community

United States Congress. (2019). *Foundations for Evidence-Based Policymaking Act of 2018, Title II: OPEN Government Data Act*. U.S. Government. https://www.congress.gov/bill/115th-congress/house-bill/4174

Urban, M., Klemm, M., Ploetner, K. O., & Hornung, M. (2018). Airline categorisation by applying the business model canvas and clustering algorithms. *Journal of Air Transport Management*, *71*, 175–192. https://doi.org/10.1016/j.jairtraman.2018.04.005

van Berkel, N., & Pohl, H. (2024). Collaborating with Bots and Automation on OpenStreetMap. *ACM Trans. Comput.-Hum. Interact.*, *31*(3), 38:1-38:30. https://doi.org/10.1145/3665326

van Loenen, B., & Grothe, M. (2014). INSPIRE empowers re-use of public sector information. *International Journal of Spatial Data Infrastructures Research*, *9*, 96–106. https://doi.org/10.2902/1725-0463.2014.09.art4

van Loenen, B., Zuiderwijk, A., Vancauwenberghe, G., Lopez-Pellicer, F. J., Mulder, I., Alexopoulos, C., Magnussen, R., Saddiqa, M., Dulong de Rosnay, M., Crompvoets, J., Polini, A., Re, B., & Flores, C. C. (2021). Towards value-creating and sustainable open data ecosystems: A comparative case study and a research agenda. *JeDEM - eJournal of eDemocracy and Open Government*, *13*(2), Article 2. https://doi.org/10.29379/jedem.v13i2.644

van Rees, E. (2024, January 8). The Case for Collaboration Between OpenStreetMap and Overture Maps. *Geo Week News*. https://www.geoweeknews.com/news/overture-maps-foundation-openstreemap-open-data-mapping

van Schalkwyk, F., Cañares, M., Chattapdhyay, S., & Andrason, A. (2016). Open Data Intermediaries in Developing Countries. *The Journal of Community Informatics*, *12*(2), Article 2. https://doi.org/10.15353/joci.v12i2.3219

van Schalkwyk, F., Chattapadhyay, S., Cañares, M., & Andrason, A. (2015). *Open Data Intermediaries in Developing Countries* [Working Paper]. World Wide Web Foundation, Washington, D.C. https://idl-bnc-idrc.dspacedirect.org/handle/10625/56288

van Schalkwyk, F., Chattapadhyay, S., Caňares, M., & Andrason, A. (2015). *Open Data Intermediaries in Developing Countries*. World Wide Web Foundation. http://hdl.handle.net/10625/56288

van Schalkwyk, F., Willmers, M., & McNaughton, M. (2016). Viscous Open Data: The Roles of Intermediaries in an Open Data Ecosystem. *Information Technology for Development*, *22*(sup1), 68–83. https://doi.org/10.1080/02681102.2015.1081868

van Veenstra, A. F., & van den Broek, T. (2015). A Community-driven Open Data Lifecycle Model Based on Literature and Practice. In I. Boughzala, M. Janssen, & S. Assar (Eds.), *Case Studies in e-Government 2.0: Changing Citizen Relationships* (pp. 183–198). Springer International Publishing. https://doi.org/10.1007/978-3-319-08081-9_11

van Veenstra, A. F., & van den Broek, T. A. (2013). Opening Moves – Drivers, Enablers and Barriers of Open Data in a Semi-public Organization. In M. A. Wimmer, M. Janssen, & H. J. Scholl (Eds.), *Electronic Government* (pp. 50–61). Springer. https://doi.org/10.1007/978-3-642-40358-3_5

Vancauwenberghe, G. (2018). Assessing Open Data. In B. van Loenen, G. Vancauwenberghe, & J. Crompvoets (Eds.), *Open Data Exposed* (pp. 149–172). T.M.C. Asser Press. https://doi.org/10.1007/978-94-6265-261-3_8

Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J. M., Loos, P., & Spann, M. (2014). Business Models. *Business & Information Systems Engineering*, *6*(1), 45–53. https://doi.org/10.1007/s12599-013-0308-y

Verhulst, S. (2013, April 25). Imagery to the Crowd. *The Living Library*. https://thelivinglib.org/imagery-to-the-crowd/

Verhulst, S., Young, A., Zahuranec, A., Calderon, A., Gee, M., & Aaronson, S. A. (2020). The Emergence of a Third Wave of Open Data: How To Accelerate the Re-Use of Data for Public Interest Purposes While Ensuring Data Rights and Community Flourishing. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3937638

Vicario, A., Garcia, D., Villar, E. A., Andal, F., Sambale, M., Tamura, M., & Tasico, R. (2020, September 3). *A Call to Correct Narratives about Geospatial Work in the Philippines* [Public statement]. https://wiki.openstreetmap.org/w/images/a/aa/A_Call_to_Correct_Narratives_about_Geospatial_Work.pdf

Vichot, S. (2025, January 14). What does Overture Maps mean for OpenStreetMap and the future of open source mapping? *Jawg Maps*. https://blog.jawg.io/what-does-overture-maps-mean-for-openstreetmap-and-the-future-of-open-source-mapping/

Visnjic, I., Jovanovic, M., Neely, A., & Engwall, M. (2017). What brings the value to outcome-based contract providers? Value drivers in outcome business models. *International Journal of Production Economics*, *192*, 169–181. https://doi.org/10.1016/j.ijpe.2016.12.008

Voigt, K.-I., Buliga, O., & Michl, K. (2017). *Business Model Pioneers*. Springer International Publishing. https://doi.org/10.1007/978-3-319-38845-8

Walsham, G. (1997). Actor-Network Theory and IS Research: Current Status and Future Prospects. In A. S. Lee, J. Liebenau, & J. I. DeGross (Eds.), *Information Systems and Qualitative Research: Proceedings of the IFIP TC8 WG 8.2 International Conference on Information Systems and Qualitative Research, 31st May–3rd June 1997, Philadelphia, Pennsylvania, USA* (pp. 466–480). Springer US. https://doi.org/10.1007/978-0-387-35309-8_23

Wang, Z., Huang, J., & Tan, B. (2013). Managing organizational identity in the e-commerce industry: An ambidexterity perspective. *Information & Management*, *50*(8), 673–683. https://doi.org/10.1016/j.im.2013.05.002

Weigand, M., Wurm, M., Droin, A., Stark, T., Staab, J., Rauh, J., & Taubenböck, H. (2023). Are public green spaces distributed fairly? A nationwide analysis based on remote sensing, OpenStreetMap and census data. *Geocarto International*, *38*(1), 2286305. https://doi.org/10.1080/10106049.2023.2286305

Weking, J., Mandalenakis, M., Hein, A., Hermes, S., Böhm, M., & Krcmar, H. (2020). The impact of blockchain technology on business models – a taxonomy and archetypal patterns. *Electronic Markets*, *30*(2), 285–305. https://doi.org/10.1007/s12525-019-00386-3

Welle Donker, F., & van Loenen, B. (2016). Sustainable Business Models for Public Sector Open Data Providers. *JeDEM - eJournal of eDemocracy and Open Government*, *8*(1), Article 1. https://doi.org/10.29379/jedem.v8i1.390

Welle Donker, F., & van Loenen, B. (2017). How to assess the success of the open data ecosystem? *International Journal of Digital Earth*, *10*(3), 284–306. https://doi.org/10.1080/17538947.2016.1224938

Whetten, D. A. (2006). Albert and Whetten Revisited: Strengthening the Concept of Organizational Identity. *Journal of Management Inquiry*, *15*(3), 219–234. https://doi.org/10.1177/1056492606291200

Wieland, H., Hartmann, N. N., & Vargo, S. L. (2017). Business models as service strategy. *Journal of the Academy of Marketing Science*, *45*(6), 925–943. https://doi.org/10.1007/s11747-017-0531-z

Wiener, M., Saunders, C., & Marabelli, M. (2020). Big-data business models: A critical literature review and multiperspective research framework. *Journal of Information Technology*, *35*(1), 66–91. https://doi.org/10.1177/0268396219896811

Wilmott, C. (2019). "Mapping-with": The Politics of (Counter-)classification in OpenStreetMap. *Cartographic Perspectives*, *92*, Article 92. https://doi.org/10.14714/CP92.1451

Witt, R., Loos, L., & Zipf, A. (2021). Analysing the Impact of Large Data Imports in OpenStreetMap. *ISPRS International Journal of Geo-Information*, *10*(8), Article 8. https://doi.org/10.3390/ijgi10080528

Wroclawski, S. (2014, January 14). Why the world needs OpenStreetMap. *The Guardian*. https://www.theguardian.com/technology/2014/jan/14/why-the-world-needs-openstreetmap

Wroclawski, S. (2018, February 16). Why OpenStreetMap is in Serious Trouble. *Emacsen's Blog*. https://blog.emacsen.net/blog/2018/02/16/osm-is-in-trouble/

