European Commission

Establishment of Sustainable Data Ecosystems

Recommendations for the evolution of spatial data infrastructures

Sébastien Martin, Prune Gautier, Slim Turki, Alexander Kotsev

2021

Joint Research Centre

EUR 30626 EN

Contact information
Name: Alexander Kotsev
Address: European Commission DG Joint Research Centre, Unit B.06, Via E. Fermi, 2749 - 21027 Ispra (VA), Italy
Email: alexander.kotsev@ec.europa.eu
Tel.: +39 0332 78 9069

EU Science Hub
https://ec.europa.eu/jrc

JRC124148

EUR 30626 EN

PDF	ISBN 978-92-76-31385-4	ISSN 1831-9424	doi:10.2760/04462

Luxembourg: Publications Office of the European Union, 2021

How to cite this report: S. Martin, Gautier, P., Turki, S., Kotsev, A. *Establishment of Sustainable Data Ecosystems. Recommendations for the evolution of spatial data infrastructures*, EUR 30626 EN, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-31385-4, doi:10.2760/04462, JRC124148

Table of Contents

Authors

Sébastien Martin, Prune Gautier, Slim Turki (Luxembourg Institute of Science and Technology)
Alexander Kotsev (Joint Research Centre, European Commission)

Acknowledgements

This publication is the result of numerous online workshops, presentations, interviews and desk research carried during 2020, in exceptional conditions.

The authors would like to thank all those who contributed directly or indirectly to the realisation of the "Data ecosystem for geospatial data" study, and in particular:

- The data ecosystem experts
    - Seen Wiid, UP42
    - Ganesh Venkatesh, UP42
    - Daniel Seidel, LiveEo
    - Bruce Winter, exactEarth Ltd.,
    - Taylor Nicholls, exactEarth Ltd.,
    - Marion Glatron, Rennes Métropole
    - Simon Saint-Georges, Rennes Métropole
    - Ghislain Delabie, OUISHARE
    - Dan Isaac, Spire
    - Rink Kruk, National Geographic Institute, Belgium
    - Janus Gohr Mørk, Agency for Data Supply and Efficiency, Denmark
    - Theo-Paul Haezebrouck, Agdatahub
    - Peter van der Veen, TripService
    - Michiel Van Dongen, Ministry of Infrastructure and Water Management, The Netherlands
    - Jean-Charles Simonin, ENEO
    - Konstantinos Plessas, University of Patras
    - Devian Diran, TNO
- Participants to the webinars for their feedback, questions and recommendations
- Colleagues from the 'Digital Economy' unit of DG JRC, and in particular Simon Vrecar, Lorena Hernandez, Francesco Pignatelli, and Michael Lutz.
- Ulla Kronborg Mazzoli and the ISA² Working *Group* on *Geospatial* Solutions.
- Robert Tomas, who initiated the work of this collaboration.
- Sven Schade, Carmelo Attardo, and Aris Tsois from DG JRC for the thorough reviews of this report.

Executive summary

"Data ecosystems for geospatial data - Evolution of Spatial Data Infrastructures" is an ELISE ISA[2] action[1] study, implemented by the Luxembourg Institute of Science and Technology (LIST) through a contract with the Joint Research Centre (JRC) of the European Commission (JRC/IPR/2019/MVP/2781).

The purpose of this study is to identify and analyse a set of successful data ecosystems and to address recommendations in support of the evolution of contemporary spatial data infrastructures that can act as a catalyst of data-driven innovation in line with the recently published European data strategy[2]. The recommendations provided here cover insights into the approaches that can be undertaken in order to ensure the evolution of contemporary spatial data infrastructures into self-sustainable data ecosystems.

Five of the identified case studies and use cases were selected for in-depth analysis based on the "Data Ecosystems Analysis Framework", presented and validated during the INSPIRE Conference webinar[3] in June 2020:

1. A **local data ecosystem** illustrated by the case of Rennes Métropole and its Rennes Urban Data interface initiative. Rennes is implementing since 2016 a collaborative and partnership-based local data strategy, targeting an inclusive and sustainable governance model for the local ecosystem, adopting the quadruple helix model. Rennes is also experimenting the concept of the City as trusted third-party allowing citizens to take back control over their personal data.
2. The **Geospatial data marketplace** is illustrated by the UP42 use case. UP42 is a marketplace and developer platform providing access to both data and analytics from multiple sources. UP42 is also offering a value distribution model that is contributing to changing the way geospatial data is accessed and analysed.
3. **Tracking technologies for supply chain** is illustrated by SPIRE. SPIRE builds and manages a constellation of nanosatellites, collecting and distributing earth observation data, Maritime data, Maritime data using AIS messages, Aviation data, using ADS-B data and weather data using Radio occultation
4. **The smart agriculture** is illustrated by API-Agro. It is a B2B data exchange platform operated by Agdatahub and a company made up of 30 partners representing the agricultural sector, from private companies and public bodies as Chambers of Agriculture, Technical agricultural institutes. It provides a

[1] https://ec.europa.eu/isa2/actions/elise_en
[2] European data strategy, February 19, 20202, https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en
[3] https://inspire.ec.europa.eu/conference2020/webinars/data-ecosystems-geospatial-data

functional, technical, commercial and legal framework for data exchange between the various stakeholders.

5. **The disaster management ecosystem** is illustrated through two case studies: The Brussels emergency services data sharing platform and the Danish Common Data on Topography, Climate and Water project preparing the country for climate change scenarios. These two cases allow to cover the emergency management phases from Mitigation, Preparedness, Response and Recovery. This highlights the importance of considering the different time dimensions: Real-time and the Historical and Simulation.

Subsequently, interviews with data ecosystems experts were complemented with desk research to enrich the analysis and to elaborate the recommendations presented in the current document. These recommendations were presented and discussed in an online stakeholder workshop on November 12th, 2020[4].

The resulting recommendations (Section 2) are presented in four complementary categories related to (i) the governance of the ecosystems, (ii) the engagement of the relevant stakeholders, (iii) the technical dimensions, and (iv) the overall economic sustainability. For each category, the main actors involved are emphasised. Furthermore, all recommendations highlight the main challenges to be overcome. As the challenges are not to be seen in isolation, other related recommendations are also indicated.

Governance

- G01. Building a collaborative governance of the ecosystem.
- G02. Identifying the most relevant actor(s) to embrace the role of orchestrator depending on the nature and evolution of the ecosystem.
- G03. Clear consideration (roles, benefits, needs and means) for all stakeholders ensure the willingness to make the data ecosystem sustainable.
- G04. The creation of a platform provides a strong incentive to structure the ecosystem.
- G05. Avoiding a fragmented landscape of stakeholders and a lack of centralised governance.
- G06. Importance of considering and aligning stakeholders' different cultures.
- G07. Exploring the role of local authorities in local data ecosystems over time
- G08. Importance of networks.

Stakeholder Engagement

- SE01. Defining and integrating the relevant stakeholders enabling the success of the ecosystem.
- SE02. Distributing value between the stakeholders.
- SE03. Considering citizen as true stakeholders.

[4] Video recording of the workshop: https://joinup.ec.europa.eu/collection/elise-european-location-interoperability-solutions-e-government/document/presentation-data-ecosystems-geospatial-data

SE04. Promoting data literacy among all stakeholders.
SE05. Organising events to increase awareness in the ecosystem and interactions frequency.
SE06. Building a data social network at the scale of the ecosystem.
SE07. Raising awareness on the incentive function played by Regulators.
SE08. Emphasizing the role of NGOs.

Technical Issues

TI01. Fostering ecosystem sustainability through problem solving approach, leading to new data cycles.
TI02. Strengthen the relationship between the data ecosystem development and the digital transformation of stakeholders.
TI03. Grasp the opportunities for data sharing by private companies as result from the GDPR entry into force.
TI04. Fostering data crowdsourcing.
TI05. Stimulate the datafication of a broader range of sectors.
TI06. Integrating data ecosystems and data cooperatives & trusts.
TI07. Put the APIs at the core of the approach.
TI08. Choosing the platform architecture (tools and capacity) based on the specific features of the ecosystem.
TI09. Integrating in the platform not only data but also services and even computational infrastructure.
TI010. Data standardisation is an enabling condition to the emergence of data ecosystems.
TI011. Identify and adopt the suitable data and metadata models, standardised where possible.
TI012. Aligning the data ecosystem with other components such as cloud and software ecosystems.
TI013. Facilitate the access to real time data and time series.

Economic Sustainability

ES01. Integration of Open Access, Open Source, Open Innovation and Open data Paradigms.
ES02. Emphasising an adaptive and agile orchestration for the ecosystem evolution, especially for data collection.
ES03. The creation of a platform is a strong enabler of business opportunities and implementation, as well as related data flows.
ES04. Synergies between individual stakeholders' business models are the key condition for the overall ecosystem sustainability.
ES05. Data ecosystems rely on long-term engagements.
ES06. Legal issues are framing the ecosystem through the definition of users' interaction rules.
ES07. Strong political and societal support facilitate the sustainability of the data ecosystem.
ES08. Extracting the value of personal data.

1. Introduction and methodology

"Data ecosystems for geospatial data - Evolution of Spatial Data Infrastructures" is an ELISE study, run by the Luxembourg Institute of Science and Technology (LIST)[5] in close collaboration with the Joint Research Centre (JRC) of the European Commission. The purpose of this study is to identify and analyse a set of successful data ecosystems and to address recommendations that can act as catalysts of data-driven innovation in line with the recently published European data strategy[6]. The work presented here tries to identify to the largest extent possible actionable items. Specifically, the study contributes with insights into the approaches that would help in the evolution of existing spatial data infrastructures (SDI), which are usually governed by the public sector and driven by data providers, to self-sustainable data ecosystems where different actors (including providers, users, intermediaries.) contribute and gain social and economic value in accordance with their specific objectives and incentives.

The overall approach described in this document is based on the identification and documentation of a set of case studies of existing data ecosystems and use cases for developing applications based on data coming from two or more data ecosystems, based on existing operational or experimental applications. Following a literature review on data ecosystem thinking and modelling, a framework consisting of three parts (Annex I) was designed. An ecosystem summary is drawn, giving an overall representation of the ecosystem key aspects. Two additional parts are detailed. One dedicated to ecosystem value dynamic illustrating how the ecosystem is structured through the resources exchanged between stakeholders, and the associated value. Consequently, the ecosystem data flows represent the ecosystem from a complementary and more technical perspective, representing the flows and the data cycles associated to a given scenario. These two parts provide good proxies to evaluate the health and the maturity of a data ecosystem.

A first set of proposed data ecosystems are provided in Annex II. Consequently, eight data ecosystems (Annex III) covering a large range of domains, where geospatial data play a significant role, are summarized based on desk research. From this, a first set of recommendation is extracted, aiming at being confirmed and completed by a second wave of use cases analyses. Finally, five of the identified case studies and use cases were selected for in-depth analysis (Annex IV) based on the "Data Ecosystems Analysis Framework", presented and validated during the INSPIRE Conference webinar[7] in June 2020:

1. **Local data ecosystem**, illustrated by Rennes Urban Data Interface (RUDI);
2. **Tracking technologies for the supply chain** ecosystem, illustrated by Spire;
3. **Geospatial Data Marketplace**, illustrated by UP42;

[5] https://www.list.lu
[6] European data strategy, February 19, 20202, https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en
[7] https://inspire.ec.europa.eu/conference2020/webinars/data-ecosystems-geospatial-data

4. **Smart Agriculture ecosystem**, illustrated by API-Agro, and
5. **Disaster management ecosystem**, illustrated by the Brussels emergency services data sharing platform and the Danish Common Data on Topography, Climate and Water project.

Semi-structured interviews with data ecosystems experts were complemented with desk research to enrich the analysis and to elaborate the recommendations presented in the current document. Interviewed experts contributed both to the co-creation of the recommendations, and the subsequent validation of those recommendations provided by the other interviewed experts. In addition, these recommendations were presented and discussed in a stakeholder workshop on November 12th, 2020.

2. Recommendations

The recommendations provided below provide insights into the approaches that can be undertaken in order to ensure the evolution of contemporary spatial data infrastructures into self-sustainable data ecosystems. In order to give a clear picture of the recommendations, this section is divided in the four main categories that highlight the gap between a data provision philosophy and an ecosystem thinking.

Ecosystem Governance	**Technical issues** *(accessibility & technical aspects)*
Stakeholders engagement	**Economic sustainability**

The recommendations inside the categories are ordered according to our perceived criticality and logic. For each of the four categories, a brief introduction of the context and the main actors is followed by recommendations which are provided in a structured manner. For each recommendation, the challenge at stake is summarised alongside the main barriers which need to be overcome. The description continues with best practice(s) which indicate how the challenges can be overcome. Recommendations are not to be considered in isolation, so relevant recommendations are also suggested.

	Ecosystem Governance
Context	The ecosystem governance designates the body of rules, procedures and practices that relate to the way interactions between the actors in the data economy are framed. An ecosystem approach that is fit for purpose arranges all interactions and exchanges issues within the data ecosystem. One of the main ecosystem governance issues, as described below, is to define which actor can embrace the role of orchestrator of the ecosystem.
Actors	Different actors are involved in governance, depending on the stage of the data life cycle / value chain. Considering geospatial data for instance, public actors still have a strong role often through a legislatively defined mandate, however the private sector is increasingly contributing to the data creation processes for example through technical capabilities or even data providers by themselves.

#	**Recommendation EG01: Building a collaborative governance of the ecosystem.**
Challenge	Governance is the key of the ecosystem sustainability. It is framing both the technical arrangement and the interactions between participants. Therefore, the key stakeholders need to agree on their roles and responsibilities beforehand.
Barriers	The governance model is difficult to adapt over time. Difficulty to get a general agreement on specific issues, e.g. when competing entities have similar objectives.
Best practices	In Smart Agriculture, the API-Agro has chosen to be established as a company (Simplified Joint Stock Company). It is the outcome of a project and an industry-wide reflection on a large agricultural data portal specifically targeting the constitution of collaborative governance. They are 30 shareholders, and the governance is distributed between many public and private actors. Moreover, this use case illustrates the importance of predefining the exchanges settings on the components to distribute and their modalities. It shows also that a common agreement is mandatory to ensure the later acceptability of the platform and therefore of the ecosystem. In the Disaster Management ecosystem, a collaborative

	governance is a prerequisite to enable the data sharing between stakeholders not always eager to share their assets. The public sector, backed up by related legislation plays a prominent role in liaising with the different stakeholders.
Related #	SE01 - TI08 - EG02

#	**Recommendation EG02: Identifying the most relevant actor(s) to embrace the role of orchestrator depending on the nature and evolution of the ecosystem.**
Challenge	The orchestration of an ecosystem is a key activity to insure its sustainability. The orchestrator is in good position to know the needs of the stakeholders. However, the role of Orchestrator is not obvious and independent, and for example we could wonder how a marketplace actor may be a platform leader. Often, this function is taken by another player of the ecosystem: a data provider, a public authority, or a data user. Moreover, the competences and skills needed to allow a stakeholder to act in the role of orchestrator can change during the life of the ecosystem.
Barriers	Difficulty to engage an actor skilled and incentivised enough to play this central role. Difficulty to assess the Return on investment of activities related to ecosystem orchestration.
Best practices	In the data marketplace ecosystem, the central position of UP42, for instance, makes it possible to orient the ecosystem development based on feedbacks of their customers. More specifically, a central actor may identify new partners to bring on board, partly based on other partners suggestions. In Spire, as there is no central orchestrator, orchestration functions are led by diverse actors in a distributed manner, based on local and specific business opportunities. In Smart agriculture, the choice of a private orchestrator is based on the hypothesis that business related issues are crucial for the sustainability of the ecosystem.
Related #:	SE01 - EG07 – TI01

#	**Recommendation EG03: Clear consideration (roles, benefits, needs and means) for all stakeholders ensure the willingness to make the data ecosystem sustainable.**
Challenge	Clear governance guidelines are required to ensure seamless interactions.

	In particular, SMEs means and perspectives have to be considered.
Barriers	Governance rules need to take into account opposite perspectives.
Best practices	The Mobility data ecosystem, as well as RUDI, highlight this aspect for public transport services. Indeed, the benefit for each stakeholder is easily understood. Users gain better services, bus companies more customers and cities public value.
Related #	ES08 - TI03 – ES04 – ES05

#	**Recommendation EG04: The creation of a platform provides a strong incentive to structure the ecosystem.**
Challenge	An ecosystem needs, as a prerequisite, to facilitate the interactions between stakeholders and the data exchanges, a platform that enables for example timesaving (e.g. data discovery) and decreases negotiation costs. In addition, it enables the creation of easily scalable products.
Barriers	Willingness of the stakeholders to collaborate. The need to standardise formats. Competing platforms. Risk of dominance of one actor (the operator) over all the rest.
Best practices	3/5rd of the analysed ecosystems are based around platforms. For Spire, they prefer for the moment to avoid any intermediaries in their business relationships. For the disaster management ecosystem, the lack of data discoverability was one of the main factors leading to the creation of the ecosystem based on a platform enabling convenient data discoverability. The data marketplace ecosystem is a particular illustration of the advantages of choosing a platform based on a marketplace model, which has implication for governance rules.
Related #:	TI08 – SE02

#	**Recommendation EG05: Avoiding a fragmented landscape of stakeholders and a lack of centralised governance.**
Challenge	In an ecosystem, stakeholders are numerous and various, and their links are not tight and obvious. Therefore, a significant risk of silos' thinking exists.
Barriers	Absence of a central actor
Best	In the disaster management ecosystem, the orchestrator sets

practices	incentives to facilitate the development of a data sharing culture, through the organisation of collaborative activities showing the long-term benefits and the value of an ecosystem perspective.
Related #	EG06 – EG01– TI08 – SE01 – TI03

#	**Recommendation EG06: Importance of considering and aligning stakeholders' different cultures and expectations.**
Challenge	To reach its goal, an ecosystem needs to encompass actors from heterogeneous backgrounds for a common purpose. They also have different ways of working and difficulty to understand each other working approaches.
Barriers	Public and private actors may have contradictory objectives and different working cultures, preventing an efficient structuration of the ecosystem.
Best practices	In RUDI, to tackle this issue and to define a common way of working, the actors decided to organise frequent meetings to identify the issues, enable the solutions and ensure the commitment of all ecosystem's stakeholders. In addition, a social sciences laboratory is working with the Metropole on this specific topic.
Related #	TI01 - EG03 - EG08

#	**Recommendation EG07: Exploring the role of local authorities in local data ecosystems over time**
Challenge	In local data ecosystems, the role of the orchestrator is often endorsed by local authorities, but their role in the long-term still need to be investigated. Another challenge is to shift from an active engagement to a distant public orchestration.
Barriers	Lack of empirical evidence. Financial issue (citizen taxes). Financial assessment of public value creation
Best practices	Within the RUDI ecosystem, local authorities may have a leading role during the emergence stage, building on previous Open Data initiatives. Empirically, the orchestration is composed of networking initiatives, creation of a data exchange platform and its governance rules and a financial role stimulating the ecosystem through public procurement.
Related #	ES01 – ES05 – TI01

#	**Recommendation EG08: Importance of stakeholder networks.**
Challenge	How to leverage state-of-the-art research output remains a challenge. To support the sustainability of ecosystems, networks can be used to share good practices, to benchmark initiatives, but also to provide insights. At the same time, it is crucial to consider existing projects instead of re-buildings ecosystems.
Barriers	Low awareness. Lack of incentives. Lack of interest to participate in networks.
Best practices	Both RUDI and the Smart agriculture ecosystems claim to be leveraging on the scalable use of research outputs. Rennes Métropole, as a member of Smart Cities Network (e.g. Eurocities Knowledge Society Forum[8]), intends to tap into the network in order to validate and disseminate the outcomes of their project, as well as to improve it with feedback and approaches shared by other smart cities.
Related #	EG07

[8] https://eurocities.eu/

	Stakeholder Engagement
Context	Stakeholder engagement is the process used within an ecosystem to engage relevant stakeholders for a clear purpose. It obliges to involve stakeholders in identifying, understanding and responding to sustainability issues and concerns. The challenge is the risk of not having enough contributors (not achieving critical mass) – leading to lower service creation.
Actors	Stakeholder engagement is one of the main functions of an orchestrator.

#	**Recommendation SE01: Defining and integrating the relevant stakeholders enabling the success of the ecosystem.**
Challenge	In less structured ecosystems, there are difficulties for companies to identify the fruitful interdependencies and thus business opportunities. This function is an essential component of the orchestrator.
Barriers	Low degree of awareness in the ecosystem.
Best practices	For data marketplace ecosystems, this function is endorsed by the main orchestrator leading the platform. Data aggregators, especially in the absence of a central platform, play an important role in the initial structuration of the ecosystem and especially the data flow. The same stand for the IT Service Providers from the Services provision side.
Related#	EG01 - EG02 – ES04

#	**Recommendation SE02: Distributing value between the stakeholders**
Challenge	How to distribute value among stakeholders in a fair and balanced way in order to ensure the stakeholders' engagement.
Barriers	Difficulty to compare tangible and intangible assets. Monetization of non-financial values such as public value, common good, quality of life, well-being.
Best practices	The data marketplace ecosystem highlights the role of the central orchestrator. Each stakeholder of the ecosystem is considered as a particular partner, and the financial remunerations are negotiated on an individual basis, considering data and/or service provided,

	as well as their engagement's value (revenue sharing model).
Related#	ES04

#	**Recommendation SE03: Considering citizen as true stakeholders**
Challenge	Citizens are more than just data objects. The goal is not to consider citizens only as data providers or end-users consuming products and services, but as central actors contributing the definition of the scope and the goals of the ecosystem, and being active participants aware of their data alongside their re-use.
Barriers	Trust (of citizen regarding public actors' concerns). Lack of citizen's interest considering data issue. Reluctance to share personal data. Perception of legal risk concerning personal data.
Best practices	In RUDI, inclusive data ecosystem governance based on co-construction of the ecosystem and its governance rules is led with participatory labs and collaborative workshops. Moreover, trusted third parties are considered as a mitigation solution to trust issues concerning citizen data sharing.
Related #	SE05 – TI01 – TI04 – SE08

#	**Recommendation SE04: Promoting data literacy among all stakeholders**
Challenge	There are heterogeneous levels of knowledge and awareness regarding the potentials and the methods to benefit from the re-use data.
Barriers	Some data producers do not realise the usefulness of their data for the creation of data-driven services providers (data re-users).
Best practices	In RUDI, the orchestrator uses workshops, trainings and panel sessions open to all stakeholders in order to raise the level of knowledge among the Quadruple Helix representatives (University, Industry, Government, and the Public). UP42, as platform leader, is organising workshops and promotes the uptake of state-of-the-art knowledge and methods on its platform, to orchestrate the knowledge dissemination in its ecosystem. Farmers are not traditionally used to utilise data in their day-to-day work, even if data-centred agriculture is their future. Therefore, API-Agro, through its role of orchestrator, has a wide mission to improve data literacy with services providers and

	associations. For logistic and tracking, there is a need to bring geospatial data literacy to a broader set of new actors of the ecosystem, including the insurance sector.
Related #	EG02 – EG06 – TI01 – D10 – ES01

#	**Recommendation SE05: Organising events to increase awareness in the ecosystem and interactions frequency.**
Challenge	Reaching the critical mass of stakeholders is not sufficient for the ecosystem to be effective: the viability of the ecosystem depends on the interactions between stakeholders.
Barriers	The data provision paradigm.
Best practices	In RUDI, the local authority intends to take benefit of external events, such as InOut[9], to organize parallel working sessions dedicated to the local data ecosystem.
Related #	SE04 - SE03

#	**Recommendation SE06: Building a data social network at the scale of the ecosystem.**
Challenge	Combining a technical layer (e.g. data catalogue) with a networking layer where actors may meet, exchange and request data, and thus, frame the ecosystem and increase the self-awareness of the different actors on the functioning of the ecosystem. This way, the ecosystem may attract new participant and ensure the current ones.
Barriers	
Best practices	For RUDI, Rennes intends to build a socio-technical component where stakeholders may fulfil their needs for data, processing and expand the network of business partners.
Related #	EG08 – SE01 – SE03

#	**Recommendation SE07: Raising awareness on the role of regulators for the creation of incentives**
Challenge	Integrate the decision made by Regulators on diverse topics within the broader context of data ecosystems.
Barriers	Regulation not enough connected to - aware of- business issues.

