Ref. Ares(2022)7915906 - 16/11/2022

e-shape

EuroGEO Showcases: Applications Powered by Europe

Diffusion of the validated model (publications) (D2.8)

ABSTRACT

This document details the actions implemented by WP2 to disseminate the co-design method it has developed in the framework of the European e-shape project.

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein.

To avoid auto-plagiarism risk, this document dissemination is limited to e-shape's community

DOCUMENT TYPE	Deliverable
DOCUMENT NAME:	e-shape-WP2-D2.8 Diffusion of the validated model (publications) (Draft)
VERSION:	Vfinal
DATE:	Oct 24th, 2022
STATUS:	
DISSEMINATION LEVEL:	PU

AUTHORS, REVIEWERS				
AUTHOR(S):	Raphaëlle Barbier, Skander Ben Yahia			
AFFILIATION(S):	ARMINES			
FURTHER AUTHORS:	Pascal Le Masson, Benoit Weil			
PEER REVIEWERS:	Lionel Menard			
REVIEW APPROVAL:	PMT			
REMARKS / IMPROVEMENTS:				

VERSION HISTORY (PRELIMINARY)			
VERSION:	DATE:	COMMENTS, CHANGES, STATUS:	PERSON(S) / ORGANISATION SHORT NAME:
V0.1	10/10/2022	1st draft version	WP2
V1.0	24/10/2022	1st draft version reviewed	PMT
vfinal	24/10/2022	Final version for submission to the EC	PMT

VERSION NUMBERING	
v0.x	draft before peer-review approval
v1.x	After the first review
v2.x	After the second review
Vfinal	Deliverable ready to be submitted

STATUS / DISSEMINATION LEVEL			
STATUS		DISSEMINATION LEVEL	
S0	Approved/Released/Ready to be submitted	**PU**	Public
S1	Reviewed	**CO**	Confidential, restricted under conditions set out in the Grant Agreement
S2	Pending for review		
S3	Draft for comments	**CI**	Classified, information as referred to in Commission Decision 2001/844/EC.
S4	Under preparation		

TABLE OF CONTENTS

ABSTRACT **2**

TABLE OF CONTENTS **5**

LIST OF FIGURES AND TABLES **5**

1 INTRODUCTION **6**

2 INTERNAL DISSEMINATION **6**

2.1 TRAINING OF PILOTS: CO-DESIGN DIAGNOSIS AND CO-DESIGN ACTIONS 6
2.2 WP2 SUPPORT ON THE SHOWCASE SUPPORT SERVICE 8

3 EXTERNAL DISSEMINATION **8**

3.1 DELIVERABLES 9
3.2 WP2 AND WP6 JOINT ACTIONS 10
3.3 ACADEMIC PAPERS 10
3.4 ACADEMIC PAPERS UNDER REVIEW 11
3.5 ACADEMIC CONFERENCES 11
3.6 OTHER CONFERENCES 13

4 INSIGHTS ON THE FUTURE OF EO CO-DESIGN ACTIVITY **12**

5 CONCLUSION **13**

6 ANNEX 1 – ACADEMIC PAPER COMPILATION AND REFERENCES **16**

LIST OF FIGURES AND TABLES

Figures:

Figure 1 Illustration of 3 slides from WP2's introduction to co-design deck section **6**

Figure 2 Representation of the "data journey" for the targeted state **7**

Figure 3 Illustrations of a self-diagnosis tool completed by an e-shape pilot **7**

Figure 4 illustration of the brochure's co-design section **10**

Figure 5 Illustrations of Raphaëlle Barbier's (WP2 leader) participation to various conferences **12**

1 INTRODUCTION

In the framework of WP2 activities within e-shape, different actions of dissemination of the method and good practices have been deployed. These actions have been realized through co-design activities with the pilots but also through more specific actions with other work packages such as WP6, in charge of the communication of e-shape and WP4, responsible for user uptake, capacity building and liaison activities.

2 INTERNAL DISSEMINATION

2.1 Training of pilots: co-design diagnosis and co-design actions

WP2 has conducted 32 co-design diagnosis where an awareness-raising work on co-design has been carried out. All diagnosis meetings started with an introduction to co-design in the Earth observation (EO) sector and an explanation of the specificities of the methodology developed by WP2.

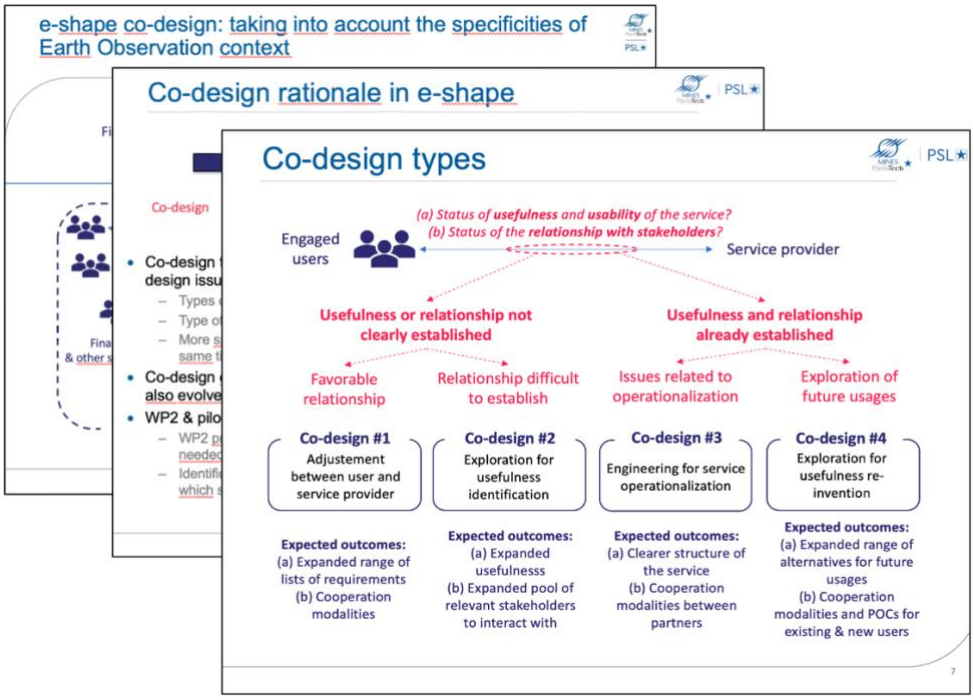

Figure 1 Illustration of 3 slides from WP2's introduction to co-design deck section

WP2 has also developed a **self-diagnosis tool for co-design needs**. The objective of this toolkit is to **allow the pilots to autonomously identify their co-design needs**. The tool takes the form of an Excel file containing 3 sheets:

1. **"Instructions"** presents all the necessary instructions to successfully conduct this work. It is composed by two sections: *co-design in e-shape* and *how to use*.
2. **"Tables"** provides the two analysis tables and additional information which guides the pilots through the diagnosis. The first table is divided in 5 parts, that correspond to the red frames represented on Figure 2 (Users communities, User competencies, Service developed by the pilot, Pilot-user relationship, and

Ability of the pilot to provide the required service (prototype/operational)). Each part is made up of different questions that need to be answered. The second table allows users to identify the type(s) of co-design that will be relevant for their pilot thanks to the answers given their answers to the different questions, compiled in the main table.

3. **"Example"** illustrates an example of diagnosis to indicate to the pilot the expected type of answer to the questions asked in the analysis tables.

This document was tested in e-shape with the on-boarded pilots. The self-diagnosis tool proved to be useful to initiate the diagnosis process, however a telco was systematically conducted afterwards to finalize it.

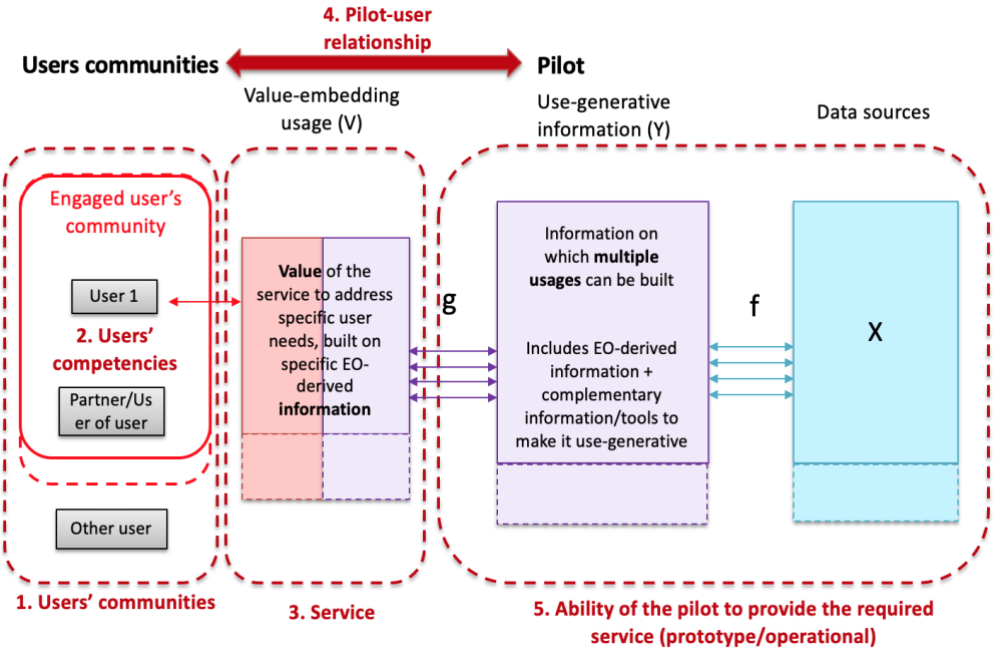

Figure 2 Representation of the "data journey" for the targeted state

Figure 3 Illustrations of a self-diagnosis tool completed by an e-shape pilot

You can download the self-diagnosis tool for co-design needs following this link: https://e-shape.eu/images/co-design/Initialassessment_questionnaire.xlsx

WP2 has also conducted five co-design actions with six different pilots. A co-design action cycle has 3 phases:

1. Preparatory meeting(s)
2. Workshop
3. Debriefing meeting(s)

A co-design action might have more than one co-design cycle, depending on the pilot's needs and constraints *(for more details regarding co-design protocols please refer to D2.6. deliverable)*.

Pilots that asked for WP2's help on co-design action have benefitted from **individual support and training** on a specific co-design type action. The co-design actions are the following:

1. **Co-design type 1** with S2-P3 pilot (*Health Surveillance Air Quality* within the *Health Surveillance Showcase*)
2. **Co-design type 2** with S3-P3 pilot (*Offshore wind resources* within the *Renewable Energy showcase*)
3. **Co-design type 3** with S3-P2 pilot (*High PV penetration at urban scale*)
4. **Co-design type 3** with S4-P2 (mySITE) and S4-P3 (myVARIABLE) pilots
5. **Co-design type 4** with S5-P4 (*Sargassum detection for seasonal planning*)

After WP2's support on their first co-design cycle, S3-P2 and S3-P3 have successfully conducted co-design workshops in autonomy.

2.2 WP2 support on the showcase support service

WP2 has also provided **continuous support to all e-shapes pilots** through the showcase support service (SSS). The ticketing system allowed pilots to ask for support on co-design matters as well as to keep WP2 on track regarding their activity and their achievements. The showcase support service has allowed WP2 to share its **co-design guidebooks** with the pilots. These user guides are based on the material developed by WP2 throughout the experimentations carried out with the five pilots aforementioned. The guidebooks are composed of **template and example slide decks** designed to help the pilots to autonomously conduct their co-design activities. Each co-design type has its own guidebook which is composed of two files:

1. "**Templates**" contains slide decks for the organization of co-design workshops and for the formalization of workshop outcomes.
2. "**Example**" holds example slide decks showcasing previous successful workshops and formalizations of outcomes.

3 EXTERNAL DISSEMINATION

WP2's external dissemination actions can be divided in three means of action: involvement in WP6's communication actions and WP4's best practice dissemination deliverable, publication of academic papers and participation to various conferences in the EO field.

3.1 Deliverables

At the end of e-shape, WP2 would have written and published **9 deliverables**, starting by a first draft of EO-specific co-design process (D2.1) and finishing by the diffusion of the validated model (D2.9). These deliverables catch the whole path followed by WP2 (theoretical background, first developments, experiments, etc.) allowing any person outside e-shape to understand the rationale and the outcomes of this research work.

WP2 was also involved in WP4's **D4.4 et D4.12** on **capacity building best practices guide**. It has shared lessons learned during e-shape (diagnosis phase, workshops, tool development, etc.), co-design's added value on capacity building and best practices for future EO co-design experiments and developments.

Deliverables:

- **D2.1 Co-design for e-shape (initial model)**: First draft of "EuroGEO-specific" co-design process. It describes the model of a co-design process adapted to e-shape's specificities based on a framework representing the development of services from Earth Observation as a relationship between data, information and usages.
- **D2.2 Co-design for e-shape (revised model)**: Updated co-design model adapted to e-shape:
    - The diagnosis of co-design needs, now includes the analysis of the adequacy between the "design environment" provided by the pilot to users and the characterization of the related users.
    - First insights on a classification of co-design actions can be outlined.
- **D2.3 Report on the experiments and feedback**: Document presenting the outcomes of the experimentations on the "diagnosis process" for all e-shape pilots. Based on these results, the analysis grid of the "diagnosis process" has been refined, especially with the classification of co-design needs in four co-design types, and a new understanding of co-design objective seems to emerge.
- **D2.4 Validated model of co-design process for e-shape (Draft)**: Updated version of the model focusing on:
    - Co-design involves the implementation of a dynamic process of specific types of co-design actions, to unlock the different blocking points occurring in the development of EO-based services over time.
    - Each co-design action aims at creating a 'resilient fit' between stakeholders.
- **D2.5 Report on the cases requiring specific co-design update (Draft)**: A complement to D2.4 deliverable compiling the outcomes of the co-design actions experimented with e-shape pilots. It especially highlights the outcomes following 'organizational' KPI and 'cognitive' KPI, and presents the frameworks proposed by WP2 to synthesize these outcomes.
- **D2.6 Validated model of co-design process for e-shape**: This deliverable aims at presenting the updated version of the co-design framework, based on latest advances of the work-package. The following results are more specifically stressed:
    - Co-design involves the implementation of a dynamic process of specific types of co-design actions, to unlock the different blocking points occurring in the development of EO-based services over time.
    - Each codesign action aims at creating a 'resilient fit' between stakeholders.
- **D2.7 Report on the cases requiring specific co-design update**: Update of D2.5 deliverable compiling the outcomes of the co-design actions experimented with e-shape pilots.

- **D2.8 Diffusion of the validated model (publications) (Draft)**: Internal and external dissemination work and insights on the future of EO co-design management.
- **D2.9 Diffusion of the validated model (publications)**: *To be written* – Update of D2.8

All available deliverables are listed here: https://e-shape.eu/index.php/resources

3.2 WP2 and WP6 joint actions

In the framework of WP6 activities, WP2 has been involved in the development of communication materials by preparing the content for the co-design sections within e-shape's website[1] and brochure. The brochure describes the co-design approach in a very short format, whereas the website allows the reader to have more details on the developed tools and guidelines.

e-shape works with and for users to maximize Earth Observation enabled benefits

Co-Design

In e-shape, a co-design model considering EO specificities is progressively designed and tested with e-shape pilots, through a dedicated work-package (WP2). A first analytical framework has been built, especially highlighting that a co-design model adapted to EO context should involve two distinct phases: (1) a critical "diagnosis process" to identify the co-design needs, classified in four main types of co-design, (2) the implementation of co-design actions to address these co-design needs.

Highlight #1: A diagnosis process to help the pilots to better structure their co-design strategy identifying relevant forms of co-design actions at different time horizons.
Based on the analysis of e-shape pilots, a certain variety of co-design needs could be identified, leading us to define four main types of co-design.

Highlight #2: Rigorous protocols to conduct co-design actions ensuring the growth of the EO ecosystem in a resilient perspective.
Co-design should be considered as a way of growing an ecosystem of efficient EO-based service designers. It is based on a continuous process involving four types of actions aiming at unlocking blocking points occurring in the development of the services.

Highlight #3: Sharing our methodology with the Earth Observation community
WP2 aims at sharing its latest achievements and discoveries beyond e-shape, with the wider EO community.

Highlight #4: Advancing academic research in design and innovation management
The development of the co-design method in e-shape also brings significant contributions to research in design and innovation management, especially investigating the specificities of a data-based regime of design, that would ensure the growth of a data-based ecosystem supporting multiple actors of various sectors in tackling grand challenges.

Figure 4 illustration of the brochure's co-design section

3.3 Academic publications and communications

The works carried out within e-shape also provide significant insights for research in innovation management and design. The following sections give the references of the different research works that were written for publication in academic journals or for communication in academic conferences.

[1] *e-shape / Co-design*. https://e-shape.eu/index.php/co-design

3.3.1 Published academic papers

1. Raphaëlle Barbier, Skander Ben Yahia, Pascal Le Masson, Benoit Weil, "**Expanding Usages Of Earth Observation Data: A Co-design approach to grow an ecosystem of efficient service designers**"; in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021, pp. 296-299, doi: 10.1109/IGARSS47720.2021.9553914. [https://hal.archives-ouvertes.fr/hal-03356299]
2. Raphaëlle Barbier, Pascal Le Masson, Benoit Weil, "**Transforming Data Into Added-value Information: The Design Of Scientific Measurement Models Through The Lens Of Design Theory**"; in *Proceedings of the Design Society: International Conference on Engineering Design*, Cambridge University Press, 2021, 1, pp.3239-3248, doi: 10.1017/pds.2021.585 [https://hal.archives-ouvertes.fr/hal-03356306]
3. Raphaëlle Barbier, Skander Ben Yahia, Pascal Le Masson, Benoit Weil, "**Co-design for novelty anchoring into multiple socio-technical systems in transitions: the case of Earth observation data**"; in IEEE Transactions on Engineering Management, 2022, doi: 10.1109/TEM.2022.3184248. [https://hal.archives-ouvertes.fr/hal-03772981]

3.3.2 Academic papers under review

4. Raphaëlle Barbier, Skander Ben Yahia, Pascal Le Masson, Benoit Weil, "**Data-push innovation beyond serendipity: the case of a digital platform strategically building up the genericity of Earth observation data**"; submitted to Creativity and Innovation Management *(under review)*
5. Raphaëlle Barbier, Skander Ben Yahia, Pascal Le Masson, Benoit Weil, "'**Resilient-fit' co-design methods: designing the integration of Earth observation data into ecosystems facing grand challenges**"; submitted to Technovation (*under review)*

3.3.3 Academic conferences

6. Raphaëlle Barbier, Benoit Weil, Pascal Le Masson, "**Creating value from data in an ecosystem: building and expanding relationships between data and seemingly distant usages**"; *R&D Management 2019*, Jun 2019, Palaiseau, France. [https://hal-mines-paristech.archives-ouvertes.fr/hal-02168086]
7. Raphaëlle Barbier, Pascal Le Masson, Sylvain Lenfle, Benoit Weil, "**Building the generativity of data to support the dynamics of multiple ecosystems: the case of Earth-observation data**", *R&D Management Conference 2021*, Jul 2021, Glasgow, United Kingdom [https://hal.archives-ouvertes.fr/hal-03356310]

3.4 Other conferences for the EO sector

WP2's **communication on co-design beyond e-shape** was supported by its participation to three conferences and webinars:

1. **e-shape General Assembly** (February 1st, 2022)
    a. Overview of the co-design approach built in e-shape
    b. Pilots' testimonies
    c. Further research perspectives beyond e-shape: trans-ecosystem co-design to support the growth of multiple socio-economic ecosystems in transitions

2. **Copernicus Horizon 2035**[2] (February 17th, 2022)
    a. Co-design to boost user acceptance of Earth observation data: a resiliency-driven co-design
    b. Presentation of e-shape
    c. Main results on co-design
    d. Pilots' testimonies
    e. Discussion on future perspectives beyond e-shape
3. **NextGEOSS Webinar – TRANSFORMING EARTH OBSERVATIONS TO KNOWLEDGE THROUGH INCLUSIVE PARTICIPATION**[3] (February 22nd, 2022)
    a. Building a co-design methodology supporting the resilient growth of the Earth observation ecosystem
    b. Introduction to co-design in e-shape
    c. Co-design experimentation results
    d. Research and operational perspectives beyond e-shape

Figure 5 Illustrations of Raphaëlle Barbier's (WP2 leader) participation to various conferences

4 INSIGHTS ON THE FUTURE OF EO CO-DESIGN ACTIVITY

Thanks to the work done in the framework of e-shape, it appears that the future development of co-design activities in the EO sector would highly benefit from additional investigations and dedicated efforts. In particular, WP2 has identified the three following paths to further support the diffusion of co-design in the EO sector:

- Guidebooks (diagnostic tool & guidelines for workshops)
- Developing co-design as-a-service (e.g., training of consultancy companies)

[2] *EU Space Conference – Copernicus Horizons 2035*. https://www.euspace-conference2022.eu/fr/

[3] *Transforming earth observations to knowledge through inclusive participation – NextGEOSS.* https://nextgeoss.eu/nextgeoss-webinar-feb22-co-design

- Establishing co-design as a critical component of EuroGEO/GEO

The **development of reference guidebooks to be shared with the overall EO community** is an important goal to achieve. It will allow to reach a certain level of standardization of the method, being an important lever for dissemination. The first feedback on these tools is positive. However, they still need to be improved to be deployed on a larger scale, including:

- Strengthening the self-diagnostic tool, making it more self-supporting
- Improving the templates for co-design action support
- Enriching guidebooks with more examples

Another interesting path would consist in **developing a dedicated co-design service or module** that would help EO projects or institutions in need to set up co-design activities for their respective situations. The role of providing such co-design service could be taken up by different form of actors. For example, some consulting firms could develop a co-design as-a-service offer to support the EO community with this task, as they are used to operationalize complex methods and making them easily deployable. A wider network of "co-design experts" could also be progressively trained. **Such initiatives are part of the "Sustainability Booster" concept, introduced by WP5.**

Finally, the EuroGEO/GEO network could stand as an important vector to raise awareness of co-design within the EO community. Establishing co-design as a critical component of the organizations would boost the co-design uptake within the EO community. Several actions could be conducted:

- Diffusion of best practices
- Setting-up co-design training for the EO community (Sustainability booster concept)
- Ensuring co-design quality (labelling system)
- Funding future research on co-design advances

5 CONCLUSION

This deliverable has presented WP2's dissemination efforts in e-shape, especially highlighting the overall ambition to work towards the diffusion of a robust and easy-to-use co-design method specifically addressing the issues faced by the EO community. WP2 has used all its channels (internal and external) to communicate on the method and reach a variety of communities in the EO field and beyond. This document also shares WP2's vision on the next steps that will allow the development and further diffusion of a routinized method helping the EO community to grow.

