

Data as a Commons Dataspace

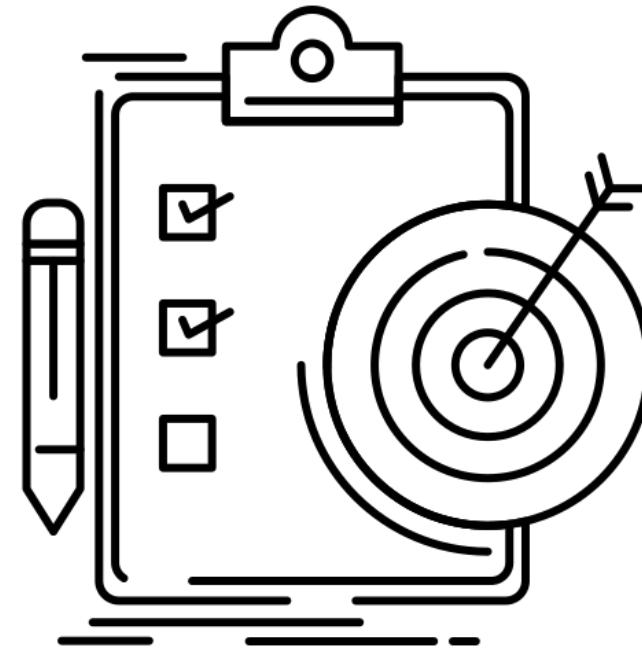
a course for the phd students of GSSI - Gran Sasso Science Institute

Maurizio Napolitano
<napolitano@fbk.eu>

Goals of today

What is a Dataspace

The Data Steward role



Today: Dataspaces

from Open Data to Dataspace

the ingredients of a Dataspace

Data Governance

Data Stewarding



Past

Freedom of information

The need of Open Data



125 years ago!

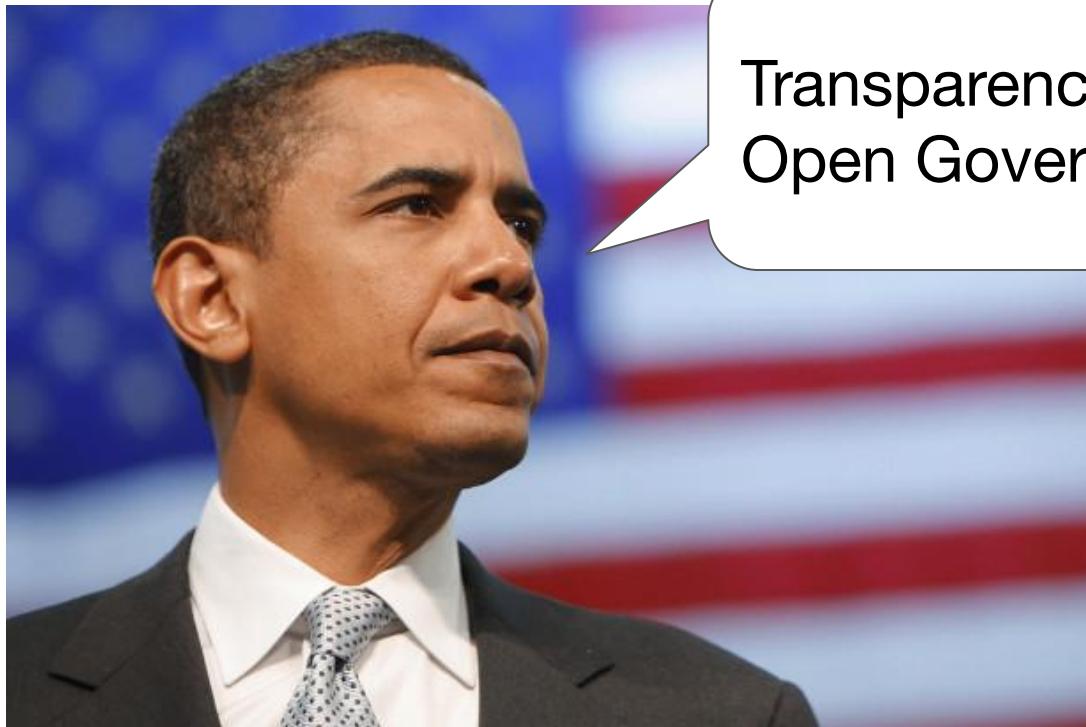
[...] I could have certainly done more, especially regarding the statistical aspects, however, in our country both private and non-profit organizations show reluctance relying on public domain for data, information and content [...] - Cesare Battisti

From the "Preface" of *Il Trentino*,
Giovanni Zippel Editore, Trento, 1898)

Present

Open Government

In 2009 this guy ...



<https://obamawhitehouse.archives.gov/the-press-office/transparency-and-open-government>

Data as a Commons - Maurizio Napolitano

The Open Data Charter - not only open government



1.

Open By Default



2.

**Timely and
Comprehensive**



3.

Accessible and Usable



4.

**Comparable and
Interoperable**



5.

**For Improved Governance
& Citizen Engagement**



6.

**For Inclusive Development
and Innovation**

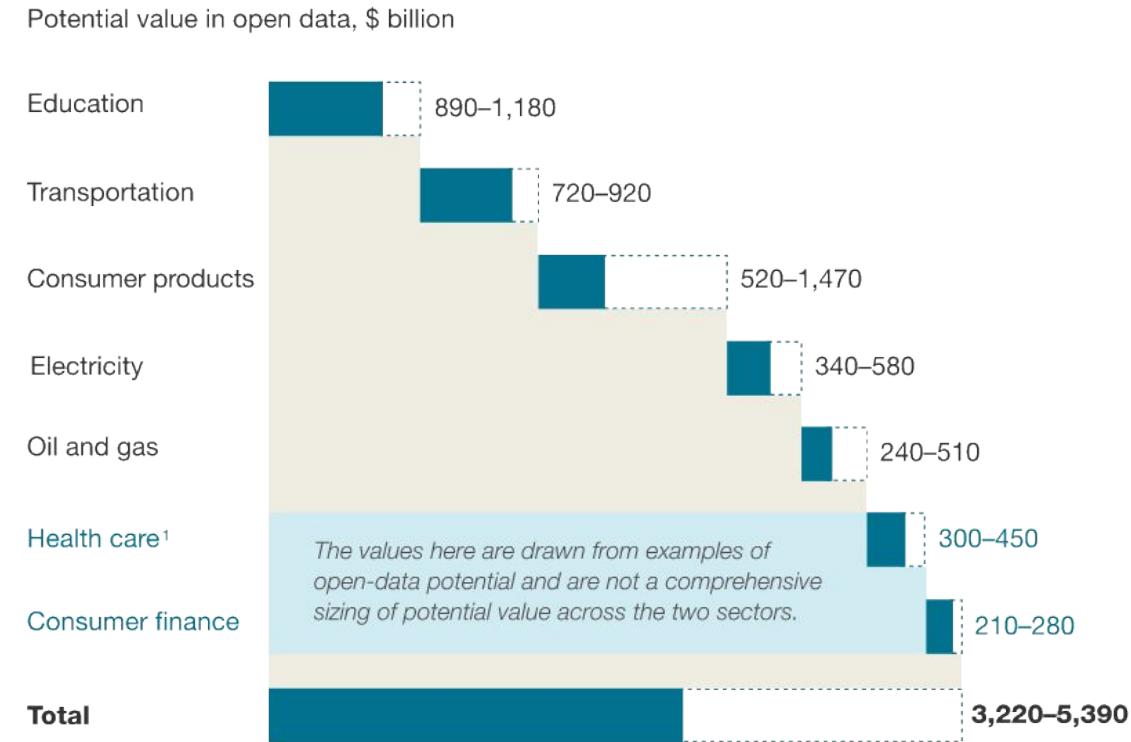
McKinsey & Company

Open data—public information and shared data from private sources—can help create **\$3 trillion** a year of value in seven areas of the global economy.



McKinsey & Company

Open data—public information and shared data from private sources—can help create **\$3 trillion** a year of value in seven areas of the global economy.

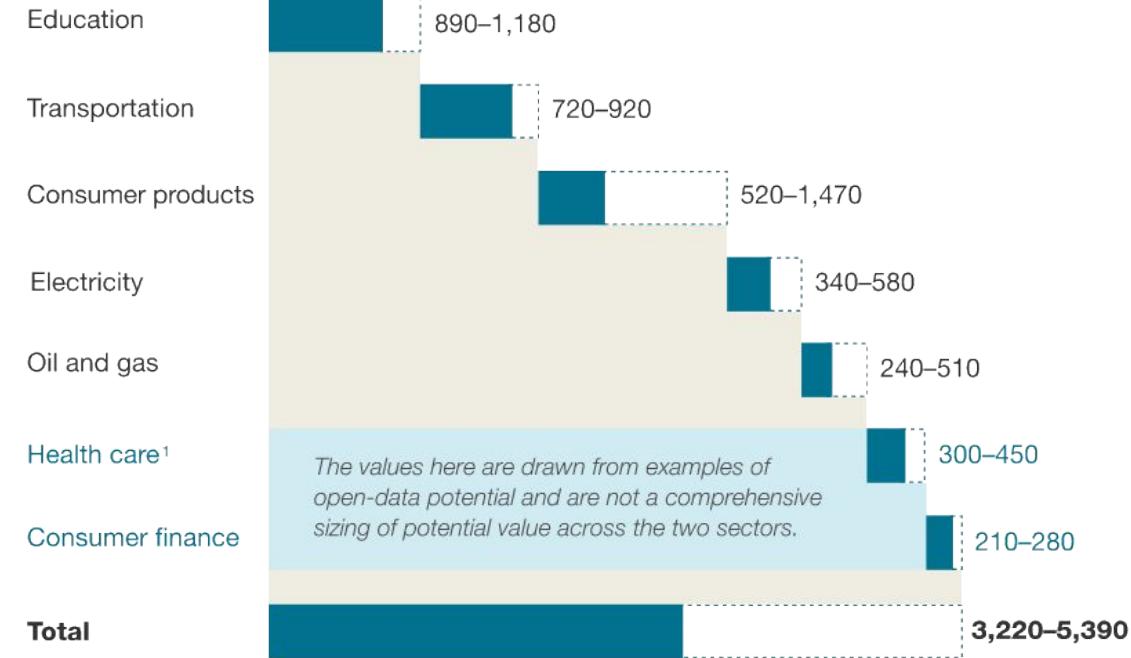


¹Includes US values only.