Xiao, Y., & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, *39*(1), 93–112. https://doi.org/10.1177/0739456X17723971

Yamashita, J., Seto, T., Iwasaki, N., & Nishimura, Y. (2023). Quality assessment of volunteered geographic information for outdoor activities: An analysis of OpenStreetMap data for names of peaks in Japan. *Geo-Spatial Information Science*, *26*(3), 333–345. https://doi.org/10.1080/10095020.2022.2085188

Yang, A., Fan, H., Jia, Q., Ma, M., Zhong, Z., Li, J., & Jing, N. (2024). How do contributions of organizations impact data inequality in OpenStreetMap? *Computers, Environment and Urban Systems*, *109*, 102077. https://doi.org/10.1016/j.compenvurbsys.2024.102077

Yap, W., Janssen, P., & Biljecki, F. (2022). Free and open source urbanism: Software for urban planning practice. *Computers, Environment and Urban Systems*, *96*, 101825. https://doi.org/10.1016/j.compenvurbsys.2022.101825

Yen, Y., Zhao, P., & Sohail, M. T. (2021). The morphology and circuity of walkable, bikeable, and drivable street networks in Phnom Penh, Cambodia. *Environment and Planning B: Urban Analytics and City Science*, *48*(1), 169–185. https://doi.org/10.1177/2399808319857726

Yi, Y., Chen, Y., & Li, D. (2022). Stakeholder ties, organizational learning, and business model innovation: A business ecosystem perspective. *Technovation*, *114*, 1–13. https://doi.org/10.1016/j.technovation.2021.102445

Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). SAGE Publications.

Yin, Y., Guo, L., & Thebault-Spieker, J. (2024). Productivity or Equity? Tradeoffs in Volunteer Microtasking in Humanitarian OpenStreetMap. *Proc. ACM Hum.-Comput. Interact.*, *8*(CSCW1), 113:1-113:34. https://doi.org/10.1145/3637390

Yoon, A., & Copeland, A. (2020). Toward community-inclusive data ecosystems: Challenges and opportunities of open data for community-based organizations. *Journal of the Association for Information Science and Technology*, *71*(12), 1439–1454. https://doi.org/10.1002/asi.24346

Yoon, A., Copeland, A., & McNally, P. J. (2018). Empowering communities with data: Role of data intermediaries for communities' data utilization. *Proceedings of the Association for Information Science and Technology*, *55*(1), 583–592. https://doi.org/10.1002/pra2.2018.14505501063

Young, A., & Verhulst, S. (2017). *Aclímate Colombia: Open data to improve agricultural resiliency*. GovLab - Open Data's Impact. https://doi.org/10.1007/978-3-319-63743-3_6

YouthMappers. (n.d.-a). *Everywhere She Maps*. YouthMappers. Retrieved 21 March 2025, from https://www.youthmappers.org/everywhereshemaps

YouthMappers. (n.d.-b). *Homepage*. YouthMappers. Retrieved 21 March 2025, from https://www.youthmappers.org

YouthMappers. (n.d.-c). *Our Story*. YouthMappers. Retrieved 21 March 2025, from https://www.youthmappers.org/our-story

YouthMappers. (n.d.-d). *YouthMappers Academy*. YouthMappers. Retrieved 21 March 2025, from https://www.youthmappers.org/academy

YouthMappers. (2024). *Leadership Fellowship*. YouthMappers. https://www.youthmappers.org/apply-leadership-fellows-initiative

Yu, C.-C. (2016). A value-centric business model framework for managing open data applications. *Journal of Organizational Computing and Electronic Commerce*, *26*(1–2), 80–115. https://doi.org/10.1080/10919392.2015.1125175

Zhang, X., An, J., Zhou, Y., Yang, M., & Zhao, X. (2024). How sustainable is OpenStreetMap? Tracking individual trajectories of editing behavior. *International Journal of Digital Earth*, *17*(1), 2311320. https://doi.org/10.1080/17538947.2024.2311320

Zhao, P., Jia, T., Qin, K., Shan, J., & Jiao, C. (2015). Statistical analysis on the evolution of OpenStreetMap road networks in Beijing. *Physica A: Statistical Mechanics and Its Applications*, *420*, 59–72. https://doi.org/10.1016/j.physa.2014.10.076

Zhou, Q., Wang, S., & Liu, Y. (2022). Exploring the accuracy and completeness patterns of global land-cover/land-use data in OpenStreetMap. *Applied Geography*, *145*, 102742. https://doi.org/10.1016/j.apgeog.2022.102742

Zhu, Z., Wulder, M. A., Roy, D. P., Woodcock, C. E., Hansen, M. C., Radeloff, V. C., Healey, S. P., Schaaf, C., Hostert, P., Strobl, P., Pekel, J.-F., Lymburner, L., Pahlevan, N., & Scambos, T. A. (2019). Benefits of the free and open Landsat data policy. *Remote Sensing of Environment*, *224*, 382–385. https://doi.org/10.1016/j.rse.2019.02.016

Zielstra, D., Hochmair, H. H., & Neis, P. (2013). Assessing the Effect of Data Imports on the Completeness of OpenStreetMap – A United States Case Study. *Transactions in GIS*, *17*(3), 315–334. https://doi.org/10.1111/tgis.12037

Zook, M., Graham, M., Shelton, T., & Gorman, S. (2010). Volunteered Geographic Information and Crowdsourcing Disaster Relief: A Case Study of the Haitian Earthquake. *World Medical & Health Policy*, *2*(2), 7–33. https://doi.org/10.2202/1948-4682.1069

Žuffová, M. (2020). *Governments in the spotlight? : On the use and impacts of freedom of information laws and proactive publication of government data* [PhD thesis]. University of Strathclyde.

Zuiderwijk, A., & de Reuver, M. (2021). Why open government data initiatives fail to achieve their objectives: Categorizing and prioritizing barriers through a global survey. *Transforming Government: People, Process and Policy*, *15*(4), 377–395. https://doi.org/10.1108/TG-09-2020-0271

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of E-Government*, *10*(2), Article 2. https://www.academic-publishing.org/index.php/ejeg/article/view/571

Appendices

APPENDIX A

Chapter 2: The search strategy

Database	Search query	Search in	Results	Notes
Scopus	TITLE-ABS-KEY ((“open data” OR “open government data”) AND (“intermediary” OR “intermediation” OR “infomediary”))	title, abstract, keywords	76	Using the singular form of a word in the search in Scopus gives the singular, plural, and possessive forms of most words
WoS	TS=((“open data” OR “open government data”) AND (“intermediaries” OR “intermediary” OR “intermediation” OR “infomediaries” OR “infomediary”))	title, abstract, author keywords, and Keywords Plus	47	
Google Scholar	allintitle: open + data + (intermediaries OR intermediary OR intermediation OR infomediaries OR infomediary)	title	44	Google Scholar only allows terms searched either in the title or in the whole publication. The latter will give about 16,900 publications, hence, the search is only done in the title.
OATD	abstract:(“open data” AND intermediaries) OR abstract:(“open data” AND intermediary) OR abstract:(“open data” AND intermediation) OR abstract:(“open data” AND infomediaries) OR abstract:(“open data” AND infomediary) OR abstract:(“open government data” AND intermediaries) OR abstract:(“open government data” AND intermediary) OR abstract:(“open government data” AND intermediation) OR abstract:(“open government data” AND infomediaries) OR abstract:(“open government data” AND infomediary)	abstract	9	In OATD, we cannot conduct the search based on abstract and title at the same time. However, based on our check, in our case, conducting the search based on the abstract will include results if we were to conduct the search based on the title.