[9] https://inout.rennes.fr/en/

Best practices	The logistic tracking ecosystem is based on AIS (Automatic identification system) data which collection and reuse are made mandatory by the International Maritime Organization (IMO). Nevertheless, this regulation was not directly intending to build a data ecosystem.
Related #	EG02

#	**Recommendation SE08: Emphasizing the role of NGOs**
Challenge	Non-Governmental Organisations (NGO) are valuable actors to engage as data providers but also to increase citizen engagement.
Barriers	None.
Best practices	In Smart Agriculture, NGOs are key actors from the data collection perspective, but also by encouraging data owners to share the data with other data ecosystems participants. They can rely on a long tradition of cooperative organisation. Moreover, in RUDI, NGOs are expected to complement the role of orchestration currently done by the local authority by becoming themselves "call for projects" leaders.
Related #	SE01

	Technical Issues
Challenge	This represents the technical side of the ecosystem. It covers a plethora of different cross-cutting technical issues related to the access to data, architectures, standards, and technologies for both the provision and use of data.
Actors	Considering the stage of the ecosystem, the leading role can be embraced by different actors, but given the specificity and technological nature of data ecosystems, all actors are to be considered impacted.

	Recommendation TI01: Fostering ecosystem sustainability through problem solving approach, leading to new data cycles
Challenge	The use and reuse of data of different kinds, originating from various sources (public, private, personal data, etc.), through new services, is expected to create a virtuous circle leading to new data cycles. In particular, it is suitable to create new (public) data intensive services, customized for the needs of different purposes. How to extract the value of combined albeit heterogeneous data remains challenging.
Barriers	Data-sharing approaches and decisions steered only by data providers. Data quality. Data formats (proprietary – not proprietary, documentation, etc.). Data standards (data models, formats, etc.).
Best practices	In RUDI, the orchestrator is organising “calls for projects” with the specific target to make the ecosystem more tangible, to define more accurate governance rules, to showcase the business opportunities. These “projects” are using already accessible data to produce data-driven services that in turn create new datasets, which are accessible and reusable, i.e. contribute to the uptake of private data for public good. In parallel, RUDI is organising dedicated hackathons. In Smart agriculture, personalised services (such as precise irrigation) combine a large range of data. In the logistic and tracking ecosystem, the ecosystem approach allows to integrate new downstream actors such as insurance companies (as data re-users and new data providers). In the data marketplace ecosystem, LiveEO is developing a

	predictive infrastructure management service for Deutsche Bahn. Through this activity, it generated new data that are made available through the data marketplace.
Related #	ES04 - ES08– D10

#	**Recommendation TI02: Strengthen the relationship between the data ecosystem development and the digital transformation of stakeholders**
Challenge	Data provision can be hindered by internal legacy IT systems. The digital transformation of public services may be enabled by data ecosystems through the provision of field-generated data instead of purely statistical data.
Barriers	Access to IoT data. Data literacy. Legacy infrastructures and systems.
Best practices	In RUDI, the local authority encourages the release of data through showcasing profitable business models and orienting call for projects leading at the same time to the update of these internal systems. The Energy ecosystem has been stimulated by the Industry 4.0 development. The Smart agriculture ecosystem showcases that the democratization of IoT devices and systems intensifies the exchanges of data and leads to completely new data cycles. Beyond the update of systems, it is also an argument to foster the participation in the ecosystem, as the disaster management and the Smart Agriculture ecosystems highlighted that organisations may capitalize on their information.
Related#	EG03 – TI01

#	**Recommendation TI03: Grasp the opportunities for data sharing by private companies as result from the GDPR entry into force.**
Challenge	The challenge for companies is to define processes allowing the release of personal data while complying with the provisions of the GDPR.
Barriers	Legal responsibility Lack of well-established approaches for sharing personal data while preserving privacy
Best practices	In RUDI, the orchestrator intends to support the stakeholders in the release of personal data with a legal support.

Related#	ES08 – SE02

	Recommendation TI04: Fostering data crowdsourcing
Challenge	Crowdsourced data are valuable but might be difficult to collect. In addition, the quality of the data that is collected needs to be ensured, as well as the sustainability of the data-collection approach.
Barriers	Personal data protection rules and privacy. Lack of consideration and long-term interest. Data literacy. Data quality and validation.
Best practice	For crowdsourced data in the Weather ecosystem, the incentives could be micro-payments or good will to contribute to the public good, but also the role of entertainment to motivate participation. For instance, in the United States, the crowdsourcing of such data is shaped by the National Oceanic and Atmospheric Administration (NOAA). NGOs play an important role in the local ecosystem of Rennes for promoting and explaining the value of personal data. In the mobility ecosystem, citizen input is proven to be useful for the sustainability of the ecosystem through the collection of new data.
Related#	SE04 – TI01 – ES08

	Recommendation TI05: Stimulate the datafication of a broader range of sectors
Challenge	Clearly, all domains are not at the same in terms of digital maturity. Therefore, they do not have the same opportunity to participate in the emergence of data ecosystems.
Barriers	Digital literacy. Data Standards. Awareness of the ecosystem.
Best practices	In logistic and tracking, the new business opportunities created by the ecosystem for insurance companies are a strong incentive for the datafication of the sector. For the Smart agriculture ecosystem, the API of API-Agro Platform is a driver to encourage all relevant stakeholders to produce new or better data.
Related #:	SE04 – ES04

	Recommendation TI06: Integrating data ecosystems and data cooperatives & trusts
Challenge	Building on data cooperatives for the benefit of other data spaces, to enrich data ecosystems. Defining the technical links between data cooperatives and data platforms.
Barriers	Data collection rules (privacy, quality, roles and responsibilities).
Best practices	Within the healthcare ecosystem, platforms exist such as OpenHumans.org which organise personal data pipelines, handling especially the consent for the data reuse. One important incentive is the awareness of contributing to public good through data creation or sharing.
Related#	ES08 – EG07 – ES04

	Recommendation TI07: Put the APIs at the core of data sharing.
Challenge	APIs act as the 'glue' of data ecosystems as they link different already existing architectures in practical terms. APIs are therefore to be considered as a mandatory component for ensuring the economic viability and sustainability of data ecosystems.
Barriers	Large dependency on pre-existing systems. Development costs. Data ownership issues.
Best practices	In RUDI, the API has to link heterogeneous systems such as the local open data platform, or companies' data. One of the achievements of UP42 is to provide such APIs for data sources but also processing APIs for running algorithms or financial clearing. For the Smart agriculture ecosystem, the API managed by the central platform is considered as a means to orchestrate internal and external APIs of stakeholders.
Related#	EG04 – TI08 – ES03

	Recommendation TI08: Choosing the platform architecture (tools and capacity) based on the specific features of the ecosystem.
Challenge	Different architectures are possible (e.g. federated, centralized, edge). No universal solution exists, and those different architectural approaches present various benefits and risks.

	Another challenge relates to the necessity for handling data within a single architecture, but with different access- and users-rights.
Barriers	Lack of data discoverability. Absence of a single solution that can satisfy the needs of all actors.
Best practices	For UP42, the choice of a centralised architecture allows to offer on their platform data quality improvement blocks at the central node of the architecture, offered as block processing services. RUDI selected a different approach, a federated, well adapted to big players of its local ecosystem. As they are often equipped with complex and legacy systems, it is easier to add a node in a federated system. Moreover, a federated architecture makes it possible to access and manipulate some data unreachable by other architecture models. Nevertheless, there is a doubt pertaining the suitability for SMEs.
Related#	EG03 – EG04 – TI07 – TI10

#	**Recommendation TI09: Integrating in the platform not only data but also services and even computational infrastructure.**
Challenge	Currently, platforms are mainly based on data provision. The challenge is to add complementary services to offer a single point of supply.
Barriers	Different expertise needed.
Best practices	UP42 has chosen this option by design. This brings costs sharing and economy of scale between the actors. RUDI is considering this option, but it implies to review a large range of component from technical ones to government rules.
Related#	TI01 – EG04

#	**Recommendation TI10: Data standardisation is an enabling condition to the emergence of data ecosystems.**
Challenge	Standardisation is heterogeneous depending on the application domain. Within the 5 in-depth analyses, geospatial data standardisation was never the main issue. However, downstream data that have to be combined are more often problematic.
Barriers	Standardisation of Data Models. Commonly agreed standards. Rapid change of technology. De facto versus de jure standards.

	Conflicting or competing standards.
Best practices	Within the logistic and tracking ecosystem, the standardisation of Automatic Identification System (AIS) messages is of paramount importance to create solid foundations for the ecosystem.
Related#	TI12 - TI01 – EG01 – ES08

#	**Recommendation TI11: Identify and adopt the suitable data and metadata models, standardised where possible.**
Challenge	One of the main challenges faced by ecosystems is the lack of common structures and semantic models (ontologies). This is also a standardisation issue. In addition, ensuring that data is provided at the right semantic granularity level thus ensuring the data ecosystem viability in essential.
Barriers	Lack of capacity. Immature standards. Lack of support by mainstream software tools. Approaches that do not follow standards.
Best practices	The disaster management ecosystem illustrates the bidirectional relationship between standardisation and ecosystem emergence, with standardisation being at the same time a prerequisite and a positive outcome of the ecosystem. In RUDI, the definition of standardised data models is not a prerequisite for the emergence of the ecosystem. The standardisation of data models is expected to evolve over time. Rennes intends to address this issue through the provision of the necessary details through the metadata of the data catalogue. For logistics and tracking purposes, the right granularity level is of paramount importance not only to comply with the data collection requirements from the regulation, but also to fulfil the needs of downstream use such as insurance. The smart agriculture ecosystem confirms the possibility to proceed a progressive alignment of the data and solutions to standards and thus to interoperable and API based ecosystems. The Legal ecosystem showcases the importance of two concepts: data provenance, and data re-use traceability. Concerning data re-use traceability, it is an important dimension of the judiciary side of the data ecosystem as some legal systems are requiring to track who accessed which piece when a case is processed, and digital documents and data make this more complicated to ensure. Data provenance is of paramount importance to ensure

	not only the trust, but also the mere validity of the analysis.
Related#	TI10 – SE04 – EG01

#	**Recommendation TI12: Aligning the data ecosystem with other components such as cloud and software ecosystems**
Challenge	Ensuring the technical evolutions over time. Ensuring the links with cloud infrastructures and the software and hardware industries.
Barriers	Resistance to change. Digital literacy. Different objectives.
Best practices	For Logistics and Tracking, the ecosystem sustainability relies on the development of up-to-date algorithms, coming from external domains such as big data and artificial intelligence. As a marketplace, UP42 also encompass the alignment of the platform with the other relevant components such as processing algorithms and infrastructures.
Related#	TI13 – EG03 – SE02

#	**Recommendation TI13: Facilitate the access to real time data and time series.**
Challenge	One of the challenges is to link the data ecosystem to external cloud infrastructures, for example to handle over-sized datasets and to benefit from the rich ecosystem of algorithm developers. Another challenge is to ensure the availability and utilisation of real-time data streams.
Barriers	Loading time. Single point of access. Outdated legacy IT systems. The IoT landscape is fragmented. Multiple competing standards. vendor lock-in. Networks latencies acting as a bottleneck.
Best practices	In the data marketplace ecosystem, some companies (e.g. Live-EO) are mature regarding data-analytics. These companies prefer to take benefit of existing commercial cloud infrastructures which at the same time have a big community of developers. In the disaster management ecosystem, the challenge of emergency data collection is overcome by streaming data.
Related#	TI01 - D12 – ES02

	Economic Sustainability
Challenge	Economic sustainability refers to practices that support long-term stability and growth of the ecosystem. The ecosystem resilience requires harmonious and evolving business models.
Actors	Economic sustainability is mainly a downstream task embraced by private companies even if also largely impacted by other topics such as governance.

#	**Recommendation ES01: Integration of Open Access, Open Source, Open Innovation and Open Data Paradigms**
Challenge	Competition between public and private actors for data creation. Reluctance of companies to open their data. In addition, some companies consider that open data policies might lead to unfair competition.
Barriers	Divergent objectives. Resistance to change.
Best practices	In RUDI, they intend to overcome the reluctance of private companies by showcasing the global added value of adopting open data approach. For the Logistic and tracking ecosystem, this challenge is still to be overcome. The Disaster Management data ecosystem witnesses the importance of combining open data and data ecosystem building as open data is recognised to facilitate a data sharing culture.
Related#	ES04 – EG03 – EG07 – SE07

#	**Recommendation ES02: Emphasizing an adaptive and agile orchestration for the ecosystem evolution, especially for data collection**
Challenge	The alignment of the ecosystem with scientific, technical and business trends is mandatory. For instance, feeding the ecosystem with accurate and new data rapidly is a success factor for the evolution of an ecosystem.
Barriers	Cost of data collection. Difficulty to clearly identify data needs.
Best practices	For the logistics and tracking ecosystem, Spire can reconfigure existing sensors or to send new sensors and satellites in a fast and cost-effective manner. New satellites costs represent 1% of

	traditional satellites costs. In addition, with this new data collection practice, Spire can address Industry 4.0 challenges.
Related#	TI01 - TI13 – EG05 – EG03

#	**Recommendation ES03: The creation of a platform is a strong enabler of business opportunities and implementation, as well as related data flows.**
Challenge	Without an ecosystem, it is challenging and costly for companies to contractualise new businesses.
Barriers	Contractual issues. Financial exchanges issues (security, etc.)
Best practices	In the Smart agriculture ecosystem, the central platform provides automatic transaction facilities. The data marketplace ecosystem provides additional insights into the advantages of a platform through providing financial clearing operations. This service is suitable to attract big players that would not be interesting in addressing SMEs in the European market.
Related#	TI08 – SE02 – EG03 – EG04

#	**Recommendation ES04: Synergies between individual stakeholders' business models are the key condition for the overall ecosystem sustainability.**
Challenge	To reach the sustainability of the ecosystem, both local and global financial profitability must be addressed. Especially for the platform structuring the ecosystem if there is one. Thus, there is a need to balance individual competition and global cooperation. The organisation of interdependencies between the stakeholders of the ecosystem is an orchestration issue of paramount importance.
Barriers	Low degree of self-awareness of the ecosystem due to a lack of contacts between the ecosystem stakeholders. Duality between competition and cooperation enhanced by public / private cultures differences and mismatch in size. Lack of trust. Fragmentation.
Best practices	In RUDI, there is a particular commitment to help companies defining their business models, for example through "Call for projects". The aim is to lead to sustainable data creation through business models. On the other hand, the local authority is planning to monitor the

	impacts leading to public value creation in order to justify their investment. In UP42, the platform leader is actively fostering "in-house" businesses and value exchanges.
Related#	SE02 – TI01 - EG07

#	**Recommendation ES05: Data ecosystems rely on long term engagements.**
Challenge	Return on Investment is not always achieved in the short term
Barriers	Trust. Absence of long-term vision. Short-term business objectives.
Best practices	In the data marketplace ecosystem, the platform leader intends to create a data business model generating long-term and cumulative effects.
Related#	EG03 – EG04 – SE02 - TI02

#	**Recommendation ES06: Legal issues are framing the ecosystem through the definition of users' interaction rules.**
Challenge	Data ownership. Rights on data and solution assets. Licensing information, when available, is not based on a common framework.
Barriers	Concerns of various actors about legal risks. Licences on data, especially through data cycles.
Best practices	For the data marketplace ecosystem, the central actor has put a lot of control in the hands of data owners in order to gain the trust and confidence of suppliers and customers in the long term, especially through end user pricing and licensing. UP42 has put a strong focus on defining clear licences, not only on the data but also on algorithms.
Related#	EG01 – EG02 - TI08

#	**Recommendation ES07: Strong political and societal support facilitate the sustainability of the data ecosystem.**
Challenge	Data ecosystems need strong political support in order to be sustainable. Moreover, data driven service creation and decision making through the ecosystem create public value, therefore provide solid arguments in support of the establishment of the

	ecosystem.
Barriers	Political agenda. Data literacy.
Best practices	In the Disaster management ecosystem, political support is reported to be a prerequisite to break the data silos built around the various political and geographical responsibilities, ensuring a minimal data sharing between different agencies. Therefore, a public sponsor push for data sharing is important. On another side, in the studied local ecosystem, the digital strategy depends on the adopted regional/local development strategy.
Related#	EG02 – EG07 – SE01 – ES05

#	**Recommendation ES08: Extracting the value of personal data.**
Challenge	Personal data are difficult to gather but also valuable to create customized services. Technically, such data requires specific tools and processes based on personal consent to extract personal data stored by companies, while respecting the EU and national legislation.
Barriers	Legal framework, especially around personal data. Technical feasibility. Privacy. Citizen reluctance for share their data.
Best practices	In RUDI, one of the objectives is to develop a personal consent module co-designed with citizens. It aims to empower the citizen through the management of their digital rights. The same was done in the Smart Agriculture ecosystem, through the informed consent module.
Related#	EG01 - TI04 - SE03 - ES04

3. Conclusions

The findings presented here, validated by the experts representing the different investigated data ecosystems, demonstrated the importance of overcoming challenges related to the governance, stakeholder engagement, technology and economic sustainability in the emergence and the self-sustainability of geospatial data ecosystems.

The recommendations presented here provide useful insights for establishing concrete approaches for modernising the way in which SDI data is shared under the INSPIRE Directive. From our perspective, ecosystem thinking is a reliable approach to understand and to steer the development of a geospatial data ecosystem. Therefore, the role of SDIs should be reconsidered, in particular in a broader framework beyond the data provision paradigm (i.e. where data-sharing approaches and decisions are not steered exclusively by those who provide the data).

A combined approach including ecosystem thinking and orchestration is crucial to ensure the ecosystem's self-sustainability and break existing silos, but also ensure that new silos will not emerge.

The recommendations provided in this document are empirically based, and some other topics should be elaborated in further works, like e.g. the certification for actors in the data economy, the particular case of decentralised governance, implication and potential of cloud computing.

In particular, there are strong relation with the European Data Strategy published in February 2020 and the consecutively launched legislative initiatives such as the Data Governance act and the forthcoming Implementing Act for High-value datasets as defined under the Open Data Directive.

References

- Arai, A., Witayangkurn, A., Kanasugi, H., Fan, Z., Ohira, W., & Pedro, S. (2020) Building a data ecosystem for using telecom data to inform the COVID-19 response efforts.
- Boguslawski, R., Valayer, C., Van Gansen, K., Keogh, D. and Pignatelli, F., European Union Location Framework Blueprint (4.0), EUR 30374 EN, Publications Office of the European Union, Luxembourg, 2020, doi:10.2760/096595 (online), JRC117551.
- European Commission (2020). A European strategy for data; COM(2020) 66 final
- Gawer, A., & Cusumano, M. A. (2014). Platforms and innovation. The Oxford handbook of innovation management.
- Hamaker, R., Jimenez-Alonso, E., Rycerz, A., Baglee, A., & Stegmaier, P. (2017). Analysis of existing data infrastructure for climate services.
- Love, N. (2004). Taking data to new depths. National Staff Development Council JSD, 25(4), 22–26.
- Lucivero, F., & Sharon, T. (2019). Introduction to the Special Theme: The expansion of the health data ecosystem–Rethinking data ethics and governance.
- Martin, S., Turki, S., & Renault, S. (2017, August). Stimulation of Open Data Ecosystems: Learnings from Theory and Practice . Chapter of Open Innovation: Bridging Theory and Practice - Digital Innovation, pp. 41-78 (2019)
- McKinsey (2016) Car data: paving the way to value-creating mobility
- Oliveira, M. I. S., & Lóscio, B. F. (2018, May). What is a data ecosystem? In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age (pp. 1-9).
- Open Data Institute (2018). Mapping data ecosystems.
- Osterwalder A, Pigneur Y (2010). Business Model Generation – A Handbook for Visionaries, Game Changers and Challengers. John Wiley and Sons, Inc., Hoboken, New Jersey.
- Perks, H., Kowalkowski, C., Wittell, L., Gustafsson, A. (2017), Network Orchestration for Value Platform Development, Industrial Marketing Management, 67, 106-121.
- Ricci F. (2020) Recommender Systems in Tourism. In: Xiang Z., Fuchs M., Gretzel U., Höpken W. (eds) Handbook of e-Tourism. Springer, Cham
- Samuel Marcos-Pablos, Alicia García-Holgado, and Francisco J. García-Pẽalvo. 2019. Modelling the business structure of a digital health ecosystem, ACM International Conference Proceeding Series, pages 838–846. Association for Computing Machinery, October.
- Vaccaroli, R., Markus, F., Danhardt, S. et al. Grand Duchy of Luxembourg: a case study of a national master patient index in production since five years. BMC Med Inform Decis Mak 20, 163 (2020).
- Van Opijnen, M., Peruginelli, G., Kefali, E., & Palmirani, M. (2017). Online Publication of Court Decisions in Europe. *Legal Information Management*, *17*(3), 136-145.

Glossary

- **Platform**: Platforms constitute a shared set of technologies, components, services, architecture, and relationships that serve as a common foundation for diverse 4 sets of actors to converge and create value (Gawer & Cusumano, 2002; Gawer, 2014)[10]
- **Data Ecosystem**: a complex socio-technical system of people, organisations, technology, policies and data in a specific area or domain that interact with one another and their surrounding environment to achieve a specific purpose. Such ecosystems evolve and adapt through a cycle of data creation and sharing, data analytics, and value creation in the form of new products, services, or knowledge, which, when used, produce new data feeding back into the ecosystem.[11]
- **Data cooperative and trust**: The notion of a data cooperative refers to the voluntary collaborative pooling by individuals of their personal data for the benefit of the membership of the group or community. The motivation for individuals to get together and pool their data is driven by the need to share common insights across data that would be otherwise siloed or inaccessible. These insights provide the cooperative members as a whole with a better understanding of their current economic, health and social conditions as compared to the other members of the cooperative generally[12].
- **Orchestration**: a dynamic set of intentional activities to promote value creation in a platform ecosystem (Perks & al., 2017)[13].
- **Data literacy**: the ability to examine multiple measures and multiple levels of data, to consider the research, and to draw sound inferences (Love, 2004)[14].

[10] https://www.researchgate.net/profile/Satish_Nambisan/publication/332241987_Global_platforms_and_ecosystems_Implications_for_international_business_theories/links/5d3e3b60a6fdcc370a694ade/Global-platforms-and-ecosystems-Implications-for-international-business-theories.pdf

[11] The term is further describe along with a more extensive glossary of relevant terms in the European Union Location Framework Blueprint https://joinup.ec.europa.eu/collection/elise-european-location-interoperability-solutions-e-government/document/report-european-union-location-framework-blueprint .

[12] https://wip.mitpress.mit.edu/pub/pnxgvubq/release/1?readingCollection=0499afe0

[13] PERKS, H., KOWALKOWSKI, C., WITELL, L., GUSTAFSSON, A. (2017), Network Orchestration for Value Platform Development, Industrial Marketing Management, 67, 106-121.

[14] Love, N. (2004). Taking data to new depths. National Staff Development Council JSD, 25(4), 22–26.

Annexes

Annex I. Analysis Framework and Methodology

Rationale and approach

First and foremost, the data ecosystem analysis framework described here is aiming at answering the goal of the study *("to investigate how Spatial Data Infrastructures (SDIs) can evolve into data ecosystems to support the goals of digital government in Europe");* and extracting and harmonizing the ecosystems concepts, whether convergent, overlapping or contradictory, from the list of identified frameworks.