McKinsey & Company

these data categories have not yet been opened

Potential value in open data, \$ billion



¹Includes US values only.

An official EU website How do you know? ▾

EUR-Lex

EN

☰ QUICK SEARCH

[Search tips](#) Need more search options? Use the [Advanced search](#)

Document 32013L0037

Directive 2013/37/UE du Parlement européen et du Conseil du 26 juin 2013 modifiant la directive 2003/98/CE concernant la réutilisation des informations du secteur public Texte présentant de l'intérêt pour l'EEE

OJ L 175, 27.6.2013, p. 1–8 (BG, ES, CS, DA, DE, ET, EL, EN, FR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV)
Special edition in Croatian: Chapter 16 Volume 004 P. 27 - 34

No longer in force, Date of [Table of contents](#) 32019L1024

Data as a Commons - Maurizio Napolitano

An official EU website How do you know? ▾

EUR-Lex

EN

QUICK SEARCH

[Search tips](#) Need more search options? Use the [Advanced search](#)

Document 32023R0138

Commission Implementing Regulation (EU) 2023/138 of 21 December 2022 laying down a list of specific high-value datasets and the arrangements for their publication and re-use (Text with EEA relevance)

C/2022/9562

OJ L 19, 20.1.2023, p. 43–75 (BG, ES, CS, DA, DE, ET, EL, EN, FR, GA, HR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV)

In force

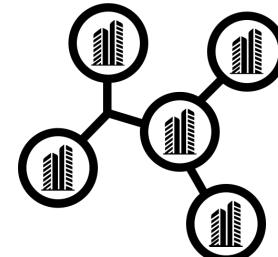
ELI: <http://data.europa.eu> Table of contents

Future

Re-use of public and private data

Benefits of private data to PA

1. **Reduce** data collection **costs** for statistical offices and administrative burdens by using already available data rather than probing companies and individuals;
2. Making **new data** available, through the provision of **real-time** data, e.g. on the mobility population
3. **Eliminating bias** from self-report data in favour of observational data.
4. The limitations of self-reported data are well known and severely limit our understanding of human life



Important

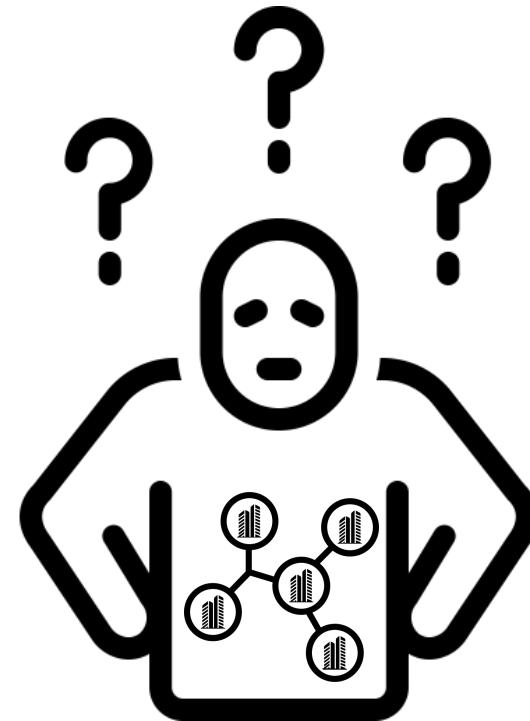
Private data are not readily available and cannot be used immediately as a substitute for official statistics usually due to their biased coverage



Concerns of Companies

renunciation of commercially sensitive data

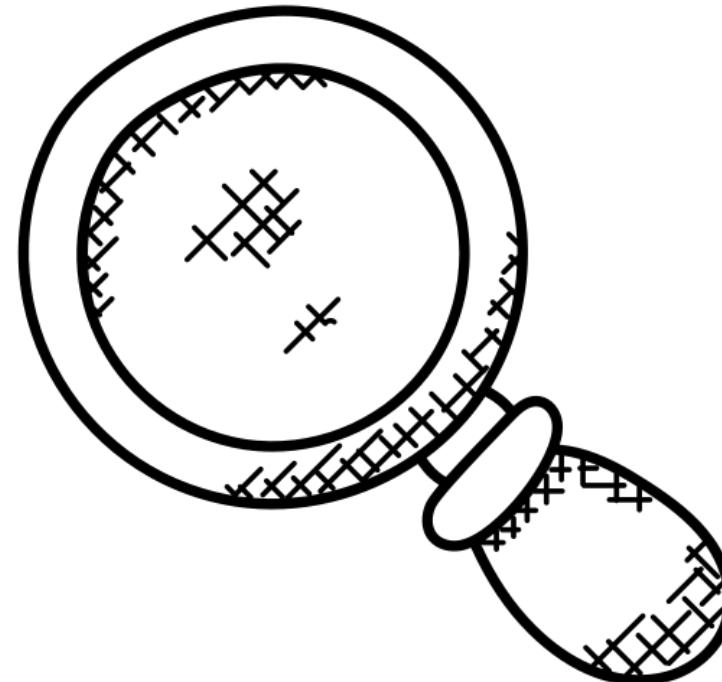
risks of privacy breaches and related costs and risks of using anonymisation techniques



Evidences about companies

Costs and **skills** are significant **barriers** on both the supply and demand side.

Companies share data on the basis of voluntary **agreements or legal **obligations**, for **free** or for a **fee**.**



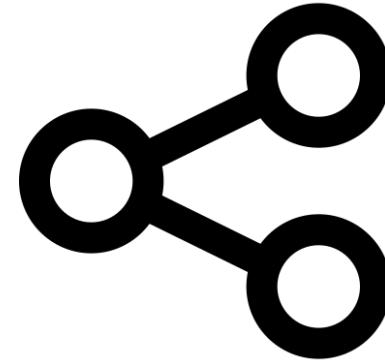
Conditions through which sharing takes place

When their data can offer advantages.

On regulatory obligations.

Covering the costs of data sharing.

Data that are not company assets.

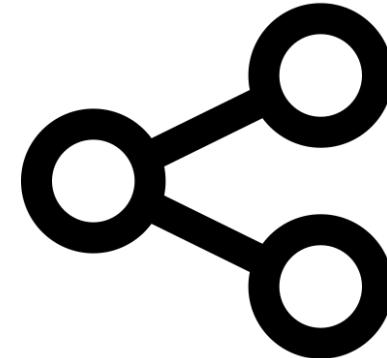


Conditions through which sharing takes place

When their data can offer advantages.

Eg

emergency situations, national security, etc., which may justify stronger political measures, versus data needed to monitor tourism

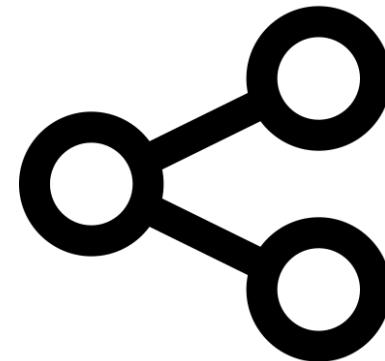


Conditions through which sharing takes place

When their data can offer advantages.

On regulatory obligations.

in regulated sectors or in cases of procurement or concessions



Conditions through which sharing takes place

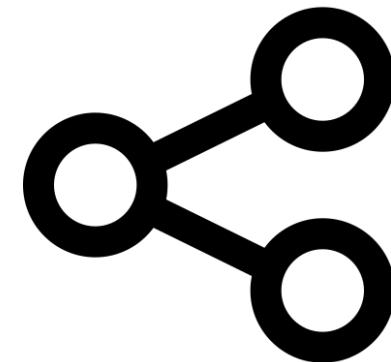
When their data can offer advantages.

On regulatory obligations.

Covering the costs of data sharing.

Eg:

real-time disaggregated anonymised data is much more expensive than aggregated bulk data. The higher the costs, the more difficult it is to justify companies' obligation to open up their data.



Conditions through which sharing takes place

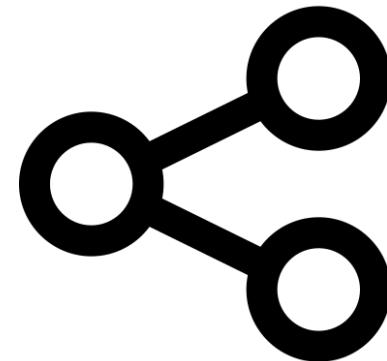
When their data can offer advantages.

On regulatory obligations.

Covering the costs of data sharing.

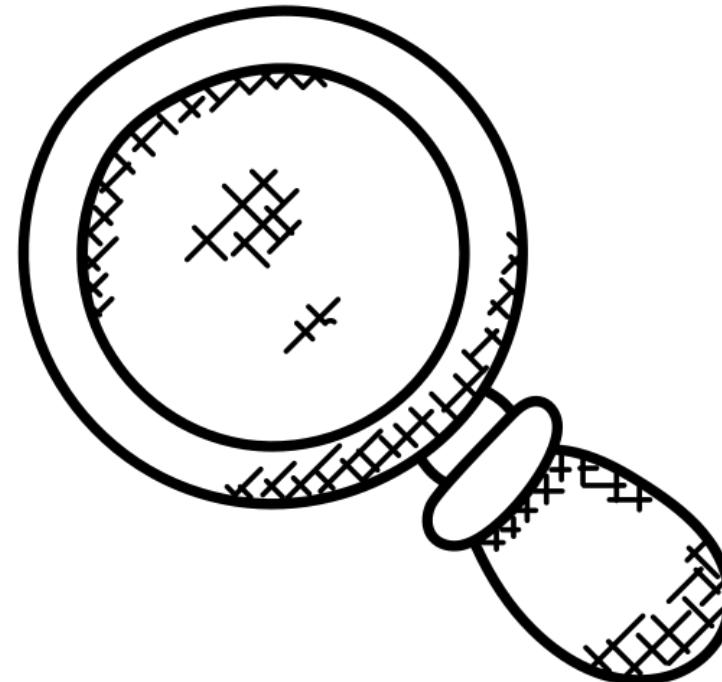
Data that are not company assets.