6 ANNEX 1 – ACADEMIC PAPER COMPILATION AND REFERENCES

Annex table

I. Raphaëlle Barbier, Skander Ben Yahia, Pascal Le Masson, Benoit Weil, "**Expanding Usages Of Earth Observation Data: A Co-design approach to grow an ecosystem of efficient service designers**"; in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021, pp. 296-299, doi: 10.1109/IGARSS47720.2021.9553914. [https://hal.archives-ouvertes.fr/hal-03356299]

II. Raphaëlle Barbier, Pascal Le Masson, Benoit Weil, "**Transforming Data Into Added-value Information: The Design Of Scientific Measurement Models Through The Lens Of Design Theory**"; in *Proceedings of the Design Society: International Conference on Engineering Design*, Cambridge University Press, 2021, 1, pp.3239-3248, doi: 10.1017/pds.2021.585 [https://hal.archives-ouvertes.fr/hal-03356306]

III. Raphaëlle Barbier, Skander Ben Yahia, Pascal Le Masson, Benoit Weil, "**Co-design for novelty anchoring into multiple socio-technical systems in transitions: the case of Earth observation data**"; in IEEE Transactions on Engineering Management, 2022, doi: 10.1109/TEM.2022.3184248. [https://hal.archives-ouvertes.fr/hal-03772981]

IV. Raphaëlle Barbier, Pascal Le Masson, Sylvain Lenfle, Benoit Weil, "**Building the generativity of data to support the dynamics of multiple ecosystems: the case of Earth-observation data**", *R&D Management Conference 2021*, Jul 2021, Glasgow, United Kingdom [https://hal.archives-ouvertes.fr/hal-03356310]

V. Raphaëlle Barbier, Benoit Weil, Pascal Le Masson, "**Creating value from data in an ecosystem: building and expanding relationships between data and seemingly distant usages**"; R&D Management 2019 [https://hal-mines-paristech.archives-ouvertes.fr/hal-02168086]

VI. Raphaëlle Barbier, Skander Ben Yahia, Sylvain Lenfle, Benoit Weil, "**Data-push innovation beyond serendipity: the case of a digital platform strategically building up the genericity of Earth observation data**"; submitted to Technovation *(under review)*

VII. Raphaëlle Barbier, Skander Ben Yahia, Pascal Le Masson, Benoit Weil, "'**Resilient-fit' co-design methods: designing the integration of Earth observation data into ecosystems facing grand challenges**"; submitted to Creativity and Innovation Management *(under review)*

ACADEMIC PAPERS

EXPANDING USAGES OF EARTH OBSERVATION DATA: A CO-DESIGN APPROACH TO GROW AN ECOSYSTEM OF EFFICIENT SERVICE DESIGNERS

Raphaëlle Barbier[(1)], Skander Ben Yahia[(1)], Pascal Le Masson[(1)], Benoit Weil[(1)]

[(1)] Center for Management Science - i3 UMR CNRS 9217 - MINES ParisTech, PSL University / ARMINES

ABSTRACT

Earth Observation data has the potential to provide significant benefits to a large variety of socio-economic stakeholders. However, creating new usages of these data is particularly challenging as it requires connecting distant data and usages ecosystems. 'Co-designing' services based on Earth Observation data appears to be a promising path to overcome insufficiencies of 'open-data' strategies. However, in this challenging context, 'co-design' cannot be limited to the mere adjustment between user demands and data supply. Based on design theory, we propose a comprehensive framework for such a 'co-design' approach, aiming at growing an ecosystem of efficient service designers. It is experimented in the *e-shape* project. First results show that: (1) such co-design involves the implementation of a dynamic process of specific types of co-design actions, to unlock the different blocking points occurring in the growth of the ecosystem over time, (2) each co-design action aims at creating a 'resilient fit' between stakeholders.

Index Terms— co-design, Earth Observation, value creation from data, open-data, data-based ecosystems

1. INTRODUCTION

Earth Observation (EO) refers to the production of information about the planet and its environment, based on different types of instruments (satellites, in-situ sensors etc). Initially produced mainly for scientific goals, EO data are now made available to every economic actor, through 'open-data' policies. Socio-economic applications of this data seem to be diverse and promising, however, in practice, developing usages from EO data seems to be particularly challenging. Indeed, this effort could be schematically described as connecting very distant socio-economic ecosystems: the ecosystem of data and the various ecosystems of potential usages. These ecosystems are called "distant" as they do not share the same dynamics, time horizons (e.g. very long cycles to develop new instruments compared to short timeline of actions in the data usage context), performance logics and competencies (e.g. data processing might require very specific technical expertise while data usages might also require specific domain expertise).

In order to connect distant data and usages ecosystems, several approaches have been promoted and implemented by the EO community in the last decades. The first one consists in having each ecosystem bridging independently half the distance, through an 'open-data' strategy [1]: on the one side, data are made available to everyone, on the other side, the different stakeholders take advantage of these resources by integrating them in their own usage context. Despite being necessary to broaden the usages of data, this approach has proved to be insufficient, as the stakeholders tend to have difficulty making use of EO data spontaneously.

This accounts for the current efforts of the EO community to operate a second approach that consists in connecting the distant ecosystems of data and usages by encouraging the development of operational services based on Earth Observation data, through specific 'co-design' activities. An important stream of literature documents the implementation of such an approach in the case of climate services (based on climate-related data, being a certain type of EO data). Co-design (also referred as 'co-production' or 'co-development' depending on the authors) mostly relates to the involvement of data users in order to adjust user demands and the supply of useful information [2]. Without appropriate processes, this might lead to ad hoc small-scale and short-lived data-based services. However, despite being implemented in several projects through dedicated processes, recent research also underlines that what is understood by 'co-design' is not systematically discussed and formalized [3].

Based on the research work carried out in e-shape, a project funded within the EU Horizon 2020 program [4], this paper aims at proposing an analytical framework for co-design in the Earth Observation context, clarifying co-design ambitions and the operational tools that could effectively contribute to the expansion of EO data usages.

2. CO-DESIGN SEEN FROM THE DESIGN AND MANAGEMENT FIELD

Literature in design and in management gives interesting insights on different approaches of co-design and its evolution over time[5]. Co-design reported in the Earth observation field seems to mainly corresponds to a first approach consisting in building specific interactions between users and service designers in order to fit the developed service to user needs. This approach of co-design, as supply and demand adjustment, has largely developed since the years 2000s [6]. However, it is interesting to notice that co-design were previously used in completely different situations, aiming at addressing other blocking points of the development of products or services concerning actors other than the user:

- First, in the 70s, for the development of embedded systems [7]: co-design referred to hardware and software integration, as the issue was to make different fields of expertise cooperate, a list of requirement being already defined
- Later in the 90s, co-design referred to reshaping collaborations between buyers and suppliers, beyond usual price negotiation, to design new required components (e.g. in the automotive industry, new modules to increase comfort and reduce pollutant emission of cars) [8].

These elements lead us to make the following proposition regarding co-design in the Earth Observation context: *co-design objective could be described as growing an ecosystem of efficient service designers by unlocking the different blocking points in the development of EO-based services, going beyond adjusting supply and demand between data users and service designers.* This paper proposes to test this proposition based on e-shape experience, and to address the two resulting questions:

1. *How to describe the blocking points occurring in the development of EO-based services?*
2. *What types of tools would be needed to unlock these blocking points?*

3. METHODOLOGY

Our research relies on the work carried out within the EU-funded e-shape project, bringing together a team of 54 experienced partners from academia, industry, institutional entities and user communities to develop 27 pilot applications, spanning 7 thematic areas (food security, health, renewable energy, biodiversity, water resources, disaster resilience and climate). Given the large number of partners, and the variety of application sectors, e-shape appears to be a particularly favorable context to have a comprehensive understanding of the issues faced in the EO field. Within a dedicated work package led by the authors of the paper, a co-design approach is being progressively designed and experimented with e-shape partners, based on recent advances in design theory [9]. The following process has been set up to assess co-design needs for each pilot:

- Questionnaire sent to each pilot
- Answers used to classify the different blocking points faced by the pilots (also called "co-design needs") and make a first diagnosis for each pilot
- Interview of one hour and a half with each pilot to validate the diagnosis of co-design needs

4. RESULTS

The outcomes of this analysis process are summarized in Table 1 for 22 pilots (anonymized), the analysis of the 5 remaining pilots being still to be validated.

Pilot	Short-term	Long-term
#1	Type 1	Type 3
#2	Type 1 with User 1 Type 2/4	Type 1 with Users 2&3 Type 3 with partner
#3	Type 1 & 2	Type 4
#4	Type 1	Type 4
#5	Type 1	Type 3 & 4
#6	Type 1	Type 2 for global scale
#7	Type 1/3/4	Type 4
#8	Type 2	Type 3
#9	Type 1	Type 4
#10	Type 1 & 4	Type 4
#11	Type 1 & Type 3	Type 4
#12	Type 2	Type 1
#13	Type 1 & Type 4	Type 4
#14	Type 1	Type 2
#15	Type 1 / type 2	Type 3/4
#16	Type 1 & 3	Type 4
#17	Type 1 & 4	Type 4
#18	Type 1	Type 3 & 4
#19	Type 3	Type 4
#20	Type 1	Type 4 / Type 3
#21	Type 1	Type 4
#22	Type 1	Type 4

Table 1: Analysis outcomes of e-shape pilots

4.1. Co-design going beyond supply-demand adjustment

A first noticeable result is the validation of the proposition according to which co-design needs are not restricted to adjusting supply and demand between data users and service designers. Indeed, this perspective only corresponds to a first type of co-design needs (referred as *Type 1* in the Table 1), whereas three other types of co-design needs have also been identified. These four types of co-design needs correspond to four types of actions, each type corresponding to a certain blocking point requiring the design of the relationship with a specific actor (cf Table 2).
For each type, the initial state, blocking point to be addressed and expected outcomes always include two dimensions:

1. A dimension related to the design of the service, described with two terms *usefulness* and *usability*, as commonly used in literature on climate services[10]. Usefulness refers to the general potential seen by users, whereas usability refers to the effective integration in users' operations.

Literature on climate services highlights that both aspects need to be addressed to successfully develop services, and that specific efforts are especially needed to move from useful to usable information, i.e. narrowing the so-called “usability gap”.

2. A dimension related to the design of a specific relationship. This second dimension is crucial as the development of a service cannot be done only through collective work phases but also requires separate work phases. Agreeing on cooperation modalities is therefore crucial to guarantee the continuation of alternate collective and separate work phases over time. In a way, ‘co-design’ has to put a strong emphasis on designing the ‘co’, and not only the service itself.

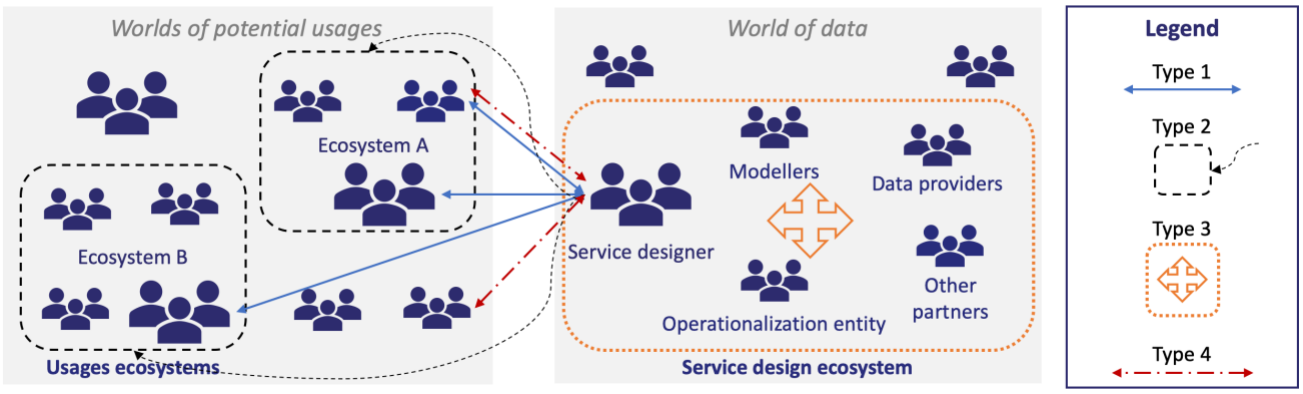

Figure 6: Schema of the four types of co-design actions aiming at connecting the distant data & usages ecosystems

	Overall context	Initial state	Blocking point to be addressed	Expected outcomes	"Quick-fit" actions	"Resilient-fit" actions
Type 1	**Adjustment between user and service designer**	(a) Usefulness already identified on a first basis but to be enhanced. Usability to be enhanced. (b) Relationship with the user to be precisely defined but at least seems favorable (user willing to devote time settling it).	Establishing **adapted relationships with specific users** for usefulness & usability assessment and enhancing	(a) Enhanced lists of requirements ensuring usefulness and usability (b) Cooperation modalities with these specific users clearly clearly formalized	Finding ONE satisfying list of requirements with one specific user	In order to end up with a robust list of requirements, exploring a range of potential lists of requirements and related cooperation modalities allows a better adaptation to surprises or external constraints
Type 2	**Exploration for usage initiation**	(a) Usefulness not well-known and/or (b) Relationship with the user appearing to be difficult to establish (uncommitted users)	Establishing **adapted interactions with user communities** for usefulness identification	(a) Expanded usefulness of the service (b) Expanded list of relevant stakeholders to interact with	Finding a new relevant user to interact with	Building relationships with a portfolio of relevant actors with a minimal usefulness established
Type 3	**Engineering for operationalization**	(a) Lists of requirements for usefulness and usability established. (b)Relationships with users established.	**Establishing adapted relationships with relevant partners** for extensive usefulness & usability realization and operationalization of the service	(a) Clarification of the service structure (parts ready to be operationalized, parts needing further exploration) (b) Cooperation modalities between R&D and operationalization entities clearly formalized	Building the technical infrastructure for an existing user	Structuring the service offer for a range of users and building all required resources through adapted partnerships
Type 4	**Exploration for usage expansion**	Usefulness, usability and relationships already established with existing users.	Establishing **adapted relationships with existing & potential new users** for usefulness reinvention	(a) Expanded range of potential alternatives for future usages (which usefulness for which actors) (b) Cooperation modalities and supports for interactions (proofs-of-concept) defined for existing and new users	Merely asking existing users what they would dream of	Setting-up a joint programme for long-term exploration of new usages (identification of obstacles, research efforts to be made, etc)

Table 2: Characterization of each co-design type and related actions

4.2. Co-design as a dynamic interplay of four types of actions to connect distant data and usages ecosystems

A second major result is the consideration of co-design as a *dynamic interplay of these four types of actions*. It is first an *interplay* of actions because at every moment, each service designer might be confronted with several co-design needs. For example, the service designer might at the same time need a *co-design type 1* to strengthen the relationship with a certain user, but also consider a *type 2* to explore a new type of user community, and prepare for the operationalization of the service through a *type 3*. This interplay is also *dynamic* because each service designer goes through different co-design types at different moments in time, depending on its evolution and the issues faced all along. This appears in Table 1 through the integration of both short-term and long-term time horizons.

4.3. Implementation of specific tools for each type of action, in a "resilient-fit" perspective

A third important result concerns the type of tools needed to support these four types of actions. It was first noticed that e-shape partners were already experienced in making some parts of these actions on their own, however they were often faced with the issue of making these efforts of connecting data and usages sustainable over time. Consequently, it appears that specific tools are required, aiming at establishing a "resilient fit" between stakeholders, rather than a "quick fit".

"Quick-fit" actions only focus on finding one type of interaction between data and usages ecosystems (single list of requirements with one user, in a punctual relationship). Whereas, "resilient-fit" actions aim at generating a range of alternatives (regarding the lists of requirements, the stakeholders involved, the types of partnerships), allowing a better adaptation to future surprises or unexpected constraints arising later in the process. These specific tools for "resilient-fit" actions are currently under experimentation within e-shape, but illustrations of "quick-fit" and "resilient-fit" actions are already given for each type of co-design in Table 2.

5. CONCLUSION

This paper introduces an analytical framework for co-design adapted to the Earth Observation context, that has been designed and experimented within e-shape. The overall ambition of such a co-design approach is to progressively connect and expand the distant ecosystems of data and usages. In this perspective, we highlighted that co-design should not be restricted to the adjustment of supply and demand between users and service designers, but should be considered as a way of growing an ecosystem of efficient EO-based service designers. It is based on a continuous process involving four types of actions aiming at unlocking blocking points occurring in the development of the services. Each type corresponds to the design of a certain committed form of relationship and should target a "resilient fit" between relevant stakeholders, in order to make sure that these efforts are sustainable over time. This approach of co-design and the related tools will be further experimented in the coming years within e-shape.

6. ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820852.

7. REFERENCES

[1] M. T. Borzacchiello and M. Craglia, 'The impact on innovation of open access to spatial environmental information: a research strategy', *International Journal of Technology Management*, vol. 60, no. 1/2, p. 114, 2012

[2] E. C. McNie, 'Delivering Climate Services: Organizational Strategies and Approaches for Producing Useful Climate-Science Information', *Weather, Climate, and Society*, vol. 5, no. 1, pp. 14–26, Jan. 2012

[3] C. M. Goodess *et al.*, 'Advancing climate services for the European renewable energy sector through capacity building and user engagement', *Climate Services*, vol. 16, p. 100139, Dec. 2019,

[4] T. Ranchin, L. Ménard, Fichaux, Nicolas, Reboul, Mathieu, and all e-shape partners, 'e-shape – EuroGEO Showcases: application powered by Europe contribution to EuroGEO and to the development of the EO industry', IGARSS 2021, Brussels, Belgium, Jul. 2021.

[5] L.-E. Dubois, P. Le Masson, B. Weil, and P. Cohendet, 'From organizing for innovation to innovating for organization: how co-design fosters change in organizations.', International Product Development Management Conference, Limerick, Ireland, 2014.

[6] M. Steen, 'Co-Design as a Process of Joint Inquiry and Imagination', *Design Issues*, vol. 29, no. 2, pp. 16–28, Apr. 2013

[7] W. H. Wolf, 'Hardware-software co-design of embedded systems', *Proceedings of the IEEE*, vol. 82, no. 7, pp. 967–989, Jul. 1994

[8] G. Spina, R. Verganti, and G. Zotteri, 'Factors influencing co-design adoption: drivers and internal consistency', *International Journal of Operations & Production Management*, vol. 22, no. 12, pp. 1354–1366, Jan. 2002

[9] P. Le Masson, B. Weil, and A. Hatchuel, *Design Theory: Methods and Organization for Innovation*. Springer International Publishing, 2017.

[10] M. C. Lemos, C. J. Kirchhoff, and V. Ramprasad, 'Narrowing the climate information usability gap', *Nature Climate Change*, vol. 2, no. 11, pp. 789–794, Nov. 2012

TRANSFORMING DATA INTO ADDED-VALUE INFORMATION: THE DESIGN OF SCIENTIFIC MEASUREMENT MODELS THROUGH THE LENS OF DESIGN THEORY

Authors: Barbier, Raphaëlle; Le Masson, Pascal; Weil, Benoit
Center for Management Science, i3 (UMR CNRS 9217), Mines Paris, PSL University, France
Contact: Raphaëlle Barbier (e-mail: raphaelle.barbier@minesparis.psl.eu)

ABSTRACT
Transforming data into added-value information is a recurrent issue in the context of "big data" phenomenon, as new sources of data become increasingly available. This paper proposes to offer a fresh look on how data and added-value information are linked through the design of specific models. This investigation is based on design theory, used as an analysis framework, and on a historical example in the Earth science field. It aims at unveiling the reasoning logic behind the design process of models combining data science and domain knowledge in specific ways, especially involving not only knowledge about the physical phenomena but also on the measuring instrument itself. More specifically, this paper shows how specific efforts on exploring the originality of the new instrument compared to existing ones can result in designing performant models to transform new sources of data into information. This also suggests several important competencies to be involved in the model- design process: (1) a detailed understanding of the limitations of existing models (2) the ability to explore both the originality of the new source of data compared to existing ones (3) the ability of leveraging independent data sources.

Keywords: Design theory, Big data, Design process, data science, information design

Cite this article: Barbier, R., Le Masson, P., Weil, B. (2021) 'Transforming Data into Added-Value Information: The Design of Scientific Measurement Models through the Lens of Design Theory',in *Proceedings of the International Conference on Engineering Design (ICED21)*, Gothenburg, Sweden, 16-20 August 2021. DOI:10.1017/pds.2021.585

INTRODUCTION

In the 1980s, within a European research project, three research teams were asked to participate to a kind of scientific competition, to make use of a new source of data to transform it into added-value information - an issue that seems very contemporary today with the rise of "big data" phenomenon and organisation of specific challenges to address it (Sitruk and Kazakçi, 2018). In the 80s project, the three teams had the same starting point and objective: *given a set of satellite data (new data source), propose the best possible model to predict solar radiation received on ground (already measured by in-situ sensors or other empirical methods), the performance criteria being clearly defined as minimizing the prediction error*. The difference between the teams thus lied in the way they designed their respective model. A synthetic document was published to shed light on each of the three models and compare them. It is particularly interesting as it opens the "black box" of the modelling activity: not only does it illustrate the classic opposition between so-called "data-driven" and "physics-driven" approaches, but it also illustrates how significantly better a model can be designed when building on the two previous approaches and going beyond each of them. Moreover, details of the approaches give hints on how to design a powerful model linking physical phenomena and data - and how to get a strong performance by creatively make use of all available knowledge, in particular knowledge of the physics of the instrument (the satellite) and not only knowledge of the physics of radiation throughout the Earth atmosphere. This example is fascinating insofar as (1) it addresses in the 1980s, almost 40 years ago, a really contemporary issue: making sense of data to get precise, unquestionably added-value information; (2) it addresses also a problem that is as old as science: the design of scientific measures (information) based on a new instrument (new source of data).

