All in? Identifying and tackling private sector's barriers to data sharing: A Perspective on geospatial data in the Netherlands

Chrysanthi Papadimitriou

First Supervisor
Dr.ir. B. van Loenen
Second Supervisor
Dr. S. Calzati

TU Delft Delft University of Technology

Co-reader
Dr. A. Rafiee

Presentation Overview

- Introduction
- Research Questions
- Related work & Literature review
- Private Sector Data Sharing Case Studies
- Methodology
- Results & Discussion
- Conclusion, Limitations & Future Work

Introduction

- Movement to create an inclusive open data ecosystem (van Loenen, 2021)
- Plethora of open government data -> PDOK with 239 open geospatial datasets
- Significant amount of data accumulated and managed by private sector (Fassnacht, 2023)
- Barriers in academic and public sector context researched extensively (Laia and Jonathan, 2020).
- EU with ODD (2019), promotes more openness in private data
- Emerging research in public undertakings, sector performs in between public and private (van Veenstra (2013); Boone and van Loenen (2022))
- Scant research on private sector data sharing

Motivation

- **Fill** the government **data gap** with private sector data
- Barriers to **private sector** data sharing **not researched enough**
- **Level of openness of a company** based on an assessment
- Barriers at the level
- How to move to a more open level to arrive in an **ecosystem with open private sector data**

Research Questions

Research Questions:

What are the barriers to private sector data sharing of geospatial data in the Netherlands? **and** *How can those barriers be approached to reach (open) data sharing?*

Sub Questions:

- *What are the different levels of openness for geospatial data?*
- *How do the different barriers relate to each level of openness of the developed model for private sector data sharing?*

Barriers to Data Sharing of the Private sector

- **Data sharing:** process of data and technologies that are used to gather, process and share data, based on a set of skills, tools, privacy and data sharing capacity
- **Public Undertaking:** a non-government party that carriers out a service, which can be a public service that is economic activity of general interest and is defined by public authorities, or it can be a service in areas that are entirely market based where the public authorities play no part
- **Open data:** data that is available in a common, machine-readable format, which anyone can access, use and share without restriction or cost for any purpose.

Security	Waterman et al. (2021)
Legal	Waterman et al. (2021); van Panhuis (2014); Sane and Edelstein (2015); Barry and Bannister (2013)
Privacy	Waterman et al. (2021)
Technical	Waterman et al. (2021); van Panhuis (2014); Sane and Edelstein (2015); Barry and Bannister (2013); Janssen et al. (2012)
Commercial	Waterman et al. (2021)
Cultural	Waterman et al. (2021); Fassnacht et al. (2023); Barry and Bannister (2013)
Strategic	Fassnacht et al. (2023)
Operational	Fassnacht et al. (2023)
Technological	Fassnacht et al. (2023)
Regulatory	Fassnacht et al. (2023)
Motivational	van Panhuis (2014); Sane and Edelstein (2015)
Economic	van Panhuis (2014); Sane and Edelstein (2015); Barry and Bannister (2013); Martin et al. (2013)
Political	van Panhuis (2014); Sane and Edelstein (2015)
Ethical	van Panhuis (2014); Sane and Edelstein (2015)
Information Technology	van Veenstra and van den Broek (2013)
Organizational and Managerial	van Veenstra and van den Broek (2013)
Legal and Regulatory	van Veenstra and van den Broek (2013)
Institutional and Environmental	van Veenstra and van den Broek (2013)
Institutional	Boone and van Loenen (2022); Janssen et al. (2012)
Task Complexity	Boone and van Loenen (2022); Janssen et al. (2012)
Technical Quality	Boone and van Loenen (2022)
Administrative	Barry and Bannister (2013)
Risk related	Barry and Bannister (2013)
Use and Participation	Janssen et al. (2012)
Legislation	Janssen et al. (2012)
Information Quality	Janssen et al. (2012)
Governance	Martin et al. (2013)
Licenses and Legal Frameworks	Martin et al. (2013)
Data Characteristics	Martin et al. (2013)
Metadata	Martin et al. (2013)
Access	Martin et al. (2013)
Skills	Martin et al. (2013)

Barriers to Data Sharing of the Private sector

Strategic	Dataset	Sector
Lack of organizational motivation to enable data sharing	geospatial data, flood risk data, transportation data, geological data	private sector, semipublic sector
Lack of management commitment and corporate strategy	geospatial data, flood risk data	private sector, Public sector
Lack of policy coherence	public health data, geospatial data	private sector, Public sector
Lack of use case identification	geospatial data	private sector, Public sector
Lack of business cases for generating revenue	transportation data, geological data	private sector, semipublic sector
Lack of a feedback process	geospatial data	private sector, public sector

Barriers to Data Sharing of the Private sector

Technical	Dataset	Sector
Lack of data standards	geospatial data, flood risk data, transportation data, geological data	Private sector, semi-public sector, public sector, public undertaking
Lack of technical infrastructure and data compatibility	geospatial data, flood risk data, utility data, ship data	private sector, Public undertaking
Lack of data accessibility and management	Transportation data, geological data, utility data, ship data	private sector, Public undertaking
Lack of data processing and validation	utility data, ship data	private sector, Public undertaking
Lack of data security mechanisms	transportation data, geological data	private sector, semipublic sector

Barriers to Data Sharing of the Private sector

Legal	Dataset	Sector
Restrictions by law	non-personal data, Transportation data, geological data, public health data	Private sector, semipublic sector, public sector
Contractual boundaries	non-personal data, public health data	private sector
Privacy constraints	personal data, public health data, utility data, ship data	public sector, private sector, semipublic sector, Public undertaking
Ownership constraints (Licensing and Copyright) liability	personal data, public health data	public sector, Private sector

Barriers to Data Sharing of the Private sector

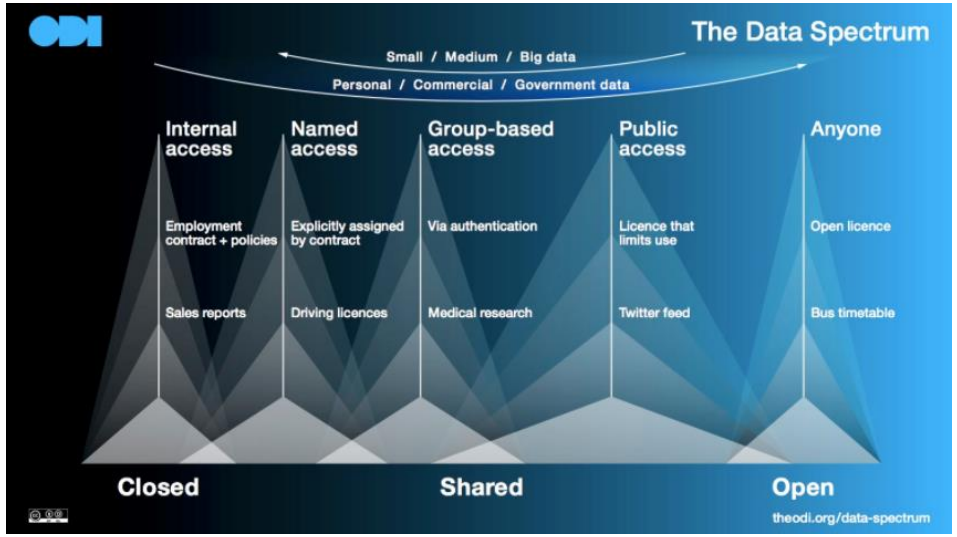
Economic	Dataset	Sector
Fear of economic damage	non-personal data, transportation data, geological data	private sector, Public undertaking
Loss of income from change in business model	utility data, ship data	public undertaking
Uncertainty about the value of data	geospatial data	private sector
Lack of resources	non-personal data, public health data, utility data, ship data	Private sector, public sector, public undertaking