E.g.
customer data



Evidences of the data sharing by companies

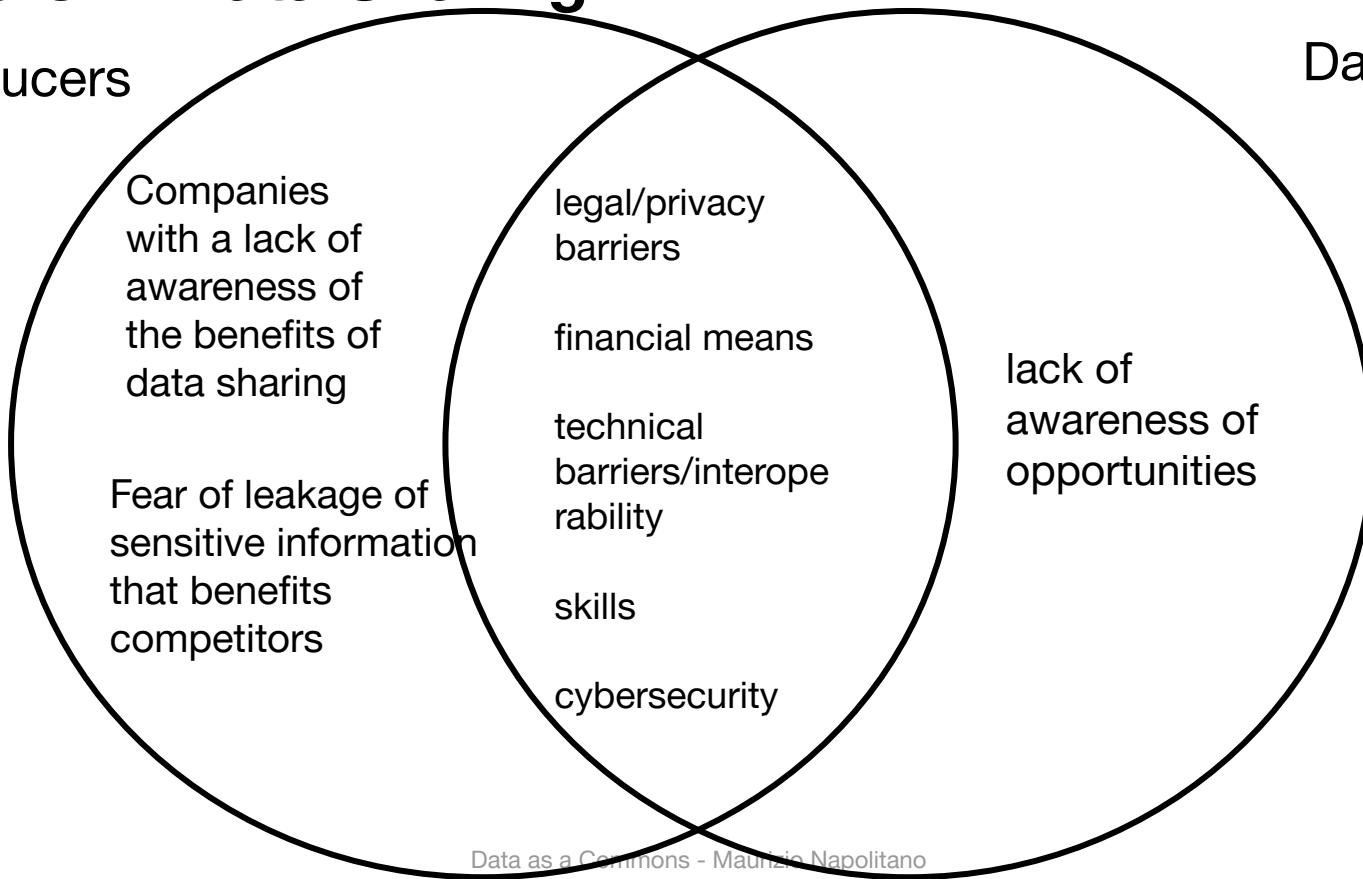
companies share data when it is aggregated and shared only with one organisation (hardly opendata)



Barriers in Data Sharing

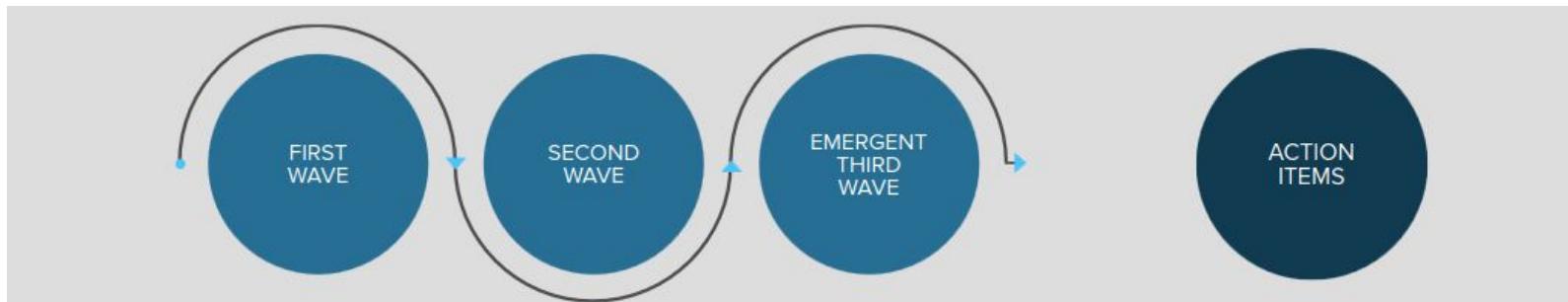
Data producers

Data re-users

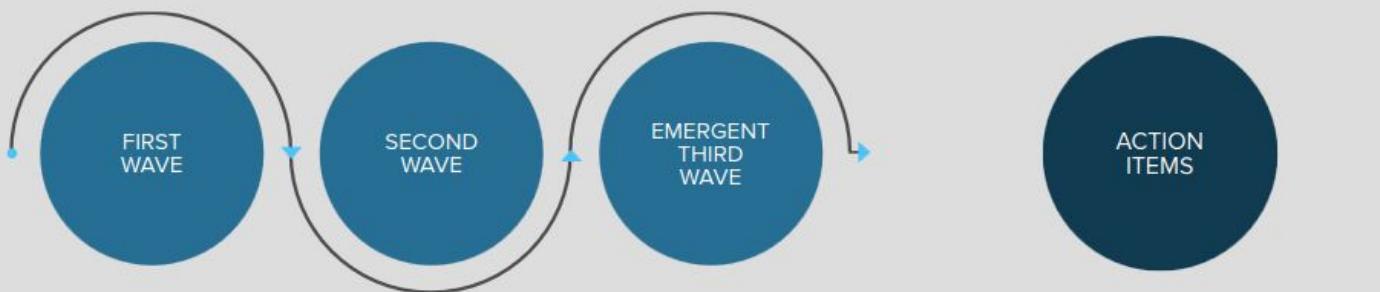








Concept:	Freedom of Information	Open Government Data	Re-use of Public and Private Data	<ul style="list-style-type: none">▶ Foster institutional data capacity▶ Increase Data Liquidity
Value Proposition:	Transparency	Transparency and Problem Solving	<ul style="list-style-type: none">▶ Evidence based policy making▶ Innovation and Entrepreneurship	<ul style="list-style-type: none">▶ Articulate value and build an evidence base of the impact of data re-use for public interest purposes
Approach:	Data on Request (Right to Know)	Open by Default (Right to Share)	Publish with Purpose	<ul style="list-style-type: none">▶ Create new data intermediaries▶ Invest in a Data Demand Identification and Assessment Methodology▶ Increase accessibility and findability of high-value datasets
Focus:	Pull Focused	Push Focused	Partnerships (Data Collaboratives)	<ul style="list-style-type: none">▶ Establish governance frameworks and seek regulatory clarity▶ Document and experiment with the number of operational models that are “fit for purpose”▶ Develop common data sharing agreements and new licensing regimes



Audience/ Demand:	<ul style="list-style-type: none">▶ Journalists▶ Lawyers and Activists	<ul style="list-style-type: none">▶ Civic technologists and "Data Geeks"▶ Government agencies▶ Corporations, technology start-ups	<ul style="list-style-type: none">▶ Community-based organizations▶ NGOs, Human Rights Organizations, and Social Justice Advocates▶ Academia and Research Institutions▶ Small businesses and startups▶ Government	<ul style="list-style-type: none">▶ Bolster public competence▶ Develop a data demand assessment and segmentation methodologies▶ Strengthen accountability mechanisms to support rights-based data re-use.
Risks and Policies:	Secrecy and obfuscation	Privacy – Mosaic effect, demographically identifiable information (DII)	Data Responsibility and Data Rights Framework	<ul style="list-style-type: none">▶ Track, monitor, and clarify decision and data provenance▶ Invest in Data Responsibility Frameworks and Technologies▶ Engage the public, especially vulnerable communities, to document public perceptions and opinions on data re-use
Institutional Responses:	Information Auditors	<ul style="list-style-type: none">▶ Data Officer▶ Open Data Portals	<ul style="list-style-type: none">▶ Chief Data Steward▶ Intermediaries	Create and empower (chief) data stewards