This paper proposes to investigate this historical example to give insights on a surprising blind spot in the successful development of data-based services, that is describing the design efforts that remain even when the type of information to be derived from data and its related value are already identified. Indeed, even when data are available (open data for instance) and when their added-value for the service is unquestionable (taking the form of valuable information), there can still be an issue in transforming available data into added-value information. This apparently small step in the chain linking data to value formally consists in using or building a model that relates data to information in a reliable way. As for model building in data science, two recent trends have emerged: on the one hand, it can be considered as a "technical" statistical black box that can be addressed by relying on the most advanced data-science algorithms (GAN, CNN..); on the other hand, more recently, scholars remind that the wealth of scientific knowledge should be leveraged in data science models (Karpatne et al., 2017; Reichstein et al., 2019). In these two trends, it is clear that models are designed, but the design process and its underlying reasoning logic that intertwines data science and domain knowledge remain allusive. Moreover, the specific role of domain knowledge related to the instrument itself is not explicitly described. Our paper thus aims at investigating the specific process of designing models to transform a new source of data into information, more specifically addressing the following question: *in which possible ways can new sources of data be leveraged in the process of designing models to transform data into added-value information?*

To investigate this question, in a first part we show how literature proposed many models relating data to added-value information but leaves a blind spot on the question of the reasoning logic behind the design process of these models, and its specific link with the new instrument. In a second part we propose to build a theoretical framework derived from design theory to analyse the design process associated to model design in the historical case mentioned above, that is then elucidated in a third part. A fourth part highlights contributions and limitations of this paper.

LITERATURE REVIEW AND RESEARCH QUESTIONS

Data-driven design: new opportunities stemming from the use of data

In recent years, the development of internet, new sensors, and computational means has dramatically increased the flow of data in almost every business, industry and research area. This phenomenon, commonly referred as “big data”, has largely been discussed in the literature, shedding light on its definition, opportunities and challenges, especially the issue of how value can be created out of this new flow of data (Gandomi and Haider, 2015; Günther

et al., 2017). Literature in design has also largely described the new opportunities arising from the use of data. (Parraguez and Maier, 2017) highlight the potential benefits of using open-data from various sources (e.g. patents, publications, business registries, company websites, social networks) for the engineering design research. The variety of usages that could be made from data is often emphasized, for example through the 20 contributions of the special issue of the Journal of Mechanical Design (Kim et al., 2017) covering topics as various as discovering future design and technological opportunities thanks to patent mining techniques, modelling complex parts of the body in new manners, giving new insights on the critical functions to be included in the design of new products and services. Literature in design also more specifically reports on the beneficial use of data in the concept development phase of the design process (Escandón-Quintanilla et al., 2018; Bertoni, 2020).

The transformation of data into usages is often described through the "Data Information Knowledge Wisdom" hierarchy (Rowley, 2007) or more recently the "data-information-knowledge" chain (Abbasi et al., 2016). Despite different definitions of these terms (Zins, 2007), they describe in a similar way the different types of design efforts to be made in order to effectively turn data into usages: (1) transforming data (that is contextualized, related to the measuring settings) into information (that should be more generally understandable, and meaningful in the context of reuse) ; (2) transformation of information into knowledge, generally referring to information used in a certain context. Thus knowledge refers to a certain "usage" or "value". These two terms will be preferred in the present paper, to avoid confusions with the term "knowledge" used in C-K design theory.

In design literature, an important stream of works has focused on the way information could be transformed into usages, through the design of appropriate content management or visualisation tools (Huron et al., 2014; Dammak and Gardoni, 2018). Regarding the transformation of data into information, scholars report on several issues. (Bertoni, 2020) notices the tendency to rather resort to relatively easy-to-use data (such as text mining of social networks) rather than building new data generation methods that would bring more valuable information (such as resorting to the use of sensors giving information on the product in use) because of the higher complexity of such approaches. The design effort to be made for extracting new types of data is also reported as an issue by (Montecchi and Becattini, 2020) in the context of using data to encourage sustainable behaviours. These issues are often related to the ability of implementing complex algorithms and specific data science techniques, that are said to be sometimes poorly understood (Parraguez and Maier, 2017). So this stream of works *suggests well the significant design effort to be made to transform data into information, however it does not fully describe the design process underlying this transformation.*

Designing models to transform data into information

Other branches of literature offer insights on how information is designed, highlighting the importance of designing appropriate models. First, literature about instrumentation design and metrology reminds us that information coming from a measuring system is always designed (even for basic direct-reading instruments). Indeed, designing specific models of the measurement process is required to adequately relate the "indication" given by the instrument and the "measurement outcome", that is information that can be attributable to the object under consideration and not to other factors related to the instrument or the environment (Mari et al., 2012; Giordani and Mari, 2012; Tal, 2017). This literature also highlights that several model-design approaches might exist, emphasizing two extreme archetypal cases: considering the system as a "black-box" where the model is derived from measures done for a number of known standard states, or considering the system as a "white box" where the model is determined by representing the physical process in details (Tal, 2017). This literature distinguishes two types of models and describes how to parameterize a (given) model in certain situations - still the question of the design of the base model remains unanswered.

Second, these considerations on how models are designed are also discussed in the literature related to the use of data for scientific activities. The question on how to build good models is common in science. To give a few examples, already in the nineties, reflexive works of several disciplines around Earth science on their modelling practices were carried out, e.g. in hydrology (Beven, 1989; Barnes, 1995), or for solar irradiance estimation for which different approaches - classified as either physical or statistical - were listed and compared (Noia et al., 1993a, 1993b). More recently, (Karpatne et al., 2017) made a similar distinction between "physics-only" (or "theory-only")

models, that are built by modelling the different underlying physical processes, and "data-only" models, that are built without using scientific theories by leveraging the large amount of available data through various data science techniques. The same authors emphasize the limitations of both approaches, calling for a new "theory-guided data science" paradigm, that would consist in combining scientific knowledge and data science. This article also gives a broad overview of the different combination possibilities by categorizing them in five main types: (1) theory-guided design of data science model families, (2) theory-guided learning of data science models given a model family, (3) theory-guided refinement of data science models outputs, (4) constructing hybrid models, (5) improving theory-based models according to observational data. Hence the authors propose two main types of models and explain how to combine them. But it remains unclear, what is the design logics in "data-only" and "physics-only" and, more importantly, whether there might another type of model-design, different from the two previous ones and their combinations. This calls for more explicitly showing the underlying reasoning logic of the design process, especially the specific manner domain and data science knowledge bases are leveraged and might interact and evolve during the process. In this perspective, a quick analysis of "physics-only" and "data-only" models reveals for instance that the role played by the domain knowledge related to the new instrument (and not only on the physical phenomena) remains unclear. The present paper will therefore address the following research question: *how are new sources of data leveraged to design an appropriate model transforming these data into added-value information?* In particular, we will wonder what is the specific role played by the domain knowledge related to the instrument and how this instrument domain knowledge could help design specific model(s) transforming instrument data into added-value information.

METHOD

As underlined in the literature presented above, transforming data into information involves designing a specific model, leveraging both domain and data science knowledge. Our methodology is twofold: first, C-K design theory is used as a theoretical framework to represent this model-design process, second this framework is applied on a specific historical case study to further unveil interesting features of what makes a relevant model-design logic.

C-K theory to represent the reasoning logic of designing models

Our investigation relies on C-K theory as it sheds light on the reasoning logic underlying a design process (Hatchuel and Weil, 2003, 2009). Such a process is indeed described as the interaction and the expansion of two spaces: a space K of knowledge and a space C of concepts. The K-space gathers all the knowledge the designers activate and progressively acquire during the design process (technical knowledge, user preferences, standards and regulations, etc). The C-space is the space where new ideas, concepts are explored. The interactions between the two spaces are represented through four different operators: K→C ("disjunction" where C-space is expanded thanks to available knowledge in K-space), C→K ("conjunction" where available knowledge in K is expanded and triggered by the concept expansion in C), K→K (self-expansion of knowledge based on logic rules, e.g. proving new theorems), C→C (expansion of concepts through partitioning of concepts).

Thanks to this framework, we can represent the problem of designing a model (M) to better estimate a certain information (Y) based on new sources of data (X) as follows (see Figure 1):

(a) The starting point is the initial concept C0 *"designing a model M for a better estimation of Y through the use of X"*

(b) This calls for investigating in K-space (C→K) what are the available models to estimate Y (knowledge base on models) and how to use X in those models (knowledge base on the instrument, i.e. existing and new sources of data). Two main types of models can be considered:

The *existing physics-driven models M_physics* based on parameters describing the atmosphere (cloud properties, aerosols, etc.), where previously existing sources of data are used to estimate the parameters of the model.

The *existing data-driven models M_data* (e.g. multiple linear regression) whose parameters are statistically estimated based on known pairs of (X,Y).

(c) Based on these knowledge base on models, a subsequent operation K→C makes appear the two archetypal approaches called "physics-driven" or "data-driven", corresponding respectively to the concepts *"M built on M_physics using X to better estimate physical parameters of M_physics"* and *"M built on M_data using X to estimate the parameters of the statistical model M_data"*. Regarding the domain knowledge on the instrument, these approaches only rely on the capacity of the instrument to be integrated in existing models, i.e. the dimension of the instrument that can be expressed relatedly to the existing sources of data (referred as the knowledge base "correlation" in Figure 1).

(d) A third approach, coined "hybrid", can be generated by leveraging both knowledge of M_physics and M_data and their respective limitations, as mentioned in literature with the approach of "theory-guided data science". This concept could be formulated as *"M built by using X to overcome limitations of existing models"*. In this third approach, the role of the domain knowledge on the instrument remains unclearly described in literature. C-K design theory helps us to formulate some first insights. Indeed, it predicts that a good design process relies on the use of independent knowledge bases: thus the best approach should investigate to what extent the new source of data is "orthogonal" to the existing ones, i.e. what are the additional independent knowledge it could bring compared to the previously used data sources. The analysis of the historical case study aims at further elucidating these elements.

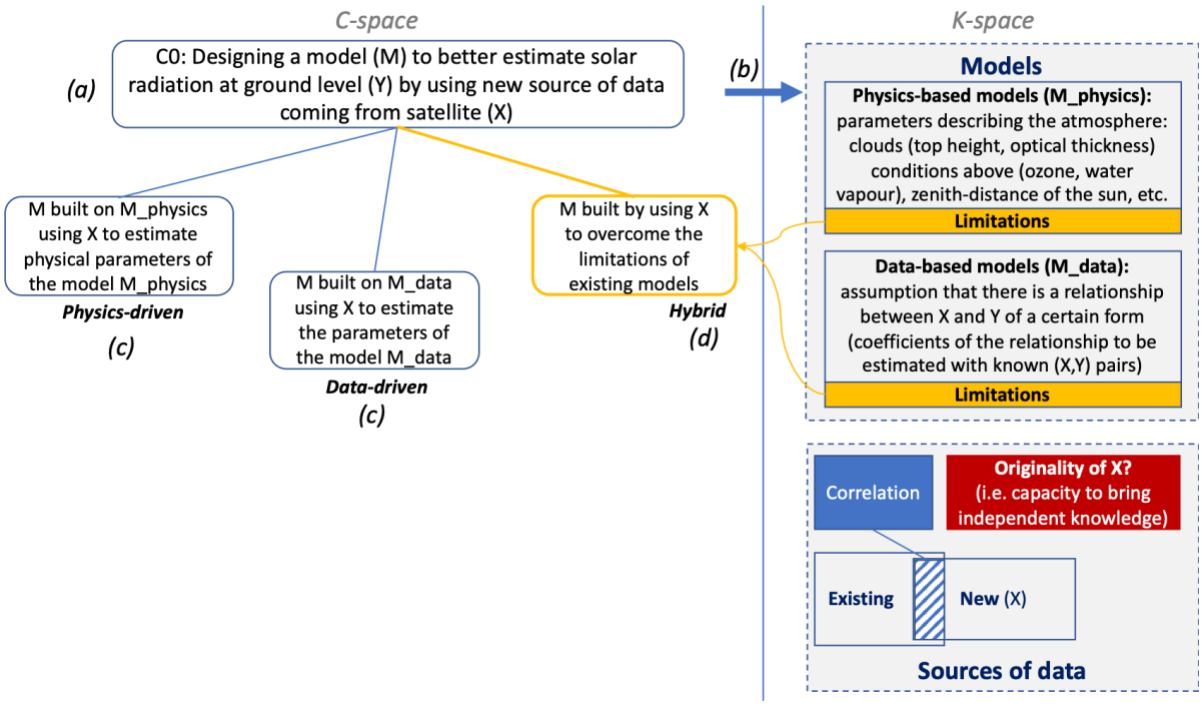

Figure 1. C-K theory used to represent the model-design reasoning logic

Case study: comparison of three model-design approaches

The case study used in this paper is particularly adapted to investigate the research question formulated above. Indeed, it corresponds to a situation where there is a new source of data with an already identified usage: it thus allows us to avoid other debates in literature related to the identification of usages for new data and to rather focus on the design efforts that remain to be made to transform data into information, already identified as useful. Our case study is based on the research work of a research organisation based in Sophia-Antipolis (France) on solar radiation estimation from satellite data, carried out in the 1980s. As mentioned in the introduction, this organisation was involved in a project supported by the European Commission's Solar Energy R&D Programme. The project aimed at assessing solar radiation more precisely and reliably, especially by integrating new data coming from satellites (whereas at the time solar radiation was mainly derived from networks of "in-situ" solar instruments, that were installed in a limited number of locations). Within this project, Sophia-Antipolis research institute, along with

two other research teams, were in charge of developing a model to link solar radiation estimates and Earth observation data including new satellite data. Each research team developed a different model, based on its respective expertise. Their different model-design approaches were compared in the final report of the project for the European Commission (Grüter et al., 1986). We also had access to the PhD thesis detailing the specific modelling approach developed by the Sophia-Antipolis team, and conducted semi-structured interviews (6 hours in total) with the researcher of this team who had been working on the development of the solar radiation methods from this European project in the 80s up to 2018.

ANALYSIS OF THE CASE STUDY: SPECIFIC ROLE OF THE DOMAIN KNOWLEDGE RELATED TO THE INSTRUMENT

Three teams corresponding to the three archetypal approaches found in literature

The three approaches developed by the different teams correspond well to the different archetypal approaches found in literature: a physics-driven approach (Cologne team), a data-driven approach (Stuttgart team) and a hybrid approach (Sophia-Antipolis team) - see also Figure 2:

Physics-driven approach (Cologne team): this team designed a model based on a "radiative transfer model" that explicitly describes the physical processes (e.g. absorption, scattering) occurring in the atmosphere. In this approach, satellite data are used to estimate existing parameters of the physical model. The limitations of this approach lie in the need of additional sources of data coming from other sources to determine some of the parameters of the model, that usually involves averaging the results over larger areas thus degrading the resolution of the final product.

Data-driven approach (Stuttgart team): this team resorted to a statistical approach: the model is based on statistical regressions between satellite data and solar radiation measurements at the Earth's surface, measured by "in-situ" stations within the considered area. Satellite data are used to estimate the coefficients of the statistical law (starting with 360 features describing each satellite image, 25 parameters were kept, being most correlated to the ground solar radiation). The limitations of the approach result from the large number of regression parameters to be estimated without considering much physical-consistency, as the parameters are mainly deduced from the texture analysis of the satellite images.

Hybrid approach (Sophia-Antipolis team): this team resorted to an approach aiming at overcoming the limitations of the two previous approaches. This approach more specifically relied on the introduction of an intermediary variable coined "cloud index", describing the level of cloudiness. This hybrid approach had proved to be the most efficient one, in terms of quality of the estimation (see error histograms on Figure 1) but also easiness of processing (almost ten times quicker than the physical one). Our empirical materials help us to further elucidate the role played by the domain knowledge on the instrument in this approach.

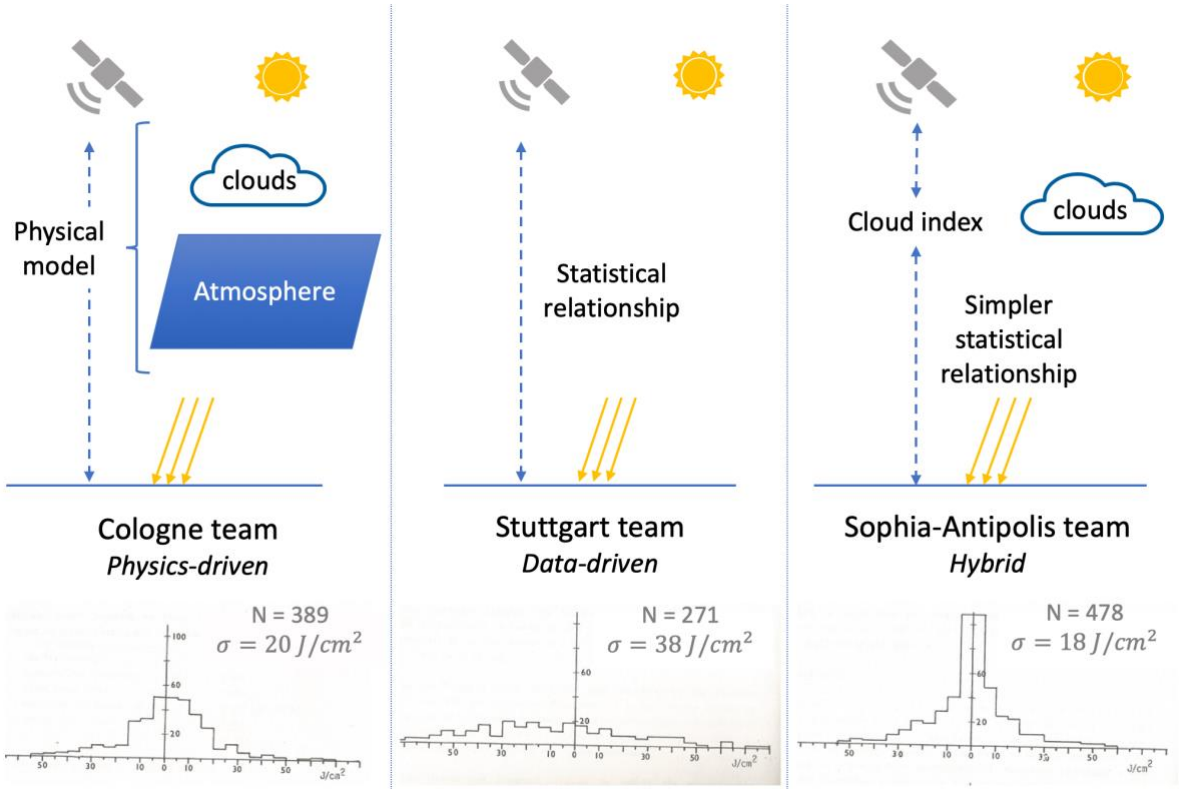

Figure 2. The three competing approaches and their respective results (histograms of errors - difference between predicted estimation and measurement at ground level)

Closer look at the domain knowledge related to the instrument: intertwined exploration of the originality of the new instrument and the way of taking advantage of it to overcome identified limitations of existing models

Based on what predicts C-K theory, the performance of such an approach could be explained by the way domain knowledge on the instrument is leveraged, especially investigating the "orthogonal" dimension of the new instrument compared to existing sources of data. The introduction of the "cloud index" by Sophia-Antipolis team corresponds to such an approach of investigating the originality of satellite data compared to existing sources of data. Indeed, (Cano, 1982) explicitly mentions that this cloud index is specifically built in order to be only determined by satellite data, without resorting to other parameters that would require to be assessed through other data sources. To do so, the estimation of this variable takes advantage of a specific property of the satellite data, i.e. the provision of time series (images of the same location at different moments in time). It is thus clear that Sophia-Antipolis team resorts to domain knowledge on the new instrument in a very specific way: rather than relying on the dimension of new sources of data correlated to existing ones, the researchers e*xplore the originality of the instrument, here exploiting a specific property of satellite data.*

A second interesting element to be noted is that *this knowledge expansion about the possibilities of the instrument is made in close interaction with the exploration of how existing models' limitations can be addressed*. Indeed, Sophia-Antipolis researchers highlight in (Grüther et al., 1986) that the cloud index can then be statistically linked to solar radiation with a simple linear relationship, thus reducing the number of regression parameters compared to fully statistical approaches, and avoiding rough estimation of some parameters of the physical approach that could not be directly estimated by satellite data. These conclusions lead us to refine the "hybrid" approach by distinguishing between two forms of "hybrid" models (see Figure 3):

"Combinatory" hybrid models that would combine parts of physics-driven and parts of data-driven models relying on a partial domain knowledge of the instrument, related to its dimensions that can be correlated to existing sources of data.

"Expansive" hybrid models that would leverage the originality of the new instrument compared to existing ones to generate expansion on how the model is designed, as highlighted in this Sophia-Antipolis case.