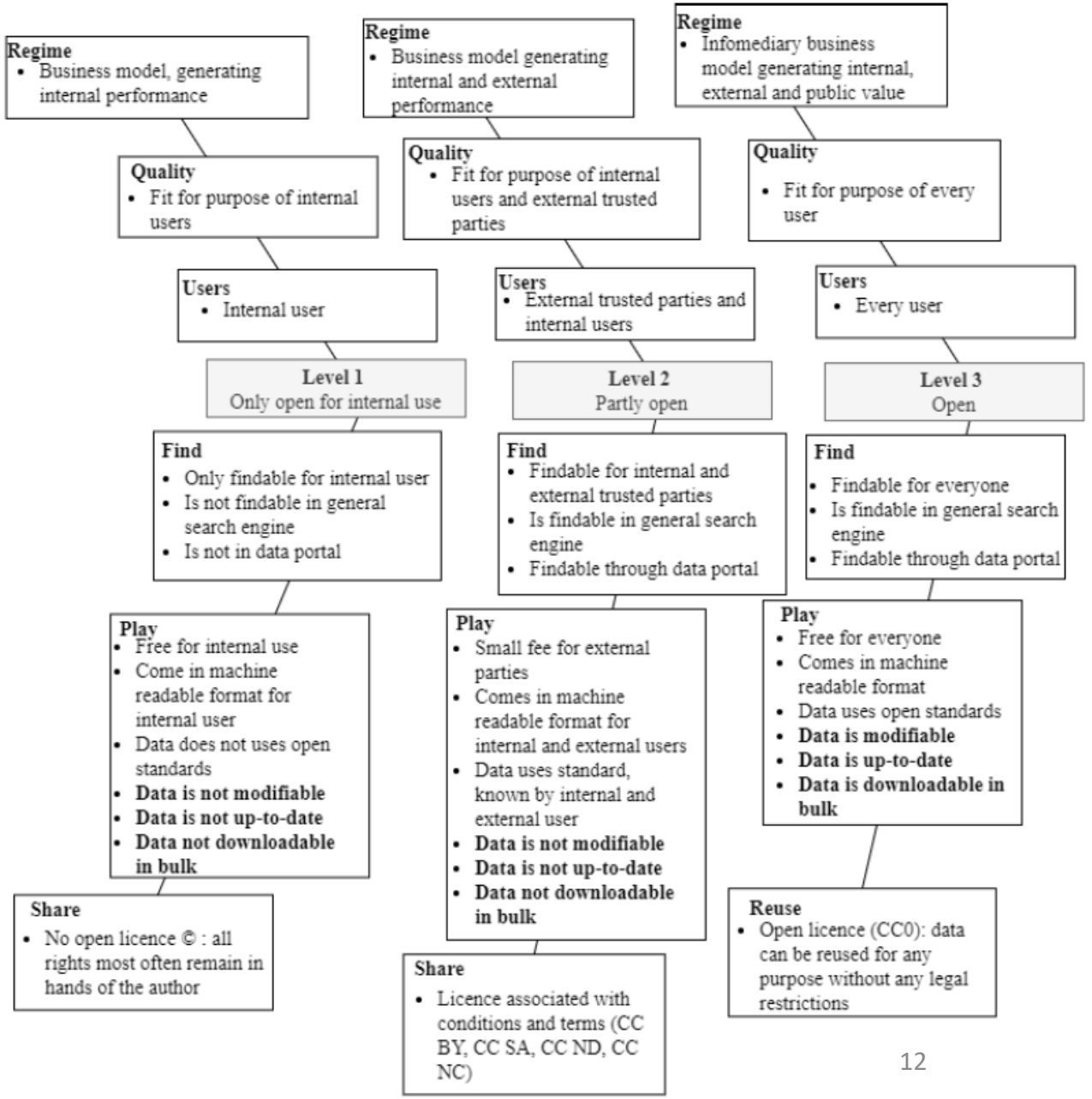
Barriers to Data Sharing of the Private sector

Cultural	Dataset	Sector
Cultural differences	transportation data, geological data	private sector, semipublic sector
Unwillingness to share data	geospatial data, flood risk data, public health data	private sector, Public sector
Lack of trust in data usage	transportation data, geological data, public health data, utility data, ship data	private sector, Public undertaking
Fear of loss of control	geospatial data	private sector, Public sector
Fear of transparency and disclosure of competitive knowledge	transportation data, geological data	private sector