The 5 stars of Open Data Engagement

- ★ Be demand driven
- ★★ Provide context
- ★★★ Support conversation
- ★★★★ Build capacity & skills
- ★★★★★ Collaborate with the community

<https://www.timdavies.org.uk/2012/01/21/5-stars-of-open-data-engagement/>

Good Examples

Barcelona

Platform:

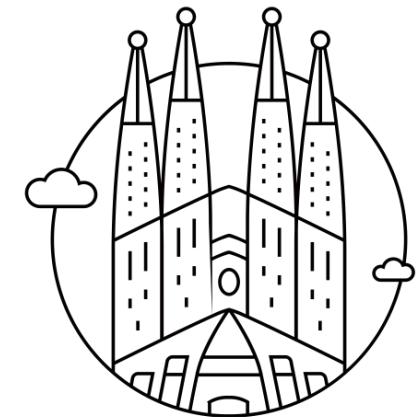
City OS - a single centralised system acting as an intermediary between the data sources and the services provided by the smart city

Goals:

Making data available to the community

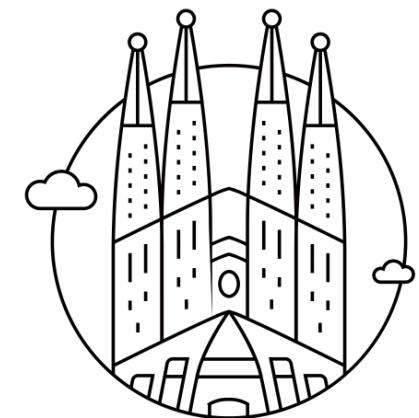
Tools:

Procurement agreements to enable the city to collect and share (non-sensitive) data also from private companies.



Barcelona - best practice

Through 'data sovereignty' clauses in public procurement, data can be turned into a commons and placed in the public domain, while preserving privacy, ethics and security.

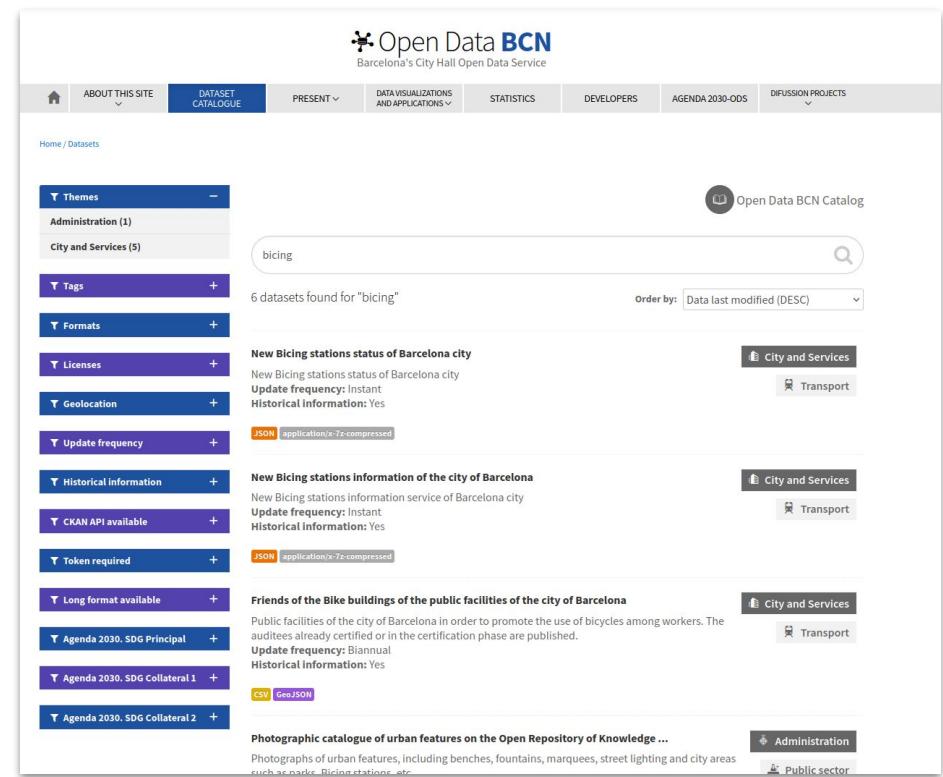


Example Bicing

urban transport service based on the shared use of mechanical and electric bicycles.

Following the contract with the municipality of Barcelona, it makes available data on stations, location and distribution of bicycles and usage statistics which are then made available on the OpenData portal

are used by the city council to enhance the viability of multimodal public transport services



The screenshot shows the Open Data BCN portal interface. The search bar at the top contains the text "bicing". Below the search bar, a message indicates "6 datasets found for 'bicing'" and "Order by: Data last modified (DESC)". On the left, there is a sidebar with various filters: Themes (Administration (1), City and Services (5)), Tags, Formats, Licenses, Geolocation, Update frequency, Historical information, CKAN API available, Token required, Long format available, Agenda 2030. SDG Principal, Agenda 2030. SDG Collateral 1, and Agenda 2030. SDG Collateral 2. The main content area displays three dataset cards:

- New Bicing stations status of Barcelona city**: New Bicing stations status of Barcelona city. Update frequency: Instant. Historical information: Yes. Format: JSON (application/x-tz-compressed). Category: City and Services, Transport.
- New Bicing stations information of the city of Barcelona**: New Bicing stations information service of Barcelona city. Update frequency: Instant. Historical information: Yes. Format: JSON (application/x-tz-compressed). Category: City and Services, Transport.
- Friends of the Bike buildings of the public facilities of the city of Barcelona**: Public facilities of the city of Barcelona in order to promote the use of bicycles among workers. The audited already certified or in the certification phase are published. Update frequency: Biannual. Historical information: Yes. Format: CSV, GeoJSON. Category: City and Services, Transport.

At the bottom, there are links for "Photographic catalogue of urban features on the Open Repository of Knowledge ..." and "Administration, Public sector".

London

- The UK has outlined the roles and actors required for strategy development and information management,
- London has set up bodies and key figures for the management and support of data governance that have developed a set of procedures, guidelines and templates to help public and private actors understand and manage all aspects of the data lifecycle in the national context.
- London has developed (and is in the process of updating) a datastore that is intended to be a central point of data collection and management also as a platform for data exchange between private or public and private parties



Collaboration with Waze

March 2021, Transport for London (TfL) partnered with Google's Waze app to provide safety messages via the Waze app to drivers and motorcyclists.

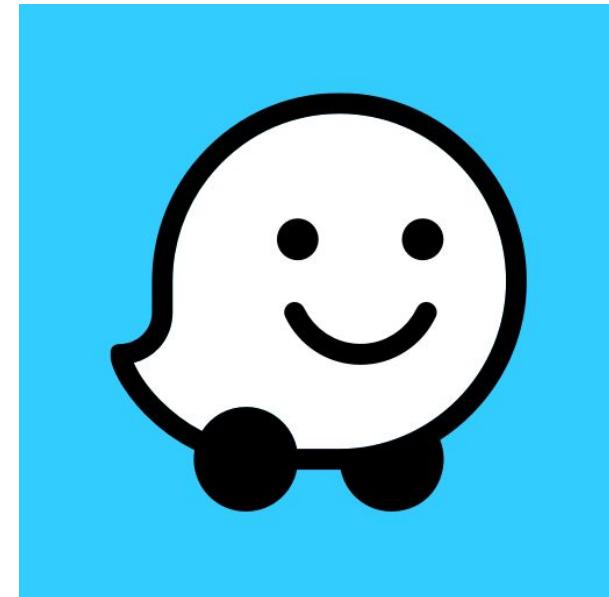
Partners in the Waze for Cities network (i.e. TfL) have access to a tool through which they can enter road closures and highlight major traffic events or emergencies.

They are also invited to participate in an online forum where strategic and operational issues can be discussed.

Waze, in return, provides all traffic data that TfL can use to improve its services.

More information on the Waze for Cities programme

<https://www.waze.com/wazeforcities>



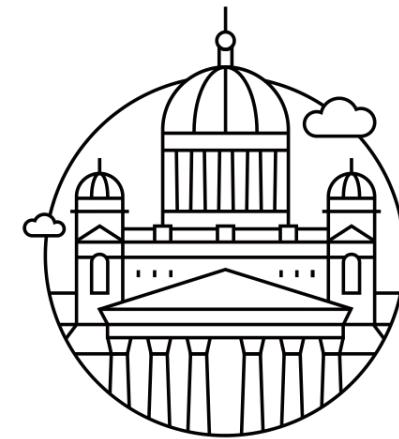
Helsinki

Strategy:

Data as enablers of innovation

Data education (MyData principles) to promote citizens' rights on personal data management

<https://www.mydata.org/about/purposes-principles/>



A non-profit organisation that promotes the rights of individuals over their personal data.

Amsterdam

Strategy:

promotion of projects in three areas: energy, mobility and circular economy

offering workshops and consultancy services to foster innovation and start-ups through the use of open data.



Tools:

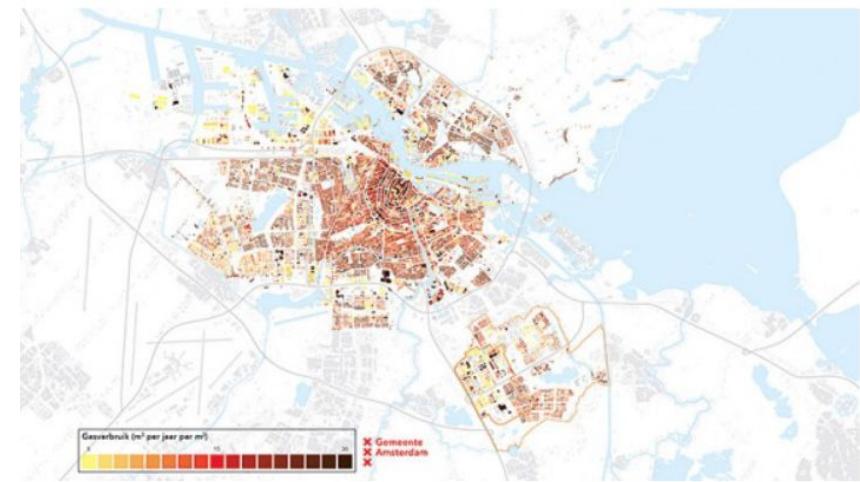
Everything published on GitHub in particular a set of processes to automate the processing, cleaning and management of data.