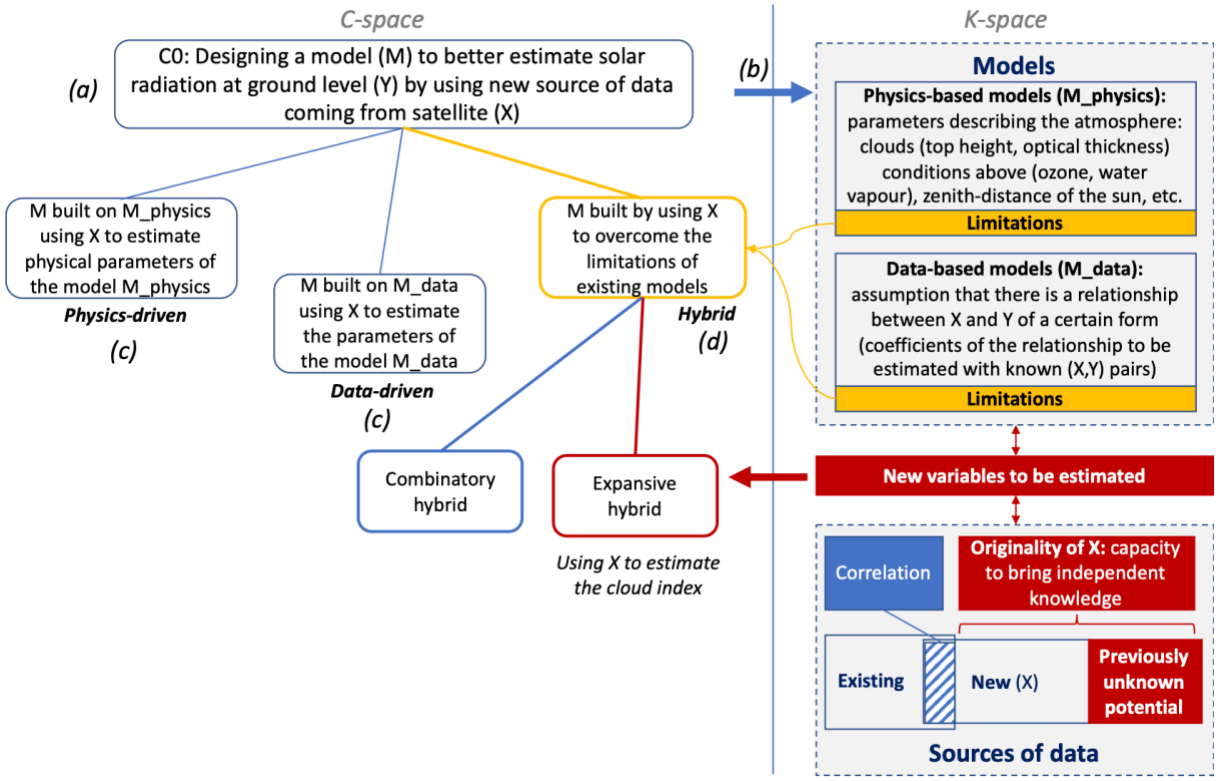

Figure 3. Representation of the model-design reasoning logic using C-K theory, completed with the case study analysis

DISCUSSION AND CONCLUSION

This paper proposes several contributions. First, it gives particular insights on the current academic discussions related to data-driven economy, where a lot of attention is either directed to the potential usages that new flows of data could create, or the new types of data to be exploited or created in order to address defined use cases. The case of solar radiation estimation offers an interesting configuration where these two considerations are already addressed: indeed, there is already a clearly identified type of information (solar radiation) to be derived from new data sources (satellite data), for a known usage (solar radiation information being of direct interest for the European Commission building at that time a European Solar Radiation Atlas). This example allows us to concentrate on the design efforts to be made to transform data into information and to show interesting features of the reasoning logic behind the design of a "hybrid" model combining domain and data science knowledge. In particular, this paper highlights the specific role of the domain knowledge related to the instrument in designing models to transform new sources of data into information. More specifically, it shows that a performant approach can result from *making specific efforts on exploring and leveraging the originality of the new instrument compared to existing ones, suggesting a specific way of building hybrid models that goes beyond a simple combinatory logic*.

These considerations can be helpful to better understand some elements found in literature. First, the difficulties of resorting to new types of data mentioned by (Bertoni, 2020) and (Montecchi and Becattini, 2020) can be better understood: our results indeed suggest that the tendency to avoid resorting to new data might result from the intricate design effort to be made with specific competencies in order to transform new data into added-value information.

The importance of *orienting the design process in a way that takes into account specific limitations of the usual models* can also be found in other contexts, e.g. in (Kazakçı, 2015) highlighting that in the context of "HiggsML challenge organised to gain insights into the study of Higgs boson in particle physics by means of machine learning algorithms", the participants who succeeded were the ones that did not simply apply the usual workflow of machine learning techniques but were able to take into account the specificity of the challenge involving an usual objective function. More interestingly, some examples of a "theory-guided data science" approach given in (Karpatne et al., 2017) can be better understood as either "combinatory" or "expansive" hybrid approaches, by considering the way domain knowledge related to the instrument is leveraged. For example, the problem of mapping surface water dynamics with satellite data starts with the *analysis of the limitations of "data-only" models*: "Remote sensing data from Earth observing satellites presents a promising opportunity for monitoring the dynamics of surface water body extent at regular intervals of time. It is possible to build predictive models that use multi-spectral data from satellite images as input features to classify pixels of the image as water or land. However, these models are challenged by the poor quality of labeled data, noise and missing values in remote sensing signals, and the inherent variability of water and land classes over space and time." From this analysis, *a way of addressing these challenges is investigated* thanks to additional domain knowledge, noticing that "locations at a lower elevation are filled up first before the water level reaches locations at higher elevations". Thus, to improve the model, *information on the elevation is identified as a new variable to be estimated* to assist classification models (it would be used as a constraint of the classifier minimizing training errors). However, such information obtained from other instruments (sonar instruments) is not available at the required granularity. Thus, *a new way of using satellite data is imagined to derive information on the elevation*, by "using the history of imperfect water/land labels produced by a data science model at every location over a long period of time", suggesting here an "expansive" hybrid approach.

This paper also contributes to practice as these results shed light on *several interesting competencies that a model designer should have to successfully develop a model combining domain and data science knowledge.* It is first highlighted that domain knowledge does not only involve understanding the physical processes but also understanding and exploring the potential of the instrument providing new sources of data. Second, the competencies of model designers should not be described as only picking in a model manual (either physics-based or data-based) given a certain situation, but should rather involve *(1) a detailed understanding of the limitations of existing models - either "physics-driven" or "data-driven" ; (2) the ability to explore both the originality of the new source of data compared to existing ones and on how it could help overcome the limitations of existing models*. This might lead to introduce new variables to be estimated from data, which might also consequently involve building new ways of considering the output data of the new instrument. Finally, these elements also suggest a third competency: *(3) the ability of leveraging together independent, potentially heterogeneous, data sources*. Indeed, the analysis of the model limitations and originality of the instrument can result in building new variables to be estimated from data, that would require looking for several new independent sources of data. This specific competency could be related to "data fusion", that had been interestingly investigated by the same Sophia-Antipolis team (Wald, 1998), as "a formal framework in which are expressed means and tools for the alliance of data originating from different sources, in order to obtain information of greater quality".

Several limitations of our research can be identified. First, the paper relies on a specific case study, highlighting the relevance of an expansive hybrid approach that leverages domain knowledge related to the new instrument by exploring its originality compared to existing sources of information. However, other interesting approaches could also result from other types of expansion that are not only related the instrument domain knowledge expansion, but that could come from the enrichment of knowledge bases related to the physics-driven or data-driven models (especially through new machine learning techniques that are currently developed). These types of expansions could be further described through other case studies corresponding to such contexts. Moreover, as highlighted in this paper, our case study corresponds to a situation where the new source of data and its usage (transformation into a certain type of added-value information) are already given. It would be interesting to investigate how the design process described in this paper could be leveraged in cases where the usefulness of information or the type of instrument to be used are still be to be explored. Finally, it is also worth noticing that our case study relies on a specific type of data (i.e. scientific or instrumental type). The relevance of our results for other types of data (such as data bases of patents or data on consumers' preferences) could be discussed. At first sight, we could consider our

case study as an extreme case that helps rediscuss basic notions. In this perspective, even with non-scientific types of data, we could assume that information is also derived from data through the use of specific models (although maybe not as complex as for Earth science). In some contexts, these models might be implicitly used and little designed, and the specific model-design competencies highlighted in this paper could open up new possibilities, by designing models that makes most use of knowledge about limitations of data-science techniques and description of the considered phenomenon (not necessarily physical processes, but for example modelling of customer behaviours), and exploration of the specificities of the data collection process (that might be different from scientific instruments). Further investigations would be interesting to further examine this question.

REFERENCES

Abbasi, A., Sarker, S., Chiang, R.H.L., 2016. Big Data Research in Information Systems: Toward an Inclusive Research Agenda. J. AIS 17, 3. https://doi.org/10.17705/1jais.00423

Barnes, C.J., 1995. The art of catchment modeling: What is a good model? Environment International, Water Modelling 21, 747–751. https://doi.org/10.1016/0160-4120(95)00082-V

Bertoni, A., 2020. DATA-DRIVEN DESIGN IN CONCEPT DEVELOPMENT: SYSTEMATIC REVIEW AND MISSED OPPORTUNITIES. Proceedings of the Design Society: DESIGN Conference 1, 101–110. https://doi.org/10.1017/dsd.2020.4

Beven, K., 1989. Changing ideas in hydrology — The case of physically-based models. Journal of Hydrology 105, 157–172. https://doi.org/10.1016/0022-1694(89)90101-7

Dammak, H., Gardoni, M., 2018. Improving the Innovation Process by Harnessing the Usage of Content Management Tools Coupled with Visualization Tools, in: Chiabert, P., Bouras, A., Noël, F., Ríos, J. (Eds.), Product Lifecycle Management to Support Industry 4.0, IFIP Advances in Information and Communication Technology.pp. 642–655. https://doi.org/10.1007/978-3-030-01614-2_59

Escandón-Quintanilla, M.-L., Gardoni, M., Cohendet, P., 2018. Improving concept development with data exploration in the context of an innovation and technological design course. International Journal on Interactive Design and Manufacturing (IJIDeM) 12, 161–172. https://doi.org/10.1007/s12008-017-0380-5

Gandomi, A., Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management 35, 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Giordani, A., Mari, L., 2012. Measurement, Models, and Uncertainty. IEEE Transactions on Instrumentation and Measurement 61, 2144–2152. https://doi.org/10.1109/TIM.2012.2193695

Grüter, W., Guillard, H., Möser, W., Monget, J.M., Palz, W., Raschke, E., Reinhardt, R.E., Schwarzmann, P., Wald, L., 1986. Solar Radiation Data from Satellite Images: Determination of Solar Radiation at Ground Level from Images of the Earth Transmitted by Meteorological Satellites - An Assessment Study, Solar Energy R&D in the Ec Series F: Springer Netherlands.

Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M., Feldberg, F., 2017. Debating big data: A literature review on realizing value from big data. The Journal of Strategic Information Systems 26, 191–209. https://doi.org/10.1016/j.jsis.2017.07.003

Hatchuel, A., Weil, B., 2009. C-K design theory: an advanced formulation. Research in Engineering Design 19, 181–192. https://doi.org/10.1007/s00163-008-0043-4

Hatchuel, A., Weil, B., 2003. A new approach of innovative design: an introduction to C-K theory. Presented at the International Conference on Engineering Design, International Conference on Engineering Design, Stockholm.

Huron, S., Carpendale, S., Thudt, A., Tang, A., Mauerer, M., 2014. Constructive visualization, in: Proceedings of the 2014 Conference on Designing Interactive Systems, DIS '14. Association for Computing Machinery, New York, NY, USA, pp. 433–442. https://doi.org/10.1145/2598510.2598566

Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., Kumar, V., 2017. Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. IEEE Transactions on Knowledge and Data Engineering 29, 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168

Kazakçı, A.O., 2015. DATA SCIENCE AS A NEW FRONTIER FOR DESIGN. Presented at the Proceedings of the 20th International Conference on Engineering Design (ICED15), p. 10.

Kim, H.H.M., Liu, Y., Wang, C.C.L., Wang, Y., 2017. Special Issue: Data-Driven Design (D3). J. Mech. Des 139. https://doi.org/10.1115/1.4037943

Mari, L., Carbone, P., Petri, D., 2012. Measurement Fundamentals: A Pragmatic View. IEEE Transactions on Instrumentation and Measurement 61, 2107–2115. https://doi.org/10.1109/TIM.2012.2193693

Montecchi, T., Becattini, N., 2020. Design for sustainable behavior: opportunities and challenges of a data-driven approach.Proceedings of the Design Society: DESIGN Conference 1, 2089–2098. https://doi.org/10.1017/dsd.2020.147
Noia, M., Ratto, C.F., Festa, R., 1993a. Solar irradiance estimation from geostationary satellite data: I. Statistical models. Solar Energy 51, 449–456. https://doi.org/10.1016/0038-092X(93)90130-G
Noia, M., Ratto, C.F., Festa, R., 1993b. Solar irradiance estimation from geostationary satellite data: II. Physical models. Solar Energy 51, 457–465. https://doi.org/10.1016/0038-092X(93)90131-7
Parraguez, P., Maier, A., 2017. Data-driven engineering design research: Opportunities using open data Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 7: Design Theory and Research Methodology, Vancouver, Canada, 21-25.08.2017.
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566, 195–204. https://doi.org/10.1038/s41586-019-0912-1
Rowley, J., 2007. The wisdom hierarchy: representations of the DIKW hierarchy , The wisdom hierarchy: representations of the DIKW hierarchy. Journal of Information Science 33, 163–180. https://doi.org/10.1177/0165551506070706
Sitruk, Y., Kazakçi, A., 2018. CROWD-BASED DATA-DRIVEN HYPOTHESIS GENERATION FROM DATA AND THE ORGANISATION OF PARTICIPATIVE SCIENTIFIC PROCESS. Presented at the 15th International Design Conference, pp. 1673–1684. https://doi.org/10.21278/idc.2018.0510
Tal, E., 2017. Calibration: Modelling the measurement process. Studies in History and Philosophy of Science Part A, The Making of Measurement 65–66, 33–45. https://doi.org/10.1016/j.shpsa.2017.09.001
Wald, L., 1998. Data fusion: a conceptual approach for an efficient exploitation of remote sensing images. Presented at the 2nd International Conference " Fusion of Earth Data: merging point measurements, raster maps and remotely sensed images ", Sophia-Antipolis (France), p. 8.
Zins, C., 2007. Conceptual approaches for defining data, information, and knowledge. Journal of the American Society for Information Science and Technology 58, 479–493. https://doi.org/10.1002/asi.20508

ACKNOWLEDGMENTS

This research work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820852.

Co-design for novelty anchoring into multiple socio-technical systems in transitions: the case of Earth observation data

Authors: Raphaëlle Barbier[1], Skander Ben Yahia[1], Pascal Le Masson[1], Benoit Weil[1]

[1]*Center for Management Science, i3 (UMR CNRS 9217), Mines Paris, PSL University, France (e-mail: raphaelle.barbier@minesparis.psl.eu)*

Pre-print version of the paper published in IEEE Transactions on Engineering Management

Abstract

Tackling grand challenges requires new forms of collaborative innovation to support intricate design processes involving heterogeneous actors. This paper specifically investigates how co-design supports the anchoring of promising novelties into multiple socio-technical systems to accelerate their respective sustainability transitions. A co-design framework adapted to this multi-system context is derived from transition research and design and innovation management research. The framework is validated empirically based on twenty-seven case studies where the novelty to be anchored corresponds to Earth observation data. Contributing to transition research, the paper shows how this multi-system co-design framework provides novelty developers with a diagnostic tool to clarify their anchoring strategy, by framing the relevant actions to conduct at different time horizons. Several enrichments of the anchoring concept are also proposed, highlighting some complementarities between different forms of anchoring and the endless property of the process. Contributing to design and innovation management research, the paper sheds light on co-design in an original perspective by considering a context crossing the usual boundaries of socio-technical systems and focusing on a diagnostic dimension preceding the organisation of collective design sessions. The co-design framework also highlights a so-called "resource-based" form of collaborative innovation aiming to build novelty-based resources for heterogeneous actors facing grand challenges. This approach complements more common "challenge-based" approaches aiming to directly address a targeted challenge.

Keywords: co-design, collaborative innovation, grand challenges, sustainability transitions, anchoring, multi-level perspective, strategic niche management, Earth observation data, digital innovation

I. Introduction

Addressing grand societal challenges appears today as a major priority, requiring deep transformations of societies. Different streams of management research underline the specific issues associated with these grand challenges, calling for new forms of technical and social innovations [1], [2], institutions and policies supporting them [3], [4], and practices within organisations [5]. Eisenhardt *et al.* [6] define them as "complex problems with significant implications, unknown solutions, and intertwined and evolving technical and social interactions", that might include among others climate change, water scarcity, poverty, or food security. Recent works particularly underline that addressing such challenges requires new forms of collective action [7]–[10], especially to stimulate intricate innovation processes beyond organisational boundaries [11]–[14], that need to be considered in a long-run and evolutionary perspective [2], [3], [11], [12], [15] to efficiently ensure deep and viable socio-technical transformations [4].

The *multi-level perspective* (MLP) framework offers an interesting analytical tool to better understand the long-term socio-technical transformations involved in addressing these grand challenges through so-called *sustainability transitions* [16]–[22]. MLP draws upon the notion of *socio-technical system*, referring to the actors, institutions and artefacts interacting to fulfil societal functions (e.g. transport, communication, nutrition) [23]–[26]. Transitions are conceptualised as non-linear processes resulting from the interactions between the three analytical levels of the socio-technical system: *niches*, defined as protective spaces for the development of radical novelties; *regime* referring to the rules and practices framing the action of the different social groups involved in transitions (e.g. engineers, users, policy makers) and accounting for the stability of the existing system; and *landscape* corresponding to the exogenous context affecting socio-technical developments (e.g. global societal trends putting pressure on the existing regime). A transition occurs when there is a shift from one regime to another following a specific interplay between these three levels. As described in early MLP works, the internal tensions of regimes or an intensified landscape pressure can generate a "window of opportunity" for niche innovations (i.e. novelties maturing in niches) to progressively agglomerate into a new socio-technical configuration, competing with the incumbent regime, and eventually establishing itself as a new regime [23].

Scholars have progressively refined the analysis of these transition dynamics by unveiling different transition pathways [24], [27]–[29], and have recently called for further research on how these transition dynamics could be accelerated [30], [31]. They especially suggest that present transition dynamics might occur through more complex and subtle interactions between niches and regimes compared to past transitions [20]. In this perspective, scholars have more precisely shed light on certain forms of mechanisms, coined "anchoring", aiming at newly or more firmly connecting a novelty to a niche or regime within a socio-technical system [32], [33]. They also underline the importance of considering the interactions between multiple socio-technical systems [29], [34]–[36], that can be especially fruitful in enhancing niche resilience, as illustrated by the case of biogas anchoring in both agriculture and energy systems [34].

Our paper aims at unveiling certain forms of managerial practices that would support novelty anchoring into multiple socio-technical systems. We especially propose to focus on so-called 'co-design' practices, supporting the implementation of an interactive design approach involving intricate learning processes [17], [37]. This analytical lens is indeed particularly relevant for our investigation as it combines two crucial aspects of grand challenges and sustainability transitions: a *design* aspect accounting for the high degree of unknown entailed in building new forms of actions to address grand challenges, a *collective* aspect accounting for the variety of actors to be involved beyond usual disciplinary and sectorial boundaries [1], [2], [6]. Our paper thus proposes to investigate the following question: *how can a multi-system co-design framework help the developers of a certain novelty steer an anchoring strategy into multiple socio-technical systems?*

A theoretical co-design framework adapted to the investigated context is first derived from transition research complemented by research in design and innovation management, leading us to distinguish between four main types of co-design depending on the nature of the learning processes in which novelty developers engage. The relevance of these co-design types is validated and discussed empirically based on twenty-seven case studies rooted in the Earth observation field, where specific actors are encouraged to use co-design to further anchor Earth observation data into various socio-technical systems that would benefit from these data to accelerate their respective sustainability transitions.

This paper offers several contributions advancing research on collaborative innovation for sustainability transitions. Contributing to transition research, our work takes a specific look at managerial practices

that can support multi-system anchoring. Our study shows that the co-design framework provides novelty developers with a helpful *diagnostic tool* to clarify their anchoring strategy by distributing their design efforts over time, thus allowing them to handle significant learning processes with broad design ambitions in a long-term perspective. Several enrichments of the anchoring concept are also derived from these outcomes.

Contributing to design and innovation management research, our work sheds light on co-design in an original perspective. Indeed, this paper does not focus on the aspects of co-design related to the organisation of specific collective design sessions or workshops, but focuses on a *diagnostic dimension* preceding the actual collective design sessions. This diagnostic dimension is especially crucial in the *multi-system context of co-design* considered in the paper, requiring specific efforts to identify the relevant actors among a complex and evolving range of heterogeneous actors. Second, in the perspective of tackling grand challenges, our co-design framework sheds light on two complementary logics of collaborative innovation, coined "*resource-based*" and "*challenge-based*". This paper eventually proposes several perspectives for practitioners, involved in novelty development activities or in policy making institutions.

II. Theoretical background

This section examines the relevance of taking a co-design perspective to support novelty developers in building an anchoring strategy into multiple socio-technical systems, and elicits a theoretical co-design framework based on transition studies and research in design and innovation management.

A. Anchoring novelties into niches or regimes to accelerate the sustainability transitions of multiple socio-technical systems: the relevance of considering a co-design perspective

The development of novelties and their wider uptake appear as important aspects in sustainability transitions of socio-technical systems. The MLP framework especially emphasises the importance of niches as the "seeds for systemic change" [22], playing the role of protective spaces where radical novelties are developed and sheltered from possible tensions with the existing regime [23]–[26]. A dedicated branch of transition research, called *strategic niche management* (SNM), has more specifically

underlined the importance of creating social networks endorsing intricate *learning processes* that concern multiple dimensions - including technical aspects and design specifications, market and user preferences, cultural and symbolic meaning, infrastructure and maintenance networks, industry and production networks, regulations and government policy, societal and environmental [38]–[41].

Drawing upon MLP and SNM insights, Elzen *et al.* [32] have introduced "anchoring" as an analytical concept to expand the organisation of these learning processes beyond the initial development of novelties towards their wider uptake in socio-technical systems. Anchoring is defined as "the process in which a novelty [i.e. a new technology, a new technical concept or a new socio-technical practice] becomes newly connected, connected in a new way, or connected more firmly to a niche or a regime". The authors precise that anchoring does not refer to the permanent uptake of the novelty in a new environment (niche or regime), but rather consists in a "continuous process of probing new connections" until their transformations into more durable links. Three forms of anchoring are distinguished: *technological anchoring* when the technical characteristics of a novelty are more specifically defined and adapted to the actors' operations and practices; *network anchoring* "when changes occur in the network of actors that 'carry' the novelty, e.g. by producing it, using it or developing it further", and *institutional anchoring* when changes occur in the beliefs, visions, or problem views of actors *(cognitive or interpretative institutions)*, formal and informal rules about what is desirable or not (*normative institutions)*, and rules and arrangements (e.g. contracts, business networks) that govern market or economic activities *(economic institutions)*.