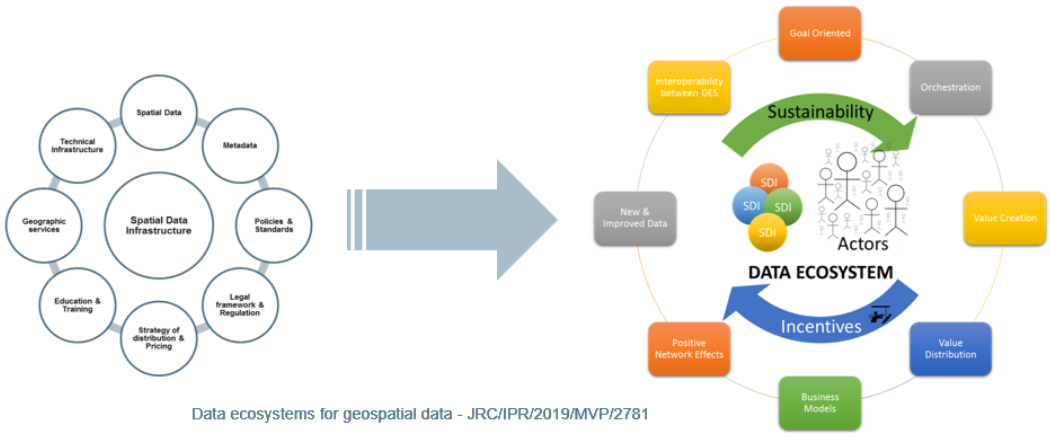

Figure 1 - Rationale of the study

The data ecosystem concept is fuzzy, and the existing initiatives are of a quite scattered nature, on top of the fact that, by definition, ecosystem is covering a large area of phenomena. Therefore, there is currently no standardized model making it possible to describe the different components of a data ecosystem. There is rather a variety of models fitted for different purposes, most often to describe a specific kind of ecosystem. As noted by (Oliveira and Al.[15]), there is also a lack of a standardized

[15] What is a data ecosystem? Marcelo Iury S. Oliveira and Al., 18: Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age - May 2018

description tools, such as the role played by the Business Process Model and Notation (BPNM)[16] in other domains.
The model used for this report is built following a literature review. As stated above, these models are often based on specific purposes and designed from the analysis of an empirical situation. Albeit the models are sometimes diverging, it was possible to align or at least to aggregate a set of core dimensions. Moreover, for the sake of this study, the model is built with the intention to put a specific focus on some dimensions, resulting mainly from a combination of other basic features:

- Reliable identification of **incentives** and **barriers**

- **Maturity**: in this study, maturity is linked with the ecosystem emergence mechanisms (Thomas[17]).

- **Orchestration**: From business ecosystems, and more accurately from platform literature, the framework introduces the concepts ecosystem leadership and of orchestration. Orchestrating an ecosystem means managing the network relations and find its foundation in innovation networks (Gawer & Cusumano, 2014[18]). It highlights the keystone actor(s) as well as actor(s) leadership. Several orchestration concepts exist: Ecosystem membership (size, diversity), Ecosystem structure (density, autonomy), Ecosystem position (centrality, status), Appropriability regime, Knowledge mobility, Ecosystem stability. As well as several kinds of orchestration like: Organizational orchestration, Technical orchestration, Standard / industry standard adoption, Internal / external interoperability.

- **Sustainability**: Sustainability is an important dimension of the ecosystems but is more the result from a combination of supportive factors than a dimension by itself. It is understood as (i) being balanced (more exactly without harmful imbalances); and (ii) having the ability to live without direct government support. An essential requirement of sustainability of the ecosystem is the financial sustainability of all the interactions led by all the ecosystem participants.

---

[16] See e.g. Samuel Marcos-Pablos, Alicia García-Holgado, and Francisco J. García-Pẽalvo. 2019. Modelling the business structure of a digital health ecosystem. In *ACM International Conference Proceeding Series*, pages 838–846. Association for Computing Machinery, October.
[17] Samuel Marcos-Pablos, Alicia García-Holgado, and Francisco J. García-Pẽalvo. 2019. Modelling the business structure of a digital health ecosystem. In *ACM International Conference Proceeding Series*, pages 838–846. Association for Computing Machinery
[18] Platforms and Innovation. Annabelle Gawer and Michael A. Cusumano, The Oxford Handbook of Innovation Management, 2014

According to the literature, an ecosystem description cannot be separated from the goal it is conveying: it may be useful not only to get a better understanding of the examined ecosystem, but it is also a tool to help the ecosystem evolving - the so-called "**ecosystem thinking**".

For these reasons, for each ecosystem, the analysis framework is divided in three parts:

1. High level analysis (Data Ecosystem Summary)
2. Data ecosystem dynamics
3. Flows of the data ecosystem

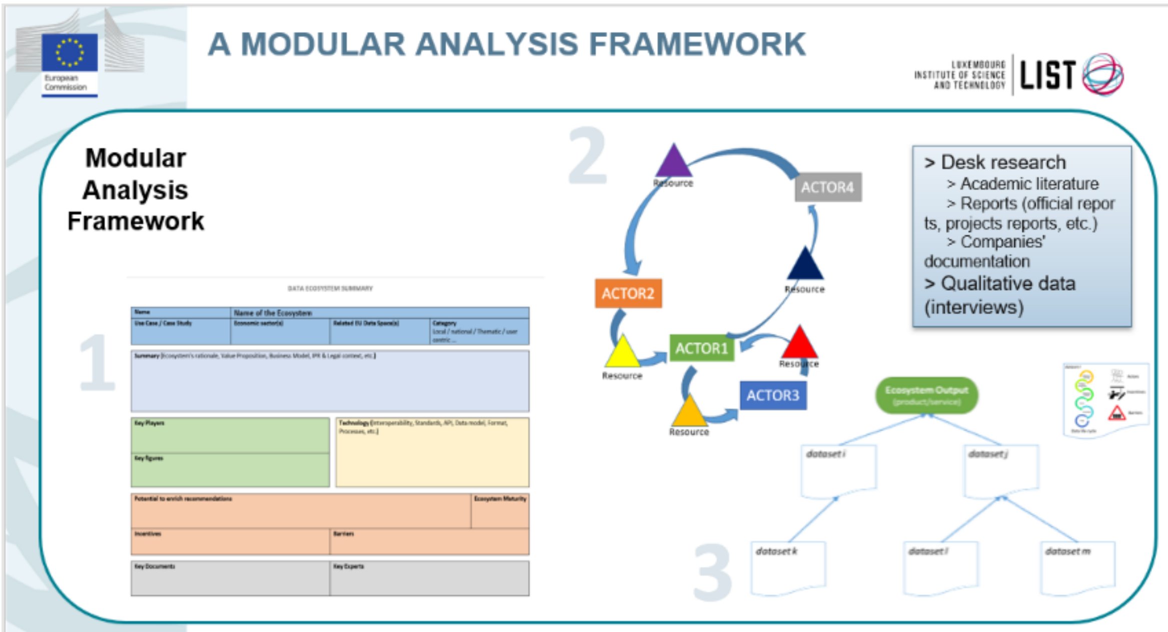

Figure 2 - A modular Analysis framework

Each view of the ecosystem is based first on narratives, highlighting the key points of the ecosystem as well as the phenomena affecting or caused by the ecosystem development, second on a graph both summarizing and contributing to understand the main features of the ecosystem.
For the high-level analysis the first view of the ecosystem is developed. Thanks to the insights extracted from the interaction with the relevant stakeholders of the selected ecosystems, the in-depth analyses will feed the second and the third views.

The selection process for the studied ecosystems has considered many criteria:

- A representative selection of domains/sectors
- A representative selection of related EU common data spaces
- The availability and accessibility of documentation
- The availability and accessibility of experts
- The spatial data importance

- The diversity of typology (local VS global, Thematic, User-generated data …)

The list of the selected ecosystems can be found in Annex 3

Data Ecosystem Summary

The first part of the framework, based partly on the insights from the two other views, is aiming at providing an abstracted or summarized view: **an overall representation of the components of the ecosystem**.
This framework focuses on Data Ecosystems key aspects like:
Goal / purposes,
Main Actors, theirs exchanges and communication,
Legal context and governance,
Technology specific aspects,
Costs and revenues / benefits,
Faced barriers and incentives.
The graphical component of the view is derived from the business model canvas[19], making it both familiar and easy to use for high level analysis and strategic decision makers.

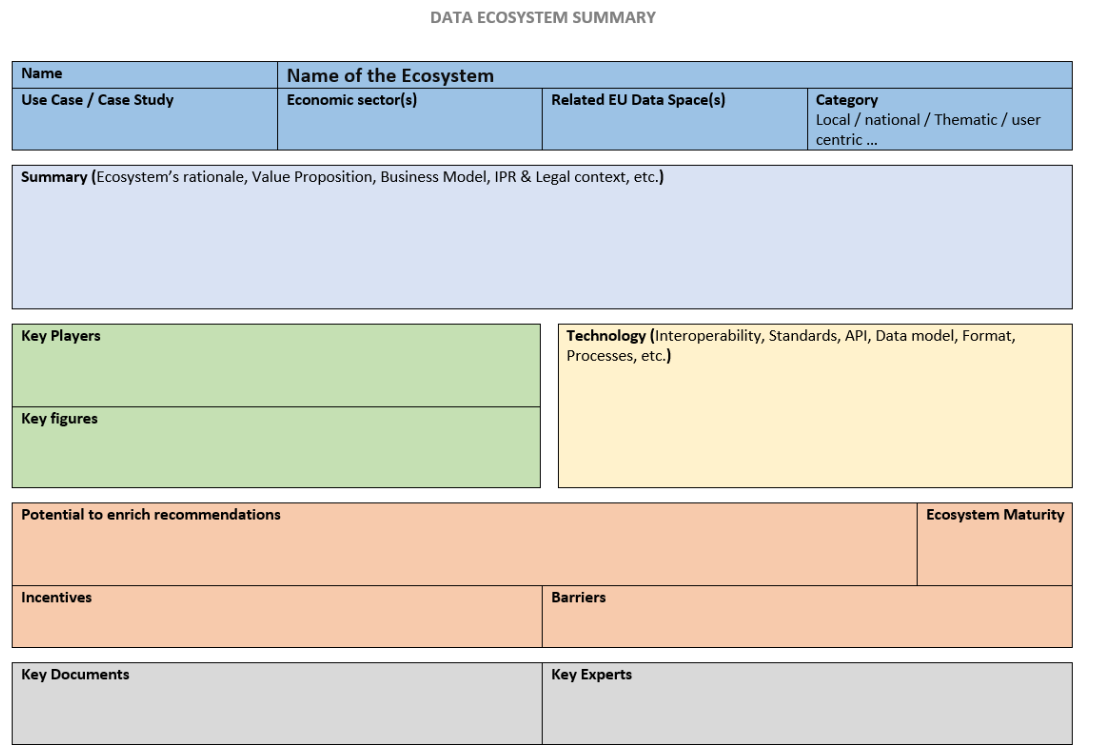

Figure 3 - Data Ecosystem Summary – Template

[19] Osterwalder A, Pigneur Y (2010). Business Model Generation – A Handbook for Visionaries, Game Changers and Challengers. John Wiley and Sons, Inc., Hoboken, New Jersey.

Ecosystem Value Dynamics

The second part, mentioned as the dynamic overview, is designed to highlight **the interactions between actors and resources**. It is aiming at understanding the structure - i.e. how the tasks are divided, what are the forms of collaboration, and value exchanges.

While a value chain traditionally relates to a single chain of activities and usually applies to one firm operating in a specific industry, a value network is a "business analysis perspective that describes social and technical resources within and between businesses. The nodes in a value network represent people (or roles). The nodes are connected by interactions that represent tangible and intangible value objects. These objects take the form of knowledge or other intangibles and/or financial value. Value networks exhibit interdependence. They account for the overall worth of products and services"[20].

Each stakeholder and each resource will be described with the following canvas:

Actor:

- Name
- Role
- Engagement (Goal and strength)
- Context (legal, competition, etc.)

Resources

- Offered resources
- Needed resources
- Incentives & barriers
- Interactions with others stakeholder's description

Graphical Representation:

The ecosystem value dynamics is represented as a network of actors creating, distributing, and consuming things of economic value. To do so, we use the "Business Model Drawing Tool[21]" constructs. The main concepts used in value modelling are:

- Actor: An actor is perceived by its environment as an independent economic (and often also legal) entity. An actor makes a profit or increases its utility. In a sound, sustainable, e-business model each actor should be capable of making a profit.

- Value object: Actors exchange value objects. A value object is a service, good, money, or experience, which is of economic value to at least one actor.

- Value exchange: A value exchange represents one or more potential, direct or indirect, trades of value objects.

[20] https://corporatefinanceinstitute.com/resources/careers/soft-skills/value-network/
[21] www.boardofinnovation.com

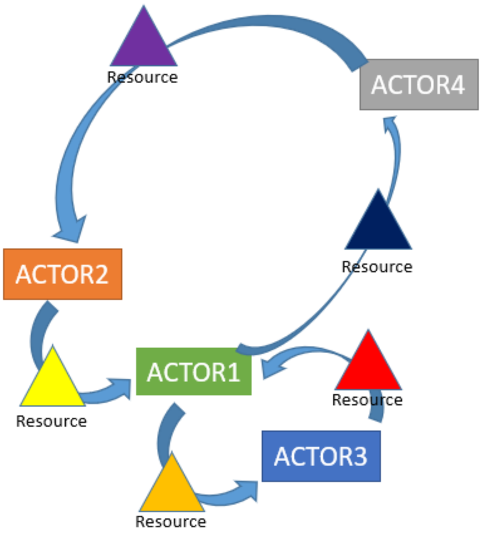

Figure 4: Representation of the actors' interactions

Ecosystem Data Flows

The third part, representing the data flows, is aiming at showcasing the technical perspective and represents the ecosystem from the perspective of the data. At least for the main outputs (services) and data of the ecosystem, the life cycles of the assets will be described. Among other objectives, the purpose of this view is to allow a clear highlighting of the number of **data cycles**. One output could be, for example, to point out that some ecosystems are not well running as data cycles are not really deep, which could be linked to other issues such as a lack of skills or unfruitful business models.

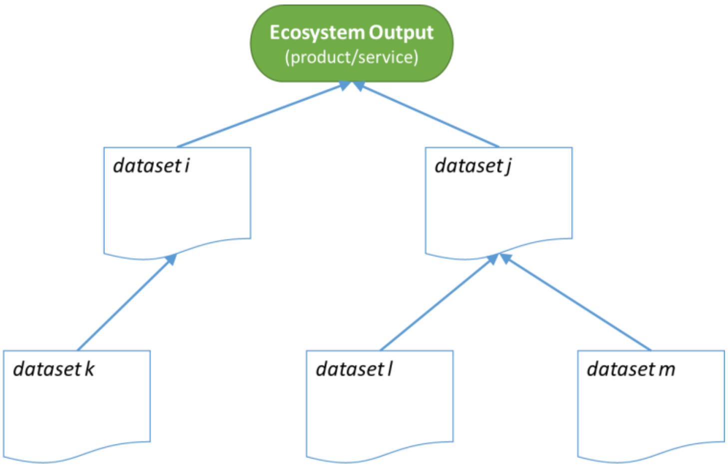

Figure 5 - Ecosystem's data cycle

1. Practically, it could be built from the output of the ecosystem, i.e. a product that is produced through the ecosystem.
2. From there, a formalism would link it to every (or a large body of the) datasets it is using.

3. This could be a (non-closed) circle representing the lifecycle of the data, from data collection to usage (list to be fixed)

    a. Each stage would include the main dimensions as well as external links (stakeholders, barriers, incentives…?) and maybe links to relevant part of the second view.
    b. To represent the data cycles, a formalism would link the relevant stage (usage, post-processing the most often we may assume) to another dataset (or sometimes the same data having known different processes).
        i. This would allow to follow the ecosystem from outputs to data, from processed data to rawer data.
        ii. Kind of insight for example: if the ecosystem has low deepness, it means there is maybe a gap to assess.

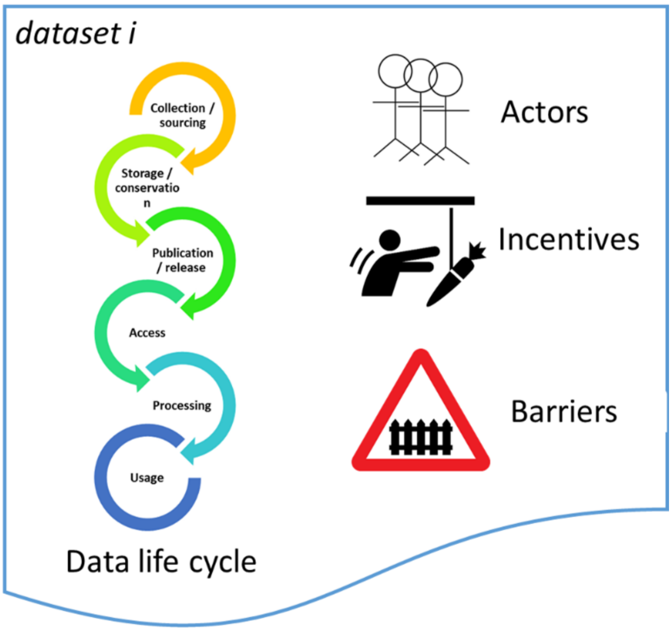

Figure 6 - Data set live cycle

Annex II. Proposed case studies and use cases

D2 - Proposed case studies and use cases

	Data Ecosystem Description				Selection criterias
	Ecosystem	Summary	Domain	Related EU Common Data Space(s)	Potential to enrich recommendations
1	**Rennes Urban Data Interface**	A pioneer in the field of open data since 2010, Rennes Métropole launched the first Metropolitan Public Data Service (SPMD), an initiative which aims to make the management and sharing of territory data a public service. The goa is to implement new types of cooperation between the city and producers and / or users of data in the territory: large companies, SMEs and start-ups, academic players, associations and users.	City	Public Administration Mobility Energy	Shift from open data provision strategy to ecosystem orchestration at the level of a city
2	**Tracking technologies for the supply chain ecosystem**	Driven by the strong development of e-commerce at least for the general public, allowed by the availability of cheap and reliable technologies of geolocation, geographic information, IoT and data analysis capabilities, the emergence of this ecosystem is allowing even small and medium business to track the goods and the vehicles at the different stages of the supply chain. On top of better information and increased customer relationship with intermediate and final users, this ecosystem allowing a fine monitoring of the fleet is a mandatory condition to enable an agile delivery system, leading in turn to cost-savings, improved security and new services to customers. As it is requesting efficient algorithms and large computation capabilities (adapting journeys is requesting exponential computation in a reasonable time...) that have to be done, it is also suitable to illustrate the interdependencies between ecosystems.	Supply-chain E-commerce Fleet tracking	Mobility Industry	Case of clear goal-oriented ecosystem with strong associated business models. Re-use of geospatial spatial data, so that the recommendations will be more easily applicable to geospatial ecosystems Total Access granted to the platform. Ecosystem oriented B2B and B2C at the same time.
3	**Data Marketplace (for geospatial imagery).**	Throught a central data marketplace, the question is to elucidate whether the artificial intelligences are a strong enough driver to turn the existing spatial infrastructures to (different) goals oriented ecosystems, benefiting in turn from the huge range of available data to provide new innovative services and to improve or create algorithms, especially in the pattern recognition area.	Machine learning, pattern recognition activities at large Variety of domains depending on the uses and the final customers of the companies using the platform	Industry Mobility Energy Agriculture	Role of marketplace's platforms to identify (or to link / combine) different kinds of actors and activities, for example those providing the data, those providing the services and ensuring the bridge between the different components of the ecosystem. So it is interesting to assess the place of platforms in the frame of ecosystems. Another interest is to understand how data availability and data analytics are supporting each other.
4	**Smart Agriculture**	Smart Farming is a farming management concept using modern technology to increase the quantity and quality of agricultural products. Farmers in the 21st century have access to GPS, soil scanning, data management, and Internet of Things technologies. By precisely measuring variations within a field and adapting the strategy accordingly, farmers can greatly increase the effectiveness of pesticides and fertilizers, and use them more selectively. Similarly, using Smart Farming techniques, farmers can better monitor the needs of individual animals and adjust their nutrition correspondingly, thereby preventing disease and enhancing herd health. Climate-Smart Agriculture (CSA) is an approach to help the people who manage agricultural systems respond effectively to climate change. The CSA approach pursues the triple objectives of sustainably increasing productivity and incomes, adapting to climate change and reducing greenhouse gas emissions where possible.	Agriculture Environment Water management Weather Industry and maintenance	Agriculture Green Deal	Interoperability of multiple data ecosystems Intensive use of Iot and Spatial data Data marketplaces Diversity of actors
5	**Disaster management**	Gathering and combination of opendata, private data and IoT generated Data etc... to solve disaster or emergency issues and help decision making regarding those topics' management.	security - mobility - environment	Public Administration	Large diversity of actors needed to be involved - paramount importance of orchestration and necessity to break thematic silos and geographical distances in order to insure an effective solution.
6	**Mobility**	Waze an app described as a community-driven GPS navigation app, which is free to download and use. Waze users ("Wazers", over 115 million people) can report accidents, traffic jams, speed and police traps, and, from the online map editor, can update roads, landmarks, house numbers, etc. Waze sends anonymous information, including users' speed and location, back to its database to improve the service as a whole. Waze launched the Connected Citizens Program (CCP) in June 2014, a free, two-way data sharing program used by over 450 governments, departments of transportation, and municipalities for traffic analysis, road planning, and emergency workforce dispatching. The app generates revenue from hyperlocal advertising to an estimated 130 million monthly users.	Mobility	Mobility	Self-sustainable business model, Service based on users generated data

	Data Ecosystem Description				Selection criterias
	Ecosystem	**Summary**	**Domain**	**Related EU Common Data Space(s)**	**Potential to enrich recommendations**
7	**Predictive maintenance**	Predictive maintenance of assets, here airplanes and vehicles, enables to the application of just-in-time principles in the field of maintenance. The benefits of these approaches are not only to reduce the costs associated to maintenance, but to improve the quality of these activities, with the goal to identify preventive interventions when required. The physical and computational infrastructures are available for long, even for road vehicles, but the emergence of this ecosystem as to face challenges related to the availability of data, that are produced and owned by a large range of actors and the necessity to comply with a strict regulation, especially for aviation.	Fleet management Fleet maintenance Predictive maintenance	Industry Mobility Energy	Links with other ecosystems (notably connected car at large): usage based insurance (UBI), road condition/traffic monitoring ; deep links with tracking ecosystem as well as A.I. ;more generally with IoT. Diversity of stakeholders at different stages (large range of interdependent actors) But competing stakeholders and solutions (maybe existence of various keystone actors, but not effective enough) Case of emerging ecosystem Interesting to evaluate how some actors such as car builders, in order not to be disrupted, find a place in the ecosystem or intend to do so. At least for the airplanes maintenance case, it is a case showing a case where regulation is having an overwhelming role in shaping the ecosystem.
8	**Pandemic Data**	Although intending to deal with data ecosystems on pandemic situation at large, this case is obviously linked to the Chinese virus crisis in 2020. Since this disruptive event, a wealth of public and private initiatives emerged to steer and facilitate the research, or to organize the response and the resilience of the populations in their daily life. The debates on the tracking applications showed the existence of different protocols and models from different kinds of stakeholders, which are competing based on different arguments, such as data quality versus privacy concerns. These heated debates shall not hide the other initiatives, such as the free release of scientific literature by scientific publishers, the improvement and the release of new public Open Data at the right scale and granularity, for example in Germany. This case should also consider some initiatives targeting a better structuration of the emerging data ecosystem, such as the EU Commission, who launched an European COVID-19 Data platform.	Health, social	Health	One interest of this case study lies precisely in its situation at the first stage of emergence. Compared to other cases established for a longer time, this one would make it possible to extract recommendations from the development of an ecosystem in its purest state, from the point of view of its definition and that of the biological metaphor, i.e. its 'bushy' character and the way it develops in a situation which is coined by emergency and according to strong and conflicting constraints. Moreover, it is at the interplay of several issues met during all ecosystems development: privacy, data sharing, data granularity, data quality, discrepancies among stakeholders from different disciplines… Not all the arrangements will be actually sustainable, but some components will remain and the overall methodologies are replicable for similar crises in the future.
9	**Accomodation Reviews (Tourism)**	The reviews for tourism accommodation is well establish and aims to increase the information flows about accommodations by guaranteeing complete and objective data concerning accommodations (offers overview and comments) to tourists, experiences feedback and (potential) customers profiles to tourism services providers, through a platform offering an accommodation reviews service	Tourism		Many actors in the place and a large variety of stakeholders in the ecosystem, with several different way of cooperating with each other
10	**Weather prediction**	From the data provision perspective, the ecosystem is established for long, first at the national scales, and even at the global with UN agency, with a strong culture of data sharing, at least among the official agencies. Depending on the countries, there are models of data re-use by third-party producing other layers of services on top on these data. These functioning models contribute to explain a large level of standardization in the field. Among the geospatial data, it is one of the most successful infrastructure, not only in terms of public-private partnerships, but also of sustainable business models for the private sector. It is also having interfaces with other ecosystems (agriculture), in particular with climate monitoring ecosystem, but also with user generated data (miniaturization, IoT…).	Weather prediction services Potentially a large range of sectors depending on the final focus adopted	Agriculture Green Deal	Good example of successful infrastructure turned to ecosystem Some of them are sustainable (at le Role of global governance (WMO agreements between countries) ensuring data re-use between meteorological agencies internationally. It is reinforced by the diversity ; different data sharing models and modalities Links with Open Source, Open Data and Open Access paradigms Generally speaking, this is an ecosystem at a pretty mature stage. New space with actors such as Spire , disrupting the field with a quasi monopoly of public stakeholders for data provision. So ecosystem mature, but tangling.
12	**Legal data ecosystems**	Legal data ecosystems are based on court decisions but also of data of economic and legal dimensions at the same time. Although being mainly of textual nature, they are rising interoperability and security issues, for example anonymisation. Intermediary actors have to enrich these data to make them re-usable by other ecosystems.	legal	transversal	public actors are providers and consummers. Private actors are enrichers. Possibility to disrupt well established business models
13	**Energy data ecosystem**	Energy represents a well standardized domain associated to strong data streams, although these are spread among different competing stakeholders. Moreover, it relies on personal data, implying privacy concerns.	Energy	Green Deal	High financial implications.
14	**e-Health**	Healthcare Digital record aims to facilitate the communication and collaboration between several practitioners and the patient in order to insure the healthcare processes	Healthcare	Health	Strong privacy issues Many actors around the topic

Selected for In-depth Analysis

Annex III. Data Ecosystems Summary

Name	**National electronic Health Record in Luxembourg**		
Use Case / Case Study **Use case**	**Economic sector(s)** Health	**Related EU Data Space(s)** **Health**	**Category** Local /Thematic

Summary ***(****Ecosystem's rationale, Value Proposition, Business Model, IPR & Legal context, etc.****)***

Electronic medical records represent a major step in promoting healthcare by providing a digital version of a patient's cross-system health data.