APPENDIX B

Chapter 2: Definitions of open data intermediaries from the literature reviewed

No.	Literature	Definition	Inspired* or adopted by
1.	Chattapadhyay (2014)	Open data intermediaries are 'organisations that share data for its access, consumption and re-usage (including re-sharing) by other organisations and individuals'	Enaholo (2017)*
2.	Janssen & Zuiderwijk (2014)	Infomediaries are those involved in 'the handling of information between information providers and consumers'	Sangiambut & Sieber (2017)
3.	da Silva Craveiro & Albano (2015)	Open data intermediaries are 'all the players (in an individual way or representatives of governments and social organisations), who are involved with public data that are released in an open format. They may or may not make use of technological, legal or structural artifacts in their activities. In making use of open data, the intermediaries aggregate value to the data to ensure that they can be understood more easily (and hence have a greater value) [by] third parties after their intervention'	Nil
4.	González-Zapata & Heeks (2015)	Open government data intermediaries are 'all actors that assist OGD [open government data] initiatives by bridging the barriers that separate public sector data producers and civil society data consumers'	Nil
5.	van Schalkwyk, Chattapadhyay, Cañares, et al. (2015) & van Schalkwyk, Cañares, et al. (2016)	An open data intermediary is 'an agent (i) positioned at some point in a data supply chain that incorporates an open dataset, (ii) positioned between two agents in the supply chain, and (iii) facilitates the use of open data that may otherwise not have been the case'	da Silva Craveiro & Albano (2017), Andrason & van Schalkwyk (2017), Maail (2017), Yoon et al. (2018), den Haan (2018), Enaholo & Dina (2020)
6.	Brugger et al. (2016)	Open government data intermediaries are 'actors who bridge gaps between data producers (governments) and data users (civil society) in that they supply essential resources and capabilities necessary to turn government data into development actions and results'	Nil
7.	Meng (2016)	A government data intermediary is 'an actor that bridges the gap between marginalized groups and OGD [open government data] by facilitating physical access, technical capacity, and value for use of information'	Nil
8.	Schrock & Shaffer (2017)	Open government data intermediaries are 'extra-institutional actors that translate, use, or otherwise mediate communication using data produced by or for government'	Nil
9.	Enaholo (2017)	Open data intermediaries are 'those who operate within the open data ecosystem by means of their contribution, in one way or the other, to the supply of open data by governments as well as to the demand for such data by citizens'	Nil

>>>

No.	Literature	Definition	Inspired* or adopted by
10.	Johnson & Greene (2017)	Infomediaries are 'specific categories of open data users who extract, aggregate, and transform data, altering it into a format that is seen as valuable, beneficial, and, most important, usable to the general public'	Gao & Janssen (2022)
11.	Robinson & Mather (2017)	A civic infomediary is 'a person or organisation that connects community members with open data so that public value can be derived from the data'	Nil
12.	Balvert & van Maanen (2019)	An open government data intermediary is 'the in-between actor standing between a government and a citizen in the process of data communication'	Nil

Note: 'Inspired' means the publication develops a new definition based on the definition offered by the source publication whereas 'Adopted by' means the publication follows entirely the definition offered by the source publication.

APPENDIX C

Chapter 3: Interview guide

Background

1 Could you please describe the function of <organisation>?

2 Could you please describe your role in <organisation>?

3 How long have you been working in this or a similar role?

4 How is your role related to open data?

Open data

5 How long has <organisation> been implementing open data?

6 What do you think is the value of open data to society from the point of view of <organisation>?

7 How does open data implementation benefit <organisation>?

8 How does open data implementation cost <organisation> (financially or others)?

9 What are the key challenges faced by <organisation> related to open data?

Open data ecosystem (defined as a network of interdependent yet self-interested open data actors)

10 What is your perception of the health or sustainability of the current open data ecosystem?

11 What do you think can be improved in the current open data ecosystem? You are more than welcome to describe more than one aspect.

12 With the development of spatial data infrastructures (SDIs) in Europe, especially since INSPIRE, what would you say are key lessons learned for other non-geo open data ecosystems?

Open data intermediaries (defined as third party actors that enhance the supply, access, and/or flow of open data and/or relationships among open data stakeholders)

13 Do you think that open data intermediaries are playing an important and positive role in an open data ecosystem right now? Kindly explain your answer.

14 How do you think open data intermediaries can play a better role in an open data ecosystem?

APPENDIX D

Chapter 3: Challenges in the ODE and a sample quote from the interviews

ID	Challenge	Sample quote from the interviews
User-driven		
C01	**Different data/ metadata standards**	'The INSPIRE directive states you should describe your metadata according to ISO standards. ISO metadata is brilliant in describing spatial metadata but basically it's non-existent for generic administrative open data, they're working with DCAT and all kind of other formats'
C02	**Different open data licenses**	'For instance, we are trying to combine governmental data with data from OpenStreetMap and [they are] different licenses. We have license on [our] governmental data which is very open, it's similar to Creative Commons. While for OpenStreetMap, you might know, it's an ODbL, the share alike, and you cannot combine open data with a very free license with share like.'
C03	**Siloed open data domains (e.g., across sectors)**	'I think another aspect which I think should be very much in focus when we speak about, well, whatever we call them, open data ecosystem or whatever, is the aspect of the domain. We speak about a spatial data ecosystem, but what we see is that – and we have years and many years of building up an infrastructure, spatial data infrastructure – but what we can see is that much of the rapid development that is actually creating value right now, is happening outside the domain. So, I think the biggest risk is that we keep on being a bit siloed, and not well connected'
C04	**High technical threshold to use open data**	'You have to be specialist in order to understand the data, to understand the services involved in accessing the data. So basically it's quite a high threshold for re-use of open data'
C05	**Unfulfilled user needs**	'[There is a] need of focusing attention on users and the use cases, the problems to be solved, rather than focusing on the data provider. Sometimes we forget the use cases and the user and their needs – I think that's an error'
C06	**Limited feedback from lay users (i.e., non-expert users)**	'[Feedback] are mostly from professional users, so to say. So, rarely from citizens, that can be really small companies like independent, […] one man business'
Circular		
C07	**Loss of open data providers' revenue**	'The income from the datasets that turned open data […] for the federal agencies in the states, it may be several millions. For them it's an issue'
C08	**Limited value return from data re-use**	'Now, they [i.e., end-users] are just harvesting a lot of value free of charge. They're not really giving anything back. We have to develop the cooperation further. So it's more also in the mind of the private sector to give back to the ecosystem'
C09	**Limited use case visibility**	'We want to work on more visibility of the re-use of data. We publish data as open data but we don't get in return what [others] did with this data? How they used it? […] This is good for [public agencies] to motivate them [to release open data]. But also to see what has [been] accomplished out of [open data]. Maybe there is something developed [by others] that [public agencies] can use for their work'

>>>

ID	Challenge	Sample quote from the interviews
Inclusive		
C10	**Limited open data from non-government sectors**	‘There are some information that we want to get from the economy [i.e., private sector] but we don’t have this open data law [in] the ecosystems. The law is just for us, for the government to do’
C11	**Lack of incentives for voluntarily publishing open data**	‘If it’s a voluntarily action, and there is no guy, no law forcing you, you need to have benefits and I’m not sure that this works in [redacted] at least’
C12	**Requiring viable business models**	‘You can’t force any of those companies to do it [i.e., be an open data intermediary], because of course they have to find a way to earn money with that as well’
C13	**Overlooked non-government open data**	‘If you want to do an analysis on heat islands within urban environments, if you have only 4% of the trees, you can’t say anything about the impact of having more or less trees on these heat islands within the city. There are at least two private initiatives, by some combinations of Lidar and all kinds of data, they created data sets of basically 99.5 or 99.9% [rough estimates] of the trees. So the data is already there. [But] I think the [typical] reaction of our government: we lack data, we should collect it’
C14	**Practical constraints in multistakeholder engagement**	‘We can always be better in touch with each other. And we are open to it, but also to a certain limit. We can’t talk every week to every software provider or intermediary or whatever’
C15	**Lack of data awareness***	Input obtained from the validation exercise
Skills-based		
C16	**Limited knowledge of open data providers**	‘Some organisations don’t know the proper regulation to apply in every case. And that happened, for example, in the case of metadata, which there is a European regulation directive. Sometimes they are difficult to implement for some local level organisations’
C17	**Limited knowledge of (potential) open data users**	‘I see the ICT domain is really developing a lot of interesting, also ecosystems you could say – [from] cloud, to edge [computing] and whatever technology, and [they’re] super fine systems. But when you then need to add the content, the data, it doesn’t fit because they [i.e., end users] didn’t have any data knowledge’

>>>

TOC

ID	Challenge	Sample quote from the interviews
Foundational		
C18	**Poor open data quality***	Input obtained from the validation exercise
C19	**Incurring maintenance costs for open data provider**	'It does cost something. We have to get the information, we have to make products and we have to service it and platform like [redacted] costs quite a lot of money too. So as long as there's enough budget to keep that running, then it's very sustainable. But it depends on the budget'
C20	**Incurring development costs for open data provider**	'Users get used to have more and more information, they demand the information to have great quality, great updating. The technology goes so fast and they asked to have it more quickly and it is a cost for us to be up to date in the newest technology and progressing day by day'
C21	**Technical difficulties in establishing open data management systems**	'The reality is that a lot of the municipalities and governmental bodies that have to work with data, they are not capable of creating everything from scratch. They don't have the fund, they don't have the staff that's capable of creating a local infrastructure on their own'
C22	**Complex and/or rigid open data standards to comply with**	'We see that OGC [Open Geospatial Consortium] standards nowadays [are] being modularised. Otherwise, the standard was huge and basically it has so many requirements'
C23	**Heterogeneous data administration**	'At the moment we have situation that approximately half of [administrative units] has open data, others have not. The perspective is that, with HVD [high value datasets under the EU Open Data Directive] from June next year, in theory, any of the states would have to. However, we notice that some of the federal states try to escape and they find gaps'
C24	**Privacy concerns**	'Now there's quite a lot of information available and people start combining this information and through combining you can draw more specific conclusions, you get more specific results. [But] which also enter into the privacy of people. We've got maps and there's open data of buildings and addresses. Combining buildings, addresses, maps, aerial images and whatever, you can easily get to where people live and how the environment is – [anything] is possible'
Broad		
C25	**Inflexible/unclear government-market boundary**	'For instance, our aerial image is open data, and there are companies who are providing services with that open data. We now encounter the situation that we want to do similar services for the whole country and we are not allowed to because we've got a law in [redacted] which says there has to be fair play between governments and companies. They make the products and services, they earn money with it, [and] the government is not allowed to give it away for free'
C26	**Reliance on a political agenda**	'If you really want to have a successful digital government, you need serious funding. It needs to be on the political agenda as well. For instance, if you have a look at [redacted] and [redacted], in [redacted], everything about data and digital transformation is a political issue. Their Prime Minister has an IT background. And he's not only the Prime Minister of [redacted], but he's also the Minister of digitalisation. So as a result, there is a completely different mindset within government'
C27	**Inflexible governance/ law (esp. with evolving technology)**	'Changing the law takes years and years and years. Technology, it's going much faster and much higher pace. So, basically, the legal part of the framework cannot keep up with the developments on the technological point of view'