Open Data Assessment Frameworks

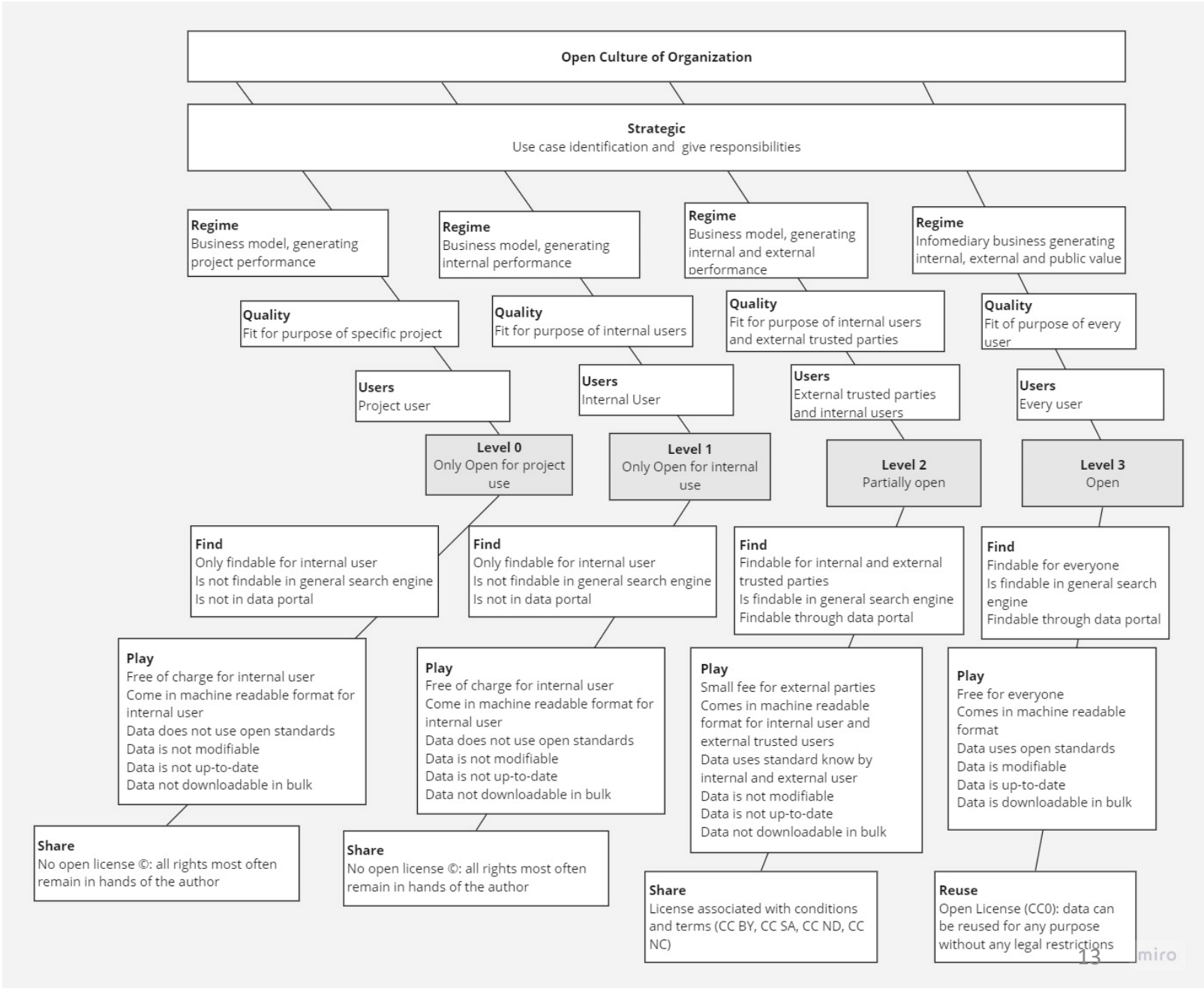
★	Available on the web (whatever format) *but with an open licence, to be Open Data*
★★	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
★★★	as (2) plus non-proprietary format (e.g. CSV instead of excel)
★★★★	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
★★★★★	All the above, plus: Link your data to other people's data to provide context





Open Data Assessment Frameworks

- **Level 0** giving an indication that even for internal access
- The Five Star model was integrated on the **Find** section of the new model were it is integrated
- **Level 2** partially as level two involves dataset that are available on the web but not with an open licence
- **Level 3** that can vary from datasets with 1 star to 5 stars.
- **Regime** of a business model that generates project performance might be more realistic



AND donates street map to OSM

2007 AND donates Streetmap of the Netherlands to OSM

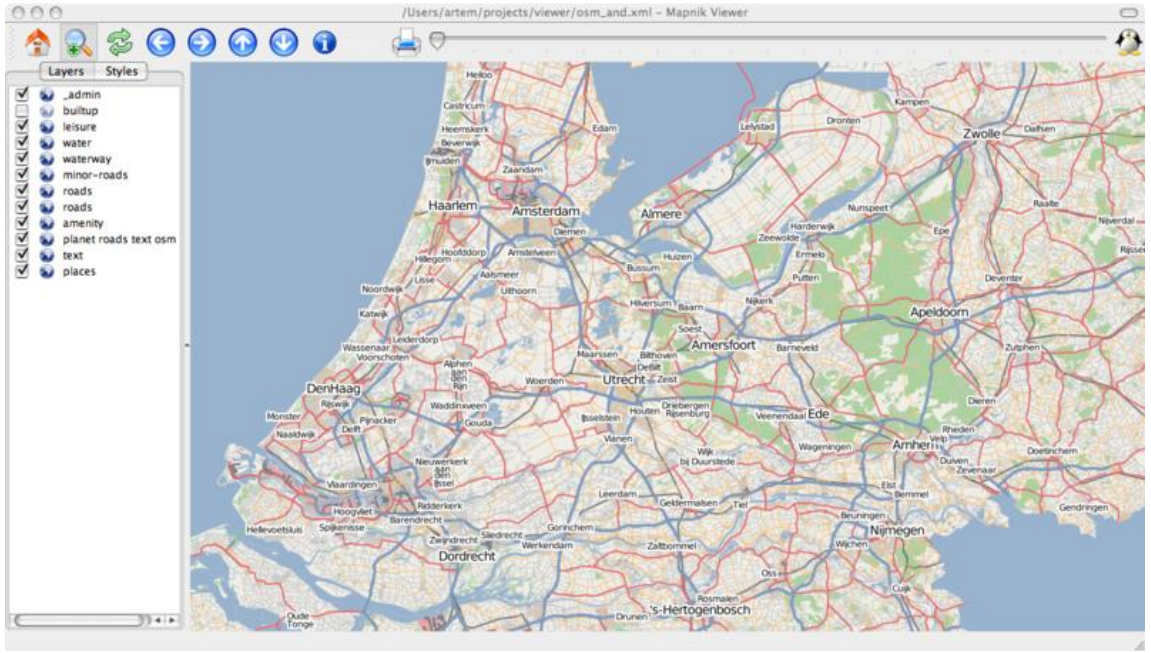
OSM community of mappers that contribute and maintain data about roads, trails

Incentives:

- create the most up-to-date map in the market
- data donated **not up-to-date** close **gap**
- National government - > open geodata INSPIRE directive 2006
- Avoid fail of business model (TOP10NL open)
- Showcase expertise, advertise products, CC-BY-SA credit

Barriers:

- up to date, quality of data
- misuse
- data science skills of users



Fugro Bathymetry data

2017 Fugro sharing bathymetry data to Seabed 2030

Incentives:
- Market OARS®
- Sustainability goals
- GEBCO database
- High skilled employees

Barriers:
- Own data (collected in transit)
- Less resources
- Data standards (GEBCO)

2 million km2

14 LIFE BELOW WATER

CONSERVE AND SUSTAINABLY USE THE OCEANS, SEA AND MARINE RESOURCES FOR SUSTAINABLE DEVELOPMENT

Microsoft Building Footprints

2018 **130 million** Building Footprints USA

- Data provided to OSM
- GitHub model
- Bing Maps API
- Geojson **open** format

Incentives:

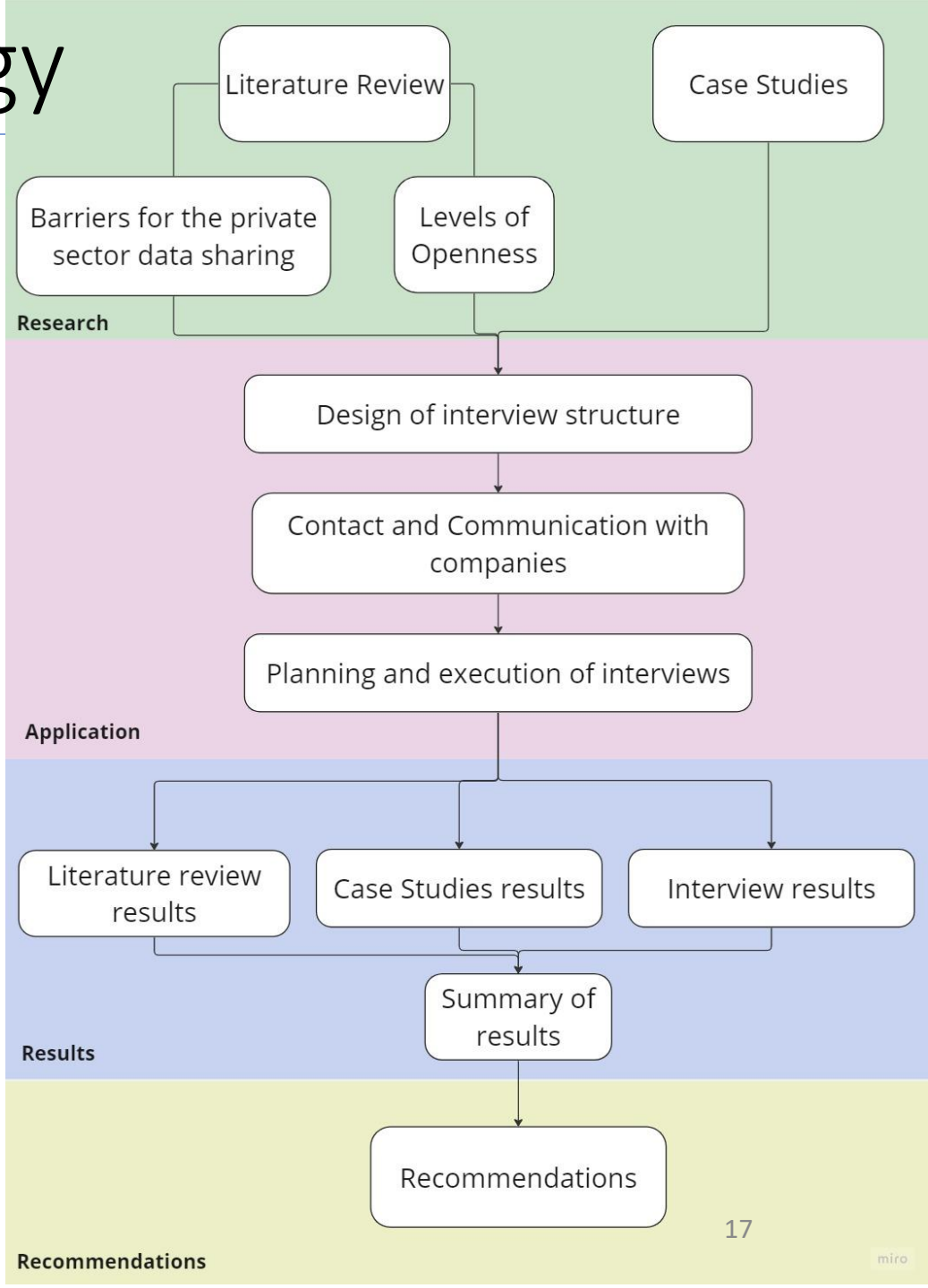
- Open Source CNTK Unified
- Microsoft Cloud, storage capacity, computational power
- HOT (Humanitarian OpenStreetMap Team)
- 15 applications: informal subdivisions, population exposed to flood hazards

Barriers:

- Licenses and conditions data, product, code
- Data science skills (GitHub doc)

Methodology

- **Exploratory** primary research to **define the problem**
- **Qualitative** analysis of the reality of geospatial data sharing of 9 companies in the Netherlands
- **Semi- Structured interviews** with **experts** in geospatial data sharing



Design of interview structure

Current Situation

1. Where do you think you are in this schema?

Past Situation

2. How did you arrive at this level in the schema?
3. What are the barriers encountered?
4. How did you resolve/overcome them?

Future Situation

5. What are the barriers you expect to move to the next level?
6. How could you overcome those barriers to move to the next level?
7. What is needed to move to the next level?

Contact and Communication with companies

- Companies produce geospatial data in the Netherlands **fit to one of the four levels**

- **Geobusiness Nederland trade association** not able to identify companies
-
- Contact as many companies as possible for interviews

- Through **mutual contact**, official **website** of the company, research and connection on **LinkedIn**

Planning and Execution of the interview

- Informed about the topic of research
- 45 minutes to 1 hour
- In person or online
- Date, time and place of the interview
- Interview questions 2 days prior
- Answer Q1 in preparation -> starting point of the interview
- Recording of the interview, name in the final report
- Follow up questions according to the topic of discussion
- 9 interviews

Results

Strategic	Company
Lack of organizational motivation to enable data sharing	Asset Insight, Cyclomedia, Geodan,
Lack of management commitment and corporate strategy	Asset Insight, Fugro, Geodan
Lack of policy coherence	-
Lack of use case identification	Asset Insight, Fugro
Lack of business cases for generating revenue	Fugro, GeoJunxion, Tensing, Geodan
Lack of a feedback process	-

Geodan -> ”if we don’t have a driver, like a contract with a company or the government, we are not urged to share data, we will not spend time on things that do not benefit us”
GeoJunxion -> ”we are a B2B company and a customer-based company that requires us to be focused on commercial data, and not making data available publicly”
Asset Insight -> five-year contract ProRail
Fugro -> ”data sharing and open data not the current strategy of the company, so there are no steps taken in that direction”

SWECO -> GIS department as knowledge exchange base shapefile
GeoJunxion -> Sourcing department
Tensing -> Good network of partners, ”Evangelists”
ESRI -> Development of software to make it easy for their users to start using GIS

Results

Technical	Company
Lack of data standards	Asset Insight, Cyclomedia, Geodan, SWECO, Imagem, Fugro
Lack of accessibility and management:	Asset Insight, Cyclomedia, Fugro, Geodan
Lack of data processing and validation	Asset Insight, Cyclomedia, SWECO
Lack of data security mechanisms	-
Lack of technical infrastructures and compatibility	Fugro, Geodan, Imagem