In Luxembourg, a shared medical record (dossier de soins partagés - DSP) is a secure, personal electronic healthcare file that is under the patient's direct control. It is provided free of charge by the eSanté Agency (Agence eSanté).

The main purpose of the DSP is to promote exchanges between healthcare professionals and to enable better coordinated patient care by centralising all of the patient's essential health-related information.

Access to a shared medical file is restricted to the patient who is its holder, and to accredited healthcare professionals attending to the patient and to whom the patient will have given the right of access. The shared medical file centralises all the patient's essential medical information that the healthcare professional would require for the coordination and continuity of care, with the patient's consent (e.g., x-ray exam results, consultation reports, lab test results, etc.).

The data ecosystem set up to implement the DSP gathers Hospitals, General Practitioners, National Registry, EU Member States, Researchers, Joint Social Security.

Key Players

Three main actors are involved in the data flow regarding the identity information that is collected in the national Master Patient Index (MPI): the National Register of Natural Persons (RNPP), the Joint Social Security (Centre Commun de la Sécurité Sociale (CCSS)) and the eSanté agency.

The different healthcare stakeholders that are involved in the ecosystem include primary and secondary care actors (general practitioners, hospitals, pharmacies, laboratories, […]), long term actors (care homes and nursing services) as well as others actors and institutions (European health-related institutions) and research

Technology ***(****Interoperability, Standards, API, Data model, Format, Processes, etc.****)***

Fundamental for a reliable functioning of the DSP are the Healthcare Providers Directory (HPD) and the Master Patient Index (MPI), both of which represent two key databases integrated within the platform's architecture.

The HPD contains information about all health professionals in Luxembourg as well as health institutions and structures. This directory ensures the identity of health professionals who wish to use the services of the platform to exchange medical information or to consult the DSP of a patient.

The MPI directory represents the national patient identity database that enables health sector stakeholders to have a single shared

organisations.	view of a patient identity, regardless of the sources of identity data. Another leading service of the platform is the pseudonymization service, which is mainly aimed at universities and research institutions. A national strategy was defined to promote interoperability between the various health ISs and to delineate all the standards to be applied in the connections between the various health structures.
Key figures At the latest update (31 December 2018), 2.418.336 identity profiles have been counted, including almost the totality of Luxembourgish residents (95.2%) as well as all the cross-border workers that are affiliated to the Luxembourgish social security system.	

Potential to enrich recommendations Healthcare professionals collect various types of data during patient care that could help clinical and academic research institutions. In this context, emerging legislations are pushing the need to protect the privacy of the patient regarding their medical data. For example, the European General Data Protection Regulation (GDPR) mirrors, through the article 25 “Data protection by design and by default” and the article 32 “Security of processing” [54], the European will to preserve confidentiality while handling the health data of patients. Different strategies allowing for an ethic use of the patient’s confidential medical data are currently materializing, among which is the approach of pseudonymization. The pseudonymization service aims to facilitate the legislative and administrative procedures of clinical and academic research.	**Ecosystem Maturity** Since 2014, when the MPI was integrated to the eHealth platform, the ecosystem is operational.
Incentives A correct patient identification represents a key prerequisite for a reliable functioning of a national eHealth platform. With such an ecosystem, health and healthcare practitioners can easily share appropriate documents related to the patient and therefore faster and better provider the appropriate treatment, with a lower risk of mistakes. Such a platform also provides practitioners services to improve their internal processes. Through this platform, patients are sure that the practitioners have all the information they need to give the right treatment, without doubling medical examinations, and a better communication with the health team in charge of him, this system enhance trust with Health	**Barriers** The Grand Duchy of Luxembourg with its 602,000 inhabitants represents one of Europe’s smallest countries. Nevertheless, the heterogeneity of citizenship is one of the key characteristics of Luxembourg. One of Grand Duchy’s particularities is the linguistic system, which is characterized by the simultaneous use of three official languages: Luxembourgish, the national language, as well as French and German. In addition to that, other languages like Portuguese and Italian, among others, are currently being used by a great part of the population. Statistics from 2017 show that 60.2% of the residents speak two or more languages. This variety had to be taken into consideration for the establishment of a

Sector.	successful identity management system. Another important characteristic of Luxembourg concerns the cross-border workers who, in addition to the residents, are also covered by the national social security system, bringing the total of affiliated people to 841,000. As such, all these people can benefit from the eHealth services and applications that are offered at the national level. Not all health and healthcare institutions are connected to the national e-santé platform, so reconciliation treatment are needed.

Key Documents	**Key Experts**
Vaccaroli, R., Markus, F., Danhardt, S. et al. Grand Duchy of Luxembourg: a case study of a national master patient index in production since five years. BMC Med Inform Decis Mak 20, 163 (2020). https://doi.org/10.1186/s12911-020-01178-y	

Name	**Energy data ecosystem**		
Use Case / Case Study **Use case**	**Economic sector(s)** Energy (heat and electricity)	**Related EU Data Space(s)** Energy mainly, strong links with green deals	**Category** Thematic

Summary ***(****Ecosystem's rationale, Value Proposition, Business Model, IPR & Legal context, etc.****)***

The energy data ecosystem, considering here both electricity and heat systems, is traditionally an important producer and consumer of data. There are many dependencies between this data ecosystem and other ones such as smart cities, or smart manufacturing… highlighting a large range of possible value creation and business models.

In parallel, there is a trend to push on the digitalization of this sector, with questions such as the nature, the conditions and the deepness of data sharing between (mainly but on only) industrial actors from different domains, the possibility to make predictions based on data through data science or more generally machine learning algorithms with the associated data needs, the development of new services based on the insights extracted from largely available data, and more general questions around data-driven decision making.

How these questions are dealt with is impacting, or requiring, the development of the energy data ecosystem.

Key Players

Traditional data ecosystem stakeholders are energy producers, as well as the producers of resources and the actors of the supply chain; regulators, policymakers, and consumers, corporations as well as households. Newcomers are the companies producing sensors and providing analytics services.

More or more, the different kinds of energy consumers, individuals or organisations, are an essential part of the ecosystem as their consumption patterns, which may be analysed once the data are aggregated, are an essential component of the ecosystem are they are

Technology ***(****Interoperability, Standards, API, Data model, Format, Processes, etc.****)***

Until now, this field is facing a lack of standardisation, or standards adoption, concerning the data models[24] as well as the metadata[25]. Nevertheless, the International Electrotechnical Commission is promoting a rich list of standards covering a large part of the data ecosystem and presented component by component on its website[26]. Concerning the personal data related to the energy domain, a report was published by the Smart Grid Task Force[27] of the European Union.

---

[24] https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8792061
[25] https://www.mdpi.com/1996-1073/13/2/444
[26] http://smartgridstandardsmap.com/
[27] https://ec.europa.eu/energy/sites/ener/files/documents/report_final_eg1_my_energy_data_15_november_2016.pdf

enabling the possibility of engaging data analytic approaches based on their data. At the same time, these data are considered as personal data, and at least as data arising privacy issues. In the US, the OpenADR Alliance intends to structure the relationships and the interactions between the energy data ecosystem's stakeholders[22]. A sharing platform is already existing in Denmark (DataHub[23]) ruled by the organisation in charge of the Danish energy infrastructure.	Moreover, existence and affordances of data sharing platforms are still lacking, despite the success of initiatives such as the platform mentioned besides. Literature reports that this ecosystem is one of the fastest growing for the data analytics approaches[28].
Key figures For the energy sector at large, in the European Union, it represents 1.6 million jobs and 4% of the non-financial business economy[29].	

Potential to enrich recommendations	**Ecosystem Maturity**
This use case represents a source of recommendations about the issues around data sharing in an industrial context. It allows to highlight the relationships between data ecosystems development and digitalization of companies, and how these concepts should be arranged together to lead to successful developments. It is also an example of ecosystem where the intra-domain specificities, there electricity and heat, are leading to different constraints and therefore are influencing the respective data ecosystem's shapes under the umbrella of the generic energy data ecosystem: the engagement of households, local authorities and business is far more important for the case of heat. The Danish platform mentioned above is an interesting example of a case where the orchestration is at least partly exercised by a company (pubic in this case) to aggregate the data collected by the diverse companies in charge of the energy production and distribution. At the same time, it is showing that an efficient orchestration may also require the enforcement of a regulation.	The roles of actors need to me more accurately defined. The platforms making tangible the interactions shaping the ecosystem are just emerging, with a high diversity among different geographic areas.

Incentives	**Barriers**
A sharing platform does not have only to ensure	There is still a lack of infrastructures

[22] https://www.openadr.org/about-us
[23] https://en.energinet.dk/Electricity/DataHub#Documents
[28] https://www.sciencedirect.com/science/article/pii/S2666546820300094
[29] https://ec.europa.eu/jrc/en/research-topic/energy-sector-economic-analysis

the communication of the data flows between industrial actors. And, on top of all organisational and legal arrangements, it has an essential orchestration role in matter of adopting and coping with standards in terms of data models and data exchanges protocols.	supporting the data sharing require for the digitalization of the sector[30], both from the organisational and technical perspectives, to ensure a convenient data sharing between (industrial) actors. Nevertheless, there is a wealth of existing projects in this domain. Unlike other domains where there is a spontaneous commitment of the citizens to contribute, such as shown for the weather data ecosystem use case focusing on crowdsourcing, where the public good may be a driver, the commitment to share these kinds of data is more questionable. Indeed, it impacts more directly people and how they are consuming. Although the environmental issues are important, it may be considered as a fraud seeking system, and thus be considered as more intrusive. A strong engagement on transparency on what is done is done with their data in the mentioned projects and platforms.

Key Documents	**Key Experts**
The "My Energy Data" report, albeit published in 2016, provides a state of play of the major issues in the European Union, as well as a useful description of the arrangements in that domain for ten countries of the EU.	

[30] https://ec.europa.eu/commission/publications/4th-state-energy-union_en#:~:text=The%20energy%20union%20aims%20to,renewable%20energy%20and%20climate%20change.

Name	Legal data ecosystem		
Use Case / Case Study **Use case**	**Economic sector(s)** Justice	**Related EU Data Space(s)** No specific data space, but partly transversal for some issues.	**Category** Thematic

Summary ***(****Ecosystem's rationale, Value Proposition, Business Model, IPR & Legal context, etc.****)***
There is a strong commitment towards the availability of better and new services based on legal information, as well as to the increased usage of data by legal ecosystem's stakeholders. As legal issues have a pervasive dimension in most of contemporaneous societies, the ecosystem should be divided between the legal data pertaining to the functioning of the judiciary system itself, and the larger scope of data with a legal dimension but which are at the same time of socio-economic nature, such as the information about companies.
There is a clear value proposition for both cases, the first one and the second one concerning the economy and the trust.
The first side clearly faces huge challenges in terms of ethical concerns around the LegalTech, especially with the discussions occurring on predictive justice, while access to the court decisions is still limited and not in line with the current technical affordances.

Key Players
There is no traditional data ecosystem around the judicial side, in the proper understanding of what is an ecosystem. However, there is a better developed ecosystem based on the data which have a direct economic value, such as the data on corporations.
Researchers, LegalTech experts and the development of artificial intelligence draw the domain to the emergence of a functioning and sustainable ecosystem.

Key figures
In the sole United Kingdom, in 2019, the legal market represents £35bn, and the investments in LegalTech are quickly growing, representing £260m that year[33].

Technology ***(****Interoperability, Standards, API, Data model, Format, Processes, etc.****)***
Albeit a low degree of maturity, a wealth of standards is already existing or being built. For the metadata, suggested standard is for example ELI[31].
Even if the basic material, at least for the judicial side of the ecosystem, is text, there are efforts to encourage practitioners and publishers to go beyond the PDF versions and to provide machine-readable formats, based on XML principles, for example LegalDocM[32].
Traditionally, legal departments are working mainly with documents, less with structured data. The availability of machine learning algorithms, the development of natural language processing approaches, represents

---

31 https://eur-lex.europa.eu/eli-register/about.html
32 https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legaldocml
33 https://blogs.thomsonreuters.com/legal-uk/2019/10/18/a-new-report-legaltech-startup-report-2019-a-maturing-market/

		an opportunity to benefit of the datafication of the ecosystem, even with a low degree of applied standardisation and without requiring skills or putting burdens on practitioners. Concerning the linkage between these data and authoritative location data is still more a research topic than an applied technology.

Potential to enrich recommendations	**Ecosystem Maturity**
This use case is primarily interesting for the diversity of the barriers faced by the development of a sustainable ecosystem, which does not prevent the existence of success stories and a progressive structuration of the ecosystem. Indeed, there are nevertheless data flows through APIs and data cycles organised thanks to the technical development and the expertise of LegalTech actors[34]. One unique feature is the case for datafication. Starting from a low tradition of data usage, this ecosystem is showing sharply how a data ecosystem may emerge combining different domain and technologies, which could be useful for other cases where the data traditions are a priori better established. This ecosystem represents a case where the orchestration activities should focus on the scope and the quality of data, i.e. focusing on the data sharing or Open Data. (Opijnen, Peruginelli, Kefali, and Palmirani, 2017)[35] analysed the European situation based on a questionnaire filled by the 28 EU member states. The recommendations are mentioning the importance of making publishing criteria publicly available, to make high courts decisions open by default, and a set of those of lower courts should be published. Licenses should not hinder the data re-use, favouring for example CC-O. Decisions should be published in machine-readable formats. It means that, put aside exceptions, judicial data should follow the general regime already defined for public Open Data, which should ensure an improvement in terms of availability, quality and overall usability. This is scalable to other kinds of data ecosystems relying at least in part on public data and confirms the importance of Open Data strategies as first building block of such ecosystems. This ecosystem showcases the importance of two important concepts: data provenance, and data re-use traceability. Concerning data re-use traceability, it is an important dimension of the judiciary side of the data ecosystem as some legal systems are requiring to track who accessed which piece when a case is processed, and digital documents and data make this more complicated to ensure. Data provenance is of paramount importance to	The data ecosystem on the judiciary side is still in its infancy for most of the countries, where the conditions of data release and the potential data structures are still being discussed or prototyped. The second one is more mature or at least more sustainable, but just begins to cope with the challenges of digitalization, especially the potential disruption of well-established business models. The set of established

[34] E.G. https://www.lexology.com/library/detail.aspx?g=694bcd8f-e479-4c4e-9f16-1002a8d80a39
[35] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3088448

ensure not only the trust, but the mere validity of the analysis. Importance of getting data at the right semantic granularity level. This use case is also showing the issues faced to develop actual data flows and data cycles when bulk download is a rare possibility and the access to data is highly fragmented, making it difficult for LegalTech or IT specialists at large to develop machine learning and data-driven services. This makes the case for of a central platform mitigating the fragmentation issues[36]. This ecosystem is still lacking a common forum to gather the views of domain experts, both from the legal questions, IT … making more tangible the ecosystem as a conscious set of stakeholders. It is also a case where ethical concerns are at their paramount, making regulation actors an important player and the policy-maker have an important orchestration role: depending on the equilibrium chosen between the expected profits and the ethical risks, it may lead to a fast development of the ecosystem or to concerns limiting its development.	standards and availability of APIs is pretty limited, leading to less efficient data flows.
Incentives There is a niche market for legal data and support from Public Authorities for Justice digital transformation.	**Barriers** Data availability remains the first and strongest barrier against the development of this ecosystem. For judicial data, especially court decisions, there is a strong fragmentation issue. In the European Union, it does not only concern the level of member States, explainable through different traditions regarding the organisation of public documents and their access. There is also a strong fragmentation inside the national boundaries, due to geographic and thematic distinctions which are hindering the potential access to data as there is no In most cases, data and metadata about court cases do not follow standards. However, this could represent a business model for companies to improve the data usability by processing the raw documents and data and improving their consistency and quality. (Marković & Gostojić, 2018) based on a comparison of European and American countries, find that the quantity and the

[36] https://journals.sagepub.com/doi/abs/10.1177/0894439318770744

	quality of data released by judiciary systems is among the lowest of the data made available by the public sector.

Key Documents	**Key Experts**
(Opijnen et al., 2017) *Online Publication of Court Decisions in Europe*	

Name	Mobility data ecosystem		
Use Case / Case Study **Use case**	**Economic sector(s)** Mobility, transports	**Related EU Data Space(s)** Mobility	**Category** Thematic

Summary ***(****Ecosystem's rationale, Value Proposition, Business Model, IPR & Legal context, etc.****)***

The mobility data ecosystem is one of the best and longstanding running ecosystems among the selected use cases.

It has a strong commitment towards better informing the final users, functioning in the most optimized way, and associated policies, such as green mobility. There is also a rich variety of business models organised around the data collection and processing, which is the matter of a large

Key Players

The list of actual and potential ecosystem participants is highly variable depending on the scale and the place considered, but in any case, it is extremely rich.

Core stakeholders are final users (individual or organisations) as users but also as data producers and providers; companies ensuring mobility service; companies which business model is on the information layer; regulation actors; policy makers (both defining the legislative context and, in large part, the infrastructures).

A larger range of stakeholders needs to be considered: industries manufacturing the transportations vehicles; those producing the sensors; insurance companies; researchers or companies analysing the data flows.

At a larger extent, employers, relatives and social communities shaping the needs and the context of mobility, with data covering not only the act of mobility, but all the societal and demographic data shaping and impacted by the mobility and its needs.

Technology ***(****Interoperability, Standards, API, Data model, Format, Processes, etc.****)***

This ecosystem relies first on a strong ground of technologies, such as RFID technology for vehicle ID and tracking

Pertaining to the standards and the provision of APIs, mobility data ecosystem was a pioneering sector and is a model on how an industry standard may be applied, with for example GTFS produced by an actor and now largely spread.

Mobility ecosystem depends on the development of a large range of technologies and their broad implementation in the actual population of vehicles[37]. The uncertainties on the technologies, models and related standards that will be used or enacted is nevertheless a barrier.

The diversity of standards proposed for the disruptive technologies and the associated data may be analysed through socio-technical approaches. For example, cities may

37 https://fpf.org/wp-content/uploads/2017/06/2017_0627-FPF-Connected-Car-Infographic-Version-1.0.pdf

Key figures		
In the European Union, the overall transport sector is representing 5% of the GDP and 10 million jobs[39].		coordinate their efforts to encourage the use of data models and legal frameworks compliant with data sharing or Open Data principles[38], taking at the same time into account the new questions arising from the evolution of mobility issues, such as bikes or scooters sharing.

Potential to enrich recommendations	Ecosystem Maturity
Mobility ecosystem is a case for which orchestration is familiar to and its usefulness is recognised for long by the stakeholders, as the diversity of actors needs to be arranged, even beyond the case of data ecosystem. Moreover, the importance of data sharing is also recognised for long and enacted by regulators, in spite of highly variable data availability and quality depending on local contexts. Secure and trusted data exchanges are a key component of an intermediary platform, as shows the case of MDM in Germany[40]. It is of paramount importance for the personal data collected It is showing how the ecosystem is adapting to the requirements to comply with legal frameworks such as the GDPR and concerns such as the management of personal data by the data producers themselves[41]. In matter of standardisation, there is a complex relationship between authorities or governments and industrials. The first ones need to adapt the infrastructure to a set or to a range of standards in order to ensure the long-term development and sustainability of the investments given to infrastructures. At the same time, they may enact regulation to foster the development and implementation of standards. The latter master the industrial rationale, and they face contradictory constraints and barriers to join the data sharing initiatives. It is also an interesting lesson learnt to reduce the concerns on the diversity of models and standards' candidates appearing for less mature ecosystem as this example is showcasing the evolution of an ecosystem, relying at the same time on a basis of adopted and proven standards, and seeking a future new equilibrium through the exploration of new models, sometimes exploring contradictories ways, that are progressively adopted by the whole ecosystem.	The ecosystem is mature from the perspective of the self-awareness of the ecosystem, for the tradition of data sharing, for the standardisation of the data models and the data flows. It does not prevent several concerns implied by disruptive technologies and is depending on

[39] https://ec.europa.eu/jrc/en/research-topic/transport-sector-economic-analysis
[38] https://github.com/openmobilityfoundation/mobility-data-specification
[40] https://www.mdm-portal.de/mdm-and-the-new-mcloud-two-platforms-one-strategy/?lang=en
[41] https://mydata.org/declaration/

This ecosystem is showing how the data availability may depend on a compromise between the objectives and sometimes contradictory interests of the different stakeholders: if some actors tend to keep the data they may capture for their own use, authorities may organise an open access to these data following the Open Data principles[42].		
Incentives - Strong political and societal support due to quality of life and environmental purpose. - Enthusiasm of citizen for soft mobility solutions ensure the collection of new data.	**Barriers** Any data collection implying the collection of data that could directly or not break anonymity of people or be harmful for privacy. In spite of the ecosystem maturity, at the local scale, on assets managed by cities, not all the municipalities have digitized data, nor they published their data in standardized formats, even if INSPIRE did a lot. At the same time, smart traffic management projects and smart cities paradigm at large are often an incentive to progress in this domain. It represents nevertheless a barrier to the participation in the implementation of disruptive technologies. Despite the important structuration of the ecosystem, there are still some conflicts on the extent to which the data sharing should be realized. For example, several public undertakings stakeholders have a different understanding of the modalities of data sharing, leading to a constant research of the best equilibrium of the ecosystem[43].	

Key Documents	**Key Experts**
Twelve example summaries of *Good practices and pledges on B2G data sharing*[44]	

[42] https://www.openmobilityfoundation.org/
[43] https://ec.europa.eu/digital-single-market/en/news/good-practices-b2g-data-sharing-smart-mobility-info-and-ticketing-system-leading-way-effective
[44] https://ec.europa.eu/digital-single-market/en/good-practices-b2g-data-sharing

Name	**Pandemic data ecosystem**		
Use Case / Case Study **Use case**	**Economic sector(s)** Health	**Related EU Data Space(s)** Health	**Category** Thematic

Summary ***(****Ecosystem's rationale, Value Proposition, Business Model, IPR & Legal context, etc.****)***

The abrupt crisis triggered by the worldwide health crisis in 2020 is an interesting lens to consider how the data ecosystems around diseases with world challenges are currently addressed. Moreover, it is a relevant case to learn how an ecosystem may emerge to face suddenly a new challenge and re-using pre-existing components and enabling new ones, without preventing conflicting perspectives shaping its development.

Key Players

Besides government, scientists, health specialists and practitioners, technology corporations and the general public are important stakeholders of this ecosystem. Telecommunication actors may also be considered as stakeholders.

Key figures

Technology ***(****Interoperability, Standards, API, Data model, Format, Processes, etc.****)***

The overall health ecosystem is already coping with the new technical capabilities, wearables, smartphones, or machine learning. Medical data require clear consent procedures and strong control on the entire life cycle of medical data. They are associated to enforcement of privacy protection and pseudonymisation technologies.

There are already platforms such as OpenHumans.org intending to build the intermediary capabilities of data cooperative for the health domain. For this specific platform, the use of standards, especially of identifiers, is encouraged, but given the current diversity of medical data needs, no specific standard is developed.