*Note: * were suggested by validators*

APPENDIX E

Chapter 3: Potential contributions of open data intermediaries and a sample quote from the interviews

ID	Potential contribution	Sample quote from the interviews
Technical		
P01	**Implement federated architecture**	'I think we now have concepts like federated architecture that we see that it's OK to have multiple platforms, to have multiple access points for data, and maybe some access points is more from geospatial perspective, and others more from administrative point of view. And it's OK that they are both there as long as they interchange as much as possible with their data. They link to each other, for instance. So there's the principle called, in some [redacted] documents, no wrong door principle. It doesn't matter where you enter as a user, the important role of the infrastructure, it should help you regardless which door you enter'
P02	**Integrate data (e.g., across sectors/ administration)**	'I think the shift that's currently topic of debate is offering services that integrate data. We have a lot of different data sources, we have different base registers or even more data sources, of course, that do not have this formal status but basically they are still organised almost like independent silos. Addresses optimise for addresses alone. Large scale topography with large scale topography. But the user is not interested in [a single dataset], [for instance,] I want to access the address registry or I want topography. [They'd say,] I want to know if they have information about buildings; and maybe if you want to know everything about this building, some of it is [from] the address registries, large scale topography, the small scale typography, the real estate, tax data set'
P03	**Transform data into open standards (esp. web standards)**	'They [i.e., intermediaries] are playing an important role since they have been adapting their tools to open data format although it is true that sometimes it has been difficult for them to get out of proprietary formats'
P04	**Customise data (based on use cases)**	'One of the things users of our open data always ask is to give it in different formats. We provide information in internationally recognised standard formats so that it's open to everybody. We don't do any industry specific formats or company specific formats. We just don't do it. It's not that we are not allowed to, but if we do one we have to do them all and it gets messy and troublesome. So what we want is we provide it in an internationally recognised standards and we want the companies, the markets, to provide it in all the industry standards because we also recognise that our format is not always the most useful for all kinds of users – [for example] architects and building companies and whatever they want it in DXF and anything. Now we've got 3D and they want it in BIM IFC and whatever'
P05	**Offer process automation**	'Once we provide the information, we have to give a step more, allowing to taking advantage of the big data based technologies and to create automatic processes through artificial intelligence and using clouds as technological support'
P06	**Develop open-source tooling**	'An intermediary like [redacted] is also really important because they're a huge driver towards open standards and open formats. Maybe even more open source tooling in the future'
P07	**Provide direct technical services***	Input obtained from the validation exercise
P08	**Offer freemium data platform***	Input obtained from the validation exercise

>>>

ID	Potential contribution	Sample quote from the interviews
Non-technical		
P09	**Foster public-private collaboration**	‘Public-private partnerships are, in my view, a key mechanism for enhancing the role that data intermediaries can play in an open data ecosystem. They could be defined as long-term contracts between a government agency and a private entity with the objective of providing a public asset or service, and in which the private party assumes a significant portion of the responsibility, risks and, generally, the potential benefits’
P10	**Foster public-civic collaboration**	‘We want to know what the civil people need so that the government can implement it. We want to know where a lift doesn’t work so we can repair these lifts. So we want to collect this information, this data, from the civil people, so we can do better services for the people’
P11	**Implement multistakeholder collaboration**	‘So our role is to create the standards and also do that always in an open process so that all stakeholders are involved and are participating from the start, because it’s really important to invest in participation from the beginning of the process because it influences the uptake of these standards afterwards. Because otherwise it’s always – ahh it’s not invented here, my standard is better, I’m already using this’
P12	**Perform open data advocacy***	Input obtained from the validation exercise
P13	**Invest in open data-based civic tech**	‘My goal as an open data officer for [redacted], is to improve the work of the civic society, the civic tech, to have certain kind of an exchange where they can tell me on which civic tech projects they’re working and which civic tech projects need to be implemented [by] the government themselves. […]. Therefore, I have these heck days where I have a good exchange [on] which ideas they’re generating to profit from it and go to the government and tell that these are projects where [government] can put money into and what we [i.e., civic organisations] can implement’
P14	**Showcase open data value**	‘Show benefits. Show countries where it works. I think <redacted> would be a good example because they have open data, they have it nationwide, they have good quality’
P15	**Promote open non-governmental data**	‘The intermediaries can work also from the economy [i.e., private] side or the civil side on working with open data, on returning open data to the government’

>>>

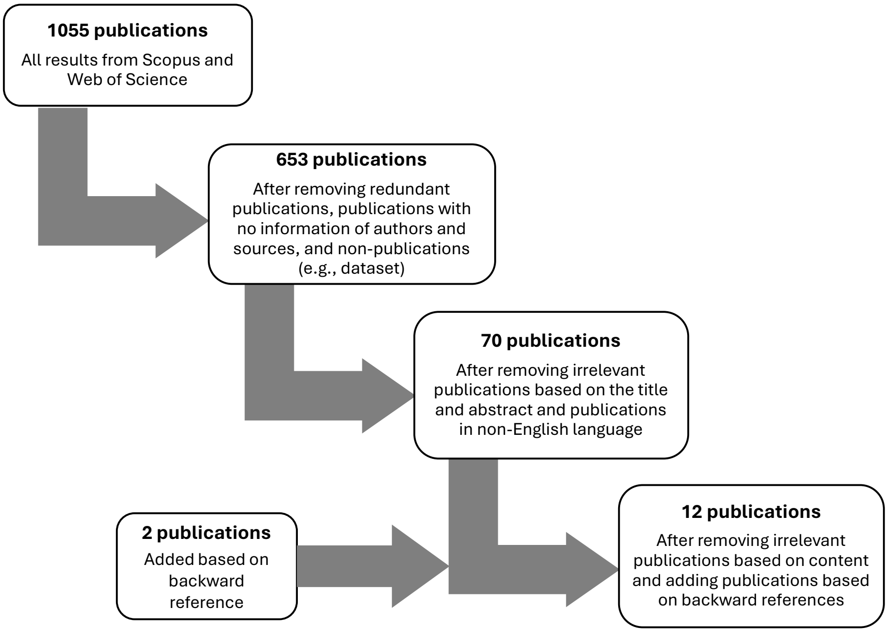
TOC

ID	Potential contribution	Sample quote from the interviews
Combination		
P16	**Provide consultancy**	'We also have a consultancy, [redacted] for example, consulting the government [redacted] [following] the law. But a lot of companies from the economy [i.e., private sector] asked whether there can be an open data consultancy for small and medium-sized companies and also for startups, showing them how they can work with data and why open data is very important'
P17	**Streamline cross-administrative processes**	'Disasters never stopped at administrative boundaries, so you should be able to combine data across borders and it need to be available because as soon as the disaster hit, you don't have time to start thinking about the creations of data or publication of data; it should be made available already. So, in the case of emergency you can just access the data immediately'
P18	**Facilitate internal re-use of open data**	'We have been running this geoportal only four years. And in four years we became the main information system within the City Council […]. Our colleagues have been giving feedback to us. They say that running their own businesses have changed since they started to use geographical information systems because -- they had numbers, […], for example, it's very clear when you see an image of what you are spending money on and how it's related to [impacts in] some boroughs'
P19	**Facilitate feedback on open data**	'Those kinds of user request always pop up on a service provider and not on the individual data provider. [From data provider's point of view:] I'm only concerned with addresses, so combining it with other data, don't ask me. So intermediaries play a really important role in the transformation [from] being a supply driven infrastructure into a much more demand driven infrastructure. Because now users have an entry point where they can post a request where they can say, this is not a nice format [and] it would really be helpful if the data is available in that [other] format as well'
P20	**Provide education on data literacy and skills***	Input obtained from the validation exercise