Asset Insight .csv and .geotiff , GDPR slows processing, temporal changes need metadata, storage
Cyclomedia storage, data standards and metadata specification, API OGC, knowledge about high-accuracy GPS
Fugro storage Amazon Cloud
SWECO data science skills
Imagem data standards
Geodan OCG standards,
Esri quality of data

Results

Legal	Company
Restrictions by law	Geodan, Asset Insight
Privacy constraints	Asset Insight, Geodan
Ownership constraints (Licensing and Copyright)	Asset Insight, Fugro, Geodan , Tensing, GeoJunxion, Imagem, Cyclomedia
Contractual boundaries	ESRI, Cyclomedia, Geodan, Fugro, Asset Insight

“We do not own the data, so it is not our decision to make”

Imagem new algorithm or model and want to share the working of that model need **agreement**
Asset liability, trainspotters open data, blur faces, compromise high quality
Fugro liability, general search engine, without experience
GeoJunxion CC SA
Cyclomedia owner of the data, no licenses, contracts for limitation of use with universities
SWECO open government data PDOK
ESRI licensing restrictions for their use of their datasets in the Living Atlas
Geodan anonymize data

Results

Economic	Company
Fear of economic damage	Asset Insight, Cyclomedia, Geodan, SWECO, Imagem, Fugro
Loss of income from change in business model	Asset Insight, Cyclomedia, Imagem
Uncertainty about the value of data	Asset Insight, Geodan
Lack of resources	Asset Insight, Geodan, Fugro, Imgem

Asset Insight: sharing data as open, means that they lose their commercial aspect
”information is knowledge, knowledge is money”

Cyclomedia: current business model is successful no intention to share their data as open”, form of funding

Imagem: only data that are created with public investments should be public data
”efforts should be linked to a business model”

ESRI: lack of standardized successful business cases to follow

Sharing becomes possibility during the project it creates additional issues, lack of human and technical resources

Results

Cultural	Company
Cultural differences	-
Unwillingness to share data	Asset Insight
Lack of trust	Fugro, GeoJunxion, Asset Insight
Fear of loss of control	Fugro, Asset Insight

Imagem and Geodan high level of open culture in the company no need to persuade the internal departments about the importance and the benefits of data sharing and open data

Fugro: ”the dataset is out of the control of the company”

GeoJunxion: misuse of the data about AND OSM streetmap, skills of the user

Asset Insight: refused to share raw data, in fear of exposing confidential knowledge.

Results

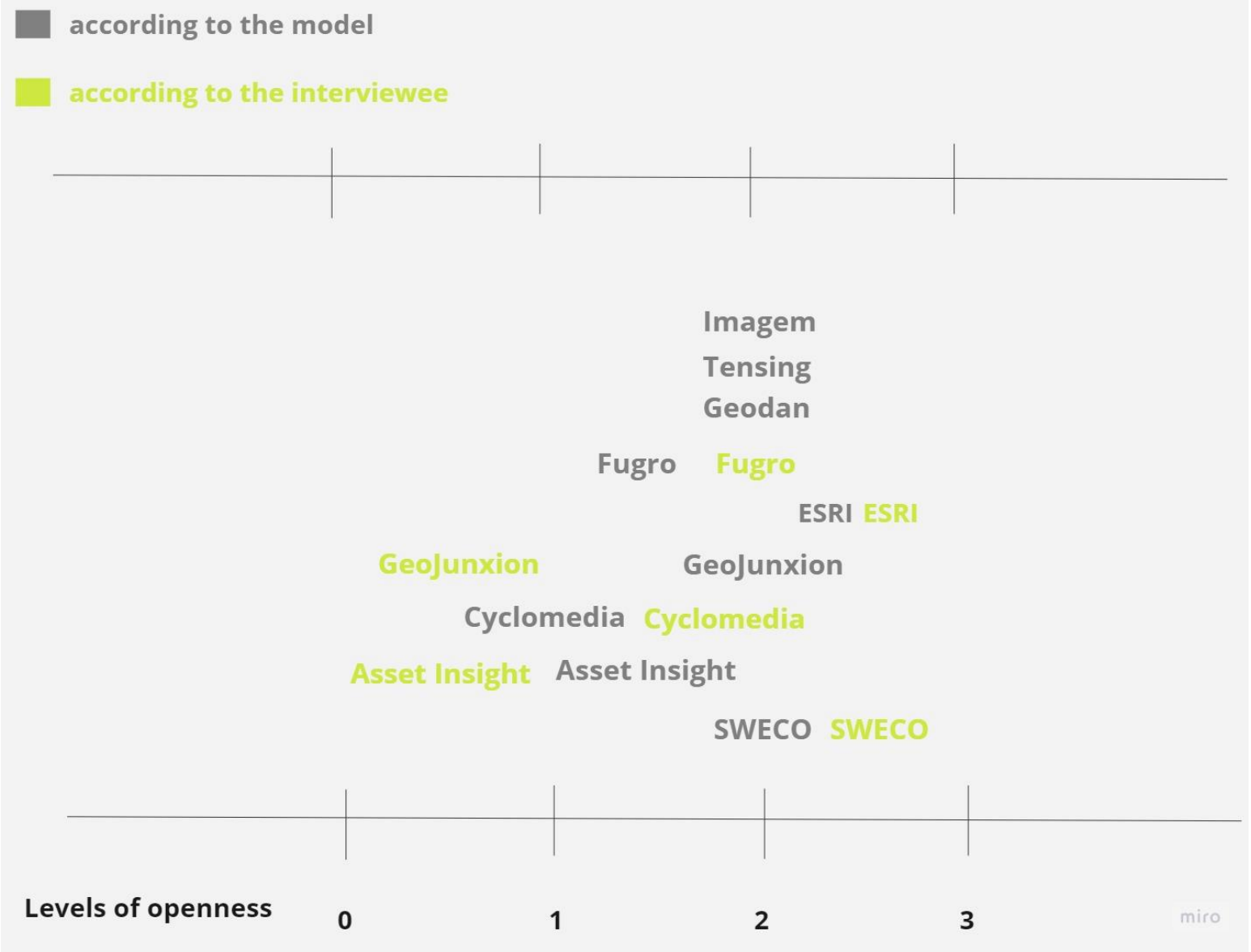
Company	Strategic	Technical	Legal	Economic	Cultural
Asset Insight	✓	✓	✓	✓	
Cyclomedia	✓	✓	✓	✓	
Fugro	✓	✓	✓	✓	
SWECO	✓	✓			
Tensing			✓		✓
Imagem		✓	✓	✓	✓
Geodan	✓	✓	✓	✓	✓
ESRI			✓		
GeoJunxion	✓		✓		✓

- Lack of data standards and license frameworks -> change on a contract basis
- Can lead to issues of access and reuse
- Unaware of existing standards for data, only software