In terms of processes, until now, and besides the medical data properly said, the health crisis led to research on data collection related to disease spreading, contact-tracing and quarantine enforcement, with large differences concerning the equilibrium between countries.

The understanding of how the disease may spread by space and time, hence the suggestions to gather data on mobility patterns of the population, which may rely on traditional socio-demographic studies. Initiatives suggest going further, especially for countries without a dense medical infrastructure, and consider the fine-grained patterns of population mobility may be collected by telecommunication companies[45]. This is highly debated, not only for privacy questions, but also for different understandings of the

45 https://www.researchgate.net/profile/Ayumi_Arai/publication/343415566_Building_a_data_ecosystem_for_using_telecom_data_to_inform_the_COVID-19_response_efforts/links/5f28ede0299bf134049ed39a/Building-a-data-ecosystem-for-using-telecom-data-to-inform-the-COVID-19-response-efforts.pdf

		stochastic nature of a pandemic and how this may be addressed by researchers. The 2020 health crisis accelerated the research already existing on the re-use of social media data to contribute to impacts modelling and decision making and gave birth to a wealth of projects. Contact tracing applications rely on the large usage of mobile phone worldwide. The technological and technical grounds are not a challenge in most of the countries, and the data cycles may not arise, as these data imply very sensitive privacy issues. Several countries intend to use these technologies to enforce social distancing and quarantine, based on the same components. There also, the question implies less technology than the respect of privacy.

Potential to enrich recommendations	**Ecosystem Maturity**
The insights and lessons learnt from this ecosystem may be extended to any other urging issue in any other domain and may be compared with those of the ecosystems around disaster management. It may be also considered to analyse how the more traditional health ecosystem reacted and re-organised itself to cope with the new challenges. It confirms at least an existing trend, i.e. the redefinition of almost any data as health-related data. The general public confirms also its role of data provider and, depending on the legal framework, of controller of the data re-use t the personal scale. It may also contribute to the collection of second range of health-related data, even if it is with little success for now. From the orchestration perspective, health data ecosystem is mostly a decentralized ecosystem and experts advocate to maintain this. Nevertheless, it may be considered as a weakness in case of emergency, especially when actors which are not directly health experts intend to contribute, either to the medical issues themselves, or to deal with the more general consequences.	This ecosystem may already rely on a strong albeit limited health data ecosystem. The relevant question is how an ecosystem emerging for emergency questions may lead to a long-terms and sustainable data ecosystem and how the lessons learnt may benefit to the overall health ecosystem on the long term.

Incentives	**Barriers**
	Goodwill of citizens to participate in the collection of a second range of health-related data is questionable. The spontaneous adoption of the contact tracing applications developed under the auspice of several governments, highly variable depending on the countries, is pretty low in general, which can be related to concerns about who will re-use these data for which purposes.

Key Documents	**Key Experts**
(Sharon & Lucivero, 2019), *Introduction to the Special Theme: The expansion of the health data ecosystem*[46] Setting-up a data ecosystem to defeat covid-19[47]	

[46] https://journals.sagepub.com/page/bds/collections/health_data_ecosystem
[47] https://council.science/current/blog/setting-up-a-data-ecosystem-to-defeat-covid-19/

Name	Predictive maintenance of vehicles' fleets		
Use Case / Case Study **Use case**	**Economic sector(s)** (partly) mobility	**Related EU Data Space(s)** Mobility; Industrial	**Category** Thematic

Summary ***(****Ecosystem's rationale, Value Proposition, Business Model, IPR & Legal context, etc.****)***
Predictive maintenance of vehicles enables the application of just-in-time principles in the field of maintenance. The benefits of these approaches are not only to reduce the costs associated to maintenance, but to improve the quality of these activities, with the goal to identify preventive interventions when required. The physical and computational infrastructures are available for long, with more details for airplanes but even for road vehicles, but the emergence of this ecosystem has to face challenges related to the availability of data, that are produced and owned by a large range of actors and the necessity to comply with a strict regulation, especially for aviation.
This ecosystem is also highly dependent on a more general ecosystem around the connected car and autonomous vehicles: by default, these trends require the development of sensors, a large part of them being useful for the data analytics needed to realize the benefits of the predictive maintenance ecosystem.

Key Players
Vehicle manufacturers remain central actors and intend to master the different stages of the value chain, gaining skills and experience in downstream domains, including analytics, and providing predictive maintenance services[48].
The standardisation efforts need to structure the ecosystem[49] are gathering a large range of participants, just from the data provision perspective. Including both traditional car manufacturing actors (manufacturers themselves, traditional embedded systems) and relatively new actors (new sensors, communication layers, analysis layers and more generally all IT, as well as insurance companies. Some actors already present in the vehicle manufacturing sector become more central, such as the European Telecommunications

Technology ***(****Interoperability, Standards, API, Data model, Format, Processes, etc.****)***
This ecosystem relies first on a preliminary set of already established standards, such as OBD.
Given the current stage of maturity of this ecosystem, different standardisation paths are followed. The car connectivity consortium[50] promotes the Car data exchange. It aims at standardizing the data types, data sets and encryption methods. However, it is still not widely implemented. The W3C is organising an Automotive Working Group, working group on a vehicle data standard[51] but it did not lead to the publication of an actual standard until now. Standardisation issues are still being explored by research and cooperation

48 E.G. https://www.en.renault-trucks.com/predict
49 E.G. https://sensoris.org/members/
50 https://carconnectivity.org/
51 https://w3c.github.io/automotive/vehicle_data/data_spec.html

Standards Institute (ETSI). The development of the ecosystem is also based on cross-industrial cooperations such as the European Automotive - Telecom Alliance (EATA).		networks with, for example, the SENSORIS project in the European Union[52]. The issues described in the report Public support measures for connected and automated driving[53] pertaining to the standardisation issues, mainly from the p. 52, are mostly relevant and shaping the ground for the predictive maintenance sector In the European Union.
Key figures For car data at large, a report published in 2016 estimated the revenues for 2030 in the range of 450 to 750 billion USD[54].		

Potential to enrich recommendations	**Ecosystem Maturity**
This use case presents strong links with other ecosystems and there is a source of scalable recommendations, including notably connected car at large), but also usage based insurance (UBI), road condition and traffic monitoring, it has also deep links with tracking ecosystem as well as artificial intelligence, more generally with IoT. The diversity of stakeholders at different stages of maintenance is implying a large range of actors, thus a rich source to analyse the interactions. It is a case, unlike the majority of other ones, where regulation actors are not at the forefront. It is interesting to evaluate how some traditional actors, such as car manufacturers, in order to no be disrupted, are struggling to find a place in this ecosystem. The potential benefits expected from the predictive maintenance paradigm are leading to a strong competition between stakeholders and their solutions which may be considered as a cause of fragmentation, and a reason why keystone actors do not emerge clearly to better structure the ecosystem which tend to appear as a set of almost parallel structures. In consequence, besides isolated initiatives, there is a low culture of data sharing. This does not prevent a functioning data sharing in the vertical view of the ecosystem, as actors are controlling them or cooperating with not competing ones, but there is little data sharing in the horizontal understanding, limiting mainly the data flows and the data cycles to isolated actors or alliances. This challenge to be solved is a potential contribution to better understand how orchestration may mitigate conflicting interests when they are not the engine of the ecosystem development, but a constraint. In other words, how this fragmented situation could at least enable the	This is an emergent ecosystem, characterized by the exploration of diverse standards and business models paths.

[52] https://sensoris.org/
[53] https://ec.europa.eu/growth/tools-databases/dem/monitor/sites/default/files/DTM_Autonomous%20cars%20v1.pdf
[54] https://www.mckinsey.com/~/media/McKinsey/Industries/Automotive%20and%20Assembly/Our%20Insights/Monetizing%20car%20data/Monetizing-car-data.ashx

emergence of a decentralized set of orchestration actions.	

Incentives	**Barriers**
As several vehicles from different manufacturers may use the same components from third parties, a shared space aggregating the data from these vehicles should be theoretically an incentive to build intermediary platforms, where a larger amount of data on given pieces could benefit to the whole sector. For the moment, this argument is not enough to balance the constraints Long-standing standardisation efforts from 1988 and getting a legislative support from the European Union in 1998[55] (interestingly for environmental concerns), with the OBD, removing a part of the constraints on data collection.	The ecosystem is highly fragmented and dependent on the evolution of bigger upstream ecosystems. The standardisation of communication protocols from the car manufacturers side, the standardisation issues a large, although explored, need to be actually and widely implemented. There are conflicting industrial logics with the concern of losing competitive advantages to competitors. In terms of business models, there are concerns about a lower possibility to control the downstream business models around vehicle maintenance.

Key Documents	**Key Experts**
(McKinsey, 2016) *Car data: paving the way to value-creating mobility*[56]	

55 https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31998L0069:EN:HTML
56 https://www.mckinsey.de/files/mckinsey_car_data_march_2016.pdf

Name	**Tourism accommodation reviews data ecosystem**		
Use Case / Case Study **Use case**	**Economic sector(s)** Tourism	**Related EU Data Space(s)** No specific data space	**Category** Thematic

Summary ***(****Ecosystem's rationale, Value Proposition, Business Model, IPR & Legal context, etc.****)***
The main purpose of the ecosystem is to increase the information flows about accommodation and improve in turn the quality of accommodation information. There are diverse business models based mainly on freemium offers. The channels are fully digital, either through web sites, mobile application, APIs. The legal framework varies according to the countries.

Key Players
The key players are tourists themselves, touristic services providers, geographic information providers, cultural heritage organisations, accommodations booking platforms, tour operators and touristic guides, researchers and market study organisations, and booking platforms

Key figures
YELP: Revenue of 943 Million €
TripAdvisor: 3300 jobs, 700 million of reviews, 1,3 million of hotels listed, valuated to 7 billion of $.

Technology ***(****Interoperability, Standards, API, Data model, Format, Processes, etc.****)***
There is no technology specific to this ecosystem, no standardized models nor generation of unstructured data or semi-structured data (perhaps due to a low level of modularity) and thus few data cycles

Potential to enrich recommendations
The potential data ecosystem around the provision of information about the tourism accommodation is an interesting case study, as it is implying a characteristic activity of the services economy, both (at least partly) based on and reaching the general public.
TripAdvisor around its platform is a major actor of this ecosystem, along with Yelp and more generic actors such as Facebook or Google.
Interestingly, despite the existence of a dominant actor and the relative high maturity of this ecosystem, there is no real structuration. It may be linked with the fact that this dominant actor does not play a role of keystone actor in the ecosystem and is not really orchestrating the ecosystem around its platform. In turn, this specificity may be linked with a relative simplicity of the basic ideas, not easily patentable, a low modularity between the resources and the actors themselves.
It is not that the ecosystem is not using advanced technology, it is actually the contrary, the actors developing and using large databases, data streaming pipelines and innovative algorithms to analyse the unstructured data. The fact

Ecosystem Maturity
The ecosystem is sustainable even if there is no real organisational structure defined.

is that the idea of aggregating reviews may be reproduced by any kind of actor, ancient or newcomer, and indeed some booking platforms are trying to cover also the downstream parts of their economic sector. The dominant actor has mainly to rely on the loyalty of the reviewers as this situation does have some features of a winner takes all case. For similar reasons, there is few modularity between the actors. On the contrary, the successful ones do not try to specialize on a part of the chain, but try to internalize minor actors or neighbours ecosystems, such as the recommendation of touristic activities, as is witnessing the case of TripAdvisor having taken control of Viator. The absence of a keystone actor orchestrating the ecosystem, on top on the use of unstructured data, contributes to explain the non-usage of standardized data model at the scale of the ecosystem, as well as a low number of data cycles after the combination of the source data.	

Incentives	**Barriers**
Involvement of not trustable actors Possibility of bad reputation / exposure for hoteliers Fake reviews partially mitigated by community-based and in-house validation processes	Trustable insights for hoteliers Possibility of good reputation / exposure for hoteliers Cross-checked information about accommodation

Key Documents	**Key Experts**
Ricci F. (2020) Recommender Systems in Tourism. In: Xiang Z., Fuchs M., Gretzel U., Höpken W. (eds) Handbook of e-Tourism. Springer, Cham.	

Name	Weather data ecosystem: the case of crowdsourcing and citizen science		
Use Case / Case Study **Use case**	**Economic sector(s)** Weather prediction; pervasive re-use	**Related EU Data Space(s)** Not directly related to a specific data space but contributes to most pf them.	**Category** Thematic

Summary (*Ecosystem's rationale, Value Proposition, Business Model, IPR & Legal context, etc.***)**
From the data provision perspective, the ecosystem is established for long, first at the national scales, and even at the global with UN agency, with a strong culture of data sharing, at least among the official agencies. Depending on the countries, there are models of data re-use by third-party producing other layers of services on top on these data. These functioning models contribute to explain a large level of standardisation in the field. Among the geospatial data, it is one of the most successful infrastructures, not only in terms of public-private partnerships, but also of sustainable business models for the private sector. It is also having interfaces with other ecosystems (agriculture), in particular with climate monitoring ecosystem, but also with user generated data (miniaturization, IoT…).
On top of a more or less dense network of fixed sensors, balloons and satellites, the weather data ecosystem is also and for long relying on the collection of data from ships and airplanes. What is new is the pervasive dissemination of communication devices among the general public, with sensors able to capture images of a given phenomenon. Moreover, at a lesser extent, there is also the development of a new trend among a part of passionate amateurs acquiring sophisticated sensors, some of which being kinds of miniaturised weather stations. There is thus a case relevant to study on the impacts of these newcomers on the traditional weather data ecosystem.
These trends are affecting the ways how data are collected, validated, aligned with authoritative data, and re-used.
The value is lying at different stages: the existence of a business model sustainable enough to support the data collection stage is a first question, more specifically if sustainable (business) models may emerge around the aggregation, the quality improvement and the transformation of these raw data in more sophisticated ones, this especially in new intermediary platforms.
A second source of value, leading to rich relationships between this ecosystem and most of the other use cases considered for this study, is the re-use of these combined data for different purposes. These purposes can broadly be divided between the monitoring of areas where the density of fixed sensors in the traditional weather systems is pretty low due to different factors, often associated with low population density areas. The second on the contrary is to use these in dense urban areas, to complement traditional monitoring with the help of the crowd and leading to the possibility of modelling fine-grained local phenomena, which are prone to happen in dense cities.
This use case is also suitable to bring insights on IPR, mostly among the new data cycles taking ground on the combination of authoritative and crowdsourced data. Data ownership is a question not only on influencing the development of the data sharing behaviour, but also the development

of the ecosystem itself. In another perspective, it may also be an interesting predictor of the willingness of citizens and people and organizations gathering data to share them when the public interest matters. Case also of accumulated kinds of legal rules bearing on the source of a combined dataset.

Key Players

Key players may be divided between traditional weather data ecosystem stakeholders and newcomers. The longstanding international cooperation efforts led to the creation of a global agency to coordinate the meteorological services, the World Meteorological Organization[57]. Most of countries have national agencies in charge of data collection. The national or local ecosystem are highly different depending on the strategic choices made, the US relying largely on private sector to complement the data collection and analysis, while other countries focused on public sector and their national agencies.

Either internally or calling to third-party, new or better developed paradigms such as machine learning algorithms are increasingly re-used and disrupt this domain, such as other ecosystems.

The disruption factors may be considered from two other kinds of stakeholders.

On one side, individuals, sometimes organising their contribution in cooperatives, are gathering data in a more and more professional and re-usable way. On the other side, other stakeholders may disrupt the ecosystem through new ways of collection data and providing new kinds of data, for example in the frame of the New space paradigm.

Technology *(Interoperability, Standards, API, Data model, Format, Processes, etc.)*

This ecosystem is first supported by a longstanding effort of standardisation at the international scale, with standards covering large part of the data collection and first stages of data cycles. Most used standards related to weather data are GRIB, NetCDF, HDF5, a report assessed their usage for atmospheric sciences[58]. The WMO is maintaining vocabulary tables, mostly identifiers. However, these standards have to be adapted depending on the new instruments available, the new questions asked to the ecosystem, the complexity of new forecasting models, and the progress of artificial intelligence.

Even with a strong standardisation effort, it remains that some of them are not well documented, others not maintained and tend to become obsolete.

Moreover, the development of crowdsourced data is rising new questions on the adaptation of complex standards by the companies producing the miniaturized stations for the general publics, but also the alignment of these crowdsourced data to the downstream data flows and their related standards[59].

---

[57] https://public.wmo.int/en

[58] https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5acd30670&appId=PPGMS

[59] https://e-space.mmu.ac.uk/974/7/Crowdsourcing_for_climate_and_atmospheri.pdf

Key figures In the European Union, the impact of weather services and related data, albeit highly variable depending on the economic or social sectors implied, is estimated at least around 3:1[60].		

Potential to enrich recommendations	**Ecosystem Maturity**
This use case is a good example of an infrastructure successfully turned to an ecosystem and having reached a pretty mature stage. This ecosystem is also showcasing the importance of global governance stakeholder, the WMO, on top of agreements between countries. Its interest in reinforced by its diversity, mostly with different data sharing models and modalities. Although its sustainable business models, it is also well embedded in Open Access, Open source and Open Data paradigms. It was already an example chosen to demonstrate the different models followed by weather services ecosystems between the European Union and the United States, the rich ecosystem developed by the latter having been an argument for the development of PSI and then Open Data policies in Europe[61]. In spite of its sustainability, this ecosystem has to face several potential disruptions. With most of the other use cases, it is sharing the challenges brought by technical developments, in particular artificial intelligence. More specifically, the emergence of citizens as true stakeholders becomes a real concern. They may be implied not only as data producers, with all the technical issues that this creates[62], but may also have a more active role. This makes this use case a relevant ground to assess the digital skills data space promoted by the European Union, combined with the citizen science paradigm. The United States weather agency is running different programs to frame the citizen's crowdsourcing. One can mention Mping[63], which organises a subset of ecosystem not only through the application itself but aggregates similar data from third-party applications and services. Another initiative consists in organising the aggregation of data from privately held weather stations, the agency built a public-private partnership to do so[64]. This way, these initiatives make this ecosystem among the pioneering ones to explore the feasibility and the success conditions of the data cooperatives and explore at the same time the orchestration role of public authorities and agencies.	The pre-existing data ecosystem is at a high degree of maturity. The integration of new data from the general public makes the specific ecosystem around crowdsourced data less mature, albeit very innovative: the sustainability of the data collection and of the business models are the key issues.

60 https://asr.copernicus.org/articles/10/65/2013/
61 http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
62 https://www.nbi.ku.dk/english/theses/phd-theses/phd_theses_2019/kasper_stener_hintzs/Kasper_Stener_Hintzs.pdf
63 https://mping.nssl.noaa.gov/
64 http://www.wxqa.com/

Incentives	Barriers
For crowdsourced data, the incentives could be micro-payments or goodwill to contribute to the public good, but also the role of entertainment to motivate participation.	There is a gap to fill given the complexity of the domain, to ensure the consistency of the ecosystem, and even more to increase citizen participation.

Key Documents	Potential Experts
(Hamaker et al., 2017) Analysis of existing data infrastructure for climate service[65]	

65 http://eu-macs.eu/wp-content/uploads/2017/07/EUMACS-D13_EXISTING-DATA-INFRASTRUCTURE-FOR-CLIMATE-SERVICES.pdf

Annex IV. In depth ecosystem analyses

The selection of case-studies and use-cases has been explained in the methodological part and is annexed to the present document.

Each view of the ecosystem is based first on narratives, highlighting the key points of the ecosystem as well as the phenomena affecting or caused by the ecosystem development, second on a graph both summarizing and contributing to understand the main features of the ecosystem.

For the high-level analysis the first view of the ecosystem is developed. Thanks to the insights extracted from the interaction with the relevant stakeholders of the selected ecosystems, the in-depth analyses will feed the second and the third views.

Therefore, the present section will present the 5 in depth analyses, namely:

1. Local data ecosystem
2. Disaster Management ecosystem
3. Logistic and tracking ecosystem
4. Smart Agriculture ecosystem
5. Spatial Data marketplace Ecosystem

1- Local Data Ecosystem

A **local data ecosystem** illustrated by the case of Rennes Métrople and its Rennes Urban Data interface[66] initiative. Rennes is implementing since 2016 a collaborative and partnership-based local data strategy, targeting and Inclusive and sustainable governance model for the local ecosystem, adopting the quadruple helix model. Rennes is also experimenting the concept of the City as trusted third-party allowing citizens to take back control over their personal data.

Interaction, structure of the ecosystem

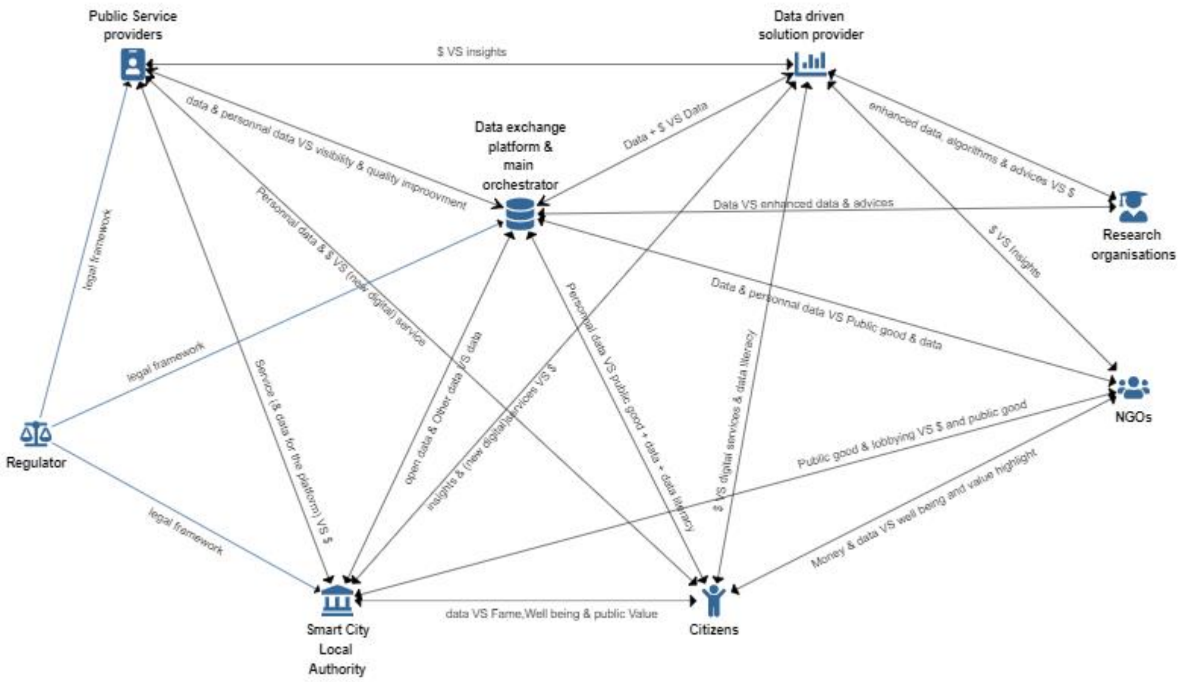

Figure 7 - RUDI ecosystem

Summary

The role of local authority is currently central. Through its project, Rennes is exploring different issues:

1) What may be the role of the local authority at the first stages of the emergence of a local data ecosystem?
2) What may be the role of a public organization in the long term? It is distant orchestration? For example, acting more on a regulation base, to ensure the continuous commitment towards public value, while keeping a healthy and sustainable ecosystem.

Private companies are actively engaged in the ecosystem for different reasons: they may have relevant data not directly created or used by other actors such as those of

[66] https://rudi.datarennes.fr/

the public sector; they may store personal data from citizens; they are central actors to create new data-driven services; more generally, they are essential participants to ensure the long-term sustainability of the ecosystem, through profitable businesses (from start-ups to multinational companies).
Researchers are also engaged as they are of paramount importance at the early stages of emergence of the ecosystem, providing
The goal is not to consider citizens only as data providers or end-users buying products and services, but as central actors contributing to define the scope and the goals of the ecosystem and being active participants aware of their data and their re-use.
NGOs are active participants having different roles in the data ecosystem. They are first actors to engage in order to increase their awareness and participation. This participation is also intended to be as project owners.