<https://github.com/Amsterdam/>

Energy Atlas

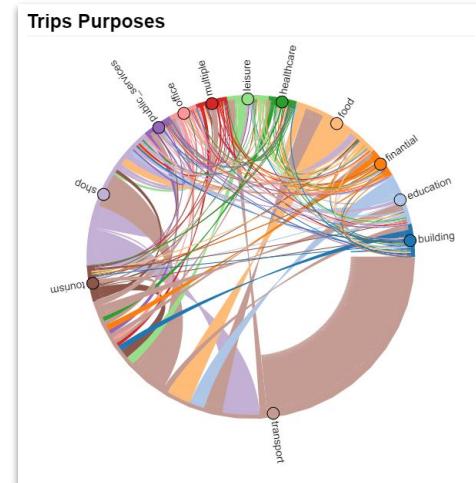
Atlas detailing information on water and energy use throughout the city, working with different public and private partners to reduce CO₂ emissions

The municipality's role as a leader and facilitator of the city administration in building the ecosystem of partners around this project has been instrumental in enabling the sharing of sensitive data.

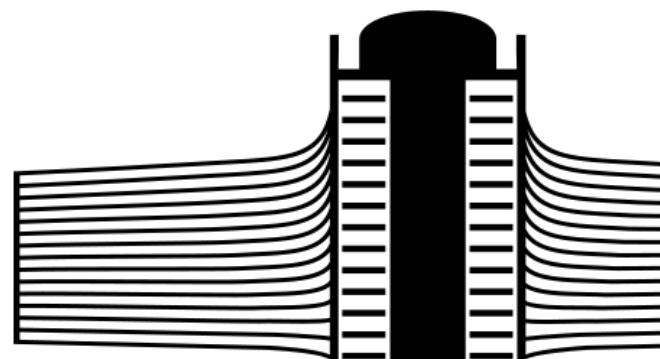


This type of partnership only works if the owners of the data see clear benefits and find ways to protect privacy.

E-scooters in Trento



European Commission



The screenshot shows the European Commission's website for the European Data Strategy. At the top, there is a navigation bar with the European Commission logo, language selection (EN English), and a search bar. Below the header, a breadcrumb trail indicates the page path: Home > Strategy and policy > Priorities > A Europe fit for the digital age > European data strategy. The main title is "European data strategy" with the subtitle "Making the EU a role model for a society empowered by data". A central text box states: "The European data strategy aims to make the EU a leader in a data-driven society. Creating a single market for data will allow it to flow freely within the EU and across sectors for the benefit of businesses, researchers and public administrations." Below this, there is a video player showing a thumbnail for a video titled "CONSCIOUS ENERGY CONSUMPTION" with a play button. To the right of the video, text reads: "Access to data and the ability to use it are essential for innovation and growth. Data-driven innovation can bring major and concrete benefits, such as" followed by a bulleted list: • personalised medicine • improved mobility • better policymaking • upgrading public services.





EN English

Search

Shaping Europe's digital future

[Home](#) | [Policies](#) | [Activities](#) | [News](#) | [Library](#) | [Funding](#) | [Calendar](#) | [Consultations](#)[Home](#) > [Policies](#) > [Data Act](#)

Data Act

The Data Act is a key measure for making more data available for use in line with EU rules and values.

The proposed Regulation on harmonised rules on fair access to and use of data — also known as the Data Act — was adopted by the Commission on 23 February 2022. [The Data Act](#) is a key pillar of the European strategy for data. It will make an important contribution to the digital transformation objective of the Digital Decade.

The new measures complement the Data Governance Regulation proposed in November 2020, the first deliverable of the European strategy for data. While the Data Governance Regulation creates the processes and structures to facilitate data, the Data



© Vadim_Ermak - iStock Getty Images Plus



Brussels, 19.2.2020
COM(2020) 66 final

**COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN
PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL
COMMITTEE AND THE COMMITTEE OF THE REGIONS**

A European strategy for data

<https://eur-lex.europa.eu/legal-content/IT/TXT/?qid=1593073685620&uri=CELEX%3A52020DC0066>



EUROPEAN
COMMISSION

Data spaces should foster an ecosystem (of companies, civil society and individuals) creating new products and services based on more accessible data.

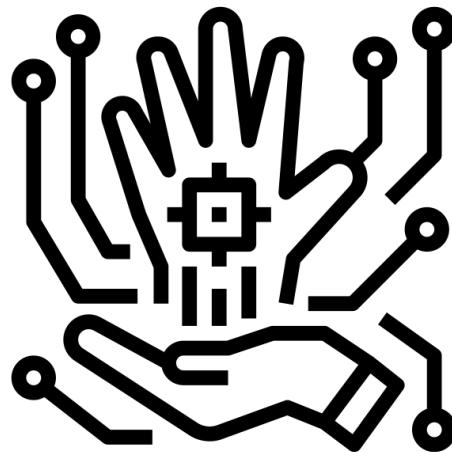
Brussels, 19.2.2020
COM(2020) 66 final

**COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN
PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL
COMMITTEE AND THE COMMITTEE OF THE REGIONS**

A European strategy for data

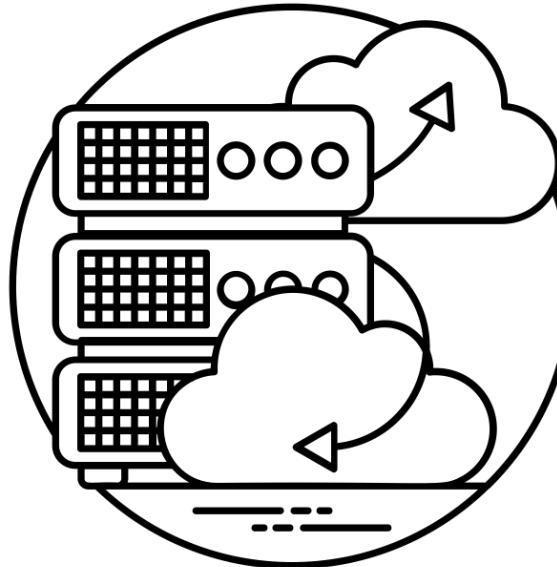
<https://eur-lex.europa.eu/legal-content/IT/TXT/?qid=1593073685620&uri=CELEX%3A52020DC0066>

Digital Skill



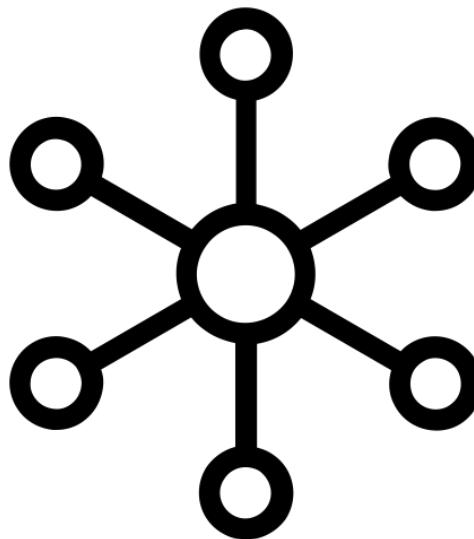
[...] improving digital skills - data literacy [...]

Technological infrastructure



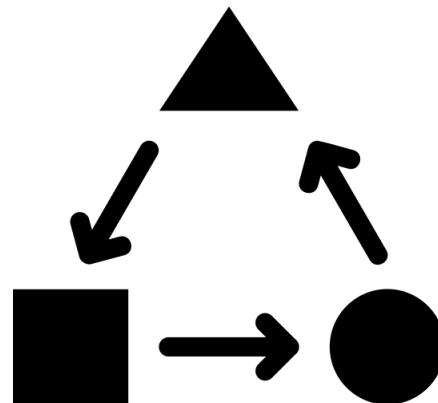
" [...] Infrastructures should support the creation of European data pools that enable Big Data analytics and machine learning, while respecting data protection and competition law, and enable data-driven ecosystems to emerge [...] "

Participants



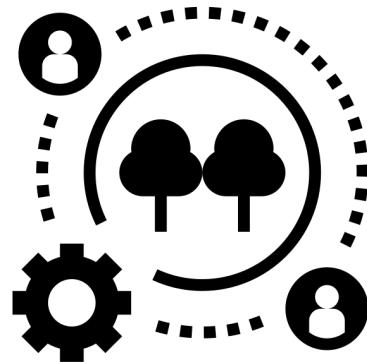
[...] Organisations contributing data would in return gain greater access to the data of other contributors, the analytical results of the data pool, services such as predictive maintenance or licensing fees. [...]

Interoperability and quality



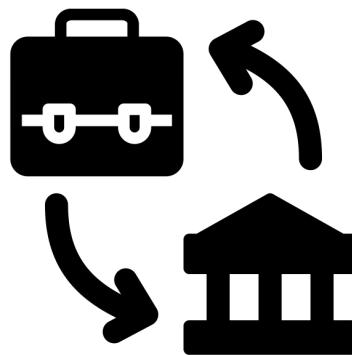
[...] Interoperability and data quality, as well as data structure, authenticity and integrity, are crucial for the exploitation of data value, especially in the context of AI deployment.
[...]

Governance

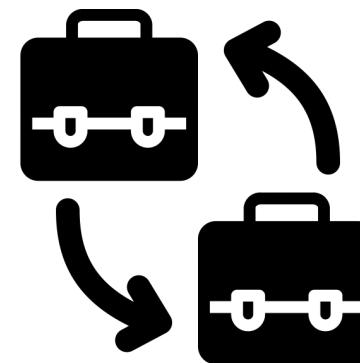


[...] further strengthen the governance of data use at societal and economic levels. For data spaces to be operational, approaches and structures are needed at the organisational level (both public and private) that make data-driven innovation possible on the basis of the existing legal framework. [...]