This anchoring concept has been further expanded to account for the interactions between multiple regimes, e.g. novelties for renewable energy production that have successfully developed by anchoring in both agriculture and electricity regimes [34]. This multi-system perspective seems especially important for different types of novelties mentioned in literature, especially so-called generic or general purpose technologies having the potential to be used in many different application domains, such as materials science, 3D printing, biological and genetic engineering, computing [30], or other digital innovations [42]. Sutherland *et al.* [34] highlight that these multi-regime anchoring processes can be supported by the emergence of a new regime, coined "fiat regime" - "fiat" meaning formal authorisation, proposition or a decree - characterised by new sets of rules and regulations facilitating the cooperation between parent regimes while preserving their own structure and dynamics (e.g. targets

set by the European Commission and implemented nationally to support renewable electricity production). A complementary focus on the underlying managerial practices could provide additional insights on how to support these multi-system anchoring processes, in line with scholars' recommendations on considering the agency involved in sustainability transitions, i.e. the actors and their micro-level actions supporting transition dynamics [16], [22], [28], [43]. Several studies already offer interesting perspectives to examine these managerial practices. Although not in a multi-system perspective, some works highlight the potential of participatory design approaches in handling the intricate multi-actor design processes supporting the sustainability transitions of the agricultural system [33], [44], [45] or the energy system [46]. For example, Beguin *et al.* [45] emphasise the role of collective and innovative 'co-design' approaches in "fostering cross learning processes amongst designers and users in order to achieve the joint building of a technology, of a desirable future, and of the activity or the collective action in which the technology will be used". Studying the animal production system, Elzen and Bos [33] have more specifically shown the fruitful combination of the anchoring concept with an interactive design approach to design a new integrally sustainable system, by especially targeting the uptake of novelties from the beginning of the design phase.

Our paper proposes to advance these works by jointly considering these two streams of recent research on anchoring, building complementarities between their respective results and remaining blind spots, i.e. on the one side investigating *multi-system* interactions with a specific attention on micro-level managerial practices, and on the other side considering the fruitful *combination of anchoring with collective and innovative design approaches* so far described within the boundaries of a single socio-technical system. We will use the term "co-design" to refer to these design approaches, as it is explicitly mentioned by Grin *et al.* [17] as one of the shared concepts of transition research, underlining that "knowledge is developed in a complex, interactive design process with a range of stakeholders involved through a process of social learning". This especially leads us to formulate the following question: *how can a multi-system co-design framework help the developers of a certain novelty steer an anchoring strategy into multiple socio-technical systems?*

Transition literature does not explicitly propose a co-design framework in the context of anchoring novelties into multiple socio-technical systems. To examine this question, we have thus been confronted with the issue of building an adapted co-design framework. Next paragraphs show how such a

framework can be theoretically derived from existing literature, leveraging transition studies to conceptualise the overall co-design setting, and additional insights from design and innovation management literature to distinguish between different types of co-design depending on the nature of learning processes.

B. Co-design framework through the lens of transition research

Transition literature provides us with several insights to conceptualise the overall setting of a multi-system co-design framework by clarifying the following elements: the considered multi-system configuration and related agency, the expected outcomes of co-design and the associated learning processes.

In the context of novelty anchoring in multiple socio-technical systems, the developers and the users of the novelties might belong to different socio-technical systems, resulting in an increasing variety of possible actors to be involved and a large heterogeneity of knowledge between these actors. We especially distinguish between the so-called *novelty-emergence socio-technical system* into which the novelty has initially developed (e.g. biogas primarily embedded into the agriculture regime to address waste management problems) and *novelty-use socio-technical systems* that might benefit from the use of the novelty to accelerate their respective sustainability transitions by better tackling the grand challenges they are facing (e.g. further anchoring of biogas into the energy regime). Looking at underlying agency, scholars have highlighted the benefits of taking an "insider" perspective, describing the strategies deployed by the advocates of niches that mobilise and create protective spaces over time through multi-actor relationships [43]. Following a similar line, our paper focuses on the view of *novelty developers* supporting the development of a certain novelty by anchoring it into various socio-technical systems.

Regarding the expected outcomes of co-design, the definition provided by Beguin *et al.* [45] suggests considering "the joint building of a technology, of a desirable future, and of the activity or the collective action in which the technology will be used". Considering the investigated context of anchoring to accelerate the transitions of multiple socio-technical systems, these outcomes can be further specified following two dimensions. First, focusing on the anchoring dimension, the expected outcomes can be characterised following the *three types of anchoring* detailed above: *technological anchoring* (in line with

the “joint technology building” aspect), *network anchoring* (in line with the “collective action” aspect), and *institutional anchoring* (in line with the “desirable future” aspect). Scholars also argue that successful anchoring seems to require an interplay of these three forms of actions [32], [33]. For example, the limited development of biogas in some countries can be associated with a lack of network anchoring or a lack of cognitive institutional anchoring resulting in altering the normative institutional support (dedicated rules and regulations) brought to renewable energy production [34]. For this reason, it seems relevant to consider co-design as potentially associated with all three types of anchoring. Second, as anchoring is investigated in the perspective of accelerating the sustainability transitions of multiple socio-technical systems, the expected outcomes of co-design can also be specified according to a second dimension related to *its interaction with transition dynamics of both the novelty-emergence and the novelty-use socio-technical systems.*

Finally, taking a micro-level perspective on co-design involves considering an additional analytical layer closer to the novelty developers’ contexts of actions. The co-design definitions provided by Beguin *et al.* [45] and Grin *et al.* [17] both suggest considering the *learning processes* underlying the co-design approach. In this perspective, SNM scholars emphasise the importance of learning processes that do not merely focus on accumulating facts and data (coined as first-order learning), but also expanding cognitive frames and assumptions (coined as second-order learning) [38]–[41]. Empirical case studies have especially highlighted that niche development might be significantly hampered when learning processes are limited to first-order learning, e.g. by restrictively perceiving users as consumers with already articulated needs [41], [47]. However, second-order learning processes appear to be particularly difficult to reach in practice and might depend on specific drivers and contexts [48], [49]. These elements suggest that *co-design might be associated with different forms of learning processes depending on the context of action*. Our investigation especially needs to consider the forms of learning processes that will address the large degree of unknown and high heterogeneity of knowledge prevailing in the context of multi-system sustainability transitions.

C. Co-design framework through the lens of design & innovation management research

Research in design and innovation management provides complementary insights to further characterise these forms of learning processes that can be specified by considering the design space of novelty developers, involving different levels of unknown related to co-evolving problem and solution spaces.

1) Expanding the design space of actors through the intertwined expansion of problem and solution spaces

Design research has shed light on a so-called 'co-evolutionary' paradigm in which creative design involves the exploration of two distinct spaces - *the problem space and the solution space* – that continuously evolve through mutual interaction [50]–[52]. Dorst and Cross [53] further explored the empirical validity of this model and elaborated on the notion of *pairing*: "creative design involves a period of exploration in which problem and solution spaces are evolving and are unstable until (temporarily) fixed by an emergent bridge which identifies a problem-solution pairing."

Similar considerations can be found in innovation management research. In particular, von Hippel and von Krogh [54] propose an original problem-solving approach, conceptualised as the discovery of viable "need-solution" pairs, linking a certain point of the need landscape (defined as the pool of need-related information) and a certain point of the solution landscape (defined as the pool of solution-related information). Indeed, the authors notice that individuals sometimes "[recognise] a problem worth solving only after encountering a potential solution worth implementing", thus contrasting with classical problem solving starting with problem formulation. For the sake of clarity, we will only keep the terms introduced in design literature, i.e. "problem-solution pairs" and "problem and solution spaces".

The context of sustainability transitions has not been explicitly investigated by these different works, that have indeed mainly considered problem-solving cases encountered by individuals in their everyday life [54], or in laboratory settings with clearly defined and delimited design tasks [55], [53]. However, the problem-solution pairing approach appears to be well adapted when problem spaces are complex and various problem-solution pairs are potentially viable, avoiding costly efforts in formulating a problem or searching an exhaustive problem or solution space [54]. Grand challenges and sustainability

transitions clearly meet these conditions, given the large amount of unknown associated with both problems and solutions [2], [6].

Moreover, although mainly based on the perspective of an individual designer that would be able to discover problem-solution pairs on its own, von Hippel and von Krogh [54] also mention situations involving multiple actors (e.g. in crowdsourcing or open source initiatives). They suggest that the approach is more likely to be successful when solver individuals or teams have expertise in aspects related to both problems and solutions. In the context of multi-system anchoring, it can be reasonably assumed that such expertise is not shared by the same actors. Basically, novelty developers might have a limited expertise on problem aspects (i.e. for what purposes the novelty could be used), and the actors that could potentially benefit from this novelty might also have a limited expertise on solution aspects (i.e how the novelty could be transformed into a promising solution). In such situations, the discovery of problem-solution pairs appears to be hardly achievable by the actors taken individually, thus requiring a dedicated collective design setting. This confirms further the interest of investigating the potential role played by co-design for actors involved in multi-system anchoring. Taking the view of novelty developers, the expected outcomes of co-design can be more precisely described according to the *expansion of the design space of novelty developers, made of two co-evolving sub-spaces (problem and solution spaces) resulting in the discovery of viable problem-solution pairs.*

2) Design space associated with various degrees of unknown related to problem and solution spaces

Recent advances in design theory are helpful to go one step further in characterising this design space. Hatchuel *et al.*[37] especially recall that the strength of design lies in its 'generativity', i.e. "the ability to conceptualize and create non-existent alternatives". It has thus been argued and demonstrated that design reasoning logic goes beyond "bounded rationality" [56], but rather involved an "expandable rationality" [57]. In this perspective, both problem and solution spaces are associated with a certain degree of unknown and can be progressively expanded through an intertwined exploration of unknown and known objects [58], [59]. Unlike usual decision-making and problem-solving paradigms, the unknown is not limited to the uncertainty on the value of well-known design parameters, but can potentially include the exploration of unknown design parameters. Hatchuel *et al.*[37] also stress that

the level of unknown (or generativity) involved in a design process determines the paradigm and social spaces in which the design process should take place: situations with a low level of unknown can be dealt with usual forms of problem-solving and social spaces, whereas situations with a higher level of unknown tend to require more generative models of design theory and the creation of original forms of social organisations. These considerations lead us to theoretically distinguish between *four types of co-design corresponding to different contexts defined by the level of unknown associated with the problem and solution spaces of novelty developers, leading to different forms of learning processes*, as defined in Table 1.

		Level of unknown - Problem space	
		Low	**High**
Level of unknown Solution space	**Low**	Co-design type 1 **Problem-related unknown:** identified problems that might need further specification **Solution-related unknown:** limited development efforts leveraging existing building blocks **Learning processes to build problem-solution pairs:** slight co-expansion of problem and solution spaces	Co-design type 2 **Problem-related unknown:** unknown or little-known problems to be identified **Solution-related unknown:** limited development efforts leveraging existing building blocks **Learning processes to build problem-solution pairs:** large on problem space, limited on solution space
	High	Co-design type 3 **Problem-related unknown:** identified problems that might need further specification **Solution-related unknown:** extensive development efforts **Learning processes to build problem-solution pairs:** large on solution space, limited on problem space	Co-design type 4 **Problem-related unknown:** unknown or little-known problems to be identified **Solution-related unknown:** extensive development efforts **Learning processes to build problem-solution pairs:** large co-expansion of problem and solution spaces

Table 3: Four types of co-design theoretically deduced from the level of unknown associated with problem and solution spaces of novelty developers

To summarise this section, recent works in transition studies have led us to raise the following research question: *how can a multi-system co-design framework help the developers of a certain novelty steer an anchoring strategy into multiple socio-technical systems?* Two main streams of literature have then been used to build a theoretical co-design framework adapted to the context of multi-system anchoring

entailing intricate learning processes. The respective insights brought by these two broad streams of literature are synthesised in Figure 1, more largely outlining the overall argument developed in the paper.

Theoretical background

Accelerating the sustainable transitions of multiple socio-technical systems… [29]-[31]

…by **anchoring** novelties to niches or regimes [32]–[34]

Anchoring in a **multi-system perspective** to enhance resiliency [34]

but need of further research on related managerial practices

Managerial practices taking the form of **interactive design processes** [33], [44], [45], [46]

but mainly studied within a single-system configuration

Focus on **agency**, i.e. the actors and their actions supporting transition dynamics [16], [22], [28], [43]

→ *Research question: how can a **multi-system co-design framework** help the **developers of a certain novelty** steer an **anchoring** strategy into **multiple** socio-technical systems?*

Co-design framework

Expected co-design outcomes (*transition research*)

Learning processes associated with different **unknown levels on problem and solution spaces of novelty developers** *(design and innovation management research)*

4 co-design types

	Multi-system anchoring (network, technological, institutional)	**Interactions with transition dynamics (novelty-emergence & novelty-use STS)**
- problem - solution		
+ problem - solution		
- problem + solution		
+ problem + solution		

Method and empiricial material

Assessment of the framework following an **inductive logic** based on **27 case studies in the Earth observation field**:

(1) Given the portfolio of case studies, how the types unfold in practice and whether certain types prevail

(2) Given one case study, whether all types are relevant and how they combine with each other

Results

(1) *Given the portfolio of case studies*, all types relevant in practice with predominance of type 1 at short-term and type 4 at long-term. Each type frames the relevant anchoring actions to conduct.

(2) *Given one case study*, evolutive combination of different co-design types distributed over time

Contributions to transition research

Co-design as a **diagnostic tool to organise anchoring actions over time with focused learning processes**

Enriching the **anchoring concept**: complementarity of forms & end-less character (grafting metaphor)

Contributions to design & innovation management

Perspectives on **co-design**: multi-system context and diagnostic dimension

Perspectives on **collaborative innovation for grand challenges**: resource-based/challenge-based

Further research

Co-design framework: typology robustness, protocols and tools supporting collective design sessions for each type

Organisational aspects: nature of the actors able to sustain multi-system anchoring processes, including novelty developers (e.g. hybrid actors [32], innovation intermediaries [79]-{82]), and providers of co-design expertise

Collaborative innovation for grand challenges: building monitoring indicators and analysing the complementarity between resource-based and challenge-based approaches

Figure 7: Synthesis of the overall argumentation developed in the paper, summarising theoretical background (precising how the co-design framework is built), method and empirical material, main results, research contributions and further perspectives (STS used for "socio-technical system")

III. Method and empirical material

Aligned with the guidelines proposed by Eisenhardt *et al.* [6] for research related to grand challenges, this research question is addressed through an empirical investigation following an inductive logic, aiming at taking advantage of rich empirical data to validate and potentially enrich the theoretical co-design framework derived from literature, drawing on multiple case studies [60] in the context of Earth observation. Assessing the relevance of the framework will especially include the two following aspects: (1) *given the portfolio of case studies*, investigating whether all four types of co-design appear to be relevant and to what extent certain types are predominant over the others, (2) *given one case study*, investigating whether the actors are concerned by one or several co-design types and how these types combine with each other.

A. Relevance of the Earth observation field as an empirical context

Earth observation (EO) data are produced by a large range of instruments (e.g. satellites, in-situ sensors such as meteorological land stations, but also more recently IoT or smartphone data), to monitor, understand, or predict the evolution of our man-made or natural environment. These data are thus a good example of a novelty that might support multiple actors in tackling grand challenges [42], e.g. building a more sustainable agriculture, building resilience to natural disasters, supporting the development of renewable energies.

The distinction between the *EO-emergence socio-technical system* and the *EO-use socio-technical systems*, into which EO data might be used to address certain grand challenges, can be justified by considering the three types of rules - regulative, normative and cognitive - defining a socio-technical regime [26]. On the regulative aspect, the EO socio-technical system is governed by laws and standards (e.g. related to satellite developments, processing and sharing of data), differing from laws and regulations followed by the potential user communities (e.g. the Common Agriculture Policy in agriculture, or the Stockholm Convention for surveillance of persistent organic pollutants). On the normative aspect, as the actors belong to very distinct profession bodies (e.g. data analyst on the one

side and farmer on the other side), they hardly share the same norms or performance logics. Finally, on the cognitive aspect, there is a large gap between the considered timelines (e.g. very long cycles to develop new instruments differing from short timelines of actions to be taken based on data uses), and the competencies (e.g. specific technical expertise related to data processing differing from specific domain expertise related to data uses).

In recent years, significant efforts have been undertaken to anchor EO data into various socio-technical systems, especially in the perspective of helping actors progress towards sustainable development goals. Considering the European context, these scientific data have been increasingly considered as a common good and made freely accessible to all potential users in an 'open-data' approach [61], [62]. However, the success of these 'open-data' policies is still limited in practice, as the different stakeholders are hardly familiar with EO data and seem to have difficulty in leveraging them on their own. In this context, significant efforts are currently carried out by the actors of the EO-emergence socio-technical system to go beyond 'open-data' policies and implement specific forms of collaborative innovation, referred as *co-design* (or similarly co-production or co-development [63], [64]), involving multiple stakeholders of the EO-emergence and EO-use socio-technical systems. This empirical context seems thus particularly adapted to investigate how co-design can support novelty developers in further anchoring the novelty in multiple socio-technical systems.

B. Empirical setting

Our empirical material is derived from our involvement in a large research project, which received a 4-year grant (2019-2023) from the European Commission under the Horizon 2020 programme. This project, called *e-shape* initially gathered a team of 54 experienced partners from academia, industry, institutional entities to develop 27 pilot applications based on EO data, organised in seven showcases (agriculture, health, renewable energy, biodiversity, water resources, disaster resilience and climate) [65]. Each pilot is in charge of developing a certain set of EO-based solutions within a specific showcase. It involves one or several organisations participating to the project, and is coordinated by one of these organisations designated as the "pilot leader". As an initial condition for project participation, each pilot must interact with at least one user organisation. These user organisations do not receive direct funding from the project and are thus considered as external actors to the project. An overview of the 27 initial

pilots is given in the appendix of the paper, describing for each: the overall pilot's rationale, the types of organisations involved in the pilot's development, and the different user groups targeted by the pilot.

All four authors of the paper are involved in e-shape, leading a work package dedicated to co-design aiming to provide the pilots with a co-design framework and guiding tools that are progressively designed and tested in interaction with the pilots. This setting is thus particularly favourable to conduct multiple case studies [60], corresponding to the different pilot cases, in the unified empirical context offered by the project.

C. Data collection and analysis

Our participation to the project enables us to have direct interactions with all the organisations involved in the project and their network of partners. On this basis, rich empirical data could be exploited from heterogeneous sources, necessary for a sound inductive approach [6]: questionnaires, interviews, observation notes taken during project meetings, and secondary sources of data on the different actors (application forms filled up by each pilot to participate to the project, websites and scientific publications of the different partners). The data used for this paper were collected between September 2018 and July 2021 with no noticeable impact due to Covid-19 situation as our main interactions were already organised on virtual platforms due to the international composition of the project. Within this timeframe, our work package activities included the validation of a co-design theoretical framework but also first experimentations of specific workshop protocols for each identified type. The present paper only focuses on the first aspect, assessing to what extent the four types of co-design provide the pilots with useful support to steer their anchoring strategy.

The validation process was designed in a collaborative research setting, involving both researchers and practitioners (here *e-shape* project members) [66]. This setting especially aims to "reduce the likelihood of drawing false conclusions from the data collected, with the intent of both proving performance of the system [of action] and adding to the broader body of knowledge in the field of management" [67]. Following guidelines for collaborative research [67], [68], the validation process consisted in *progressively building a shared interpretation of empirical data and findings between researchers and practitioners*. This involved the rigorous formalisation of a shared understanding of each pilot's context and the associated assessment of relevant co-design types for each pilot, through a sequence of steps

detailed below and synthesised in Table 2. As the pilots were not familiar with the 'anchoring' concept, our approach consisted in first assessing the relevance of co-design types based on their definitions related to the nature of unknown and learning processes, and deriving the associated anchoring strategy from the complementary points of analysis discussed with the pilots and detailed below in step 3.

Step 1 (September 2018 - May 2019) – Preliminary data analysis: building an overall understanding of the empirical context and the research tools supporting the validation process

This first step corresponds to a preliminary strategy for data analysis allowing us to become more familiar with the empirical context. Following case study guidelines [60], this involved the manipulation of empirical data supported by the creation of dedicated visual displays and templates, facilitating the triangulation of data by organising heterogeneous materials, in a more synoptic and comparable form [69]. We indeed built a specific template to synthetically represent each pilot as a chain linking data sources, the information derived from these data, their expected uses and associated actors. This template was first tested, discussed and validated in a dedicated one-day meeting organised with one pilot in April 2019. The outcomes were formalised in a deliverable report, reviewed and validated by the participants and by two external reviewers, and made publicly available on the project website [70]. The template was considered as a useful tool to create a shared understanding between researchers and practitioners, and was thus consequently used to support the validation process undertaken for each pilot by systematically (1) drafting a first version of the pilot template based on secondary sources of data and observation notes taken during the project kick-off meeting; (2) sharing this pre-filled template with the pilot through the online management platform of the project and updating it based on the pilot's feedback.

Step 2 (November 2019) – Shared validation of the co-design framework with project coordinators

The theoretical framework involving four types of co-design was first discussed with the project members having an official coordinating role, i.e. the project management team, showcase coordinators and work package leaders. A dedicated meeting was organised by our team in November 2019, where the co-design framework was presented, using the definition of the four types of co-design related to the nature of the associated learning processes. These different forms of learning processes were

illustrated on concrete examples of pilots based on the preliminary empirical data gathered in step 1. The relevance of the framework was discussed and approved by all participants of the meeting.

Step 3 (November 2019 - July 2021) – Assessment of the co-design framework for each pilot

To assess the relevance of the co-design framework for the 27 pilot cases, an interview of one hour and a half was conducted remotely with each pilot (video calls based on zoom application). The participants to this interview included at least two members of our research team and the pilot leader, and also in some cases the showcase leader and additional members of the pilot when judged relevant by the pilot leader. The profiles of the interviewees for each pilot are detailed in the appendix of the paper. Prior to the interview, a preliminary report was written by our team and shared with the pilot through the online management platform, formalising the current status of understanding on the pilot's context based on the information gathered in the previous steps. This report was structured in different points of analysis listed below, related to the anchoring activities of the pilot and transition dynamics of both EO-emergence and EO-use socio-technical systems, and associated with different parts of the template built in step 1 (see Figure 2):

- *Pilot's understanding of the transition dynamics of targeted EO-use socio-technical systems (a)*, i.e. overall regime organisation (specific rules and regulations, regime actors), the actors that would potentially benefit from EO data and their position within the socio-technical system;
- *Status of EO data anchoring into the targeted EO-use socio-technical systems*:
    - Network anchoring by specifying *the actors identified as potential users by the pilot (b)*
    - Technological anchoring by specifying *the expected EO-based solution to be built (c):* problems expected to be taken by would-be users based on EO data, lists of requirements when identified;
    - Cognitive and normative institutional anchoring by specifying the *current capacity of niche/regime actors to handle EO data on their own (d)*, i.e. identifying their familiarity with EO data, and their design and development capacities, and the potential rules and standards that could potentially encourage the use of EO by these actors;

- Economic institutional anchoring by specifying the *nature of the relationships built with these actors so far (e),* i.e. the history of the relationship, the forms and intensity of the interactions, whether these interactions have been contractually formalised;

- *Integration of the pilot in the transition dynamics of EO-emergence socio-technical system*:

- *Overview of the pilot members' history and expertise (f)*: main research fields and expertise of pilot members, involvement in previous projects, role of e-shape in their overall trajectory;
- *Ability of the pilot to build a solution addressing a certain problem once specified (g)* for a first prototype and its further operationalisation: identifying potential development challenges, and the relationships to be created or reinforced to overcome these challenges.