Discussion

Tensing, Imagem, Geodan could not identify -> too provider driven
Providers 0, 1 and 2
The rest 2 and 3

Provider (Asset Insight, Cyclomedia, Fugro)
User (SWECO, Tensing, Imagem, Geodan, GeoJunxion, Esri)
Intermediaries (ESRI, Imagem, SWECO, Geodan, Tensing, GeoJunxion)



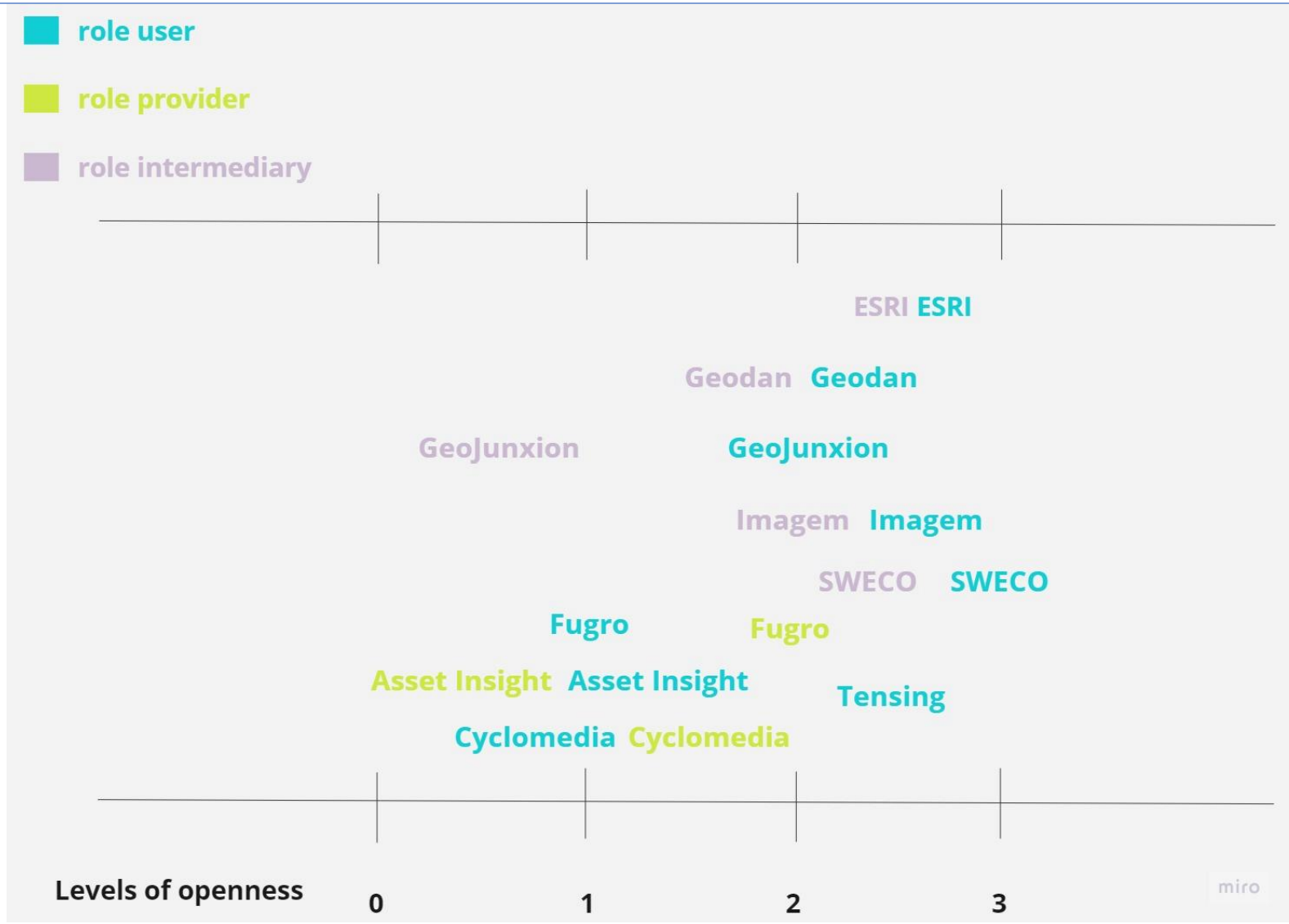
Discussion

Providers -> more barriers to overcome from level 1 to level 2

- Technical barriers: data standards, interoperable software, educate employees data skill
- Legal: licensing framework not contracts
- Economic: better allocation of resources

Intermediaries -> from level 2 to level 3

- Strategy not open data, but contribute to data sharing
- Less technical, legal, economic barriers, open culture contributes
- Legal barriers on all except SWECO that used open data



- Department based Asset Insight
- Project based for level 1 and level 2 (owner decision)
- Dataset based (case studies)

Recommendations

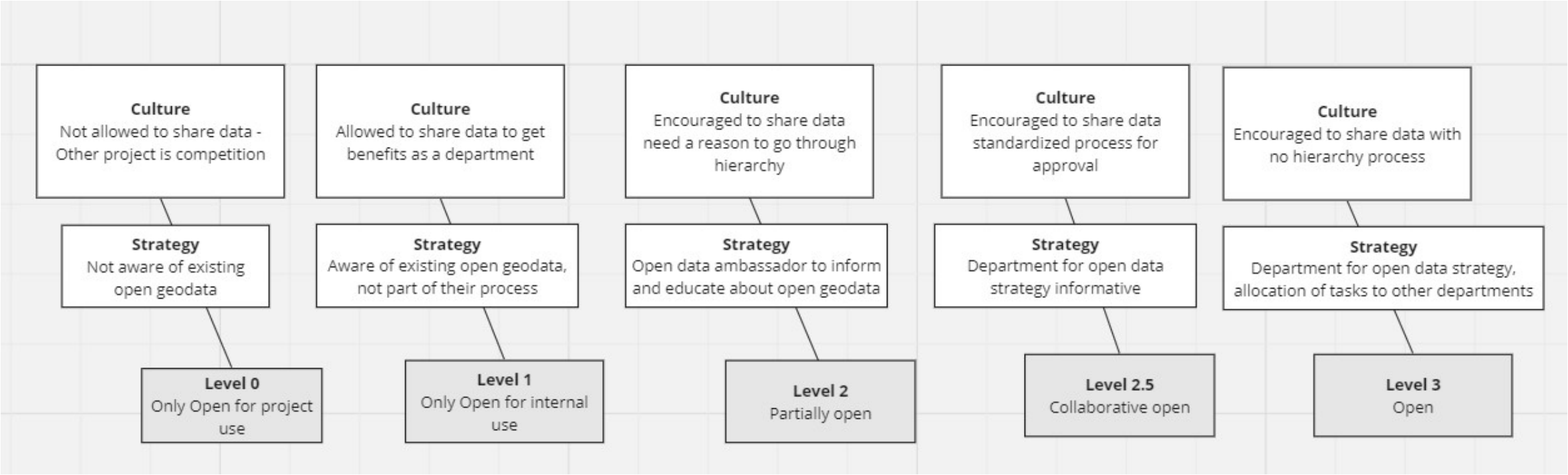
Government

- Opportunity to **integrate current legal and technical** European Interoperability Framework, created for public administration, on going conversation, private sector in mind, need to run through the entirety of the model for smooth transitions.
- Create **catalogue of existing open data standards**
- Add **five-star system** for open datasets to facilitate limited data science skills, no need to hire more skilled personnel.
- Collaborative funding
- Level 3 not the end goal