*Note: * were suggested by validators*

APPENDIX F

Chapter 4: Filtering decisions of the systematic literature review



APPENDIX G

Chapter 4: Codebook by the end of Stages 1, 2 and 3

Initial codebook by the end of Stage 1	Codebook by the end of Stage 2	Codebook by the end of Stage 3
Value proposition		
Type of open data products	**Type of main open data-based product**	**Type of main open data-based product**
Single-purpose app	D-D Aggregated/refined data: fully open	D-D Aggregated/refined data: fully open
Interactive app	D-D Aggregated/refined data: with restrictions	D-D Aggregated/refined data: with restriction
Information aggregator	D-D Data repository: fully open	D-D Data repository: fully open
Comparison model	D-D Data platform: fully open	D-D Data platform: fully open
Data repository	D-D Data repository: with restrictions	D-D Data repository: with restrictions
Data platform	D-D Data platform: with restrictions	D-D Data platform: with restrictions
Data refining	D-INFO Single-purpose app	D-INFO Single-purpose app
Advocacy	D-INFO Interactive app	D-INFO Interactive app
Consultancy	D-INFO Insights	D-INFO Insights
Business intelligence	D-K Consultancy	D-K Consultancy
Process optimisation	D-K Intelligence service	D-K Intelligence service
Product/service improvement	D-K Journalism	D-K Journalism
Research and development	SU Product/service development	SU Product/service development
	SU Training/courses	SU Training/courses
	SU Advocacy	SU Advocacy
	Source of data	**Source of data**
	Only open data	Only open data
	Combined with non-open data	Combined with non-open data
	Product components	**Product components**
	Single product	Single product
	Multiple units	Multiple units
	Other open data-based product	**Other open data-based product**
	Offered by the organisation	Offered by the organisation
	Not offered by the organisation	Not offered by the organisation
	Non-open data-based product	**Non-open data-based product**
	Offered by the organisation	Offered by the organisation
	Not offered by the organisation	Not offered by the organisation
	Linking of other product(s) to the main open data-based product	**Linking of other product(s) to the main open data-based product**

>>>

Initial codebook by the end of Stage 1	Codebook by the end of Stage 2	Codebook by the end of Stage 3
Value proposition		
	Necessary	Necessary
	Complementary	Complementary
	Unrelated	Unrelated
	Not applicable	Not applicable
Offering	**Offering**	**Offering**
Newness	Newness	Newness
Performance	Performance	Performance
Customisation	Customisation	Customisation
Getting the job done	Getting the job done	Getting the job done
Design	Design	Design
Price	Price	Price
Cost reduction	Cost reduction	Cost reduction
Risk reduction	Risk reduction	Risk reduction
Accessibility	Accessibility	Accessibility
Convenience/usability	Convenience/usability	Convenience/usability
Brand	Brand	Brand
Security and privacy	Security and privacy	Security and privacy
Responsiveness/feedback	Responsiveness/feedback	Responsiveness/feedback
Channel	**Channel**	*Deselected*
Direct	Direct	
Indirect	Indirect	
	Direct and indirect	
Consumer segment	**Consumer segment**	**Consumer segment**
Data users: specific	Data users: highly skilled	Data users: highly skilled
Data users: general	Data users: generic	Data users: generic
Data providers: specific	Data providers: highly skilled	Data providers: highly skilled
Data providers: general	Data providers: generic	Data providers: generic
	Data providers and users: highly skilled	Data providers and users: highly skilled
	Data providers and users: generic	Data providers and users: generic

Note: Bold items are codebook categories, with codebook elements listed below them. Definitions for each element of the codebook by the end of Stage 3 are provided in Appendix H.

Initial codebook by the end of Stage 1	Codebook by the end of Stage 2	Codebook by the end of Stage 3
Value creation		
Key partners	**Critical partner**	*Deselected*
Open coding	Data providers: text (e.g., minutes)	
	Data providers: mix of structured and unstructured data	
	Data providers: geospatial data	
	Data providers: structured data	
	Funder/supporter: joint public organisations	
	Funder/supporter: public organisation	
	Funder/supporter: multiple public organisations	
	Funder/supporter: private donors/ foundations	
	Commercial service providers	
	Non-commercial service provider	
	Subject matter experts (not in-house)	
	Third-party/end-distributor	
	Brokerage partner	
	Sector-wide partners/community	
	Parent company	
	University/research institution	
Key activities	**Critical stage of the open data lifecycle**	**Critical stage of the open data lifecycle**
Open coding	Identification	Identification
	Preparation	Preparation
	Publication	Publication
	Re-use	Re-use
	Evaluation	Evaluation
Key resources	**Critical resources (other than financial)**	*Deselected*
Open coding	Physical	
	Specialised skills: data scientists	
	Specialised skills: subject matter experts (in-house)	
	Specialised skills: journalists	
	Specialised skills: geospatial specialists	
	Specialised skills: copywriting	
	Specialised skills: software engineers	
	Specialised skills: designers	
	Institutional: partnership with data providers	
	Institutional: human resource	

>>>

Initial codebook by the end of Stage 1	Codebook by the end of Stage 2	Codebook by the end of Stage 3
Value creation		
	Human: volunteers/community	
	Infrastructure: external service/platform	
	(Near) real-time data	
Customer relationship	**Customer relationship**	**Customer relationship**
Personal assistance	Personal assistance	Personal assistance
Dedicated personal assistance	Self-service	Self-service
Self-service	Communities	Communities
Automated services	Co-creation	Co-creation
Communities		
Co-creation		
Value capture		
Cost structure	Cost structure	*Deselected*
Fixed costs	Fixed costs	
Variable costs	Variable costs	
Economies of scale		
Economies of scope		
Revenue streams	**Main revenue stream**	**Main revenue stream**
Asset sale	Asset sale	Asset sale
Usage fees	Usage fee/service delivery	Usage fee/service delivery
Subscription fees	Subscription fee	Subscription fee
Lending/renting/leasing	Brokerage fees	Brokerage fees
Licensing	Advertising	Advertising
Brokerage fees	Sponsorship: private	Sponsorship: private
Advertising	Sponsorship: public	Sponsorship: public
Sponsorship	Sponsorship: public and private	Sponsorship: public and private
Dual licensing	Freemium	Freemium
Freemium	Membership fees	Membership fees
Membership	Volunteer	Volunteer
Voluntary/community	Crowdfunding	Crowdfunding
Royalty	Cross-subsidy	Cross-subsidy
Razor and blades		
Crowdfunding		
	Source of revenue	**Source of revenue**
	Solely from (augmented) open data	Solely from (augmented) open data
	Other sources besides (augmented) open data	Other sources besides (augmented) open data
	Not relevant (for main revenue stream: volunteer)	Not relevant (for main revenue stream: volunteer)

Note: Bold items are codebook categories, with codebook elements listed below them. Definitions for each element of the codebook by the end of Stage 3 are provided in Appendix H.

APPENDIX H

Chapter 4: Definitions for elements of the codebook by the end of Stage 3

Codebook categories/elements	Description
Type of main open data-based product	
D-D Aggregated/refined data: fully open	Pre-processed open data, available to anyone without a fee.
D-D Aggregated/refined data: with restriction	Pre-processed open data, available to fee-paying consumers.
D-D Data repository: fully open	Online portal to access data without restrictions (including fees) on the data or the portal.
D-D Data platform: fully open	Online portal to access and share data without restrictions (including fees) on the data or the portal.
D-D Data repository: with restrictions	Online portal to access data with restrictions (e.g., fees) on the data or the portal.
D-D Data platform: with restrictions	Online portal to access and share data with restrictions (e.g., fees) on the data or the portal.
D-INFO Single-purpose app	Mobile or web application with limited—often unidirectional—functionalities that transform data into ready-to-consume information.
D-INFO Interactive app	Mobile or web application with more interactive functionalities that transform data into ready-to-consume information.
D-INFO Insights	Report or similar forms of material providing key information on certain items (e.g., business registration) based on (augmented) open data.
D-K Consultancy	Customised service facilitating the process of using or providing open data.
D-K Intelligence service	Customised service to acquire and apply knowledge based on open data (and other data).
D-K Journalism	Journalism products such as news content based on (augmented) open data.
SU Product/service development	Support products/services to facilitate open data publication or re-use.
SU Training/courses	Training/courses related to open data publication or re-use.
SU Advocacy	Open data advocacy, campaigning, or lobbying.
Source of data	
Only open data	Only open data are used.
Combined with non-open data	Open data are used in combination with non-open data.
Product components	
Single product	The product analysed is a single product.
Multiple units	The product analysed has multiple modular units within it.