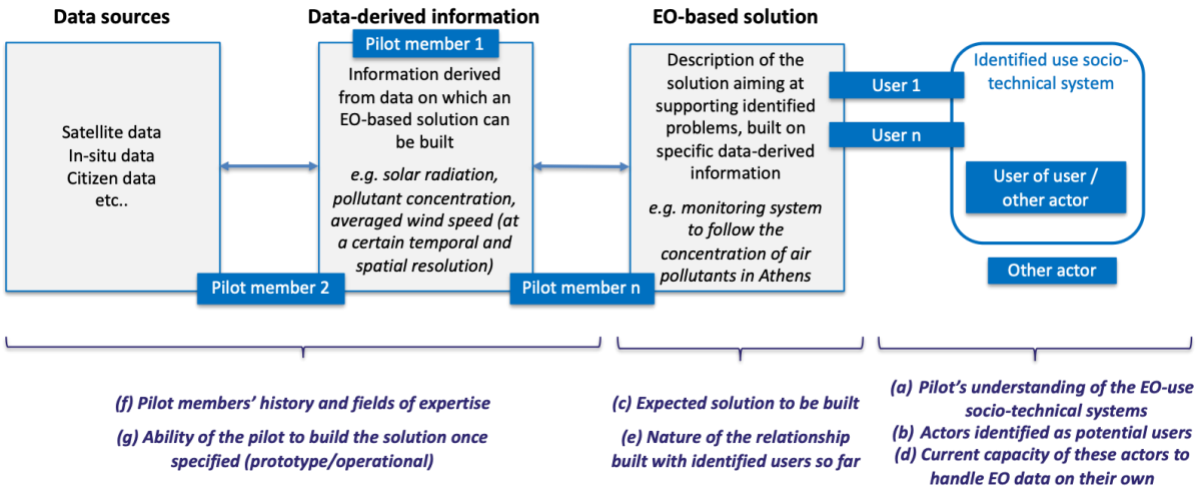

Figure 8: Template built for preliminary data analysis and link with the points of analysis guiding the interviews

Each interview was then organised following a semi-guided process. In a first phase, one member of our team made a brief reminder of our research goals, presented the synthetic template of the pilot and the associated preliminary analysis on the five main points mentioned above. A second phase was dedicated to a thorough discussion with the pilot to further understand its context following the points of analysis defined above. A final phase consisted in confirming the relevant types of co-design for the pilot context. To take into account the dynamics of socio-technical systems, two time horizons were considered: short-term for co-design types assessed as currently relevant for the pilot, and long-term for co-design types that the pilot expected to be relevant in the future.

After each interview our team updated the report written prior to the meeting, completing the description of the pilot's context and formalising the conclusion on the relevant co-design types at short-term and longer-term time horizons, occasionally resorting to the recordings of the interviews in cases where it was harder to understand the participants (e.g. due to technical or language issues). This report was systematically shared through the online project management platform and validated by all participants of the meeting.

	Data collection	**Data analysis**
Step 1 (Sept 2018-May 2019) Building an overall understanding of the empirical context	**Secondary sources** (pilots' application forms to the project, academic publications, websites) **Field notes** (kick-off meeting, informal interactions with project members)	**Construction of a template** for heterogeneous data compilation, **tested and validated by researchers and practitioners** (one-day meeting with a pilot and formalisation in a reviewed and shared deliverable report)
Step 2 (Nov 2019) Shared validation of the co-design framework with project coordinators	**2h meeting** with the seven showcase coordinators, the project management team and the work -package leaders	**Validation of the framework of four co-design types** by all participants of the meeting, and agreement on undertaking further assessment on all case studies
Step 3 (Nov 2019-July 2021) Assessment of the co-design framework for each pilot (in total 27 case studies)	**1h30 semi-guided interview with each pilot** (see table in Appendix for the details on the participants)	**Assessment of each pilot's context and identification of relevant co-design types validated by all researchers and practitioners:** **- Report written by researchers before each interview**, including specific points of analysis and filled-up template **- Report updated by researchers, shared and validated after each interview by all participants**, updating the description of the pilot's context and the conclusion on relevant co-design types at two time horizons

Table 4: Overview of the process followed for data collection and analysis

IV. Results

This section presents the findings that emerged through the analysis of the twenty-seven case studies. The outcomes consist in assessing the relevance of the co-design framework by (1) considering the portfolio of case studies to examine the respective relevance of all types; (2) considering case studies separately to analyse how the different co-design types combine with each other within a single case study.

A. Specifying the relevance of the four co-design types considering the portfolio of case studies

The outcomes of the co-design diagnosis for each pilot are shown in Table 3, confirming that all four types of co-design are relevant in the context of supporting EO data anchoring in heterogeneous socio-technical systems. The respective contributions of each type to multi-system anchoring and transition dynamics are detailed in the following paragraphs, synthesised in Table 4 and represented graphically in Figure 3.

Pilot	Short-term	Long-term	Pilot	Short-term	Long-term
#1	Type 1	Type 3	**#15**	Type 1 & 3	Type 4
#2	Type 1 with user 1 Type 2 for new user group	Type 1 with users 2 & 3 Type 3 with partner	**#16**	Type 2	Type 1
#3	Type 3	Type 4	**#17**	Type 1	Type 4
#4	Type 1 & 2	Type 4	**#18**	Type 1	Type 2
#5	Type 1	Type 4	**#19**	Type 1 or 2	Type 3 & 4
#6	Type 1	Type 3 & 4	**#20**	Type 1 & 3	Type 4
#7	Type 1 for user group 1	Type 2 for user group 2	**#21**	Type 1	Type 3 & 4
#8	Type 1 & 3 & 4	Type 4	**#22**	Type 1 & 4	Type 4
#9	Type 3 & 4 for user group 1 Type 1 & 3 for user group 2	Type 4	**#23**	Type 1	Type 3 & 4
#10	Type 2	Type 3	**#24**	Type 3	Type 4
#11	Type 1	Type 4	**#25**	Type 1	Type 3 & 4
#12	Type 3	Type 4	**#26**	Type 1	Type 4
#13	Type 3	Type 4	**#27**	Type 1	Type 4
#14	Type 1 & 4	Type 4			

Table 5: Validated assessment of co-design types for the twenty-seven case studies (pilots)

1) Co-design type 1

It is theoretically defined as a situation where the level of unknown associated with problem and solution spaces is considered as relatively low by the pilot. This situation applies when the pilot has already identified potential uses of EO by specific actors that are willing to interact further. A slight expansion of problem and solution spaces is thus needed to build problem-solution pairs.

In this perspective, co-design type 1 aims at supporting a certain form of *technological anchoring*, consisting in further detailing the specifications of potential EO-based solutions addressing problems that have been identified on a first basis. This entails enhancing some forms of *network anchoring*, by establishing a robust relationship between the identified users and the relevant members of the pilot. On the pilot's side, it appears that all members of the pilot are not necessarily involved: some of them might indeed focus on the development of a certain building block as defined by the actors interacting with users. On the users' side, the identified users proved to be either niche actors (e.g. a start-up willing to integrate EO data to estimate solar energy potential on building roofs), or regime actors (e.g. a national health agency having the project of building a data observatory for health-related issues). This

co-design type also involves some forms of institutional anchoring. Creating robust interactions between actors might indeed require the enrichment of the actors' respective perceptions and visions *(cognitive institutional anchoring).* The EO developers might for example need to push the users towards considering different ways of using EO data, such as for monitoring purposes (e.g. to assess the concentration of air pollutants), decision-support purposes (e.g. triggering existing pollution mitigation actions when a threshold is exceeded), or design-support purposes (e.g. designing new pollution mitigation actions by using EO data to build and assess various scenarios). This can also involve *economic institutional anchoring* by reshaping the existing forms of contracts or value chains (e.g. a public agency contracting with unusual types of actors), and *normative institutional anchoring* if integrating EO-based solutions in users' existing workflows and procedures requires the introduction of specific rules.

In terms of socio-technical systems' dynamics, co-design type 1 can be described as *supporting the identified dynamics of a given EO-use socio-technical system* by anchoring EO data to a niche or regime actor of this system (e.g. supporting the development of a niche related to solar resource self-consumption in urban areas aiming to accelerate the sustainability transition of the energy system).

2) Co-design type 2

It is theoretically defined as a situation where the level of unknown is considered high for the problem space and low for the solution space. This can especially occur when the pilot does not have sufficient knowledge on the EO-use socio-technical system to identify the potential of EO data for specific actors and/or the relationships with these actors seem difficult to establish (e.g. if actors are not willing to devote time to the interactions, or if previous interactions have been limited to one-shot exchanges). A significant exploration and expansion of the problem space is thus needed until problem-solution pairing can be initiated.

In this perspective, co-design type 2 implies a form of *technological anchoring* consisting in building and sharing the legitimacy of potential EO-based solutions. This might involve extensive efforts to identify the added-value of this new source of information compared to existing sources (e.g. correcting specific errors of existing instruments or capturing new physical phenomena), to ensure its technical validation (i.e. taking into account the limits of the measuring instrument, indicating a trustworthiness index associated with the provided information, precising if specific corrections have already been made), and

to establish its legitimacy within a given community (e.g. by facilitating the comparison of this new source of information compared to commonly used ones). This involves a certain form of *network anchoring* consisting in building relationships between relevant pilot members and actors of the EO-use socio-technical system, targeting a better understanding of this socio-technical system. It seems especially important to consider both regime and niche actors that provide complementary insights and collaboration opportunities. For example, in the offshore wind industry, regime actors such as utility companies developing and operating wind farms prove to be reluctant to use EO data but can share precious knowledge on the existing socio-technical system's rules and dynamics; whereas certain niche actors, such as specialised consultants, appear to be interested in using EO data to improve their wind-resource analysis workflows but struggle to make such data broadly accepted by the industry. This co-design type also entails some forms of *institutional anchoring*: *cognitive* as it consists in making emerge robust and reliable promises that might be associated with EO data, *economic* by creating specific forms of contracts (e.g. partnership to undertake a specific exploratory study), *normative* by encouraging the introduction of rules or standards to be shared in the EO-use socio-technical system (e.g. making EO data accepted as a legitimate source of information by banks assessing the expected performance of wind offshore projects).

In terms of socio-technical systems' dynamics, co-design type 2 consists in *identifying and linking up with the ongoing dynamics of an EO-use socio-technical system that are partly unknown* by assessing and creating the promising anchoring points.

3) Co-design type 3

It is theoretically defined as a situation where the level of unknown is considered low for the problem space and high for the solution space. This especially occurs when the pilot does have sufficient knowledge on the EO-use socio-technical system to be able to target a specific problem that EO data could address, but faces a number of issues related to the operationalisation or long-term maintenance of an EO-based solution addressing this identified problem. A large expansion of the solution space is thus needed until a point where the viability of problem-solution pairs can be ensured by EO-based solution developers.

In this perspective, co-design type 3 aims at establishing a certain form of *technological anchoring*, consisting in building the engineering required for operationalising EO-based solutions addressing identified problems. This involves *network anchoring* consisting in building specific relationships between relevant representatives of solution developers with other actors of the EO-emergence socio-technical system (e.g. providers of technical infrastructures, other data providers, organisations taking charge of commercialisation aspects). As noted for co-design type 1, all pilot members might not be concerned by this action but only the ones that need to reshape their partner network to sustain the required efforts towards long-term operationalisation of the solution. This co-design type also entails some forms of *institutional anchoring*: *cognitive* as it consists in elucidating new visions related to the engineering infrastructure required to sustain the EO-based solutions, *economic* by creating or reshaping specific forms of contracts (e.g. between a research lab and a spin-off supporting engineering and commercialisation aspects), *normative* by establishing certain forms of standards to be shared in the EO-emergence socio-technical system (e.g. standards related to the release of in-situ measurements).

In terms of socio-technical systems' dynamics, co-design type 3 consists in *leveraging ongoing dynamics of the EO-emergence socio-technical system* (in case ongoing dynamics bring new resources for supporting the operationalisation efforts, such as the emergence of cloud computing infrastructures), *and potentially influencing these dynamics* (for example by reshaping the existing network of actors through the creation of new forms of partnerships).

4) Co-design type 4

It is theoretically defined as a situation where the level of unknown is considered high for both problem and solution spaces. This situation might appear as particularly challenging as it involves the most substantial learning processes. Our empirical investigation reveals that this type of co-design appears as relevant only when the pilot already has significant knowledge on the EO-use socio-technical system through previous developments of EO-based solutions. A large exploration of both problem and solution spaces can thus be reasonably handled by the pilot with the objective of building new problem-solution pairs.

Co-design type 4 thus aims at enhancing a certain form of *technological anchoring* that consists in exploring the specifications of future EO-based solutions, by taking advantage of the existing ones. This

involves *network anchoring*, between relevant pilot members belonging to the EO-emergence socio-technical system and niche or regime actors of the EO-use socio-technical system, willing to take part in such a joint exploration effort. This can be illustrated by the pilot involved in building EO-based solutions to better predict the influx of sargassum algae on Caribbean beaches having negative environmental and economic impacts for local actors. The pilot already provides local actors with a 6-month ahead prediction bulletin of algae influxes, and aims to sustain and expand this solution, potentially by exploring several ways of stimulating the emerging actors involved in tackling the negative impacts of algae influxes. This co-design type also entails some forms of *institutional anchoring*: *cognitive* as it consists in elucidating new visions and promises associated with the future uses of EO data (e.g. exploring how sargassum forecasts could be provided following the model of weather forecasts), *economic* by creating or reshaping specific forms of contracts (e.g. extending the scope of partnership with existing users or creating new partnerships with others), *normative* by introducing certain forms of standards on future uses of EO data or on the related production and maintenance infrastructure.

In terms of socio-technical systems' dynamics, co-design type 4 can thus be described as a way of *identifying and stimulating future promising dynamics of both EO-emergence and EO-use socio-technical systems* (e.g. further mobilising actors involved in collecting or transforming algae).

5) Occurrences of the different types

Table 3 shows that several types appear more frequently than others, especially type 1 in the short term and type 4 in the long term. The predominance of type 1 over the other types in the short term can be largely explained by the initial configuration of the project as the pilots were expected to have identified at least one potential user organisation to join the project. But interestingly, the existence of other types shows that the pilots face heterogeneous issues beyond further specification of EO-based solutions for identified actors. This especially suggests that the identification of relevant users might be actually more complex than expected (type 2), or that the pilots are also concerned with other issues related to the long-term sustainability of developed EO-based solutions (type 3) and their further expansion (type 4). The predominance of type 4 in the long term seems consistent as this type requires stringent conditions that can only be met after primary problem-solution pair developments and associated learning on the EO-use socio-technical systems.

B. Analysing the combination of the four co-design types relevant for one case study

Another order of outcomes consists in analysing how the different types of co-design might combine within the context of one single pilot. First considering *a given time horizon* (short-term or long-term), it appears that several co-design types are relevant for one single pilot, hence underlining that *co-design types are not exclusive.* This seems pretty much consistent as each co-design type corresponds to complementary anchoring objectives concerning different actors. For example, at the same time, the pilot might be willing to strengthen a collaboration with identified actors (type 1), while also willing to explore other use cases of EO data in different EO-use socio-technical systems (type 2). Second, the results show that *the relevant co-design types are distributed over time*, confirming the usefulness of considering different time horizons. According to the definition of these types, a certain temporal trajectory could be expected: (1) *co-design type 2* to learn on partly unknown EO-use socio-technical systems and find relevant actors that would ensure linking up with the system's dynamics; (2) *co-design type 1* to build the adapted relationships with the relevant actors identified in type 2; (3) *co-design type 3* to build the engineering and infrastructure of the EO solution, in order to meet the lists of requirements identified in type 1; (4) *co-design type 4* to explore future uses and associated solutions based on the first uses built through previous co-design types.

However, the analysis of the pilots shows that this temporal trajectory cannot be systematically followed. Indeed, the pilots appear to regularly face unexpected changes within the EO-emergence or EO-use socio-technical systems, leading to a switch between different types of co-design. Several pilot cases give telling examples of this phenomenon. Indeed, some pilots had to transform the initially planned co-design type 1 or type 4 into a type 2, because the actors initially identified as relevant users had changed their priorities, declining their initial interest for collaboration (due to COVID-19 crisis in one case, due to the internal restructuration of the company in another case). Moreover, in some cases, a type 3 can be launched without being preceded by a thorough type 1. Indeed, the identified problem (required as a starting point of type 3) is not necessarily derived from specific user requirements, but might also result from the dynamics of the EO-emergence socio-technical system (e.g. to adapt to the identified competitors going towards a certain direction).

These results highlight that *the combination of relevant co-design types for a given case study might evolve over time*, especially to adapt to the continuous evolution of the different socio-technical systems.

Type	Design space	Multi-system anchoring	Link with STS dynamics
1	**Problem-related unknown:** identified problems that might need further specification **Solution-related unknown:** solutions requiring limited development efforts leveraging existing building blocks **Learning processes to build problem-solution pairs:** slight co-expansion of problem and solution spaces	**Technological:** enhancing the specifications of novelty-based solutions addressing identified problems **Network:** relevant novelty developers and identified users (niche/regime actors of novelty-use STS) **Institutional:** new articulation of technical and user-related aspects (cognitive), reshaping existing forms of contracts (economic), introducing rules limited to the identified users (normative)	**Supporting identified novelty-use STS dynamics** by anchoring the novelty into relevant niche/regime actors of this STS
2	**Problem-related unknown:** problems unclearly identified **Solution-related unknown:** solutions requiring limited development efforts leveraging existing building blocks **Learning processes to build problem-solution pairs:** large on problem space, limited on solution space	**Technological:** building and sharing the legitimacy of novelty-based solutions **Network:** relevant novelty developers & niche/regime actors of newly targeted novelty-use STS **Institutional:** building new visions and promises associated with the novelty uses (cognitive), building new forms of contracts (economic), introducing shared standards on the novelty uses (normative)	**Identifying and linking up with ongoing novelty-use STS dynamics** by assessing and creating the favourable entry points for anchoring
3	**Problem-related unknown:** identified problems that might need further specification **Solution-related unknown:** solutions requiring extensive development research efforts **Learning processes to build problem-solution pairs:** large on solution space, limited on problem space	**Technological:** building the engineering required for operationalising novelty-based solutions addressing identified problems **Network:** relevant novelty developers & niche/regime actors of initial novelty-emergence STS **Institutional:** building new visions for the novelty production and maintenance infrastructure (cognitive), building new forms of contracts (economic), introducing shared standards on the novelty production and maintenance infrastructure (normative)	**Leveraging or influencing initial novelty-emergence STS dynamics** to strengthen the anchoring viability based on strong engineering and operationalisation efforts
4	**Problem-related unknown:** problems unclearly identified **Solution-related unknown:** solutions requiring extensive development or research efforts **Learning processes to build problem-solution pairs:** large co-expansion of problem and solution spaces	**Technological:** exploring the specifications of future novelty-based solutions based on existing ones **Network:** relevant novelty developers & niche/regime actors of novelty-emergence and novelty-use STS **Institutional**: expanding visions and promises associated with the novelty based on existing uses (cognitive), reshaping/building contracts (economic), introducing shared standards on future novelty uses or production and maintenance infrastructure (normative)	**Identifying and stimulating future promising dynamics** of both initial novelty-emergence and novelty-use STS

Table 6: Synthesis of the relevance of each type of co-design according to the characteristics of the design space, the effects from a niche development perspective, and the contribution to the dynamics of EO and usage STS (STS used for "socio-technical system")

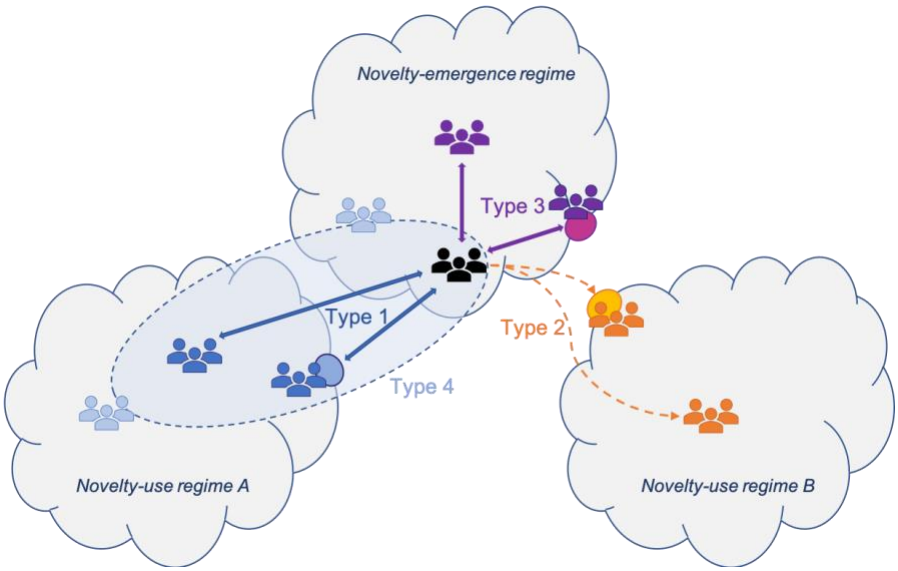

Figure 9: Graphical representation of the different co-design types based on the associated forms of network anchoring, inspired by the simplified representation of socio-technical systems (STS) proposed in Elzen et al. [32]. Regimes are represented by irregular forms to underline their constant transformations. Niches are located at the edges of regimes with a partial overlapping zone (e.g. accounting for shared technical components, actors operating both in the regime and the niche). Novelty developers are represented in black. ***Type 1*** *consists in enhancing network anchoring with identified relevant niche/regime actors of a novelty-use STS (in dark blue).* ***Type 2*** *consists in identifying niche/regime actors that might be promising anchoring points in a newly targeted novelty-use STS (in orange).* ***Type 3*** *consists in reshaping novelty anchoring with niche/regime actors of the novelty-emergence STS to sustain the engineering of identified novelty-based solutions (in purple).* ***Type 4*** *consists in expanding existing anchoring with existing or new niche/regime actors to explore future novelty-based solutions based on existing ones (in lighter blue).*

V. Discussion

In this section, we discuss key contributions of the paper, for both transition research and design and innovation management research, especially advancing research on collaborative innovation for grand challenges and sustainability transitions.