Governance issues
At the basis of Rennes project, is the observation of a fragmented data landscape, especially at the local scale. The data ecosystem perspective may complement the Open Data policies to ensure the availability and the usability of relevant data at the right scale.
Rennes intends to build a collaborative governance of the ecosystem. The consortium of the project reflects this principle.
The ecosystem of a smart cities requires to gather a large range of actors of different kinds (public / private) and whom objectives are not always aligned or convergent. As they may be considered as specific kind of PPP. Workers themselves are also having different disciplinary backgrounds, and different working cultures. This issue is particularly important at the emergence stage of a data ecosystem, it can prevent the participation of participants. Not solved early, it may also create discrepancies at later stages. It goes beyond the mere case of the ecosystem perspective, but it may be harmful to the development of an ecosystem. For example, people may have different views on the role of top management to validate any decision, or to give more autonomy to a sub-component of a running project.
Jeopardize the sustainability of the ecosystem. Actors need to define a common way of working, have frequent meetings to identify the issues and enable the solutions, granting the motivation of all the relevant stakeholders.
As they target a shared governance, stakeholders have to co-design the features of the platform. On the long term, this requires also to build a “data social network”. From the technical perspective, this network is built on a platform, where to access the data, re-use them, but also to meet and thus to frame the ecosystem. This platform is also a socio-technical component, especially from the meeting between stakeholders. This way, they can express and fulfil their needs in terms of data, processing, ideas, research of partners.

Users and re-users engagement
In terms of ecosystem orchestrations, it consists in increasing the links and the intensity of the relationships, or in other words, the self-awareness of the ecosystem. To attract new participants and ensure the commitment of existing ones.

It has to overcome the risk of having to few contributors, which would jeopardize the creation of high-quality services and would be specifically harmful at the early stages of the ecosystem. (and the reluctance of citizens to share their data)
At the early stage, citizens are recruited to form a panel and contribute in the co-construction of the platform supporting the data ecosystem.

Creation of new data
Rennes is organising a call for projects targeting different objectives : to make the ecosystem activities more tangible for stakeholders; to define more accurate governance rules through these collaborations; to foster the re-use; to showcase new business opportunities and the possibility to build sustainable business models based on the re-use of data; to incentive the actors to share their data; same for the exploitation of their data; to enable new nodes in the ecosystem (as it is a federated approach)
These projects are quite open but have a specific interest for the solutions focusing on personal data issues, or for projects taking into account the public interest.

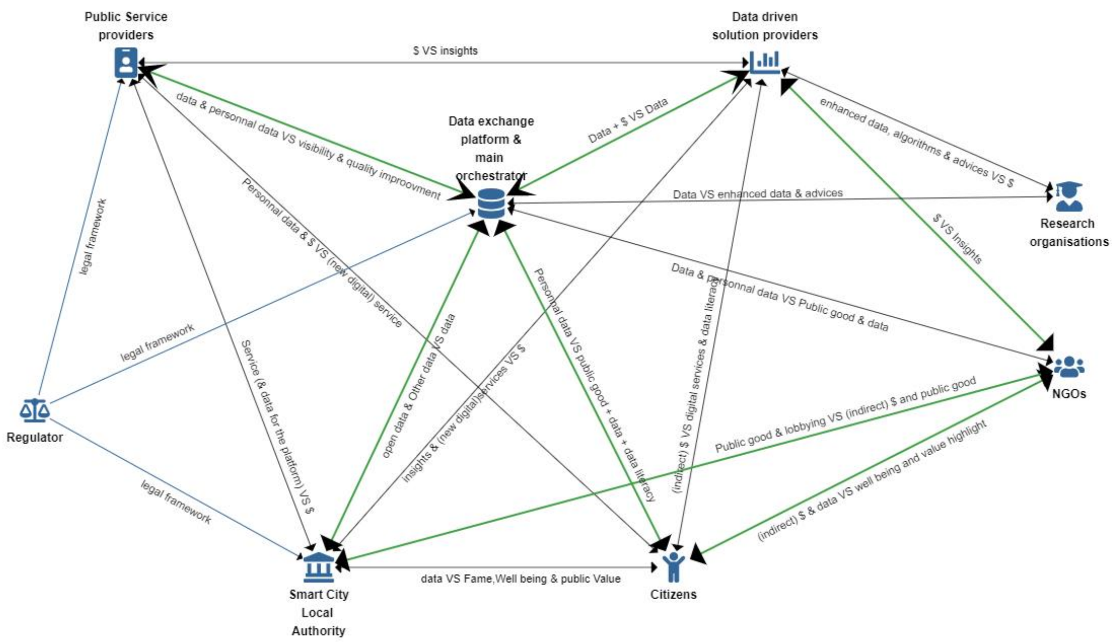

Figure 8 - RUDI call for projects ecosystem

Same stands for the citizens side. Citizens are experts on the usage (of public services), they are an essential component of the ecosystem, even if their interest for the data issues is not obvious nor general. Their contribution to get better personalised services is reflected in Figure 3.

Data sharing (commitment)
Until now, Rennes engaged mainly big companies, so-called public or para-public companies, or at least in charge of large public service deliveries. For some companies, putting a mandatory commitment to Open Data some data represents, in their perspective, a risk to their business models, as it is depriving them from what they may consider as a competitive advantage. Moreover, not all the companies report this as a priority concern.

A large obstacle to data sharing reported from the Rennes data ecosystem is the risk represented by the GDPR rules, as companies built strict (and expansive) internal systems and processes to comply with these rules. Any ecosystem based on the re-appropriation of personal data should take into account the concerns about the risks brought by these issues.
More generally, the provision of data may be hindered by the internal processes and the historical features of the companies' information systems, e.g. ENEDIS and its legacy systems.

Digital transformation of the public sector and of public services
One challenge of the Rennes project is to understand if and how the ecosystem may enable the redesign of public services provision, as part of the digital transformation of the public services.
Moreover, the project of Rennes is based on the finding that there is a competition between traditional public service providers and new stakeholders to provide data-driven services. The services must be able to adapt the ways citizens are consuming, behaving. They need also to be more customized, individualized, adapted to real-time needs.

Business models & sustainability of the platform
Rennes platform is still at an early stage. More than business models, business opportunities have to be explored by the calls for projects. The project currently led by Rennes Métropole intends to show the economic benefits and the overall sustainability of business models based on the production and re-use of shareable data.
Rennes Métropole, at this stage, links the sustainability of the ecosystem to two questions: finding a business model for the platform itself (depending of course of the business models of the actors implied in the ecosystem), finding a satisfying governance model for this platform. Pertaining to the business model of the platform, it needs to cover the cost implied by the infrastructure supporting the data ecosystem, to enable compensations of the processes developed by different kinds of actors to make the data available.
An ongoing action of the project developed by Rennes Métropole is to build tools to score the services created and monitor the impacts.
As the platform brings the benefits – that can be linked to those of a data marketplace – of negotiation costs and time savings, they may be included in the scoring of the impacts of the platform, and be considered as arguments of the sustainability and the economic benefits of the platform.
The current project aims at exploring the willingness of the different actors to participate in financing the infrastructure. The demonstration of sustainable business models is an important dimension.
Also, data requests are mainly considered for big companies, or at least to local satellites of these companies. As a recommendation, it should be explored in more details regarding the consequences of these approaches and policies on the companies

Data literacy
One outcome expected from these activities is to increase the data literacy, the acculturation on data issues, to increase the awareness of the different actors on the importance of data for the creation of new services.

Data flows and data cycles

Summary

The data flow illustrated in the figure below shows how Rennes intends to foster the re-use of personal data, held by citizens themselves (including crowdsourcing) but mainly by companies. There the public transport services having collected personal data through the course of their service delivery, are providing back these data on citizens request. These data may be combined with other personal data voluntary provided by citizens, with crowdsourced data and other kinds of data such as statistical ones. It enables first the enrichment of the datasets, then the identification of usage patterns as well as the precise identification of needs. This may be used in turn to feed algorithms targeting different new or improved services, such as service personalization.

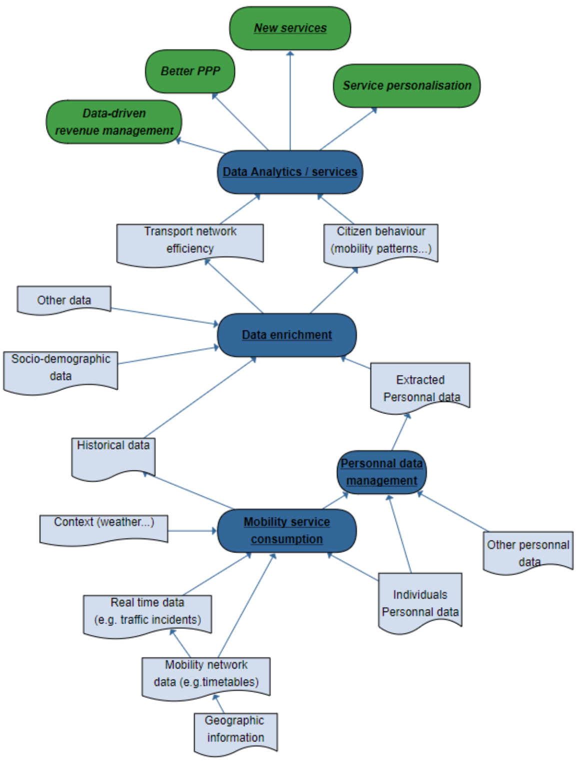

Figure 9 - RUDI New/better mobility services data flow

Architecture of the platform

In Rennes, the chosen federated architecture is well aligned with the current features of the big players, such as ENEDIS. For smaller actors, difficulties may arise to participate, for skills and resources issues. In Rennes, the choice of a federated architecture enables to access, manipulate or combine data in a way that would not always be possible in other architecture models. For example, it may be possible to access some health data and not to download them. Thus, it enables to combine data that could not be combined with other architecture choices.

Even if the question is still to be decided, Rennes is considering the inclusion in the platform of the services developed by third-parties of the ecosystem. For that, they need to be integrated in the core of RUDI. This issue also has governance implications.

Data access, API and data standards

APIs need to link different architectures already existing:

1) The data catalogue of the platform itself, it is one important outcome expected from the project, being an interface to access data provided by various producers/providers;

2) The Open Data platform of the Métropole;

3) Other Open Data portals

4) Citizens and company's data

5) Services already using the platform

The services built on the data provided through the ecosystem need also to cover a multitude of spatial scales (from local to global).

Concerning the standards (and the models, the current stage of the platform in Rennes recommends not to focus on a priori alignment as a mandatory condition to participate in the ecosystem. On the contrary, the expected technical maturity of the datasets is expected to be lower than in other projects, such as the Open Data release. The ecosystem is perspective brought by the platform and the related projects is considered as a way to enable a progressive re-structuring of the data, either by the data producers themselves or by other strata of the ecosystem with specific business models. Data interoperability (and quality issues) do not arise at the same importance for different projects. To explore this, the call for projects initiated by the Métropole are considered as a suitable (and varied enough) approach to trigger these transformations.

In Rennes, the heterogeneity of the data leveraged by the ecosystem has to be handled by the catalogue, relating also to handling the management of different levels of access rights to the datasets.

Data sharing issues

The project is partly based from the finding that it is difficult to produce shareable data (skills, time, production costs). Another hypothesis is that this ecosystem could provide a broader scope provided as Open Data, even from a different perspective: not only the very largely aggregated data that are currently representing the majority of available Open Data, but data such as those issued from IoT, even with restrictions such as a certain level of abstraction (security issues). This could benefit to the local researchers and students needing different kinds of data to test and build their models. As part of the digital transformation, it could lead to a greater use of these kinds of data by public sector actors, and later on their re-use by the different users communities.

One benefit of the ecosystem approach, combined with the focus of data usually not accessible is to broaden the scope of available (Open) Data. At the same time, this leads to the re-use of different kinds of data in the ecosystem, with different kinds of licenses and re-use conditions. Even if this situation is not so exceptional for the public sector, some GIS data are already not disclosed in the traditional Open Data release, especially when they imply personal data.

Personal data

Different kinds and challenges emerge concerning the personal data and their geospatial dimension. The main goal of the project led by Rennes Métropole is to give citizen a way to access their data and to new services based on these data. This relies on a citizen empowerment approach, to make them aware and capable of managing their digital rights. Another goal is to make available and re-usable a range of personal data while remaining respectful of privacy.

From the technical perspective, it means the platform needs digital rights management functionalities. It required also a (dynamic) digital consent to shape the data sharing of personal data, built in cooperation with researchers from the LTSI[67].

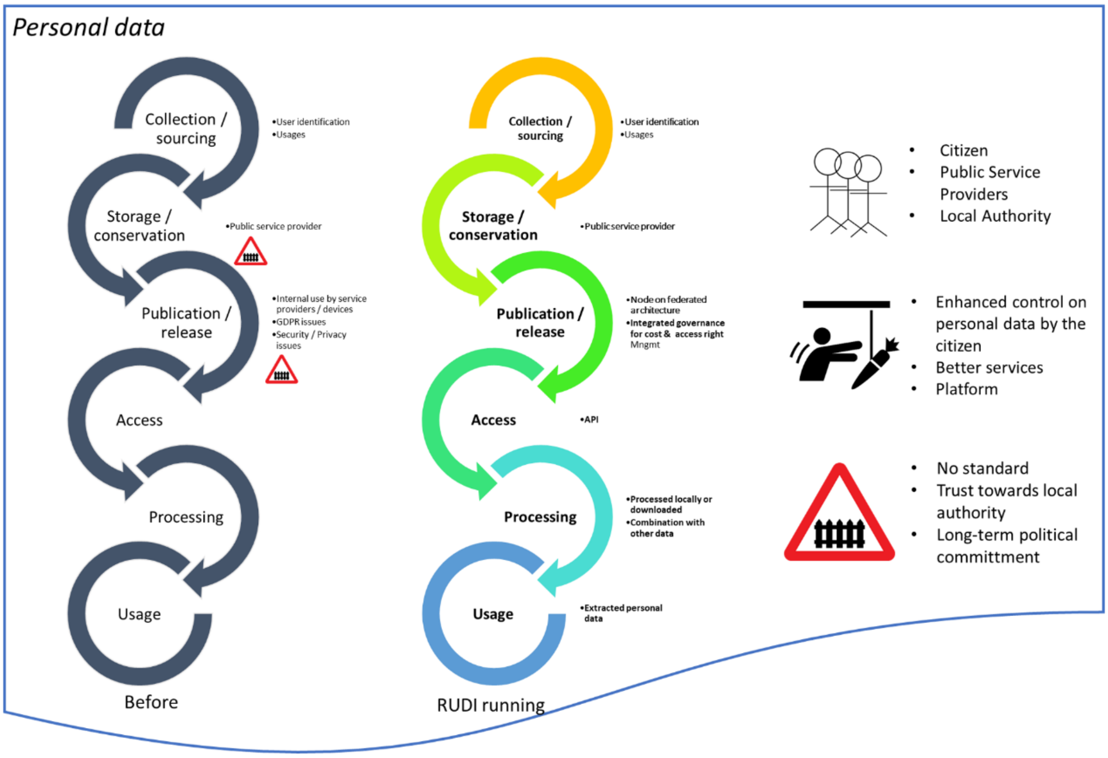

Figure 10 - RUDI personal data set

Several kinds of personal data are identified by the project:

1) Personal data directly provided by citizens to any organization for a specific purpose, which is a sub-set of the already well documented crowdsourcing principles;

2) Personal data collected as such by third parties, e.g. by NGOs

3) Personal data gathered by companies or other third parties for a given objective, e.g. delivering a public service related to mobility, implying the

[67] https://www.ltsi.univ-rennes1.fr/

collection of personal data. These data may then be reclaimed by citizens leveraging the GDPR framework and the personal data rights management module of the platform, then again be re-used for other purposes.

2- Disaster Management ecosystem

The disaster management ecosystem is illustrated through two case studies: The Brussels emergency services data sharing platform and the Danish Common Data on Topography, Climate and Water project preparing the country for climate change scenarios. These two cases allow to cover the emergency management phases from Mitigation, Preparedness, Response and Recovery. This highlights the importance of considering the different time dimensions: Real-time and the Historical and Simulation.

Interaction, structure of the ecosystem

Summary

This data ecosystem is drawn based on the inputs from two quite different although complementary use cases. The first one is based on the data sharing issues for first response among the security services of Brussels. In opposite, the second use case is more focused on the long-term preparation of infrastructures against climate change in Denmark. This diversity allows to illustrate the different challenges faced by the actors in these two cases, but also the need to orient the ecosystem to the creation of data at different temporal granularities and in the long-term.

In this data ecosystem, there is no central actor orchestrating the interactions between stakeholders and the data flows. It is a case where a distributed kind of orchestration prevails, based at the same time on legal and more often on initiatives of actors. Compared to similar cases without a central actor, there is no overwhelming role of commercial value, but a strong potential of public value.

There is a large range of actors involved in this ecosystem. Government and government public services and agencies are of paramount importance, and are involved at different scales, from the European Union level – at least for the strategic policies – to local authorities, depending also on how the first response and the related social challenges are dealt with in a given country.

At a city scale, first response implies a tight collaboration between public safety organizations (fire forces, medical services, and police). A lack of coordination leads often to a fragmented landscape, without systematic data sharing between actors.

At the interface between public and private sectors, there is the need to prepare the collection in the long term of data allowing to mitigate disasters consequences and increase long-term preparation, which may be a combination of regulation and incentives to collect and store the relevant data, with a central role for scientific actors to help organizing the collection and long-term storage and conservation of relevant data.

The private sector is also an important actor of this ecosystem. Even if commercial value is not the main driver of the ecosystem, private companies are first contributing to first response, depending on the responsibilities affected to them, and are also providing products and services increasing the efficiency of the ecosystem. Considering only the insurance companies, they represent one million employees in Europe (2014).

Citizens, beyond benefiting from the ecosystem and funding it through taxes (and funding NGOs also), are contributing to it as volunteers for some services, but more generally may participate in the data collection through crowdsourcing approaches.

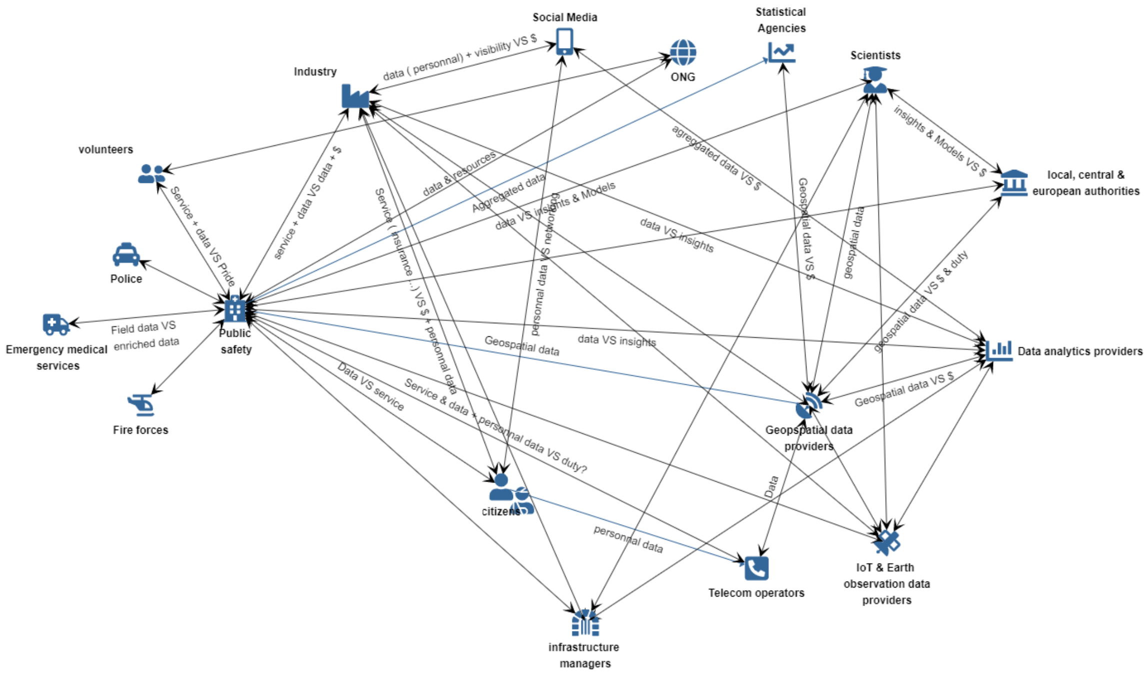

Figure 11 – The disaster Management ecosystem

Governance issues

This gives an extremely fragmented ecosystem's landscape, leading to numerous bottlenecks, a low degree of awareness and few truly collaborative processes at the scale of the whole ecosystem. This is in part a governance problem as there is no central actor to coordinate and stimulate the ecosystem.

Mitigating policies are considering a combination of regulation commitments and incentives. In particular, the absence of data and knowledge sharing mechanisms hinders the emergence of the ecosystem. What the example of Brussels is showing is also the importance of local initiatives to identify the bottlenecks, organize small-scale demonstrations of the interest of engaging data sharing initiatives, breaking the silos and addressing the related technical issues. Previous experiences report data sharing and Open Data as a prerequisite to introduce a culture change mechanism and to build on that to make the data ecosystem emerging. A strong political support and internal sponsorship for data sharing is of paramount importance for data sharing initiatives sustainability.

Business models, sustainability of data provision and of the ecosystem at large

Literature and isolated experiments illustrate the large scope of business models and benefits linked to the emergence of a data ecosystem around disaster management. As noted above, most of them rely on public good or public value first, even if these benefits may also be scored in terms of risks and loss avoidance. Among the mentioned areas expected to benefit of a data ecosystem are the smart safety, the data analytics applied to crime historical data, the data-driven prioritization allocation of limited resources, and more generally the data-driven decision making, Moreover, from a pure economic viability, insurance companies may be directly impacted by such trends, for example for pay per use insurance

Data literacy

The use cases considered currently report data literacy as one limiting factor of the ecosystem emergence, for almost all the dimensions usually considered for that domain. The awareness about data sharing, the value of data re-use has still to be advocated. Long-term conservation are also a concern, especially as there is no direct and obvious return on investment. Same stands for the knowledge of standards, data and metadata models and formats, especially for the open source ones. This is linked to a shortage of skills and resources. Mitigation may be engaged through small steps initiatives, but need also large scale training policies.

Data flows and data cycles

Summary

The data flows described in the figure below illustrate one scenario of this ecosystem where data are enriched, combined and analysed to build a more resilient system. The fact is that it requires a large range of data from different sources. The data sharing occurring in Brussels is a first step to gather a unified layer of data, and these need to be cleaned and made understandable by third parties in order to enable their combination with other kinds of data with different geospatial dimensions, from weather data, which are pretty well standardized, to the data crowdsourced from citizens. Each of the disaster management recovery process requires good quality data from the previous stage, to extract their value and combine them with other data to ensure a relevant extraction of the lessons learnt and the validity of the extracted insights for the long-term preparedness. In turn, a healthy ecosystem shall enable a feedback mechanism using past phenomena insights to be introduced as input of the previous stages in the data processing pipeline.

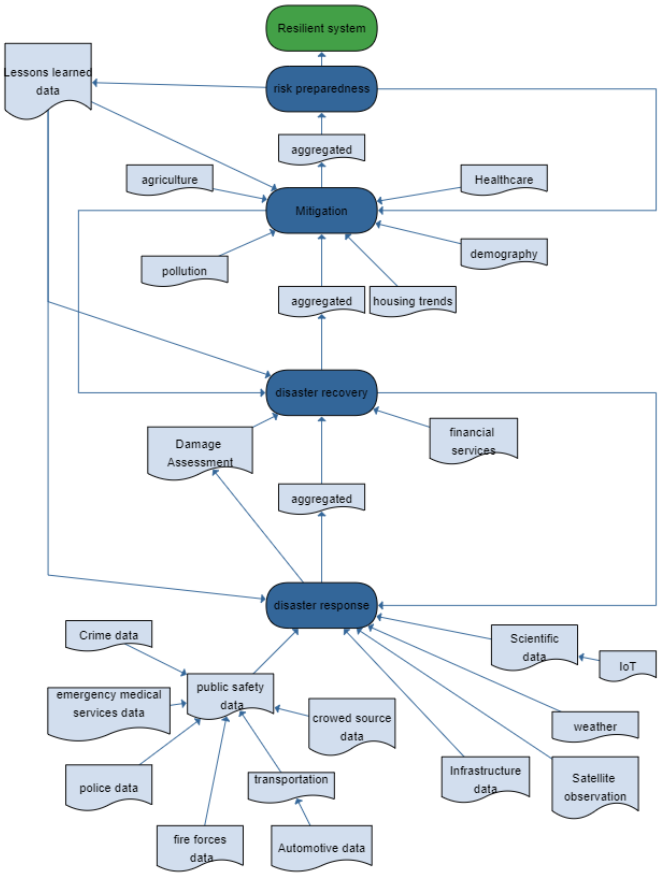

Figure 12 - Resilient system dataflow

Architecture issues

The use case analysed for this ecosystem does not rely on a central platform. Even if it is not always a mandatory criterion this use case shows the difficulties that may occur to structure the ecosystem in the absence of a system bringing technical and processes facilities. This ecosystem requires an architecture needs allowing the collection of heterogeneous data in real-time or near real-time and to make them accessible to a large range of actors (Klaus-Dieter et al., 2017). This requirement is not met until now.