>>>

Codebook categories/elements	Description
Other open data-based product	
Offered by the organisation	The organisation offers other open data-based products.
Not offered by the organisation	The organisation does not offer other open data-based products.
Non-open data-based product	
Offered by the organisation	The organisation offers non-open data-based products.
Not offered by the organisation	The organisation does not offer non-open data-based products.
Link of other product(s) to the main open data-based product	
Necessary	The other open data-based or non-open data-based product offered by the organisation is essential to the product analysed.
Complementary	The other open data-based or non-open data-based product offered by the organisation complements the product analysed.
Unrelated	The other open data-based or non-open data-based product offered by the organisation is not related to the product analysed.
Not applicable	The organisation does not offer other open data-based or non-open data-based products.
Offering [The elements and their descriptions in this category are largely adopted from Osterwalder and Pigneur (2010)]	
Newness	Satisfy an entirely new set of needs that consumers did not previously perceive.
Performance	Improve processes and organisational performance.
Customization	Tailor products or services to the specific needs of consumers.
Getting the job done	Help consumers get certain jobs done.
Design	Product offers a more seamless or user-friendly design.
Price	Offer products or services at a lower price than others.
Cost reduction	Offer products or services that reduce the development or operational costs of an organisation.
Risk reduction	Offer products or services that reduce certain risks.
Accessibility	Make data, products, or services available to consumers who previously lacked access to them.
Convenience/usability	Offer products or services that make certain aspects more convenient.
Brand	Product or service is well known for its brand or popularity.
Security and privacy	Enhance data security or privacy.
Responsiveness/feedback	Facilitate the channelling of responses or feedback.
Consumer segment	
Data users: highly skilled	Highly skilled data users, typically from professional domains.
Data users: generic	Generic data users, including highly skilled and lay people.
Data providers: highly skilled	Highly skilled data providers, typically from professional domains.
Data providers: generic	Generic data providers, including highly skilled and lay people.
Data providers and users: highly skilled	Highly skilled data providers and users, typically from professional domains.
Data providers and users: generic	Generic data providers and users, including highly skilled and lay people.

>>>

TOC

Codebook categories/elements	Description
Critical stage of the open data lifecycle	
Identification	Activities include demanding open data, facilitating stakeholders' interactions, and identifying the risks of opening data.
Preparation	Activities include compiling data, building data capacity, and augmenting data (by combining with non-open data).
Publication	Activities include curating data and improving the technical openness of data.
Re-use	Activities include building data capacity, contextualizing data, developing products and services, interpreting data, visualising data, and facilitating stakeholders' interactions.
Evaluation	Activities include validating data and channelling feedback.
Customer relationship	
Personal assistance	This relationship is based on human interaction. The consumers can communicate with a real representative to get help during the sales/consultancy process.
Self-service	The organisation provides all the necessary means for consumers to help themselves.
Communities	All members, theoretically have a more or less equal opportunity to influence the governance mechanisms and processes.
Co-creation	A lead body facilitates the contribution and use of data, products, and services.
Main revenue stream	
Asset sale	Sell a product such as software.
Usage fee/service delivery	Charge fees based on product use or service delivery.
Subscription fees	Charge regular (e.g., monthly or annual) fees.
Brokerage fees	Obtain third-party commission (e.g., from train operators by facilitating the sale of train tickets).
Advertising	Charge fees for advertisement.
Sponsorship: private	Obtain sponsorship from the private sector, including private foundations.
Sponsorship: public	Obtain sponsorship from government organisations.
Sponsorship: public and private	Obtain sponsorship from the private sector and government organisations.
Freemium	Offer initial or minimal functionalities of the product or service for free, with the extended version requiring a fee.
Membership fees	Obtain regular (e.g., monthly or annual) fees from registered members.
Volunteer	Resources are provided for free by volunteers.
Crowdfunding	Obtain funds by asking for donations.
Cross-subsidy	Funds are obtained from another unit within the organisation or its parent or sister organisation.
Source of revenue	
Solely from (augmented) open data	Revenue or funding is obtained solely from/for the (augmented) open data.
Other sources besides (augmented) open data	Revenue or funding is also obtained from other sources apart from/for the (augmented) open data.
Not relevant (for main revenue stream: volunteer)	No revenue or funds were obtained since volunteers provided resources for free.

APPENDIX I

Chapter 4: Sample cases and their elements based on K-means clustering [with defining categories indicated by D]

Cluster	(No. of cases) Cases with positive salient scores ≥ 0.04	Categories: Elements (No. of cases) – D indicates a defining category – The bold text indicates the common (also via combination) element(s)	
C1	(9) OpenActive, Clear-HealthCosts, active//choice, Wikidata, OpenLitterMap, LG Inform, wheelmap.org, Confiscati Bene, OpenStreetMap, Open-CycleMap.org	D	*Type of main open data-based product*: **D-D Data platform: fully open** (8), D-INFO Interactive app (1)
		D	*Source of data*: **Only open data** (9)
			Product components: Single product (6), Multiple units (3)
		D	*Other open data-based product*: **Offered by the organisation** (7), Not offered by the organisation (2)
		D	*Non-open data-based product*: **Not offered by the organisation** (8), Offered by the organisation (1)
		D	*Link of other product(s) to the main open data-based product*: **Complementary** (7), Not applicable (2)
			Offering: Accessibility (6), Newness (2), Getting the job done (1)
		D	*Consumer segment*: **Data providers and users: generic** (7), Data users: generic (1)
		D	*Critical stage of the open data lifecycle*: **Preparation** (7), Publication (2)
		D	*Customer relationship*: **Co-creation** (5), Self-service (2), **Communities** (2)
		D	*Main revenue stream*: **Crowdfunding** (4), **Sponsorship: public** (2), **Sponsorship: private** (2), **Sponsorship: public and private** (1)
		D	*Source of revenue*: **Only from (augmented) open data** (9)
C2	(8) Opencorporates, CommoPrices, infoem-presam Goolzoom, Aircheckr, Waar-neming, TransportAPI, IPlytics		*Type of main open data-based product*: D-D Data repository: with restrictions (4), D-D Aggregated/refined data: with restrictions (3), D-D Data platform: with restrictions (1)
		D	*Source of data*: **Combined with non-open data** (6), Only open data (2)
		D	*Product components*: **Single product** (7), Multiple units (1)
		D	*Other open data-based product*: **Not offered by the organisation** (8)
		D	*Non-open data-based product*: **Not offered by the organisation** (8)
		D	*Link of other product(s) to the main open data-based product*: **Not applicable** (8)
			Offering: Accessibility (4), Performance (3), Convenience (1)
		D	*Consumer segment*: **Data users: highly skilled** (5), **Data users: generic** (2), Data providers and users: generic (1)
		D	*Critical stage of the open data lifecycle*: **Preparation** (8)
		D	*Customer relationship*: **Self-service** (6), Personal assistance (2)
			Main revenue stream: Freemium (5), Subscription (2), Sponsorship: Private (1)
		D	*Source of revenue*: **Only from (augmented) open data** (6), Other sources besides (augmented) open data (2)

>>>

Cluster	(No. of cases) Cases with positive salient scores ≥ 0.04		Categories: Elements (No. of cases) – D indicates a defining category – The bold text indicates the common (also via combination) element(s)
C3	(22) Golemio, Mime, Populate, GIS4tech, Piperlab, Terranea, Gispo, Basedig, TPXimpact, FixMyCity, Geomatic, Open-datasoft, Synapta, Smartvel, Guadaltel, Whythawk, Air & Space Evidence, MAPEGY, OpenMove, DEPP, dataninja, Waterjade		*Type of main open data-based product*: SU Product/service development (10), D-K Consultancy (8), SU Training/courses (1), D-K Intelligence service (1), D-INFO Interactive app (1), D-D Data platform: with restrictions (1)
		D	*Source of data*: **Combined with non-open data** (16), Only open data (5)
		D	*Product components*: **Multiple units** (21), Single product (1)
		D	*Other open data-based product*: **Not offered by the organisation** (22)
		D	*Non-open data-based product*: **Not offered by the organisation** (22)
		D	*Link of other product(s) to the main open data-based product*: **Not applicable** (22)
			Offering: Customization (7), Performance (6), Getting the job done (6), Responsiveness (1), Newness (1), Cost reduction (1)
			Consumer segment: Data providers and users: generic (7), Data users: highly skilled (4), Data providers and users: highly skilled (4), Data users: generic (3), Data providers: generic (3), Data providers: highly skilled (1)
			Critical stage of the open data lifecycle: Reuse (13), Preparation (7), Publication (1), Evaluation (1)
		D	*Customer relationship*: **Personal assistance** (21), Self-service (1)
		D	*Main revenue stream*: **Usage fee/service delivery** (16), Subscription (3)
			Source of revenue: Only from (augmented) open data (15), Other sources besides (augmented) open data (7)
C4	(14) EMAIV, Bo-er&Bunder, Zapmap, Farm Dog, UC, LocalFocus, EOS Data Analytics, regionaal energieloket, Hermitage Maps, PlantVillage Nuru, Ex Machina, Geniasis, BlindSquare, FiveThir-tyEight	D	*Type of main open data-based product*: **D-INFO Interactive app** (12), D-K Journalism (1), D-INFO Insights (1)
			Source of data: Combined with non-open data (9), Only open data (3)
		D	*Product components*: **Single product** (12), Multiple units (2)
		D	*Other open data-based product*: **Offered by the organisation** (10), Not offered by the organisation (2)
		D	*Non-open data-based product*: **Offered by the organisation** (12), Not offered by the organisation (0)
		D	*Link of other product(s) to the main open data-based product*: **Complementary** (12), Unrelated (2)
			Offering: Performance (6), Getting the job done (4), Customization (2), Accessibility (2)
			Consumer segment: Data users: highly skilled (6), Data users: generic (7), Data providers: highly skilled (1)
		D	*Critical stage of the open data lifecycle*: **Reuse** (14)
		D	*Customer relationship*: **Self-service** (10), Personal assistance (4)
			Main revenue stream: Subscription (4), Sponsorship: Public (2), Asset sale (2), Usage fee/service delivery (1), Sponsorship: Public and private (1), Cross-subsidy (1), Brokerage fees (1)
		D	*Source of revenue*: **Only from (augmented) open data** (9), Other sources besides (augmented) open data (3)