A. A multi-system co-design framework used as a diagnostic tool to identify relevant anchoring actions at different time horizons

Considering transition research, our empirical investigation of twenty-seven case studies suggests that the framework of four co-design types plays an important role in helping novelty developers clarify their anchoring strategy. This multi-system co-design framework indeed *provides novelty developers with a diagnostic tool to identify and sequence their anchoring efforts by focusing on certain aspects at once (as delimited in each co-design type)*. Each co-design type indeed frames the relevant forms of anchoring actions to conduct. The distribution of different types over time especially allows novelty developers to undertake ambitious and intricate learning processes by progressively addressing a reasonable amount of unknown at each step. An initial situation with a high level of unknown on both problem and solution

spaces could be addressed by first focusing efforts on problem expansion while limiting solution expansion (type 2), or the other way round (type 3), then followed by other types until viable problem-solution pairs can be reached. To be noted that the assessment of relevant types for each pilot only mirrors the 'ex-ante' vision of the pilot on the relevant forms of anchoring actions, it thus does not account for how these actions will be effectively implemented in reality. It especially appears that the pilot might actually change its strategy compared to what was initially planned, in reaction to potential unexpected developments (see for the example the switch from type 1 to type 2 due to a decreased interest of previously identified users). However, these changes do not undermine the guiding effect provided by the framework: having an explicit framework indeed enhances the ability of actors to more easily react to unexpected developments by switching from one frame of anchoring actions to another one, as defined in the different co-design types.

Our investigation also leads us to propose several enrichments of the anchoring concept as defined by Elzen *et al.* [32]. First, the co-design framework enriches our understanding on the possible interplay between the three forms of technological, network and institutional anchoring activities, that seems to play a critical role in ensuring a successful anchoring process [32], [33]. In each co-design type, the three forms of anchoring appear to be highly complementary. More specifically, it is worth noting that technological anchoring is considered as the driver of the anchoring strategy reflected on by the pilots, defining their main objectives according to the targeted developments of the technology. Interestingly, this technological anchoring takes broader forms than making the technology more specific to given user needs as mainly described in literature similarly to type 1 [32]–[34]: it can also involve building and sharing the legitimacy of the novelty at the larger scale of the socio-technical system (type 2), building the engineering required for operationalising novelty-based solutions addressing identified problems (type 3), or building the specifications of future novelty-based solutions by taking advantage of the existing ones (type 4). Although being the initial driver of the pilots' considerations, each co-design type also underlines how technological anchoring needs to be supplemented by specific forms of network anchoring and institutional anchoring.

A second enrichment of the anchoring concept concerns the nature of anchoring mechanisms. Elzen *et al.* [32] highlight the continuous and long-term efforts involved in newly or more firmly connecting a certain novelty to its environment, but consider that anchoring ends up when durable links are created.

Our study tends to suggest that such links cannot be actually considered as durable once and for all, as the environment into which the novelty anchors might evolve itself, potentially following unexpected new dynamics. To account for this evolutionary character of the environment into which the novelty anchors, we thus propose that anchoring should not be considered as a temporary process followed by durable links, but rather as an *ever-running process*, in which a novelty becomes newly connected, connected in a new way, or connected more firmly to a certain environment (regime or niche actor of a socio-technical system), *specifically considering that this environment is constantly evolving either stimulated by the connected novelty or due to other external factors*. Anchoring could thus be compared to a *grafting* process, underscoring the 'living' feature of the process drawing parallels with the biological world of plant grafting where a tissue of plant (the novelty) is added to growing plants (niche or regime of a socio-technical system considered as a living body) to make the plants further grow (transition dynamics) by taking advantage of the characteristics of the grafted tissue. In this perspective, anchoring does not only concern the newly targeted novelty-use socio-technical systems, but rather *jointly concerns the novelty-use socio-technical systems and the novelty-emergence socio-technical system*, as the latter also undergoes continuous transformations, thus requiring repeated efforts to reshape the connections of the novelty with its initial emergence environment.

By taking a closer look at managerial practices, our research also offers complementary insights on how to sustain fruitful interactions between different socio-technical systems through this process of anchoring. Sutherland *et al.* [34] especially unveil the benefits of creating a so-called 'fiat' regime supporting the interactions of two pre-existing regimes while preserving their respective structures and dynamics, but also underscore the inherent difficulty in ensuring its long-term sustainability. In the EO context, the investments of the European Commission in projects supporting the development of specific managerial practices such as co-design could be interpreted as the creation of such a 'fiat' regime between the EO-emergence regime and various EO-use regimes. However, differing from the cases reported by Sutherland *et al.* [34], this 'fiat' regime does not only consist in setting new rules and regulations guiding transition dynamics, but rather aims at developing the capacities of relevant actors of establishing robust and sustainable interactions between socio-technical systems. This gives another perspective to the creation of such a 'fiat' regime, that could be rather considered as a temporary

support to experiment, implement and embed good practices that could be later sustained by the actors themselves.

Regarding the issue of accelerating sustainability transitions, the impact of an anchoring strategy supported by co-design on transition dynamics is hard to directly assess. However, this paper proposes an improved understanding of how speeding up transitions could be operationally supported by specific managerial practices, "shying away from merely describing the temporal dynamics transitions" as encouraged by Sovacool and Geels [31]. The co-design framework indeed proves to be helpful in *developing the ability of actors to further interact with transition dynamics*, by continuously identifying, adapting to, enhancing and provoking dynamics at regime and niche levels of the different socio-technical systems. Recent works have also highlighted the role of novelty users that might have different profiles and contributions in shaping sustainability transitions [42]–[44]. Our research has taken the view of novelty developers but also suggests that the nature of considered users plays an important role in the anchoring process. It seems especially crucial to identify the type of users that will be adapted to a given design objective (e.g. considering a large variety of users with various competencies in co-design type 2, but a tendency to focus on relevant actors with sufficient novelty-related competencies for co-design type 1, or with broad exploration competencies for co-design type 4).

B. Enriching the forms of collaborative innovation for sustainability transitions: diagnostic dimension and multi-system perspective of co-design and resource-based vs. challenge-based collaborative innovation

Taking the perspective of design and innovation management literature, this research also contributes to deepening our understanding of the possible forms of collaborative innovation, especially in the context of addressing grand challenges.

First, our study sheds an original light on co-design. Literature largely reports on co-design by considering the protocols and range of possible toolkits to organise collective design sessions involving multiple actors such as probes, demonstration tools, or visual displays [45], [46], [71]–[73]. In these approaches, the nature of the involved actors is identified by the team implementing co-design and is considered as an initial input of the process. Our paper shows that co-design does not only consist in the actual organisation of collective design sessions but might also include a *diagnostic dimension* to identify what

are the relevant actors to be involved and for what purposes (as defined by the different co-design types of the framework developed in the paper). This diagnostic dimension intervenes as a preliminary phase prior to the actual implementation of design sessions, but it should also be regularly reassessed to take into account possible evolution of the socio-technical systems. The elicitation of such a diagnostic dimension is actually connected with the general context in which co-design takes place. Indeed, the co-design theoretical framework elaborated in this paper corresponds to a situation where co-design does not occur within one single socio-technical system - e.g. focusing on the transitions of the agricultural system [45], or the energy system [46] – but in a *multi-system* perspective involving interactions between multiple and evolving socio-technical systems. The heterogeneity between these socio-technical systems complexifies the range of potential actors to be involved in the design process, thus requiring supplementary efforts to identify and frame the relevant setting for subsequent collective design sessions. This multi-system perspective on co-design also echoes recent advances in other streams of works in innovation management, especially calling for further research on open innovation processes expanding the concept of openness (initially related to knowledge exchange across organisational boundaries) towards openness at an industry or larger societal scale [12], [14], [74], in which digital technologies play a specific role in crossing existing boundaries [75], [76].

Second, our research also leads us to better distinguish between different forms of collaborative innovation supporting sustainability transitions. The co-design approach described in this paper could be indeed categorised as a so-called *"resource-based"* form of collaborative innovation, differing from a so-called *"challenge-based"* form that literature more largely focuses on. The difference between resource-based and challenge-based lies in the nature of the trigger and driver of the collaborative process. In the resource-based case, the collaborative process starts from a specific novelty (in the sense given by Elzen *et al.* [32], i.e. a new technology, a new technical concept or a new socio-technical practice) that has been initially developed by an initial pool of actors (e.g. EO data) and attempts at transforming it into an actionable resource for a larger number of actors facing grand challenges. Taking a problem-solving perspective, the actors steering this form of collaborative innovation should not be reduced to mere problem solvers: they rather act as resource providers to support others in their own problem-solving processes. It is worth highlighting that this novelty is not directly a resource for grand challenges: such a resource is actually built through the collaborative innovation process. The objective

of *resource-based* collaborative innovation could thus be summarised as *creating the infrastructure and the conditions into which multiple actors might better tackle their own challenge-related problems*. By contrast, a *challenge-based* collaborative innovation process is triggered by challenge-related objectives and aims at *organising a joint exploration of solution paths responding to or progressing towards these challenge-related objectives*. Considering a problem-solving perspective, this would consist in formulating problems (although not clearly defined) and searching for solutions given this problem. This last perspective seems to prevail in the last management studies on collaborative innovation for grand challenges, e.g. depicting the involvement of advocacy groups in search consortia for joint search of solutions in EU-funded projects [9], NGOs transforming the social interactions of local groups to tackle social inequality [7], the creation of local ventures in response to the 2010 Haiti earthquake [10], or the implementation of dedicated open innovation approaches to enhance sustainability in the food and beverage industry [13], or to respond to the recent Covid-19 crisis [12]. In transition research, the anchoring mechanisms described so far also correspond to a challenge-based approach, as the novelty is anchored to address a certain identified challenge, such as designing an integrally sustainable animal production system [33], or the development of biogas to address waste management issues and renewable energy production [34]. It is worth noting that "challenge-based" does not necessarily mean that the problem is considered as fixed and clearly defined once and for all. Ferraro *et al.* [2] indeed underline that the objective of collaboration should be "repeated participation, inscription, and experimentation, continuously generating novelty and sustaining engagement", rather than "reaching some final conclusion".

C. Limits and perspectives for further research

Several limitations and perspectives for further research can be highlighted. First, it will be worth further testing and enriching our co-design framework based on additional empirical contexts (either in EO or in other fields). To be noted that we do not claim for exhaustivity with the typology of co-design. Indeed, it might be relevant in some contexts to refine the four co-design types in other sub-types to better address specific aspects of the design process under consideration. Nevertheless, although our results are derived from an investigation in the particular EO context, we can reasonably assume that they might also be insightful for other empirical contexts facing similar issues, especially considering how other forms of novelties (such as generic technologies with a large range of potential applications, but also

potentially new kinds of practices) could also be transformed into resources for various actors facing grand challenges.

Moreover, the co-design framework focuses on aspects of co-design preceding the implementation of dedicated collective design sessions. Further research is thus needed to explore the protocols and tools that could support the implementation of such collective design sessions in the context defined by each co-design type. This will involve examining existing co-design tools and practices more closely [46], [71], [73], [77], and potentially extend them or build new ones to handle the specificities and complexities of the design processes at stake. Beyond co-design literature, several works related to collaborative innovation for grand challenges already indicate that these protocols should take care of several important dimensions. This includes creating a certain form of "participatory architecture", providing the necessary structure and rules of engagement to ensure long-term involvement of the actors, as proposed by Ferraro *et al.* [2] and as recently advocated for the organisation of hackathon sessions in the context of COVID-19 crisis [12]. These protocols should also involve creating specific drivers and guidelines to go beyond the mere accumulation of facts (e.g. by merely collecting expressed user needs) and rather target second-order learning involving the expansion of cognitive frames, as encouraged by transition scholars [40].

Furthermore, the successful implementation of the co-design framework relies on specific conditions that would deserve further investigation. It indeed first relies on specific kinds of actors (corresponding to the "relevant pilot members" mentioned above) that are able to sustain anchoring processes in a multi-system perspective by circulating among the different socio-technical systems and articulating the variety of actors involved in building EO-based solutions and the actors identified as potential users of the solutions. Elzen *et al.* [32] have already identified specific forms of "hybrid actors" having a crucial role in bringing about anchoring processes through their ability to circulate between niche and regime. Hence, it will be worth wondering whether these crucial actors identified in the paper could be compared to such hybrid actors in a multi-regime perspective, or if they could be comparable to other figures of actors described in innovation management literature, such as innovation intermediaries [78]–[81], or "cross-application managers" involved in the development of generic technologies [82]. Second, it is also worth reflecting on the role we had as researchers in making the co-design framework operational for novelty developers. Indeed, the pilots found extremely useful to have a third-party actor providing them

with an external look at their activities, thus encouraging them to clarify their anchoring strategy beyond what they would spontaneously do on their own. This element calls for further research on how this co-design framework could be integrated in novelty developers' workflows on an operational basis beyond the project timeline and through which organisational forms.

Finally, more largely considering the issue of tackling grand challenges, how to monitor the progress of collaborative innovation towards this objective remains eminently challenging, be it in a challenge-based or resource-based perspective. In a challenge-based approach, Ferraro *et al.* [2] especially highlight the difficulty in accounting for the complex and heterogeneous visions of worth that could be potentially relevant to measure the progress towards an evolving target. In a resource-based perspective, the exact effect of the collaborative innovation process cannot be easily expressed in terms of quantified challenge-related targets (e.g. reducing emissions of n %). Our co-design framework, however, suggests that a resource-based approach could be monitored in terms of quality of the anchoring processes, by monitoring how technological specifications have been enriched, the new forms of partnerships that have been initiated, and the institutional rules that have been further entrenched (e.g. expansion of cognitive frames thanks to learning processes, introduction of new standards or economic relationships). Nevertheless, significant efforts are still needed to build operational indicators based on these considerations. In addition to exploring such indicators, it could also be interesting to further investigate how challenge-based and resource-based collaborative innovation might complement each other in tackling grand challenges.

VI. Conclusion

This article investigates how a multi-system co-design framework could help the developers of a certain novelty steer an anchoring strategy into multiple socio-technical systems. A framework of four co-design types has been derived from transition research and design and innovation management research. Each type corresponds to specific learning processes entailing various levels of unknown, and can be described according to the form of technological, network and institutional anchoring it aims to enhance, and the related interactions with the transition dynamics of the novelty-emergence and novelty-use socio-technical systems. The framework has been tested and enriched empirically in an inductive approach,

drawing on 27 case studies undertaking co-design efforts to anchor Earth observation data into multiple socio-technical systems.

Contributing to transition research, we have argued that this co-design framework provides novelty developers with a *diagnostic tool supporting them in clarifying their anchoring strategy by considering an evolving combination of different co-design types distributed over time*. This especially allows them to better handle the complexity of learning processes involved in sustainability transitions by sequencing their design efforts. We have also proposed several enrichments of the anchoring concept, shedding light on specific complementarities between the three forms of technological, network and institutional anchoring, and underscoring the continuous and endless character of anchoring, illustrated by the 'grafting' biological metaphor accounting for the constant evolution of socio-technical systems to which a novelty might come to be connected.

Contributing to design and innovation management research, our paper endeavours to enrich the current understanding of possible forms of collaborative innovation. Compared to existing literature on co-design, the co-design framework elaborated in this paper especially includes two original aspects: (1) it does not occur within the boundaries of a single socio-technical system but occurs across *multiple socio-technical systems*, echoing recent works in open innovation calling for further considerations on innovation processes at a large societal scale; (2) it does not focus on the actual organisation of collective design sessions but sheds light on a preliminary *diagnostic dimension*, that appears to be crucial in a multi-system perspective to identify the relevant actors to be involved in subsequent design sessions and for what purposes. Furthermore, concerning collaborative innovation for grand challenges, the paper introduces a distinction between *challenge-based* collaborative innovation organising collective action directed towards a targeted challenge, and *resource-based* collaborative innovation organising collective action to create the infrastructure and the conditions into which multiple actors might benefit from a certain resource (e.g. Earth observation data) to better address their respective challenge-related problems on their own.

These considerations open up interesting perspectives for practitioners. For policy-makers, our research especially encourages them to consider specific forms of innovation policies supporting the anchoring of promising novelties into multiple socio-technical systems. These policies could go beyond usual funding or regulation instruments, limited in terms of ensuring anchoring sustainability, *by focusing on enhancing*

the ability of actors to identify and interact with transition dynamics of these socio-technical systems on their own. That might include encouraging these actors to build a certain expertise in co-design as suggested in our paper, but also certainly other forms of competencies that could be further assessed. This echoes recent scholar discussions, suggesting that innovation policies for grand challenges should not consist neither in mere demand-oriented or supply-push instruments but rather in policies that would move away from orchestration towards creating conditions for others to self-organise and experiment around grand challenges [4]. Finally, our results offer insights for practitioners considering how Earth observation data, or other kinds of novelties with significant use potential, could contribute to tackling grand challenges. In this respect, our research suggests that building interactions between the novelty-emergence and novelty-use socio-technical systems might be beneficial but also requires intensive efforts that should not be overlooked. In this regard, navigating across heterogeneous socio-technical systems appears as a crucial capacity that might need to be strengthened and supported by specific actors. Further research in different empirical contexts could confirm the relevance of such approaches, possibly enrich them, and develop adapted guiding tools.

Appendix

SC	Pilot's rationale	Pilot's members (pilot leader in bold)	Targeted user groups	Interviewees for framework validation
Agriculture	Supporting global agricultural monitoring	**Independent research institute (Belgium)** Public research institute (Netherlands) Independent research institute (Austria) National meteorological institute (Germany) Non-profit public-private network (Greece) Public research institute (Israel)	National, regional and global agricultural organizations and administrations	- Team leader of the Agricultural Applications group (pilot leader organization) - Researcher in the Agricultural Applications group (pilot leader organization)
	Supporting farmers for CAP (Common Agricultural Policy) compliance and farm performance	**Public research institute (Greece)** Private ICT company (Greece) Independent research & technological organization (Belgium) Public research institute (Netherlands) Independent research institute (Austria) National meteorological institute (Germany)	Paying agencies Agriculture cooperatives Agro-consultants Insurance companies	Researcher in the institute of astronomy, astrophysics, space applications and remote sensing (pilot leader organization)

		Non-profit public-private network (Greece)		
	Supporting farmers with crop insurance services	**Public research institute (Netherlands)** Independent research institute (Austria) Public research institute (Ethiopia)	Digital finance & payment services' provider Insurance Companies Micro-Finance Institutions, Various key-government agencies	- Lead of the Spatial Agriculture and Food Security research theme (pilot leader organization) - Senior researcher in the department of natural resources (pilot leader organization)
	Supporting agriculture activities at farm level	**Independent research institute (Belgium)** Independent research institute (Austria) National meteorological institute (Germany) Non-profit public-private network (Greece) Public research institute (Israel)	Agro-consultants Policy Makers Agricultural cooperatives Agro-industries Farmers	- Team leader of the Agricultural Applications group (pilot leader organization) - Researcher in the Agricultural Applications group (pilot leader organization)
Health surveillance	Surveilling mercury pollution	**Public research institute (Italy)** Public research institute (Germany) Public research institute (Italy)	Health communities Conference of Parties (UN Minamata Convention) Local and regional authorities	- Research director of the institute (pilot leader organization) - 3 senior researchers in atmospheric pollution & remote sensing (pilot leader organization)
	Surveilling persistent organic pollutants	**Public research institute (Czech Republic)** Public research institute (Italy)	Policy makers Regional organisation groups and Conference of Parties (UN Stockholm Convention)	- Director of the Centre of toxic compounds & of the Stockholm Convention Regional Centre (pilot leader organization) - Head of the Data services core facility (pilot leader organization)
	Improved monitoring of air quality and related health issues, to support public health assessment and urban planning.	**Public research institute (Greece)** Private ICT company (Greece) National space agency (Germany) National meteorological institute (Finland) Public research institute (Italy) Independent research institute (Austria)	International organizations (UN, WHO) National authorities Cities and municipalities Private sector (insurance, real estate, industrial companies)	- Research director of the department for environmental research & sustainable development (pilot leader organization) - Researcher in the same department (pilot leader organization)
Renewable energy	Nowcasting and short-term forecasting of solar energy	**Public research institute (Greece)** Public research institute (Switzerland) Public research institute (France) Private company (France) International inter-governmental organization (Egypt)	Ministries of Electricity and Renewable Energy Power generation operators Power distribution and transmission operators	- Researcher in the department for environmental research & sustainable development (pilot leader organization) - Senior researcher in remote sensing & energy (Swiss pilot member)