Data access & standard issues

The use cases analysed are showing a lack of data discoverability, even for the available data. The data flows are also hindered by the data silos (fragmented and specific data), the absence of standardized structures and of a common semantic model to describe the data, which are moreover not always collected in a constant and consistent way. This domain is also facing the absence of a consistent geospatial data platform. It means that raw data are facing huge quality problems, jeopardizing their ease of understanding and thus their usability for a larger re-use. There is then a lack of data at the convenient granularity level for downstream usage. There, the main challenge is to ensure the quality and the sustainability of the data provision, the current situation showing the difficulty to get data trusted and convenient data for decision-making.

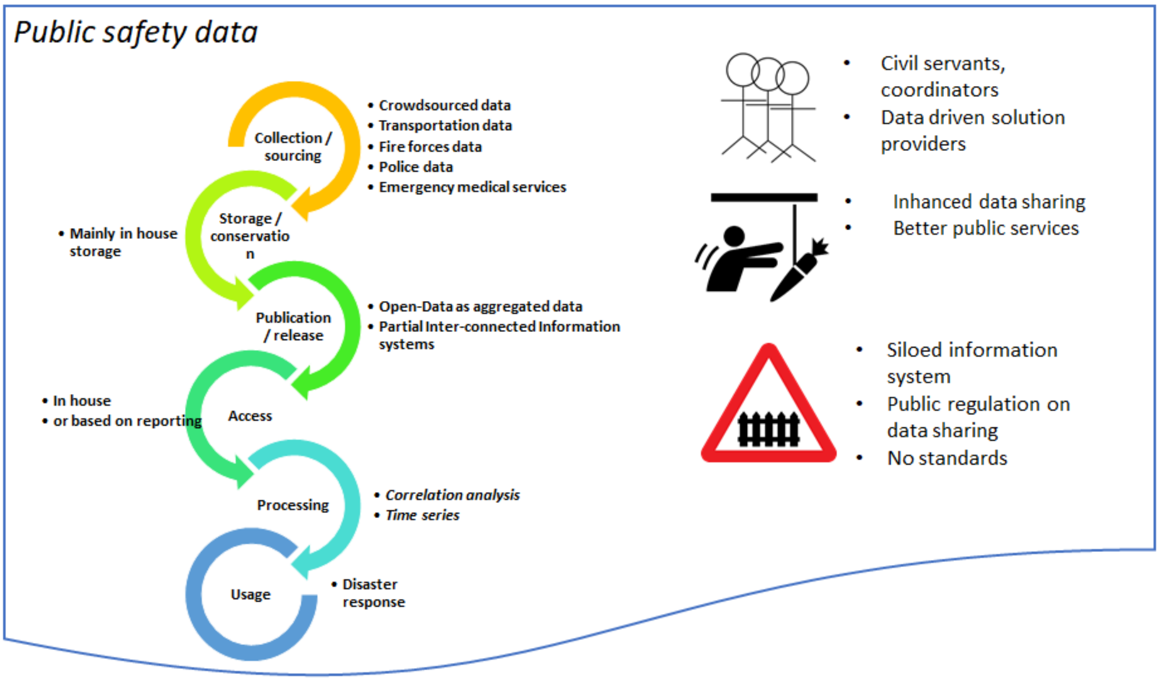

Figure 13 - Public safety dataset

3- Logistic and tracking ecosystem

Tracking technologies for supply chain is illustrated by Spire[68]. Spire builds and manages a constellation of nanosatellites, collecting and distributing earth observation data, Maritime data, Maritime data using AIS messages, Aviation data, using ADS-B data and weather data using Radio occultation

Interaction, structure of the ecosystem

Summary

Data collection is ensured by different stakeholders.

The regulator may have a strong influence on the ecosystem. For example, there is a strong commitment brought by the regulators of the maritime traffic sector (see the third section) to gather data.

Maritime traffic companies are thus equipping their vessels with the mandatory devices. They have at the same time a role of data provision to the downstream stages of the ecosystem and of consumers, either directly or indirectly, and of the products and services based on their data, for example a service of optimized and reactive maritime traffic.

Intermediary companies are gathering the data from different sources, such as on-shore stations or satellites (as done by Spire) and make them accessible for the data ecosystem.

Another role in this ecosystem is to provide complementary data based on the needs of the ecosystem. This role can be endorsed by public (e.g. Copernicus) or private companies (e.g. Spire, see the section on the "business models" and "new data creation").

IT solution providers are adding value combining the data produced by vehicle location tracking and enrich them with complementary data to provide products and services to other actors, downstream or upstream in the ecosystem. They enable the ecosystem to benefit of state-of-the-art solutions, pertaining for example to change detection and overall pattern recognition algorithms, as well as accessing effective infrastructure to handle big data ad extract value from them.

Insurance represents an actor ensuring the commitment of the ecosystem to the collection of suitable data and providing a market representing a broader scope for maritime tracking data.

Other actors of the supply chain, such as retailers, also have several roles to ensure the sustainability of the ecosystem, representing a market for transporters and a customer demanding data and insights to the other actors, to ensure the competitiveness of the whole supply chain.

[68] https://spire.com/

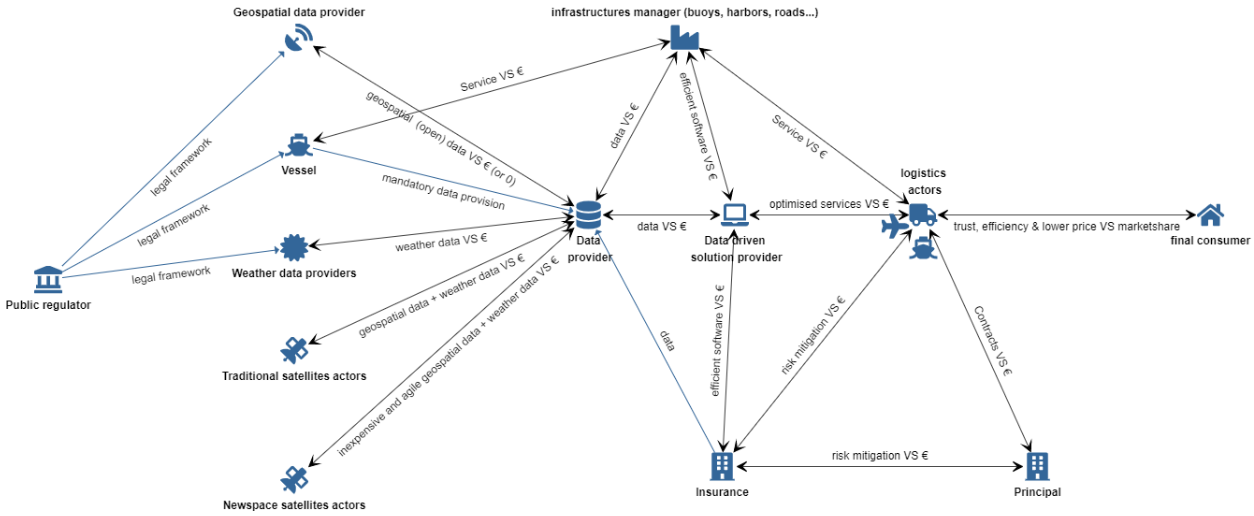

Figure 14 - Logistic and Tracking Ecosystem

Governance and structure of the ecosystem

The main characteristic of this ecosystem is the absence of a central actor, such as the platform identified for other use cases who could have a leading role in the orchestration of the ecosystem. This has consequences on the structure, on the dynamics of the ecosystem. On top of the initial role of the regulators, the ecosystem is mostly driven by business opportunities. It means that actors are leveraging more traditional business opportunities strategies to engage suppliers or customers.

This may lead to a low degree of self-awareness of the ecosystem. In terms of orchestration, a hypothesis is to consider less a fragmented ecosystem than a decentralized ecosystem, as actors tend to have mostly bilateral relationships, with incentives and influences flowing from actor to another, with less overall measures of integration.

It is important to gather and disseminate success stories about data re-use and successful business models, to attract new ecosystem's participants.

Business models

In this ecosystem, for example concerning the maritime tracking business, companies may endorse different roles, being at the same time a role of data producer but also of data aggregator. As data aggregator, intermediaries may gather the vehicles information coming from various sources. A company such as Spire also provides a kind of on-demand data provision model. Based on the new space paradigm, it is able to reconfigure existing satellite devices or even to launch new satellites in a cost-effective way (see "new data creation" section), enabling the efficient creation of complementary data which may moreover be re-used for other purposes in other data ecosystems.

The decentralization of the data ecosystem does not mean it is unbalanced, as the actors have strong albeit indirect relationships depending on their contractual relationships.

Experts report that Geo-location is an essential component enabling the development of Industry 4.0 and their supply chains. There is a large agreement among experts on the profitability of this, which in turn is making the business case for the data collection (and related hardware/devices) layers.
The overall business case, and the related business models, are those of a more reactive supply chain, assuming a deeper and more frequent collaboration and alignment between providers, manufacturers and consumers, to build for example a control tower based on event management based on alert, or even more on analytical analysis of events.

Sustainability of data provision and of the ecosystem at large
The main business model exemplified for this use case is maritime tracking for SC. It consists in gathering and providing data in real-time or near real-time about the locations of ships and their movements. Industry 4.0 requires integrated information flows, for which tracking technologies are an essential component. The benefits of the business models based on vehicle monitoring. Vessel route optimization may benefit in turn to the profitability of the whole logistics and SC domain.
As the commercial data created in this data ecosystem may face the competition of similar data produced by the public sector, there is a debate on the different responsibilities of public and private actors for data provision. From the perspective of private actors, who have to invest in the infrastructure and their operation, it may be perceived as an unfair competition.

Data flows and data cycles

Summary
The scenario examined for the data flows is considering a complex case where AIS data are at the basis of a chain leading to a better rescue system in case of accident, a better management of the fleet in general, and a more efficient and data-driven maritime insurance system.

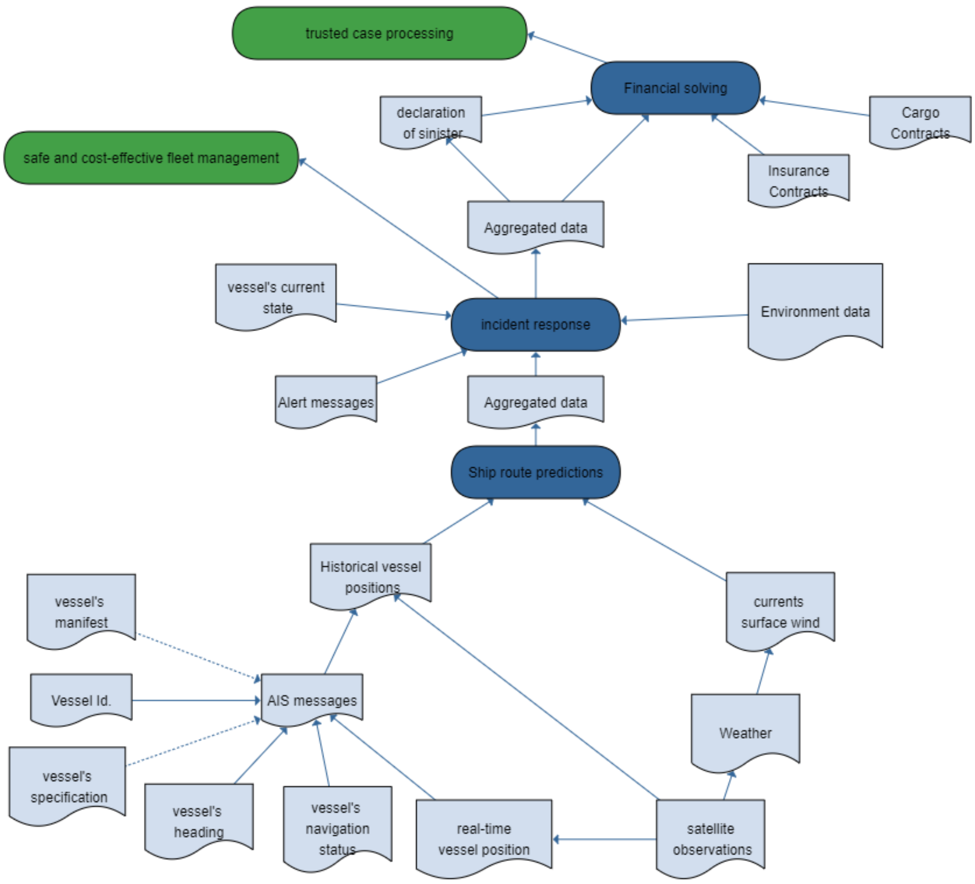

Figure 15 - Spire dataflow

The figure represents first the AIS messages collection, the various variables are mentioned with more details in the “data standards” section. They may be collected through satellites.

In parallel, other data depending on the context can be leveraged. For example, taking the accident in Mauritius, earth observation data could be useful, provided they have suitable temporal granularity and high-resolution images. There, it could be ensured by public or private earth monitoring systems, with actors such as Planet. For other incidents, weather data could also be leveraged. In particular, the models providing currents and surface winds data may be used to for downstream tasks, for example to establish the difference between human errors and external factors.

In case of incident, especially when human lives are endangered, AIS messages make the case for more accurate rescue organization. AIS data may be enriched with other data related to the rescue operations and with complementary data.

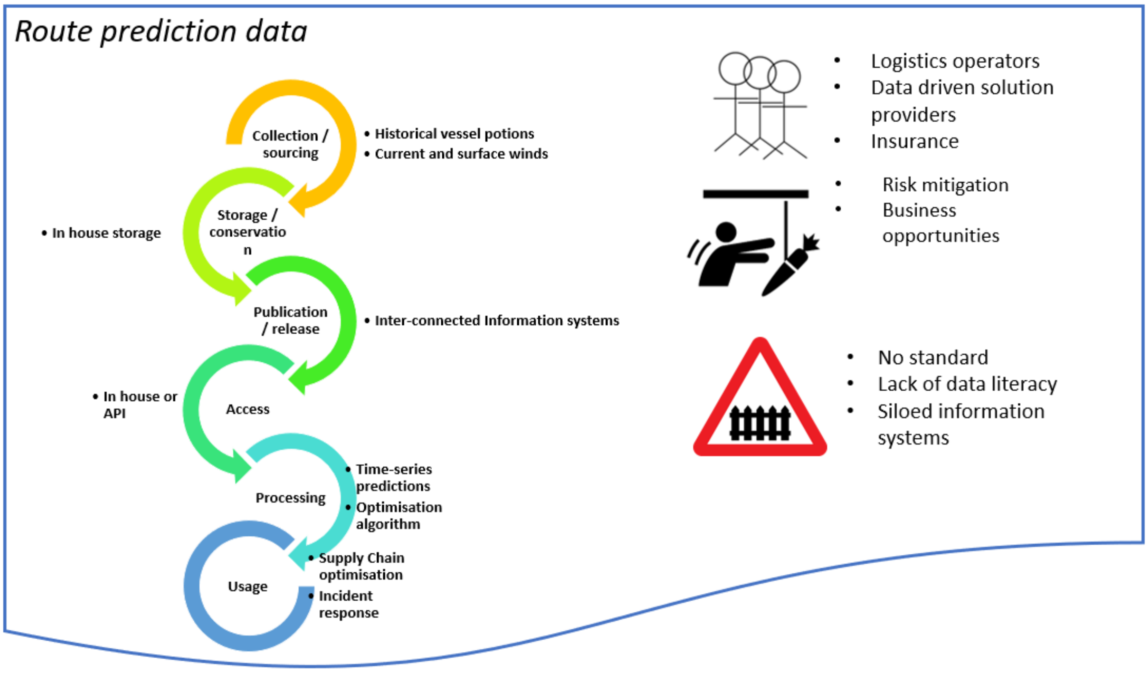

Figure 16 - Route prediction data

Considering the case where a part of the freight is lost, the interface between the data created to maritime tracking and the insurance system needs to be well-designed. Insurance systems may address specific requests to the upstream data collection stages, for example pertaining to their suitability pertaining to their re-use in a judicial context, or more generally in a context where conflicts could emerge. For that, as a hypothesis, to ensure their contractual validity, these data may be gathered and shared through a Blockchain system, thus relying on the development of the software ecosystem at large. The datafication and the digital transformation at large of the insurance sector is still not completely achieved, which goes with a low degree of standardisation pertaining to the data. However, smart contracts are a first step in this direction. Other challenges are the adaptation of current processes, the knowledge and understanding of digital opportunities by insurance workers and decision-makers in the industry. The benefits of an approach integrating the freight contracts (and data), the insurance contracts (and data also), the records of the damages, with maritime tracking are the automation of the cases resolution processes, or at least faster ones, data-grounded cases resolution, leading thus to a larger trust among stakeholders, and to a reduction of the costs. Moreover, this scenario provides a greater scope for the re-use of AIS and location data and may represent a complementary incentive for their collection and storage in a suitable way.

Architecture of the platform

The use case analysed for this ecosystem does not rely on a central platform. The architecture needs to allow the collection of heterogeneous data that in turn need to be accessible in real-time or near real-time to a large range of actors. The data flows have to be organised to make it possible for the participants to use the benefits of the overall software ecosystem, especially of artificial intelligence tools, in a timely manner.

Data models & data standards issues

Maritime tracking is mainly based on the Automatic Information System (AIS), which is made mandatory by the International Maritime Organization (IMO) for vessels of 300 or more gross tonnage. In the European Union, this is currently covered by the Commission implementing regulation (EU) 2019/838 of 20 February 2019 on technical specifications for vessel tracking and tracing systems and repealing Regulation (EC)No 415/2007[69]. It is based on transceivers-equipped vessels, receivers relying mainly themselves on VHF and satellite positioning. The signature may be collected by a variety of systems, such as other vessels (for example for collision avoidance), shore-based stations, buoys, and satellites (S-AIS).

AIS messages include four kinds of information:
a) static information;
b) dynamic information;
c) voyage related information;
d) inland navigation specific information (when relevant).

a) Static vessel information (every 6 minutes)
- ID
- name
- Call sign of the vessel
- IMO number
- Type of vessel
- Overall length
- Overall beam
- Unique European vessel identification number (ENI) (if relevant)
- location of antenna

b) dynamic information (variable)
- Position according to World Geodetic System from 1984 (WGS 84)
- Speed Over Ground
- Course
- Heading
- Rate of turn
- Position accuracy (GNSS/DGNSS)
- Time of electronic position fixing device
- Navigational status
- Status of Blue sign

c) voyage related information (every 6 minutes)
- Type of cargo
- Destination (ISRS location code)
- Category of dangerous cargo
- ETA

[69] https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32019R0838

- Maximum present static draught

This ecosystem benefits first of the long-term efforts to standardize the basic data it is reusing, including geographic information (e.g. WGS 84) and coordinates – and especially satellite positioning (e.g. GPS) –, or radio frequencies.
This ecosystem is an interesting case showing how a system of geo-location information built for an initial purpose, i.e. collision avoidance to complement radar systems, may be leveraged for other purposes and businesses.
It shows the role of an organization, here the IMO, acting as regulator, to keep the AIS) data standardized and usable. Beyond the mere standardisation, the IMO also requests the exchange of data between vessels and with on-shore facilities.
At the same time, from the perspective of data re-users, this information is not enough semantically standardized, it is dirty and inconsistent for their need, hence the need to invest resources to clean the data and make them usable.

Creation of new data
The business needs related to the SC have consequences for the data expected from the ecosystem, their access modes and their quality. In particular, these require the collection of data at the right spatial scale, gathered, stored and made accessible in real-time. As stated above, this ecosystem proved its suitability to gather specific data, the needs of downstream businesses may be efficiently addressed by the upstream actors through the ecosystem perspective, and in this case, thanks to the new space paradigm, where new satellites, new sensors, new software updates, may address the specific needs of business in a time-efficient and cost-effective way.

4- Smart Agriculture ecosystem

The smart agriculture is illustrated by API-Agro[70]. It is a B2B data exchange platform operated by Agdatahub[71] and a society made up of 30 partners representing the agricultural sector, from private companies and public bodies as Chambers of Agriculture, Technical agricultural institutes. It provides a functional, technical, commercial and legal framework for data exchange between the various stakeholders.

Interaction, structure of the ecosystem

Summary
The platform illustrated through the case of API-Agro is at the centre of the data ecosystem, with an actor understanding is role mainly under the network animation perspective, both from the technical and organizational perspectives. This platform, Agdatahub, is the output of a European project, built mainly around research organizations related to agriculture and intending to make their strengths and assets

[70] https://api-agro.eu/
[71] https://agdatahub.eu/

converging for the sake of the general agriculture ecosystem development. Compared to other use cases analysed for this study, the platform around API-Agro is among the largest, encompassing different domains such as food production, but also, at least in part, the agro-food supply chain. Thus, it requires the involvement of a large body of stakeholders. Moreover, this ecosystem is covering two different systems, agriculture as such and agro-food industry.

One rationale for the creation of the platform is the absence in the former data ecosystem of easy and even more integrated contractualisation system, an affordance easily brought by a platform (see UP42 use case).

The platform also intends to overcome one barrier identified and justifying its creation, i.e. the fact that data are available, although in a scattered way, but a lack of maturity hinders their complete (and convenient) release, representing an awareness issue.

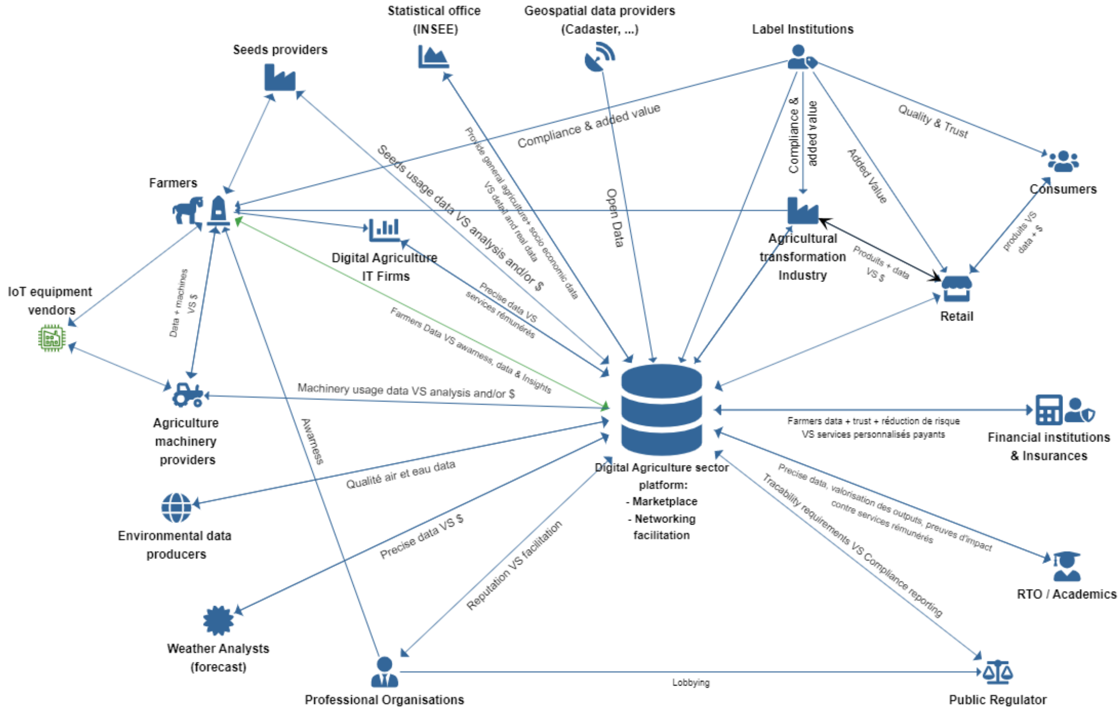

Figure 17 - Smart Agriculture Ecosystem

Governance issues

The use case analysed presents an original approach, as the main orchestrator, or the hub firm, has been created explicitly following a collaborative and even a co-construction approach. It takes the legal form of a business created under the French law as a Société par actions simplifiée (SAS). Including thirty shareholders at the beginning, the platform and the governance were designed following a collective governance model, with a responsibility distributed among a large number of public and private actors. It implies the risk of a loss of focus and the risks associated to objectives that are potentially divergent between the actors, but it enables at the same time the platform leader to get a deep knowledge of the objectives, working culture, business needs and opportunities, or skills of a broad range of participants,

and to be more eager to include their concerns in the decisions made by the platform, taking into account the acceptability of these decisions.
In terms of users and re-users engagement, it facilitates the common understanding between the platform and the actors it is targeting to embed, beyond other more traditional users engagement mechanisms, such as the organization of hackathons and a specific attention brought to the different data literacy levels in the ecosystem.