Recommendations

Model

need a harmonized culture and strategy



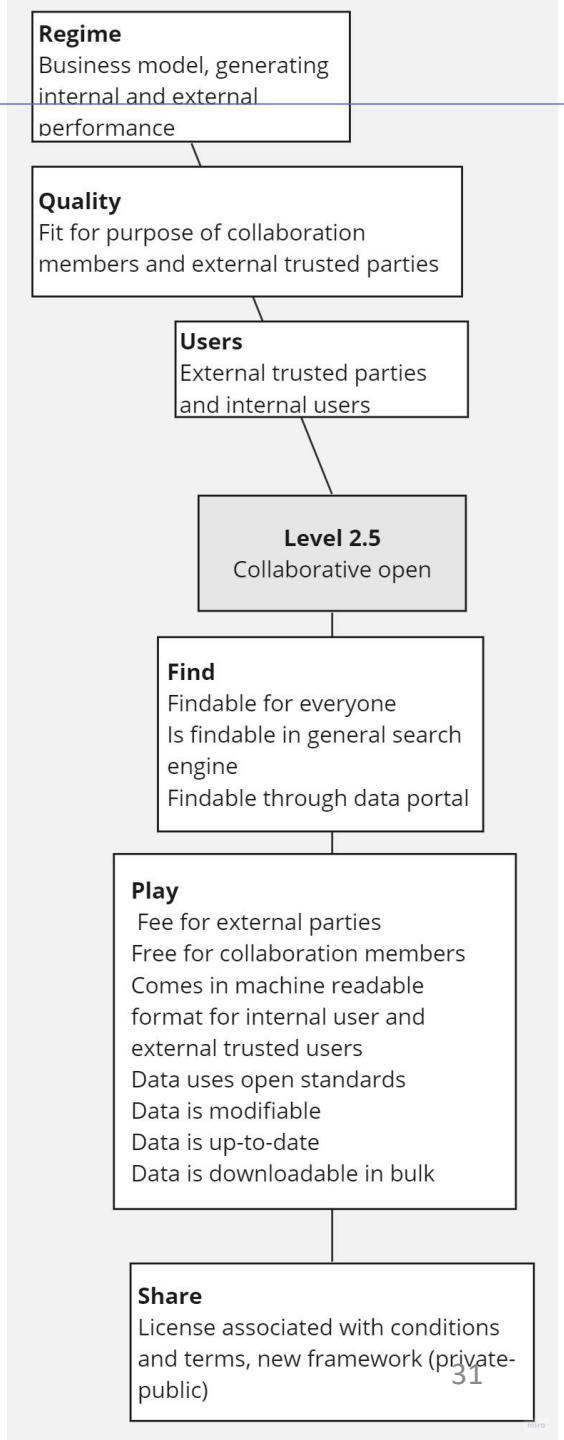
- Strategy: ESRI, Tensing, SWECO, GeoJunxion
- Culture: Asset Insight, Cyclomedia, Fugro, ESRI slow process -> 2.5

Recommendations

- **Fugro US public-private collaborative funding**, using investment from both public and private sector for the development and maintenance of shared datasets would be more appealing for the private sector

- **Imagem** failed attempt of the involvement of the private sector in the Galileo project with not benefits for the private sector

- **Regime** end data with be used both for economic (private) and social (public) benefits
- **Find** at level 3
- **Play** part of **level 3** of the schema, no free of charge, need to **maintain business model** for some of their projects, or datasets
- **Share** licensing conditions of **new framework**