>>>

Cluster	(No. of cases) Cases with positive salient scores ≥ 0.04	Categories: Elements (No. of cases) – D indicates a defining category – The bold text indicates the common (also via combination) element(s)	
C5	(20) Tutki Hankintoja, German Intensive Care Availability Register, Dblp Computer Science Bibliography, Haridussilm, scanR, FRIS Onderzoeksportaal, basemap.at, Statsregnskapet, Medicatio, SNS Transparency Portal, Wegweiser Demokraphie, European Statistical Systems, Offene Vergaben, GlobalStat, IDEA, Geoportal Malopolski, Open Maps for Europe, InfluencAir, Semantic Finlex, Open Spending	D	*Type of main open data-based product*: **D-D Data repository: fully open** (18), D-INFO Interactive app (1), D-D Data repository: with restrictions (1)
		D	*Source of data*: **Only open data** (19), Combined with non-open data (0)
		D	*Product components*: **Single product** (18), Multiple units (2)
			Other open data-based product: Not offered by the organisation (5), Offered by the organisation (4)
			Non-open data-based product: Not offered by the organisation (6), Offered by the organisation (2)
		D	*Link of other product(s) to the main open data-based product*: **Not applicable** (15), Complementary (4), Unrelated (1)
		D	*Offering*: **Accessibility** (16), Convenience (4)
		D	*Consumer segment*: **Data users: generic** (16), Data users: highly skilled (2), Data providers: highly skilled (2)
		D	*Critical stage of the open data lifecycle*: **Preparation** (17), Publication (2), Identification (1)
		D	*Customer relationship*: **Self-service** (20)
		D	*Main revenue stream*: **Sponsorship: public** (15), **Sponsorship: private** (3), Membership fee (1), Asset sale (1)
		D	*Source of revenue*: **Only from (augmented) open data** (14), Other sources besides (augmented) open data (1)
C6	(5) Nasdaq Data Link, Spend Network, ArcGIS, Enigma, Infocif		*Type of main open data-based product*: D-D Data repository: with restrictions (3), D-D Data platform: with restrictions (2)
		D	*Source of data*: **Combined with non-open data** (5)
		D	*Product components*: **Multiple units** (5)
			Other open data-based product: Offered by the organisation (3), Not offered by the organisation (2)
			Non-open data-based product: Offered by the organisation (3), Not offered by the organisation (2)
		D	*Link of other product(s) to the main open data-based product*: **Complementary** (4), Unrelated (1)
			Offering: Accessibility (3), Performance (1), Convenience (1)
		D	*Consumer segment*: **Data users: highly skilled** (3), Data users: generic (1), **Data providers and users: highly skilled** (1)
		D	*Critical stage of the open data lifecycle*: **Preparation** (5)
		D	*Customer relationship*: **Self-service** (4), Personal assistance (1)
			Main revenue stream: Subscription fee (3), Asset sale (2)
			Source of revenue: Other sources besides (augmented) open data (3), Only from (augmented) open data (2)

>>>

TOC

Cluster	(No. of cases) Cases with positive salient scores ≥ 0.04	Categories: Elements (No. of cases) – D indicates a defining category – The bold text indicates the common (also via combination) element(s)	
C7	(18) QEdu, GaugeMap, Shock, Kannattaako kauppa, UV Index Widget, FloodAlert, Know Your Hoods, Mapa delinquencial, BrokenLifts, Swis-sTrains, LondonTu-be, DinGeo, Vektis, openRaadsinformatie, OpenTrees.org, Romanian Railways, Umweltzone, ÖPNV	D	*Type of main open data-based product*: **D-INFO Single-purpose app** (15), D-D Aggregated/refined data: fully open (2), SU Product/service development (1)
		D	*Source of data*: **Only open data** (16), Combined with non-open data (0)
		D	*Product components*: **Single product** (17), Multiple units (1)
		D	*Other open data-based product*: **Offered by the organisation** (14), Not offered by the organisation (0)
		D	*Non-open data-based product*: **Offered by the organisation** (15), Not offered by the organisation (2)
			Linking of product(s) to the main open data-based product: Unrelated (11), Complementary (7)
			Offering: Convenience (7), Accessibility (4), Design (3), Newness (2), Getting the job done (2)
		D	*Consumer segment*: **Data users: generic** (17), Data providers: generic (1)
		D	*Critical stage of the open data lifecycle*: **Reuse** (12), Preparation (6)
		D	*Customer relationship*: **Self-service** (17), Communities (1)
			Main revenue stream: Cross-subsidy (7), Volunteer (3), Freemium (3), Sponsorship (Private) (2), Sponsorship (Public) (1), Membership fees (1), Advertising (1)
			Source of revenue: Other sources besides (augmented) open data (11), Only from (augmented) open data (5), Not relevant (1)
C8	(25) Bomstasjon.no, UpDownVienna, Tutiempo, Hooikoorts-radar, Urbonaut, IrCELine, GetThere, NCSE School Informa-tion Map, Geburtss-pital, Open Court, Kanarek, FindToilet, ISS detector, Rijden de treinen, Tree of Truth, Cyprus by Bus, Leefbaarometer, Alim confiance, Atlatszo, HDScores, Transfer-muga, 1848, Future Readiness Index, GapMinder, Civio	D	*Type of main open data-based product*: **D-INFO Single-purpose app** (23), D-K Journalism (2)
		D	*Source of data*: **Only open data** (21), Combined with non-open data (1)
		D	*Product components*: **Single product** (22), Multiple units (3)
		D	*Other open data-based product*: **Not offered by the organisation** (15), Offered by the organisation (0)
		D	*Non-open data-based product*: **Not offered by the organisation** (15), Offered by the organisation (0)
		D	*Linking of product(s) to the main open data-based product*: **Not applicable** (25)
			Offering: Convenience (10), Accessibility (10), Newness (3), Performance (1), Design (1)
		D	*Consumer segment*: **Data users: generic** (25)
		D	*Critical stage of the open data lifecycle*: **Reuse** (24), Preparation (1)
		D	*Customer relationship*: **Self-service** (25)
			Main revenue stream: Sponsorship: public (6), Sponsorship: private (5), Freemium (5), Volunteer (4), Crowdfunding (2), Subscription (1), Advertising (1)
			Source of revenue: Only from (augmented) open data (14), Other sources besides (augmented) open data (3), Not relevant (1)