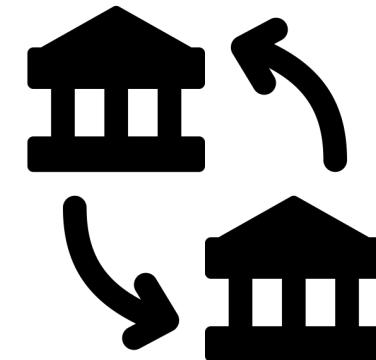
Scenarios



G: availability of
opendata
B: offering data for
public policy

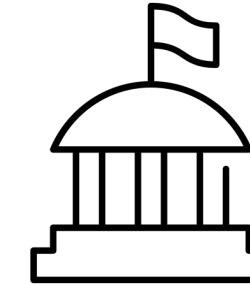
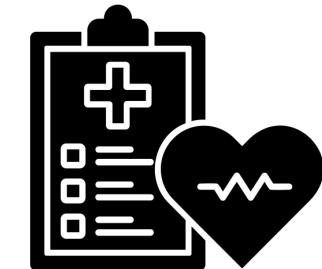
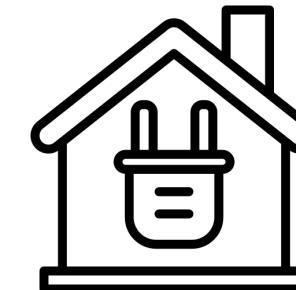
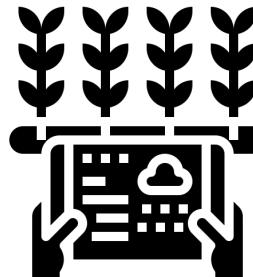
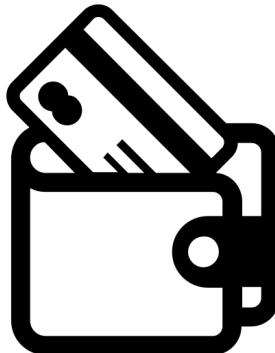


mutual trust
economic
incentives
legal clarity

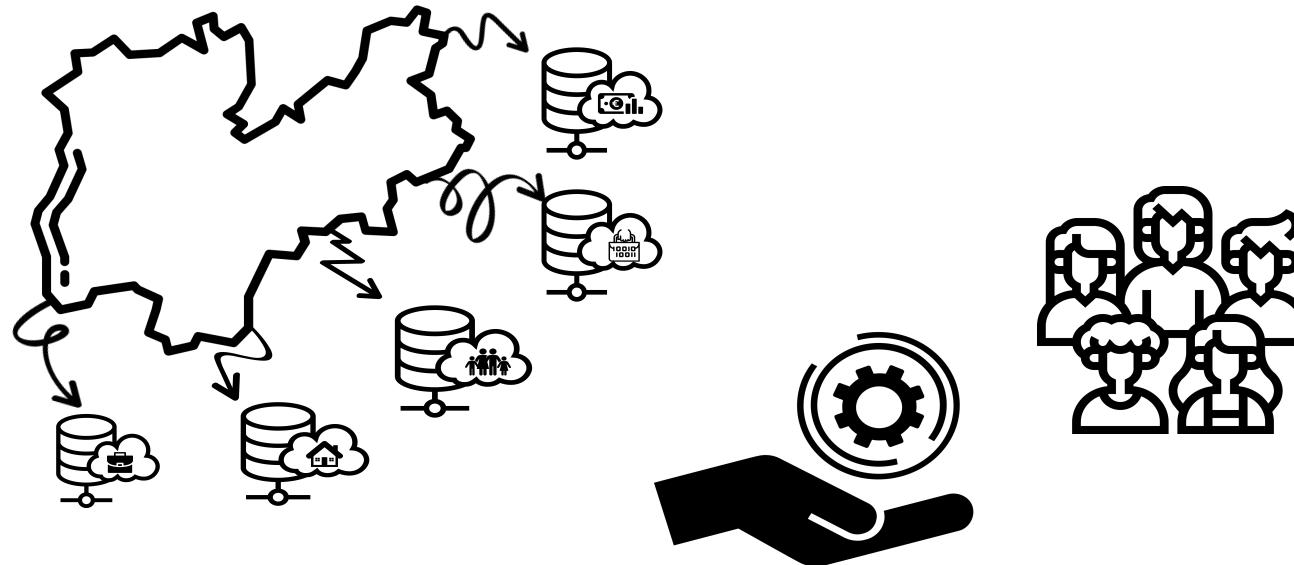


improving public
policies

Sectors for the dataspace



data as service



need to aggregate data between different sources

Master Data Management



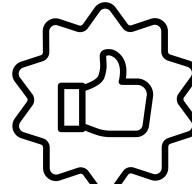
provider



privacy



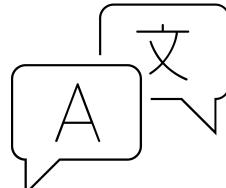
policy



quality



security



vocabularies

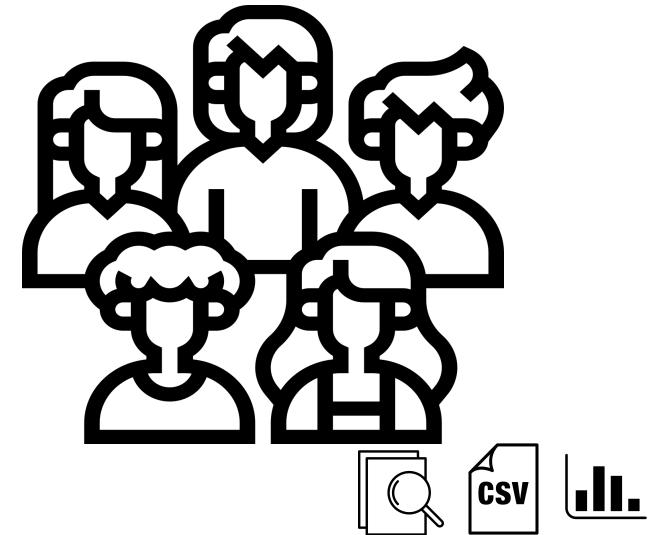
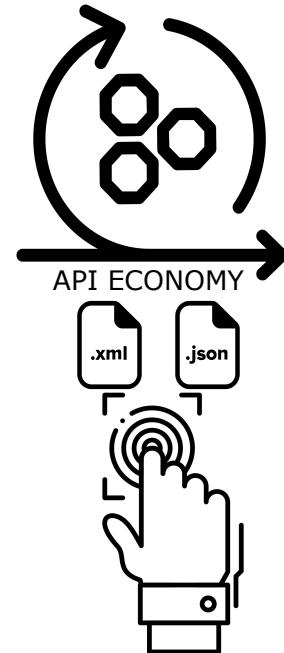
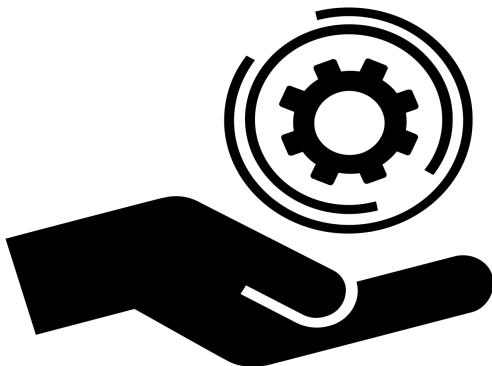


updates

Necessary steps for the creation of data that guarantees the service

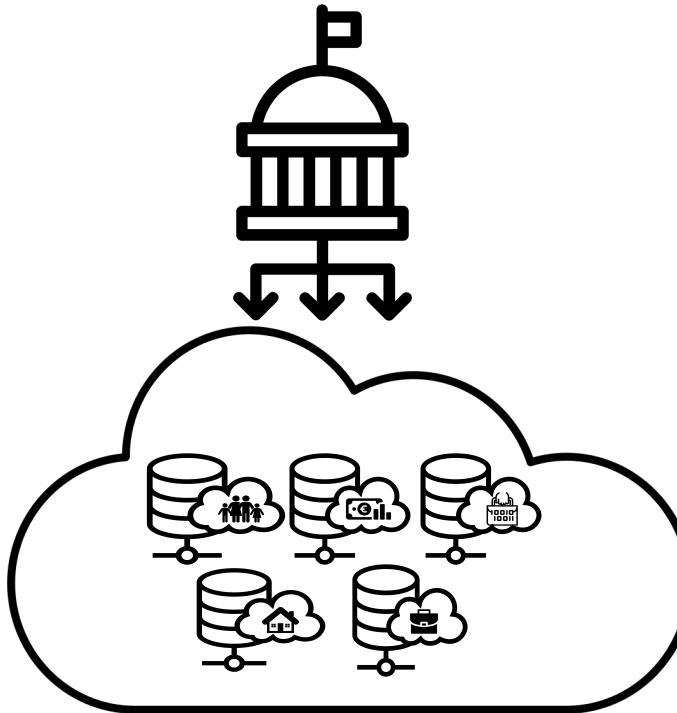


from data to services



The need of **data governance**

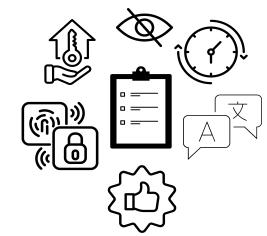
providers
constraints
new stakeholder
data stewards



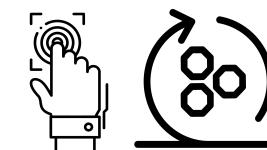
goals

foster the growth of the dataspace in terms of:

- new data
- new stakeholders
- new services



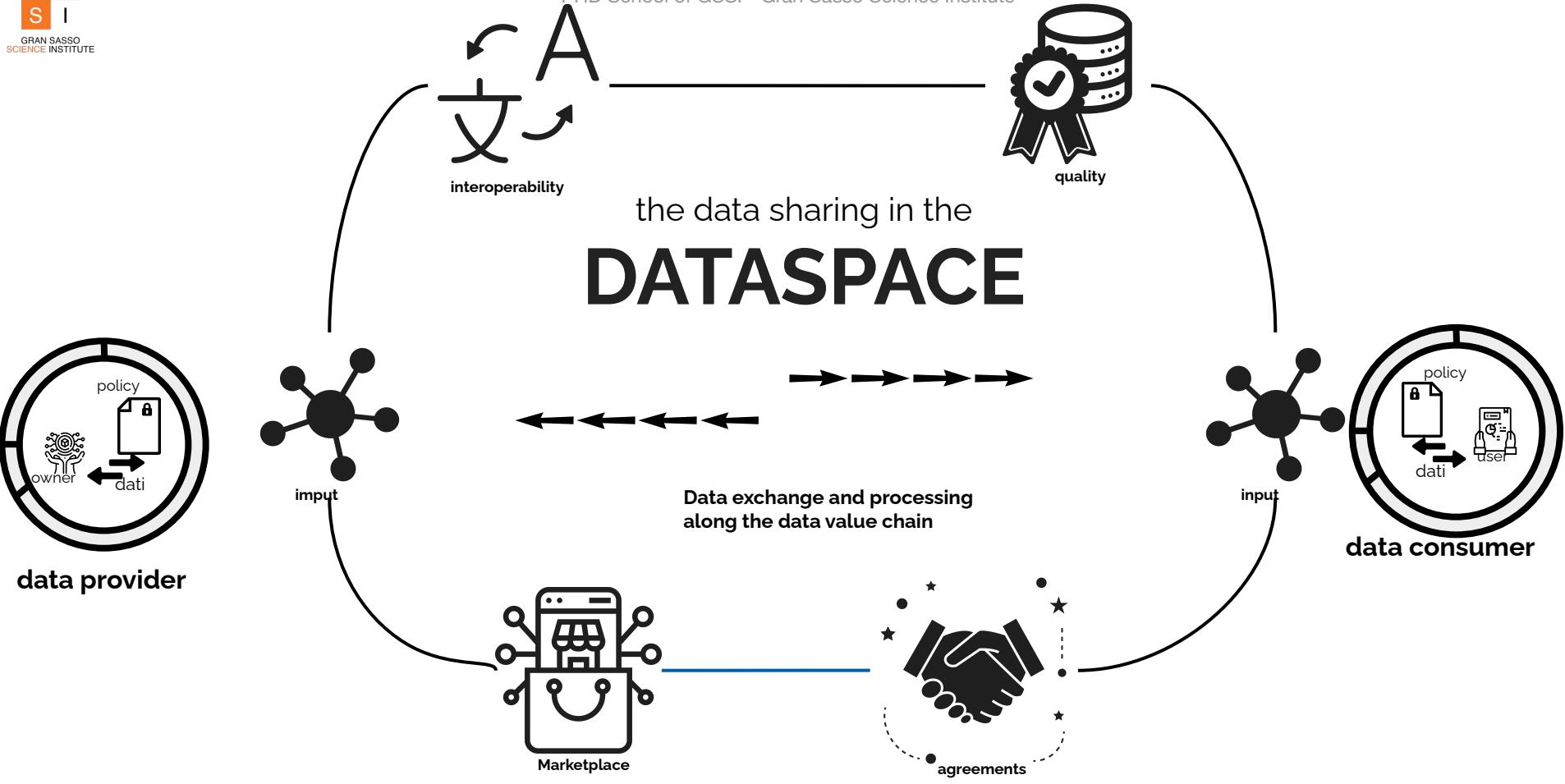
Data **quality** services and enabling **reuse** in compliance with the **constraints** set by the data supplier



Access management tools and **reuse** between machines



Data enhancement **tools** such as **monitoring** (based on **origin** and **traceability**), data **catalog** and publication of services



Open Data Hub - a good example of implementation

The screenshot shows the homepage of the Open Data Hub. At the top, there's a navigation bar with links for Project, Datasets, Services, Events, and Community. To the right of the navigation are links for Help and Start now, along with the logo for Technische Universität Südtirol.

Mission
Open Data Hub supports the implementation of an **innovative platform**, where talents can **use data and create innovative solutions** for a SMART Green Region.

[Discover our Services →](#)

Share Data!
Discover the benefits of sharing your data and how to become a Data Provider for the Open Data Hub

[Data Sharing service →](#)

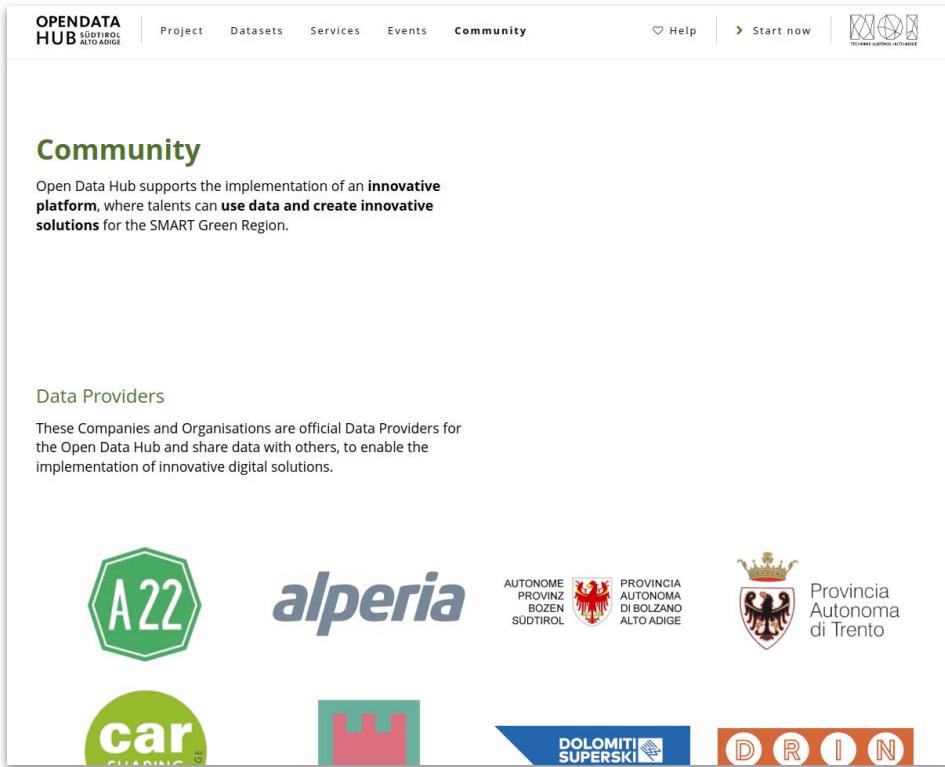
Access Data!
Discover how to access data from the Open Data Hub as a Data Consumer, to build your own apps and services

[Data Access service →](#)

CC BY

Data as a Commons - Maurizio Napolitano

Open Data Hub - the stakeholders



The screenshot shows the 'Community' section of the Open Data Hub website. The page header includes links for Project, Datasets, Services, Events, Community, Help, Start now, and a logo for 'TECHNOLOGY & INNOVATION'. The 'Community' section features a green header with the title 'Community'. Below it, a text block states: 'Open Data Hub supports the implementation of an **innovative platform**, where talents can **use data and create innovative solutions** for the SMART Green Region.' A 'Data Providers' section follows, which includes logos for A22, alperia, Provincia Autonoma di Trento, car sharing, DOLOMITI SUPERSKI, and DRIN.

Community

Open Data Hub supports the implementation of an **innovative platform**, where talents can **use data and create innovative solutions** for the SMART Green Region.

Data Providers

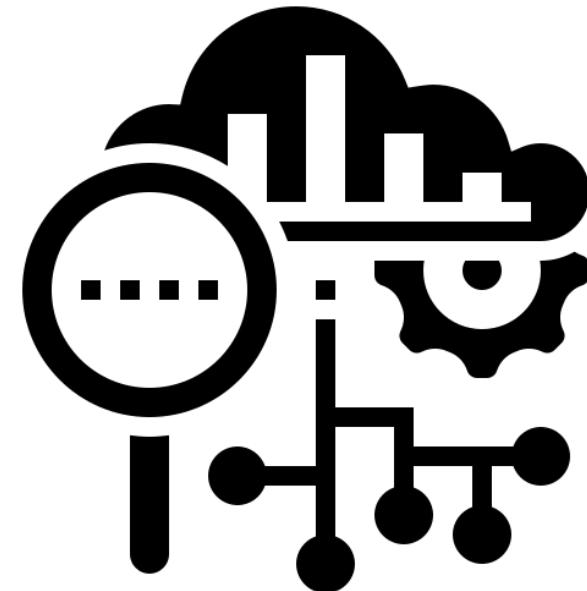
These Companies and Organisations are official Data Providers for the Open Data Hub and share data with others, to enable the implementation of innovative digital solutions.