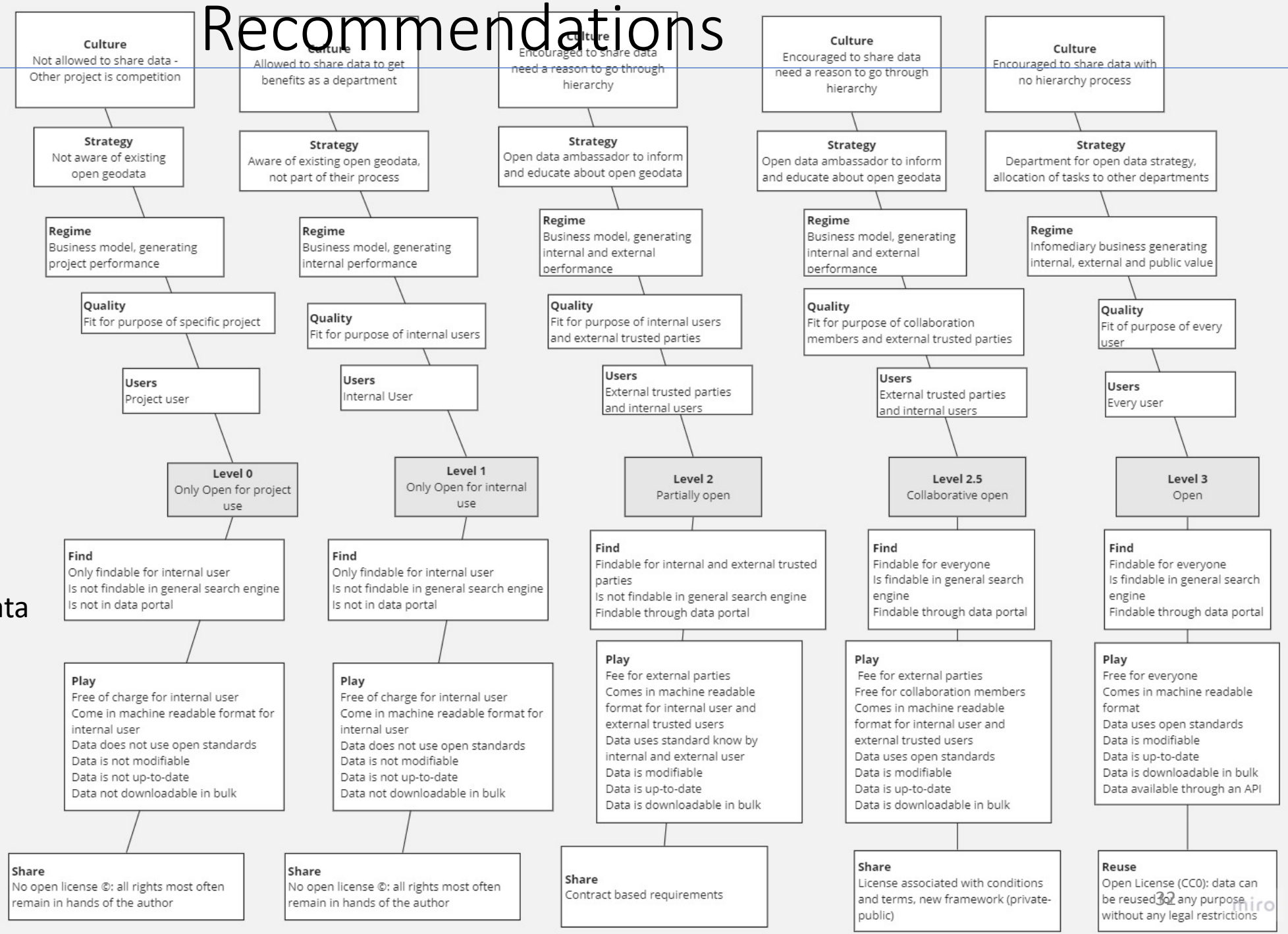


Recommendations

Private Sector

Providers:
level 1 to level 2
lack of data standards
software interoperability
data science skills

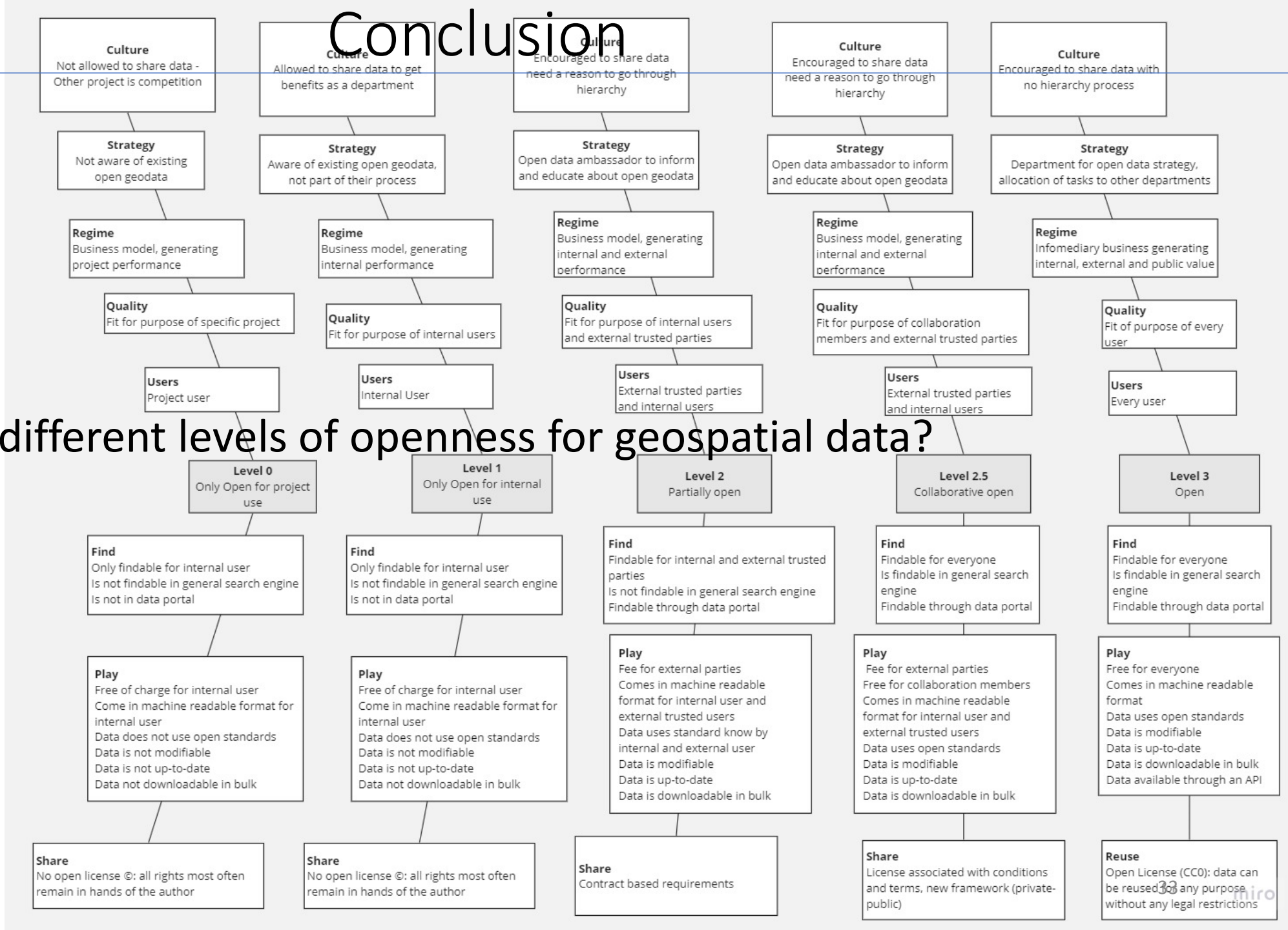
Intermediaries:
Level 2 to 2.5 to 3
Clear strategy for open data
Use existing open government data
Use case identification



Conclusion

What are the different levels of openness for geospatial data?

5 levels



Conclusion

How do the different barriers relate to each level of openness of the developed model for private sector data sharing?

- **Level 0** and **1 most barriers** to overcome, limited open standards, **not interoperable** software, limited **data science skills**, contract restrictions, not licensing frameworks, delays in data processing, **lack of resources**
- **Level 2** and **3** more open culture, **strategy** that aids data sharing, **less technical** barriers, **data standards** to support, **ownership** issues, **business model** that allows open datasets

Conclusion

What are the barriers to private sector data sharing of geospatial data in the Netherlands?

- Depend on the level
- Level depends on role of the company (user, producer, intermediary) – double levels for each company
- Level depends on dataset, project, department

How can those barriers be approached to reach (open) data sharing?

- Collaboration public-private sector on funding pool, technical and legal interoperability layer, integrate existing systems
- Government change metadata and quality indication – 5 star
- Case identification catalogue by the government with success stories

Limitations

Validation of the new model

One interview per company

Bias role of the interviewee

New schema considers the role of the company

Higher level position of interviewee

Future Work

- Benefits of data sharing
- Create catalogue of success cases, used to show value added to data for private companies

- Develop level of openness schema for role of the company
- Make it circular, lacks feedback
- Level of openness of a user, hard to define

- Open participation of private sector in the development of standards, not only testing of existing standard (OGC)
- Open data standards that are used, which need to stop being supported to have a clear view

Thank you!

TU Delft Delft University of Technology