>>>

Cluster	(No. of cases) Cases with positive salient scores ≥ 0.04		Categories: Elements (No. of cases) – D indicates a defining category – The bold text indicates the common (also via combination) element(s)
C9	(16) The Smartfiles Network, PravoSud, Taranis, GREENSPIN, Sun Energia, Agroknow, LOCATA Analytics, OmgevingsAlert, TopPlace, HogeNood, Hipcamp, geoFluxus, 9292, Sud na Doloni, Bike Citizens, Waar is mijn stemlokaal	D	*Type of main open data-based product*: **D-INFO Interactive app** (15), D-D Data repository: with restrictions (1)
			Source of data: Only open data (8), Combined with non-open data (3)
		D	*Product components*: **Single product** (13), Multiple units (3)
		D	*Other open data-based product*: **Not offered by the organisation** (15)
		D	*Non-open data-based product*: **Not offered by the organisation** (15)
		D	*Linking of product(s) to the main open data-based product*: **Not applicable** (16)
			Offering: Convenience (10), Performance (3), Customization (3)
			Consumer segment: Data users: generic (7), Data users: highly skilled (6), Data providers and users: generic (2), Data providers: highly skilled (1)
		D	*Critical stage of the open data lifecycle*: **Reuse** (16)
			Customer relationship: Self-service (11), Personal assistance (5)
			Main revenue stream: Subscription fee (8), Brokerage (2), Asset sale (2), Usage fee/service delivery (1), Sponsorship: public and private (1), Freemium (1)
		D	*Source of revenue*: **Only from (augmented) open data** (16)
C10	(9) Publish What You Pay, Open Data Watch, Open State Foundation, K-Monitor, Code for Ghana, Bellingcat, The Institute for Development of Freedom of Information (IDFI), SocialBoost, Open Data Institute	D	*Type of main open data-based product*: **SU Advocacy** (7), **SU Product/service development** (1), D-K Journalism (1)
		D	*Source of data*: **Only open data** (7), Combined with non-open data (2)
		D	*Product components*: **Multiple units** (9), Single product (0)
		D	*Other open data-based product*: **Not offered by the organisation** (9), Offered by the organisation (0)
		D	*Non-open data-based product*: **Not offered by the organisation** (9), Offered by the organisation (0)
		D	*Linking of product(s) to the main open data-based product*: **Not applicable** (9)
			Offering: Getting the job done (4), Accessibility (3), Responsiveness (1), Newness (1)
			Consumer segment: Data providers and users: generic (4), Data users: generic (3), Data providers: generic (1), Data providers and users: highly skilled (1)
			Critical stage of the open data lifecycle: Identification (5), Reuse (2), Evaluation (2)
			Customer relationship: Communities (5), Self-service (2), Personal assistance (1), Co-creation (1)
		D	*Main revenue stream*: **Sponsorship: public and private** (4), Usage fee/service delivery (2), **Sponsorship: private** (2), **Crowdfunding** (1)
		D	*Source of revenue*: **Only from (augmented) open data** (9)

Curriculum vitae

Ashraf Shaharudin

1992	Born in Johor Bahru, Malaysia
2012 – 2015	Bachelor of Engineering (Electrical and Electronic) The University of Western Australia, Perth, Australia
2016 – 2018	Master of Economics Universiti Putra Malaysia, Serdang, Malaysia
2018 – 2021	Research Associate Khazanah Research Institute, Kuala Lumpur, Malaysia
2022 – 2025	PhD Researcher Faculty of Architecture and the Built Environment, Delft University of Technology, Netherlands

List of publications

Peer-reviewed publications

- Shaharudin, A., van Loenen, B., & Janssen, M. (2023). Towards a Common Definition of Open Data Intermediaries. *Digital Government: Research and Practice*, 4(2), 6:1-6:21. https://doi.org/10.1145/3585537

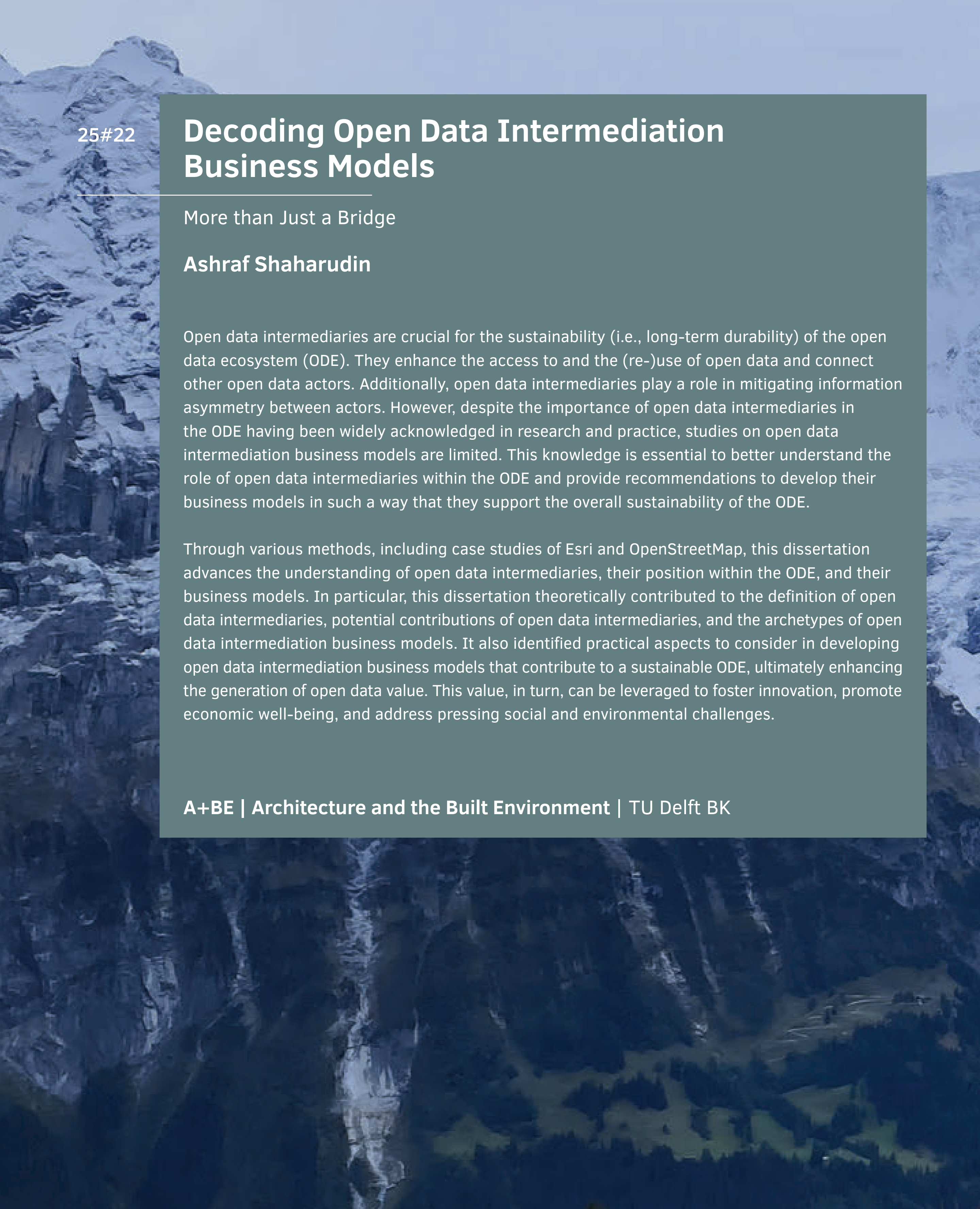
- Shaharudin, A., van Loenen, B., & Janssen, M. (2024). Exploring the contributions of open data intermediaries for a sustainable open data ecosystem. *Data & Policy*, *6*, e56. https://doi.org/10.1017/dap.2024.63

- Shaharudin, A., van Loenen, B., & Janssen, M. (2025). Developing an Open Data Intermediation Business Model: Insights From the Case of Esri. *Transactions in GIS*, 29(1), e13304. https://doi.org/10.1111/tgis.13304

Other publications

- Shaharudin, A., van Loenen, B., & Janssen, M. (2023). Identifying Business Models of Open Data Intermediaries: A Review. International Conference on Open Data: Open Data Challenges and Opportunities in Times of Crisis and Growth (ICOD 2022), Zagreb, Croatia. Zenodo. https://doi.org/10.5281/zenodo.8063619

- Shaharudin, A., López Reyes, M. E., Pantazatou, K., Storm, I., Larsen, B., van Loenen, B., & Kronborg Mazzoli, U. (2024). Workshop on Data Ecosystems and Spatial Data Infrastructure - Facilitators for Data Value Creation. European Spatial Data Research (EuroSDR). https://doi.org/10.5281/zenodo.11148721



25#22

Decoding Open Data Intermediation Business Models

More than Just a Bridge

Ashraf Shaharudin

Open data intermediaries are crucial for the sustainability (i.e., long-term durability) of the open data ecosystem (ODE). They enhance the access to and the (re-)use of open data and connect other open data actors. Additionally, open data intermediaries play a role in mitigating information asymmetry between actors. However, despite the importance of open data intermediaries in the ODE having been widely acknowledged in research and practice, studies on open data intermediation business models are limited. This knowledge is essential to better understand the role of open data intermediaries within the ODE and provide recommendations to develop their business models in such a way that they support the overall sustainability of the ODE.

Through various methods, including case studies of Esri and OpenStreetMap, this dissertation advances the understanding of open data intermediaries, their position within the ODE, and their business models. In particular, this dissertation theoretically contributed to the definition of open data intermediaries, potential contributions of open data intermediaries, and the archetypes of open data intermediation business models. It also identified practical aspects to consider in developing open data intermediation business models that contribute to a sustainable ODE, ultimately enhancing the generation of open data value. This value, in turn, can be leveraged to foster innovation, promote economic well-being, and address pressing social and environmental challenges.

A+BE | Architecture and the Built Environment | TU Delft BK