Data Governance

"a system of decisions, rights and responsibilities for information-related processes, executed according to agreed models that describe who can take what action with what information and when, under what circumstances, using what methods"

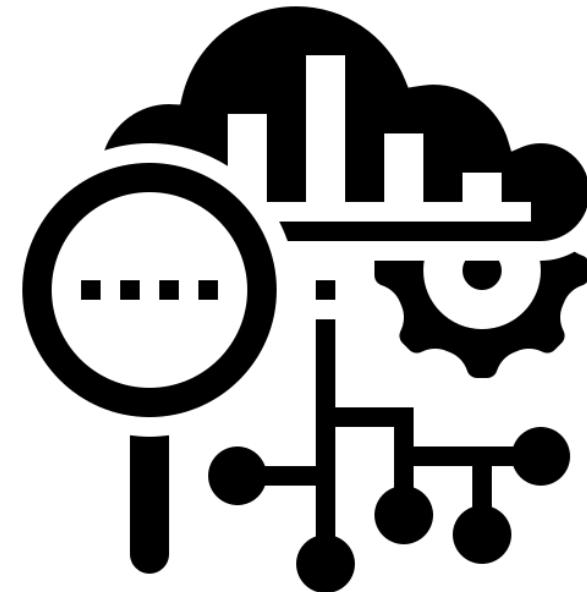
- Data Governance Institute



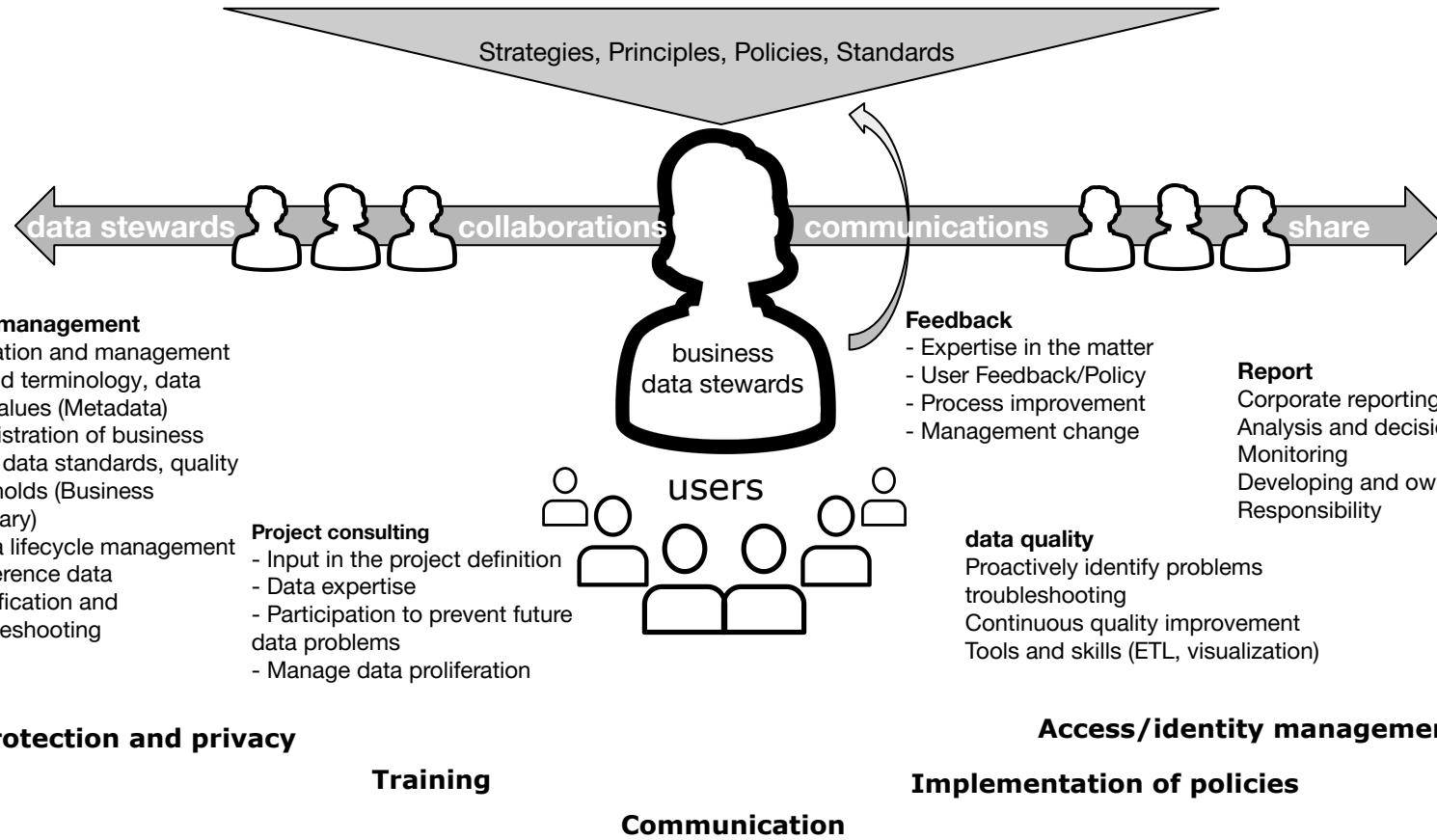
Data Governance

"a system of decisions, rights and responsibilities for information-related processes, executed according to agreed models that describe who can take what action with what information and when, under what circumstances, using what methods"

- Data Governance Institute

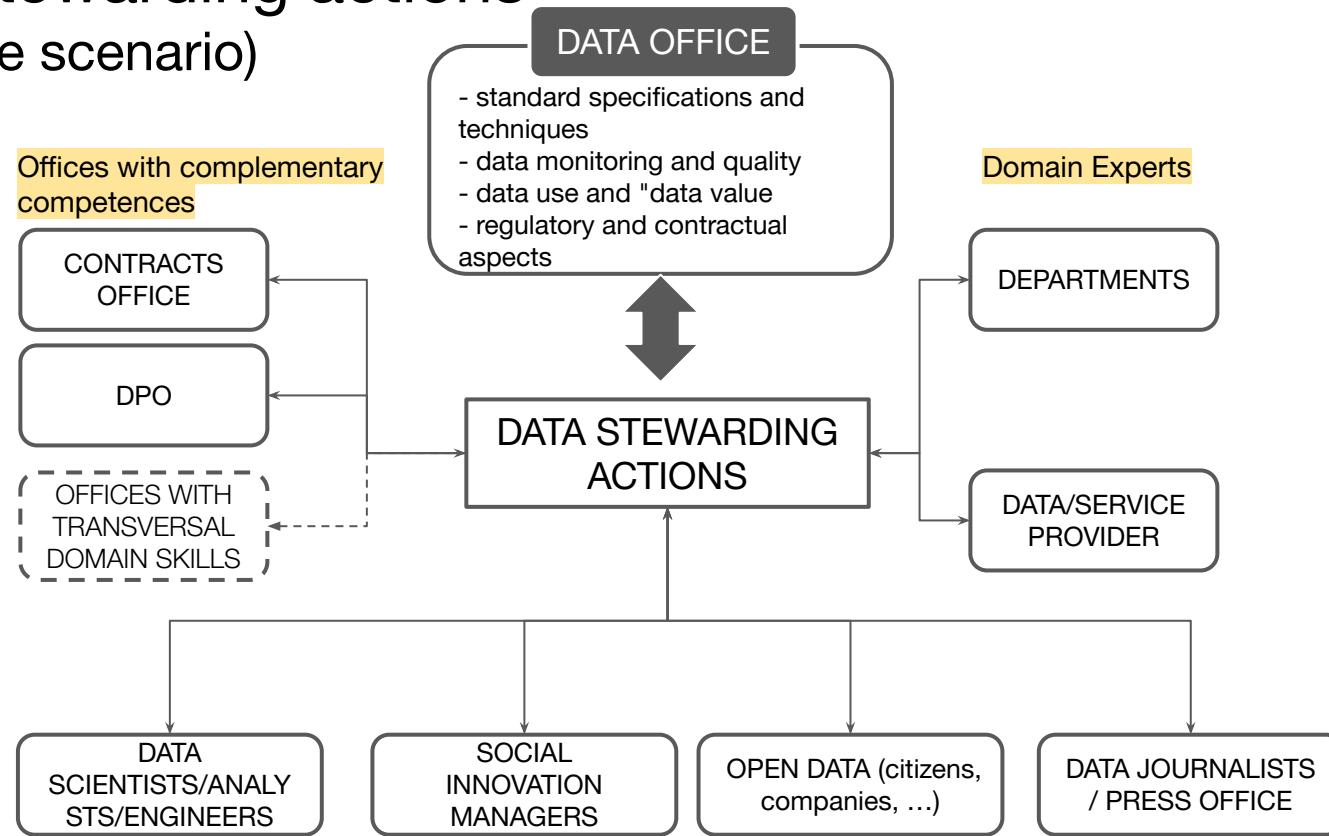


DATA GOVERNANCE



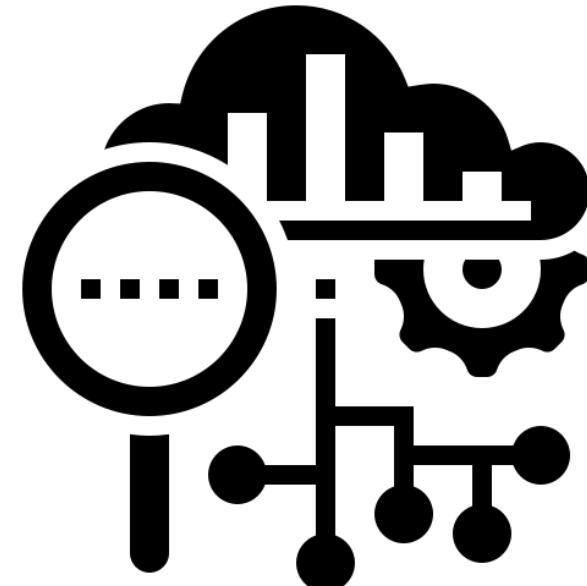
data stewarding actions

(possible scenario)



Conclusions

- the open data movement has helped put data at the centre
- the re-use of public and private data needs help from both sides
- Dataspaces are the way forward but require strong data governance
- data stewarding is the proactivity needed





for contacts:
Maurizio Napolitano



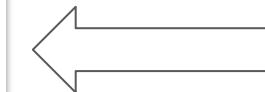
napolitano@fbk.eu



napo



Special thanks
Martina De Sanctis
Riccardo Nanni
Lihua Ma
The Noun Project



The presentation is
here

<https://bit.ly/datacommons02>