Business models

The platform is attached to the differences between a data platform and a farmers' platform. The business model of the platform itself was explicitly designed following the concept of open innovation, following the principle that companies may take benefit of internal and external processes to be more innovative, the interaction between different perspectives brought by different actors being more suitable to enhance the extraction of the value from available information.
It is enabling or facilitating different kinds of business models, for example for a greater development of Decision Support System. More generally, it supports the dissemination of analytics service providers and, through network effects, provides new customers opportunities and larger opportunities for partnerships.

Data flows and data cycles

Summary

The scenario illustrating the data flows of this ecosystem is based on a data-driven precise irrigation scenario. It requires mostly static data such as those provided by the land registries and dynamic data, if not real-time data, to assess the current situation of an agriculture parcel in terms of humidity, its needs depending on the plant species that are planted there, and the configuration of the irrigation system, including the machines and the practices of the farmers. Each stage requires to access a large number of heterogeneous data, some being Open Data, others being farmers data, including IoT data. The data flows may then converge to a decision support system on precise irrigation, leading to water and financial savings, and is thus contributing to a greener food supply chain.

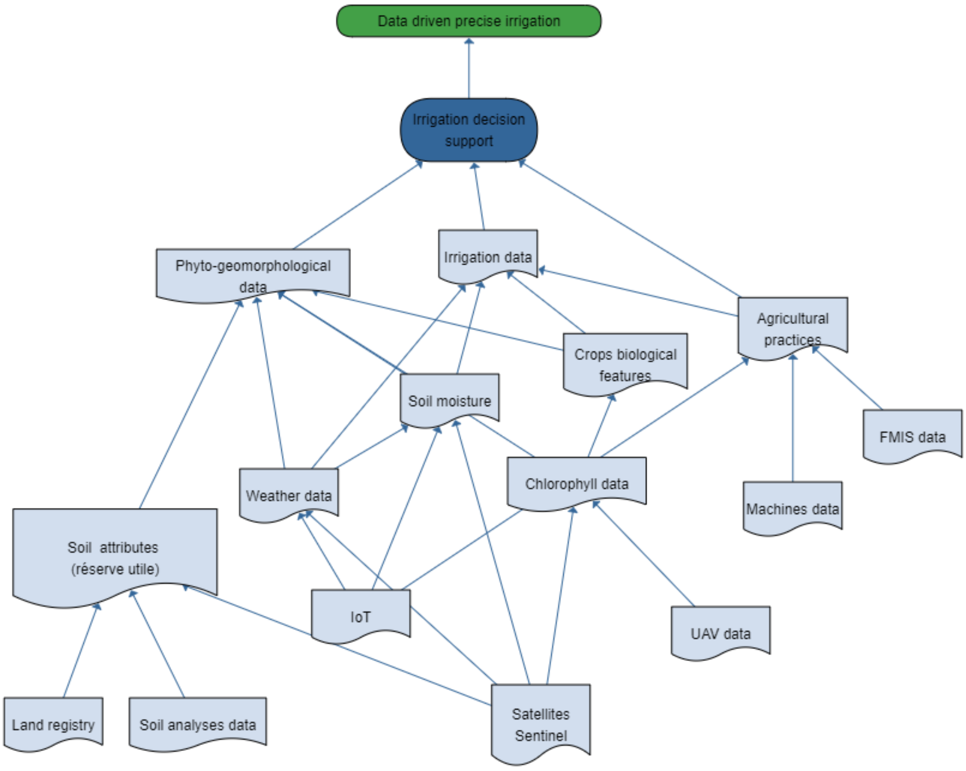

Figure 18 - Data driven precise irrigation dataflow

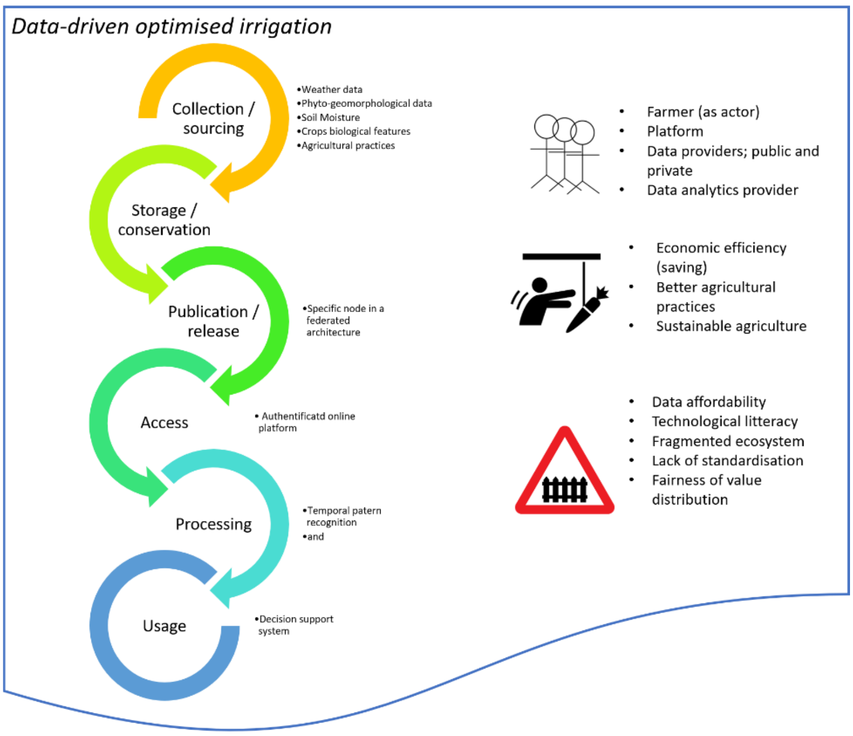

Figure 19 - Data driven optimised irrigation dataset

Architecture of the platform & data access

The platform is based on a cloud architecture and is following a classic SaaS model, to ensure a real-time data provision. It is also handling the technical side of the informed consent module by which farmers may share their data.

This ecosystem illustrates the best the potential of APIs at all the stages of the data life cycle and of the value chain, they are at the basis of the approach. In the example above detailed, they are linking different databases on the ground (soil databases), on the plant varieties and other data such as local weather data, structuring the successive data cycles and shaping technically the ecosystem, and even from the organization standpoint, as most of APIs rely on a contract. As clear rules of usage and dissemination are required by APIs, it contributes to build the ecosystem structure and make it more tangible for its participants. APIs and the platform itself require several socio-technical arrangements and tools, such as APIs keys, traffic analysis or financial clearing infrastructure. There is actually a combination of APIs, some public, other privates, and most often a combination of both. This diversity requires in turn a platform architecture orchestrating – in the IT meaning of the word – the different processes.

Several users reported they updated their own infrastructure and information systems architecture to be able to cope with the APIS, to interface themselves with them, and take benefit of the overall approach, and thus confirming the links between successful data ecosystems and digital transformation of their actors.

Data models and standards

As the data sources are heterogeneous, the metadata system needs to be at the same time versatile to manage this diversity, and accurate enough to allow a minimum interoperability between the data. That is why the platform adopted the DCAT model for its metadata. Moreover, it is providing visualization tools, including a cartography module, to enhance the data discoverability.

There is a contrasted landscape pertaining to the standardisation of data, not only between the kinds of data, but also on the sector. The field observation is that there are standardized data in agro-industry sector, but a very low culture of volunteer data sharing. Conversely, farmers have an established tradition of data sharing, without having always standardized data, which can be the root of conflicts between the participants of these two sub-ecosystems. That is why, even before its creation, the platform stakeholders fostered the design and implementation of specific data models covering the needs of the sector, i.e. GIEA & GIEA2.

This use case also illustrated the benefits of an ecosystem: it can be considered as a way to foster the implementation of interoperable solutions, benefiting in turn to all the participants and ensuring the connection to other ecosystems.

5- Geospatial data Marketplace ecosystem

The **Geospatial data marketplace** is illustrated by the UP42[72] use case. UP42 is a marketplace and developer platform providing access to both data and analytics from

[72] https://up42.com/

multiple sources. UP42 is also offering a value distribution model that is contributing to changing the way geospatial data is accessed and analysed.

Interaction, structure of the ecosystem

Summary

UP42 may be considered as an archetype of a company building a platform to give a structure to the related data ecosystem. Among the five use cases, it is the closest of the models developed by literature around hub firms, presenting thus most of the functions of an ecosystem orchestrator. It is at the centre of ecosystem and of the value chain.

In this ecosystem, the marketplace ensures diverse roles that are otherwise spread among several stakeholders. First and foremost, it is a central access point intending to be the most comprehensive possible pertaining to data, algorithms and computation facilities (see the business models section). UP42 is currently offering 33 data blocks and 62 processing blocks. To ensure its own sustainability, the marketplace is committed to be the most comprehensive possible. The platform intends to be of general purpose and to serve the earth observation needs of any customer. The platform is also providing cloud services, especially storage facilities. Moreover, it is offering all the services of a marketplace, including contractual framework or financial clearing (see business models). The platform intends to develop its exchanges with traditional data providers not directly involved in the platform, to report the issues and needs of its current partners. It is also intending to consider the contextual changes of the ecosystem, for examples actors or trends not directly involved as such in the ecosystem, but having a potential impact on it in the long-term. This is for example the case for the hardware industry, with which the marketplace manages deep connections as they may enable better or new data collection, in a more affordable way.

Other stakeholders such as data creators or aggregators are actively encouraged to put their data sources in the platform, to benefit from a number of customers expected to be higher. Same stands for companies providing analytics of the broadest range possible and looking for customers. Data-driven service providers may benefit of the central as a central access point to find data sources. They have an important role to ensure the dynamics of the ecosystem, while seeking for customers of the products they develop including at least some products or services of the platform. Even if these companies are – for now – mainly geospatial solution providers, the versatility of this kind of products is impacting indirectly a very large range of sectors, generating new possibilities to enrich the data ecosystem.

The case analysed shows a platform mainly organised around business-to-business (then B2B) upstream, but also downstream as the data-driven service providers are also using the platform for b2B purposes - without excluding targeting the general public in the future – and thus confirming that the economic sustainability and long-term growth of the ecosystem is ensured by the network effects and the interplay between complementary business models.

The platform has an essential role to ensure the ecosystem structure and membership. Through its position, it is easy to know the needs of the ecosystem participants, to get their feedback. Thus, the platform is well positioned to keep the current partners contributing, attract new ones and benefit of basic network effects, as ecosystem participants may mention third parties great to bring on board.

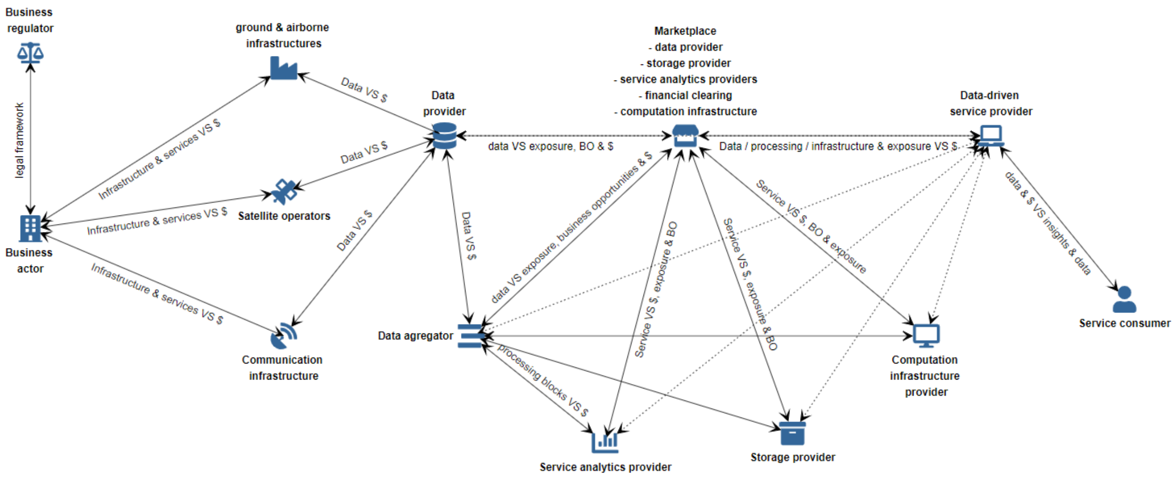

Figure 20 - UP42 marketplace ecosystem

Governance issues

This use case is showing how a marketplace may embrace the role of a platform leader in a data ecosystem. This central actor is committed to be generalist to ensure an open ground for the evolution of the ecosystem, while most of the other participants are bringing a very specialized expertise and contribute to reach other domains and other ecosystems.

UP42 is an example of a central actor contributing to structure the ecosystem. In this case, it is both a platform-centric and a marketplace model. It is largely taking in charge the orchestration functions, especially as it contributing to shape the connections between the participants and to manage the membership of the ecosystem. In particular, it has a central role to engage potential users to actively engage in the network. As noted above, other actors also have a role to include new participants and to ensure the economic sustainability through finding customers.

The platform and the marketplace are putting the ground of contractual relationships, which represent the tangible links in the data ecosystem, along with the data flows exchanged.

Value sharing is directly impacted by the governance settings. It is following a revenue-shared model, partnerships are different between actors, defined through bilateral negotiations. This allows for example to bring a greater share to the provider of a very strategic data source.

Business models

The business model of the platform is based on the concentration of different kinds of products and services (data blocks, processing blocks, including cloud services) to attract customers, most having themselves B2B models, to ensure its own economic viability, which highly depending on the vitality of its participants. Customers are buying credits, spent depending on the use of the marketplace blocks. Data markets are a competitive business, with actors quickly entering and exiting the field. The

position of the specific marketplace studied is to be at the same time the most comprehensive possible, and to provide also processing blocks and cloud services, which means the possibility of more integrated workflows.
Data-driven service providers, such as Live-EO, find in the platform a transparent business model through its contractual framework, to build their own business. Selling data is reported to be a difficult business model as such. It is more perceived as a complementary or enabling factor providing a competitive advantage for selling intermediary products or insights based on data.
The data business model supported by the platform is aiming at creating long-term effects, where small partners put their assets. The platform allows cumulative effects and is an avenue for long-term growth.
The financial clearing facilities brought by the platform may be the ground of exploratory business models, where companies may put data or services in the platform with having to invest first on the expensive customer relationships and engagement processes. It is for example the case of Exact Earth which can this way reach new markets at low cost and thus at low risk.

Sustainability of data provision and of the ecosystem at large
The interdependencies between the partners are an important driver of the sustainability of the ecosystem. The platform owner is committed to ensure the long-term growth of its partners, from which all the participants are benefiting. The ecosystem system is particularly well-fitted for the cases where there is a need to build scalable application based on a research-intensive environment.
The objective of the hub-firm is also to engage each participant in the success of the whole ecosystem, which requires to increase their level of awareness and may lead to distributed co-marketing efforts

Data literacy
The platform leader intends to enable knowledge sharing in the ecosystem through different means, such as a blog which articles may enlighten different topics of interests for the ecosystem, e.g. the explanation of some data types as well as their manipulation, innovative re-uses of data, or more strategic notes that are important for the evolution of the ecosystem. This way, this actor is at the same time demonstrating the capabilities of the overall ecosystem, aiming at attracting new participants, and empowering the current ones.
For the case analysed, UP42 is strongly committed to bring IT knowledge to the participants and it implies to actively monitor the available technologies, then to raise awareness, knowledge and skills of the participants.

Data flows and data cycles

Summary
The processes represented in the figure below are an example of the data flows enabled by the ecosystem, around the services provided by a geospatial data analytics

company, Live-EO[73], to enable the predictive maintenance of the vegetation and more generally the ground around the infrastructures of a transportation company, here the Deutsche Bahn. The data marketplace, one of the functions embraced by UP42, is enabling the fast and cost-effective access to a large of data produced by different data producers or aggregators and, for this example, focus on radar data. Currently, the data driven service provider, i.e. Live-EO, is using for a part the platform, for another external sources. From the platform in particular, it is using the available data cleaning algorithms. From its customers, in this case the Deutsche Bahn, it is gathering data related to the railroad infrastructure. Then, adding its own analytics layers, Live-EO provides a service allowing the infrastructure manager to deal more properly and in a cost-effective way of its infrastructure maintenance (pruning, ground stability monitoring…).

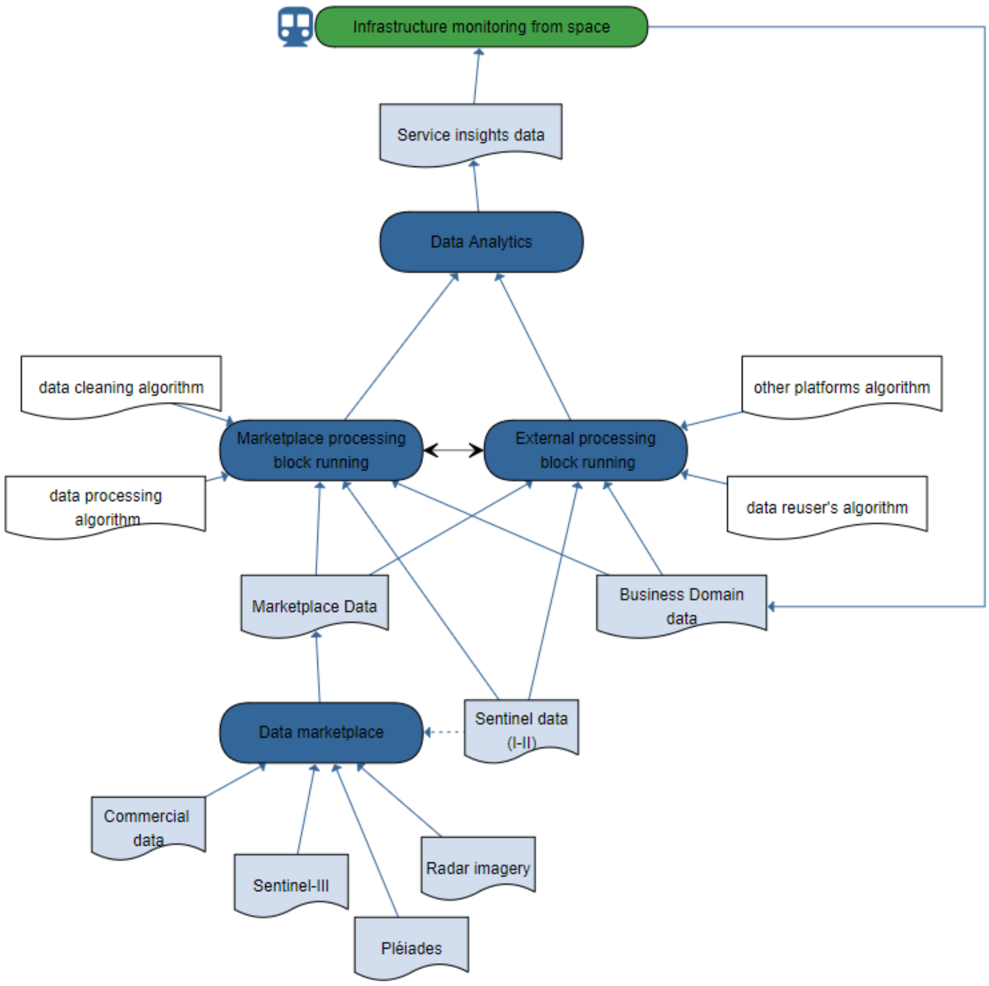

Figure 21 - LiveEO Data Flow

This demonstrates the usefulness of satellite data provided by various public (including Open Data and FAIR data) or private actors and how they may be re-used along the

[73] https://live-eo.com/

data processing and the value chains. It illustrates also the importance of complementary business models, based on the re-use of existing data (Open Data and commercial data), in combination with the implementation of state-of-the-art analytics (machine learning).
More generally, based on the variety of data included in the platform, the platform is including different purposes algorithms, some being focused on something very specific, such as data quality improvement in the frame of a larger data processing pipeline, which is allowed by the architecture of the platform, others being closer to direct customer's usage. Based on the platform architecture also, customers are also empowered to bring their own data and algorithms on the platform. The customizability of using the platform enables them in turn to deliver tailored products and services to their own customers.

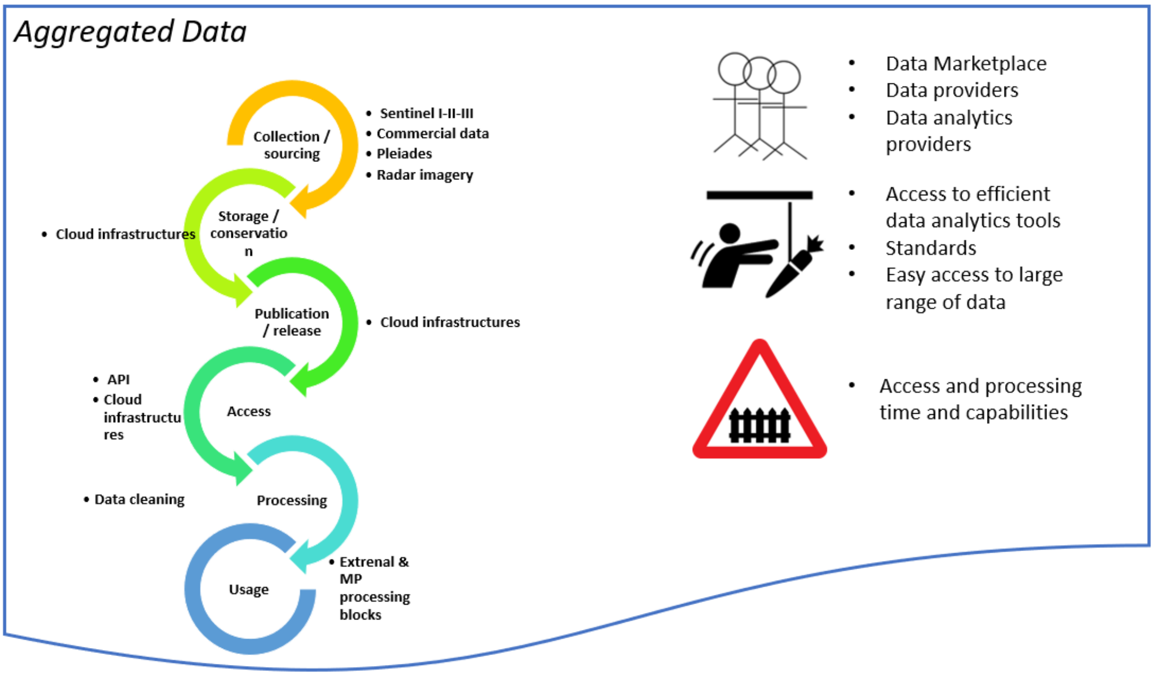

Figure 22 - UP42 Aggregated Dataset

This data flow is also an illustration of the data cycles, leading to the creation of new data sets, enabled by healthy data ecosystems. Based on the needs of its customer, its own assets and the platform environment, Live-EO is intending the creation of an improved dataset put on UP42 platform, making possible to access the geospatial data, which are difficult to handle given their size, at the convenient time and cost.
In this case, existing data are processed to be re-used more easily throughout the ecosystem. In other cases, tailored data may also be created for the needs of the ecosystem and provided to ecosystem customers. Alternatively, based on stakeholders' agreements, data – or at least a subset of them - provided by the end-customers might also be provided as new inputs for the ecosystem.
More generally, this use case shows few commitments towards Open Data, data sharing being based on business needs.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-30626-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office of the European Union

doi:10.2760/04462
ISBN 978-92-76-31385-4