	Encouraging high photovoltaics penetration in urban areas	**Public research institute (France)** Private company (France) National Space Agency (Germany)	Energy providers Citizens Collectivities & urban planners	- Director of the research institute (pilot leader organization) - Senior researcher in remote sensing & energy sector (pilot leader organization) - Research engineer expert in databases and web services (pilot leader organization) - 2 senior researchers in remote sensing & solar energy (German space agency)
	Providing wind resource assessment tools for the offshore wind industry	**Public research institute (Denmark)**	Offshore wind farm developers Offshore wind farm operators Consultants Research, academia, educators	- 2 senior researchers in remote sensing & wind energy (pilot leader organization) - Technical lead of wind resource assessment applications (pilot leader organization)
Biodiversity	Monitoring & modelling the states of ecological ecosystems by integrating remote sensing & in-situ data	**Public research institute (Italy)** Public research institute (Spain) Independent non-profit research organization (Netherlands) Independent non-profit research organization (Greece) Public research institute (Germany)	Technical staff and managers of European Protected Areas (PAs)	- Director of the department of geosciences and Earth resources (pilot leader organization) - 4 senior researchers in remote sensing & ecological ecosystems (pilot leader organization & representatives of the other pilot's organizations)
	Building a common registry of observation & experimentation facilities of ecological ecosystems	**Public agency (Austria)** Public research institute (Germany) Public research institute (Serbia)	Research communities Technical and scientific staff of project groups European and national conservation agencies	- Head of department for ecosystem research and environmental information management (pilot leader organization) - Data engineer in the same department (pilot leader organization) - Researcher in remote sensing & ecological ecosystems (German organization) - Manager of the product development center (Serbian organization)
	Providing harmonized sets of variables for biodiversity observation and conservation	**Public research institute (Germany)** Public research institute & agency (Finland) Public research institute (Netherlands)	Research communities Monitoring agencies that inform ministries	4 senior researchers in remote sensing & ecological ecosystems (pilot leader organization & representatives of each other pilot's organization)
Water	Providing historical and near-real time	**National meteorological & hydrological institute (Sweden)**	Geological institutes Water and marine authorities	Senior researcher in hydrology, leading research in forecasting of water

	information for a number of hydrological variables	Public research institute (Luxembourg)		variables (pilot leader organization)
	Estimating flood hazard at a large-scale	**Public research institute (Luxembourg)** National meteorological & hydrological institute (Sweden)	Members of the Global Flood partnerships: Research and meteorogical institutes R&D Companies Governmental authorities	Senior researcher, leading the group on remote sensing & natural resources modelling (pilot leader organization)
	Providing a near-real time visibility score for specific diving locations	**Independent research institute (UK)**	Diving centers, commercial divers	- Senior researcher in remote sensing (pilot leader organization) - Linux data analyst & support engineer (pilot leader organization) - Data and web services engineer (pilot leader organization)
	Predicting the landing areas and severity of the sargassum algae season	**Private company (France)**	Local authorities Research community Private sector (insurance companies, tourism, algae valorization)	Project manager in the environmental applications department (pilot leader organization)
	Improving monitoring and regulation of fishing activities in the Northeast Atlantic	**Public agency (Portugal)** Private company (Portugal)	Fishermen associations NGOs International organizations scientific communities Regional and national authorities	- 2 senior researchers in marine science (pilot leader organization) - 2 project engineers (private company)
Disaster resilience	Improving monitoring of volcanic eruptions	**Public research institute (Italy)** Public research institute (Greece) Public research institute (Italy) Public research institute (France) National meteorological institute (Finland) National meteorological institute (Iceland)	Special units in meteorological institutes Air traffic managers Civil protection authorities Local authorities Policy makers	Researcher in remote sensing & environmental analysis, leading an international working group on aerosol research (pilot leader organization)
	Developing a tool for fire and risk assessment and supervision scenarios	**Public research institute (Italy)** European body – SatCen Public research institute (Greece)	National and regional civil protection authorities Hydro-Meteorological Agencies	Research director - expert in atmospheric modelling and statistical analysis of extreme events (pilot leader organization)
	Assessing geohazard vulnerability of cities and critical infrastructures	**Public agency – geological survey (Spain)** European body – SatCen Private company (Italy)	Urban managers and civil protection authorities Energy and infrastructure companies Policy makers	Researcher in remote sensing & geosciences (pilot leader organization)

	Assessing geo-hazard vulnerability of agriculture	**Public research institute (Greece)** Private ICT company (Greece) Independent research institute (Austria)	Insurance company Farming cooperatives	- Senior researcher in remote sensing for hydrology, floods, natural disasters (pilot leader organization) - Researcher in remote sensing and atmospheric modelling (pilot leader organization)
Climate	Providing territorial and ocean carbon and greenhouse gas fluxes information to support the Global Carbon Project	**Research Infrastructure (headquarters in Finland)** Public research institute (Germany) Public research institute (Norway) Public research institute (UK) Public research institute (UK) National meteorological institute (Finland) Public research institute (Finland) Non-profit research institute (Italy)	Global Carbon Project Research and international organizations related to GHG and ocean carbon emissions	Researcher with specific focus on data analysis for ecology (pilot leader organization)
	Providing information to municipalities on heat waves, heavy precipitations and extreme weather events	**National meteorological institute (Germany)** National meteorological institute (Finland) National meteorological institute (Austria)	Cities and municipalities Consultancy company for urban climatology and wind research	- Researcher in the department of climate and environment consultancy (pilot leader organization) - 2 senior researchers in seasonal and climate applications (Finnish meteorological institute) - Head of the data center for climate change (Austrian meteorological institute)
	Providing seasonal forecasting of forest harvest conditions to support forest industry	**National meteorological institute (Finland)** Public research institute (Finland)	Private companies in forest management support R&D company	Senior researcher in seasonal and climate applications (pilot leader organization)
	Supporting hydropower companies by better predicting hydrological conditions	**National meteorological institute (Finland)**	Hydropower or energy companies	Senior researcher in remote sensing and geoscience (pilot leader organization)
	Helping transportation and tourism sectors to better prepare for seasonal changes	**National meteorological institute (Finland)** Public research institute (Greece)	Tyres companies Tourism stakeholders	- 2 senior researchers in seasonal and climate applications (pilot leader organization) - 2 researchers (senior & research assistant) in remote sensing & atmospheric modelling (Greek organization)

Acknowledgement

This research work has been carried out within the *e-shape* project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 820852. The authors warmly thank all *e-shape* partners for their contributions, and especially Thierry Ranchin, scientific coordinator of the project.

References

[1] G. George, J. Howard-Grenville, A. Joshi, and L. Tihanyi, 'Understanding and Tackling Societal Grand Challenges through Management Research', *Academy of Management Journal*, vol. 59, no. 6, pp. 1880–1895, Dec. 2016, doi: 10.5465/amj.2016.4007.

[2] F. Ferraro, D. Etzion, and J. Gehman, 'Tackling Grand Challenges Pragmatically: Robust Action Revisited', *Organization Studies*, vol. 36, no. 3, pp. 363–390, Mar. 2015, doi: 10.1177/0170840614563742.

[3] D. Foray, D. C. Mowery, and R. R. Nelson, 'Public R&D and social challenges: What lessons from mission R&D programs?', *Research Policy*, vol. 41, no. 10, pp. 1697–1702, Dec. 2012, doi: 10.1016/j.respol.2012.07.011.

[4] S. Kuhlmann and A. Rip, 'Next-Generation Innovation Policy and Grand Challenges', *Science and Public Policy*, vol. 45, no. 4, pp. 448–454, Aug. 2018, doi: 10.1093/scipol/scy011.

[5] I. Montiel, A. Cuervo-Cazurra, J. Park, R. Antolín-López, and B. W. Husted, 'Implementing the United Nations' Sustainable Development Goals in international business', *J Int Bus Stud*, vol. 52, no. 5, pp. 999–1030, Jul. 2021, doi: 10.1057/s41267-021-00445-y.

[6] K. M. Eisenhardt, M. E. Graebner, and S. Sonenshein, 'Grand Challenges and Inductive Methods: Rigor without Rigor Mortis', *Academy of Management Journal*, vol. 59, no. 4, pp. 1113–1123, Aug. 2016, doi: 10.5465/amj.2016.4004.

[7] J. Mair, M. Wolf, and C. Seelos, 'Scaffolding: A Process of Transforming Patterns of Inequality in Small-Scale Societies', *Academy of Management Journal*, vol. 59, no. 6, pp. 2021–2044, Dec. 2016, doi: 10.5465/amj.2015.0725.

[8] J. P. Doh, P. Tashman, and M. H. Benischke, 'Adapting to Grand Environmental Challenges Through Collective Entrepreneurship', *Academy of Management Perspectives*, vol. 33, no. 4, pp. 450–468, Nov. 2018, doi: 10.5465/amp.2017.0056.

[9] A. Ø. Olsen, W. Sofka, and C. Grimpe, 'Coordinated Exploration for Grand Challenges: The Role of Advocacy Groups in Search Consortia', *Academy of Management Journal*, vol. 59, no. 6, pp. 2232–2255, Sep. 2016, doi: 10.5465/amj.2015.0730.

[10] T. A. Williams and D. A. Shepherd, 'Building Resilience or Providing Sustenance: Different Paths of Emergent Ventures in the Aftermath of the Haiti Earthquake', *Academy of Management Journal*, vol. 59, no. 6, pp. 2069–2102, May 2016, doi: 10.5465/amj.2015.0682.

[11] J. M. Ahn, N. Roijakkers, R. Fini, and L. Mortara, 'Leveraging open innovation to improve society: past achievements and future trajectories', *R&D Management*, vol. 49, no. 3, pp. 267–278, 2019, doi: 10.1111/radm.12373.

[12] A. Bertello, M. L. A. M. Bogers, and P. De Bernardi, 'Open innovation in the face of the COVID-19 grand challenge: insights from the Pan-European hackathon "EUvsVirus"', *R&D Management*, p. 10.1111/radm.12456, Feb. 2021, doi: 10.1111/radm.12456.
[13] M. Bogers, H. Chesbrough, and R. Strand, 'Sustainable open innovation to address a grand challenge : Lessons from Carlsberg and the Green Fiber Bottle', *British Food Journal*, vol. 122, no. 5, pp. 1505–1517, Jan. 2020, doi: 10.1108/BFJ-07-2019-0534.
[14] J. Sims, J. Gichoya, G. Bhardwaj, and M. Bogers, 'Write Code, Save Lives: How a Community Uses Open Innovation to Address a Societal Challenge', *R&D Management*, vol. 49, no. 3, pp. 369–382, 2019, doi: 10.1111/radm.12338.
[15] A. J. Porter, P. Tuertscher, and M. Huysman, 'Saving Our Oceans: Scaling the Impact of Robust Action Through Crowdsourcing', *Journal of Management Studies*, vol. 57, no. 2, pp. 246–286, 2020, doi: 10.1111/joms.12515.
[16] J. Köhler *et al.*, 'An agenda for sustainability transitions research: State of the art and future directions', *Environmental Innovation and Societal Transitions*, vol. 31, pp. 1–32, Jun. 2019, doi: 10.1016/j.eist.2019.01.004.
[17] J. Grin, J. Rotmans, and J. Schot, *Transitions to Sustainable Development: New Directions in the Study of Long Term Transformative Change*. Routledge, 2010.
[18] M. Zolfagharian, B. Walrave, R. Raven, and A. G. L. Romme, 'Studying transitions: Past, present, and future', *Research Policy*, vol. 48, no. 9, p. 103788, Nov. 2019, doi: 10.1016/j.respol.2019.04.012.
[19] J. Markard, R. Raven, and B. Truffer, 'Sustainability transitions: An emerging field of research and its prospects', *Research Policy*, vol. 41, no. 6, pp. 955–967, Jul. 2012, doi: 10.1016/j.respol.2012.02.013.
[20] A. Smith, J.-P. Voß, and J. Grin, 'Innovation studies and sustainability transitions: The allure of the multi-level perspective and its challenges', *Research Policy*, vol. 39, no. 4, pp. 435–448, May 2010, doi: 10.1016/j.respol.2010.01.023.
[21] J. Farla, J. Markard, R. Raven, and L. Coenen, 'Sustainability transitions in the making: A closer look at actors, strategies and resources', *Technological Forecasting and Social Change*, vol. 79, no. 6, pp. 991–998, Jul. 2012, doi: 10.1016/j.techfore.2012.02.001.
[22] F. W. Geels, 'The multi-level perspective on sustainability transitions: Responses to seven criticisms', *Environmental Innovation and Societal Transitions*, vol. 1, no. 1, pp. 24–40, Jun. 2011, doi: 10.1016/j.eist.2011.02.002.
[23] F. W. Geels, 'Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study', *Research Policy*, vol. 31, no. 8, pp. 1257–1274, Dec. 2002, doi: 10.1016/S0048-7333(02)00062-8.
[24] F. W. Geels, 'The dynamics of transitions in socio-technical systems: A multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860–1930)', *Technology Analysis & Strategic Management*, vol. 17, no. 4, pp. 445–476, Dec. 2005, doi: 10.1080/09537320500357319.
[25] F. Geels, 'Co-evolution of technology and society: The transition in water supply and personal hygiene in the Netherlands (1850–1930)—a case study in multi-level perspective', *Technology in Society*, vol. 27, no. 3, pp. 363–397, Aug. 2005, doi: 10.1016/j.techsoc.2005.04.008.
[26] F. W. Geels, 'From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory', *Research Policy*, vol. 33, no. 6, pp. 897–920, Sep. 2004, doi: 10.1016/j.respol.2004.01.015.

[27] F. W. Geels and J. Schot, 'Typology of sociotechnical transition pathways', *Research Policy*, vol. 36, no. 3, pp. 399–417, Apr. 2007, doi: 10.1016/j.respol.2007.01.003.
[28] S. Lenfle and J. Söderlund, 'Project-oriented agency and regeneration in socio-technical transition: Insights from the case of numerical weather prediction (1978–2015)', *Research Policy*, vol. 51, no. 3, p. 104455, Apr. 2022, doi: 10.1016/j.respol.2021.104455.
[29] G. Papachristos, A. Sofianos, and E. Adamides, 'System interactions in socio-technical transitions: Extending the multi-level perspective', *Environmental Innovation and Societal Transitions*, vol. 7, pp. 53–69, Jun. 2013, doi: 10.1016/j.eist.2013.03.002.
[30] B. K. Sovacool, 'How long will it take? Conceptualizing the temporal dynamics of energy transitions', *Energy Research & Social Science*, vol. 13, pp. 202–215, Mar. 2016, doi: 10.1016/j.erss.2015.12.020.
[31] B. K. Sovacool and F. W. Geels, 'Further reflections on the temporality of energy transitions: A response to critics', *Energy Research & Social Science*, vol. 22, pp. 232–237, Dec. 2016, doi: 10.1016/j.erss.2016.08.013.
[32] B. Elzen, B. van Mierlo, and C. Leeuwis, 'Anchoring of innovations: Assessing Dutch efforts to harvest energy from glasshouses', *Environmental Innovation and Societal Transitions*, vol. 5, pp. 1–18, Dec. 2012, doi: 10.1016/j.eist.2012.10.006.
[33] B. Elzen and B. Bos, 'The RIO approach: Design and anchoring of sustainable animal husbandry systems', *Technological Forecasting and Social Change*, vol. 145, pp. 141–152, Aug. 2019, doi: 10.1016/j.techfore.2016.05.023.
[34] L.-A. Sutherland, S. Peter, and L. Zagata, 'Conceptualising multi-regime interactions: The role of the agriculture sector in renewable energy transitions', *Research Policy*, vol. 44, no. 8, pp. 1543–1554, Oct. 2015, doi: 10.1016/j.respol.2015.05.013.
[35] R. Raven, 'Co-evolution of waste and electricity regimes: Multi-regime dynamics in the Netherlands (1969–2003)', *Energy Policy*, vol. 35, no. 4, pp. 2197–2208, Apr. 2007, doi: 10.1016/j.enpol.2006.07.005.
[36] R. Raven and G. Verbong, 'Multi-Regime Interactions in the Dutch Energy Sector: The Case of Combined Heat and Power Technologies in the Netherlands 1970–2000', *Technology Analysis & Strategic Management*, vol. 19, no. 4, pp. 491–507, Jul. 2007, doi: 10.1080/09537320701403441.
[37] A. Hatchuel, P. Le Masson, Y. Reich, and E. Subrahmanian, 'Design theory: a foundation of a new paradigm for design science and engineering', *Res Eng Design*, vol. 29, no. 1, pp. 5–21, Jan. 2018, doi: 10.1007/s00163-017-0275-2.
[38] F. Geels and R. Raven, 'Non-linearity and Expectations in Niche-Development Trajectories: Ups and Downs in Dutch Biogas Development (1973–2003)', *Technology Analysis & Strategic Management*, vol. 18, no. 3–4, pp. 375–392, Jul. 2006, doi: 10.1080/09537320600777143.
[39] R. Kemp, J. Schot, and R. Hoogma, 'Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management', *Technology Analysis & Strategic Management*, vol. 10, no. 2, pp. 175–198, Jan. 1998, doi: 10.1080/09537329808524310.
[40] J. Schot and F. W. Geels, 'Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy', *Technology Analysis & Strategic Management*, vol. 20, no. 5, pp. 537–554, Sep. 2008, doi: 10.1080/09537320802292651.
[41] R. Hoogma, R. Kemp, J. Schot, and B. Truffer, *Experimenting for Sustainable Transport*. Taylor & Francis, 2002. doi: 10.4324/9780203994061.

[42] G. George, R. K. Merrill, and S. J. D. Schillebeeckx, 'Digital Sustainability and Entrepreneurship: How Digital Innovations Are Helping Tackle Climate Change and Sustainable Development', *Entrepreneurship Theory and Practice*, p. 104225871989942, Jan. 2020, doi: 10.1177/1042258719899425.
[43] A. Smith and R. Raven, 'What is protective space? Reconsidering niches in transitions to sustainability', *Research Policy*, vol. 41, no. 6, pp. 1025–1036, Jul. 2012, doi: 10.1016/j.respol.2011.12.012.
[44] E. T. Berthet, C. Barnaud, N. Girard, J. Labatut, and G. Martin, 'How to foster agroecological innovations? A comparison of participatory design methods', *Journal of Environmental Planning and Management*, vol. 59, no. 2, pp. 280–301, Feb. 2016, doi: 10.1080/09640568.2015.1009627.
[45] P. Beguin, M. Cerf, and L. Prost, 'Co-design as an emerging distributed dialogical process between users and designers', in *System innovations, knowledge regimes, and design practices towards transitions for sustainable agriculture*, Paris, France: INRA-Département Sciences pour l'Action et le Développement (SAD), 2012.
[46] S. Hyysalo, T. Marttila, S. Perikangas, and K. Auvinen, 'Codesign for transitions governance: A mid-range pathway creation toolset for accelerating sociotechnical change', *Design Studies*, vol. 63, pp. 181–203, Jul. 2019, doi: 10.1016/j.destud.2019.05.002.
[47] F. W. Geels and J. Schot, 'The Dynamics of Transitions A Socio-Technical Perspective', in *Transitions to Sustainable Development: New Directions in the Study of Long Term Transformative Change*, Routledge, 2010, pp. 11–101.
[48] P. Harborne, C. Hendry, and J. Brown, 'The Development and Diffusion of Radical Technological Innovation: The Role of Bus Demonstration Projects in Commercializing Fuel Cell Technology', *Technology Analysis & Strategic Management*, vol. 19, no. 2, pp. 167–188, Mar. 2007, doi: 10.1080/09537320601168060.
[49] H. S. Brown, P. J. Vergragt, K. Green, and L. Berchicci, 'Bounded socio-technical experiments (BSTEs): higher order learning for transitions towards sustainable mobility', in *System innovation and the transition to sustainability: theory, evidence and policy*, Edward Elgar Publishing, 2004, pp. 191–219.
[50] M. Maher and H.-H. Tang, 'Co-evolution as a computational and cognitive model of design', *Res Eng Design*, vol. 14, no. 1, pp. 47–64, Feb. 2003, doi: 10.1007/s00163-002-0016-y.
[51] M. L. Maher, 'Creative design using a genetic algorithm', in *Computing in Civil Engineering*, 1994, pp. 2014–2021.
[52] M. L. Maher, J. Poon, and S. Boulanger, 'Formalising design exploration as co-evolution', in *Advances in formal design methods for CAD*, Springer, 1996, pp. 3–30.
[53] K. Dorst and N. Cross, 'Creativity in the design process: co-evolution of problem–solution', *Design Studies*, vol. 22, no. 5, pp. 425–437, Sep. 2001, doi: 10.1016/S0142-694X(01)00009-6.
[54] E. von Hippel and G. von Krogh, 'CROSSROADS—Identifying Viable "Need–Solution Pairs": Problem Solving Without Problem Formulation', *Organization Science*, vol. 27, no. 1, pp. 207–221, Dec. 2015, doi: 10.1287/orsc.2015.1023.
[55] R. M. Stock-Homburg, S. L. M. Heald, C. Holthaus, N. L. Gillert, and E. von Hippel, 'Need-solution pair recognition by household sector individuals: Evidence, and a cognitive mechanism explanation', *Research Policy*, vol. 50, no. 8, p. 104068, Oct. 2021, doi: 10.1016/j.respol.2020.104068.
[56] H. A. Simon, *Models of Thought*. Yale University Press, 1979.
[57] A. Hatchuel, 'Towards Design Theory and expandable rationality: The unfinished program of Herbert Simon', *Journal of management and governance*, vol. 5, no. 3/4, pp. 260–273, 2001.

[58] A. Hatchuel and B. Weil, 'C-K design theory: an advanced formulation', *Research in Engineering Design*, vol. 19, no. 4, pp. 181–192, Jan. 2009, doi: 10.1007/s00163-008-0043-4.
[59] P. Le Masson and B. Weil, 'Design theories as languages of the unknown: insights from the German roots of systematic design (1840–1960)', *Res Eng Design*, vol. 24, no. 2, pp. 105–126, Apr. 2013, doi: 10.1007/s00163-012-0140-2.
[60] R. K. Yin, *Case Study Research: Design and Methods*. SAGE, 2009.
[61] M. T. Borzacchiello and M. Craglia, 'The impact on innovation of open access to spatial environmental information: a research strategy', *International Journal of Technology Management*, vol. 60, no. 1/2, p. 114, 2012, doi: 10.1504/IJTM.2012.049109.
[62] M. Zotti and C. L. Mantia, 'Open Data from Earth Observation: from Big Data to Linked Open Data, through INSPIRE', *Journal of e-Learning and Knowledge Society*, vol. 10, no. 2, May 2014, Accessed: Nov. 19, 2020. [Online]. Available: https://www.learntechlib.org/p/148274/
[63] J. Vogel, E. McNie, and D. Behar, 'Co-producing actionable science for water utilities', *Climate Services*, vol. 2–3, pp. 30–40, Sep. 2016, doi: 10.1016/j.cliser.2016.06.003.
[64] C. M. Goodess *et al.*, 'Advancing climate services for the European renewable energy sector through capacity building and user engagement', *Climate Services*, vol. 16, p. 100139, Dec. 2019, doi: 10.1016/j.cliser.2019.100139.
[65] T. Ranchin, L. Ménard, Fichaux, Nicolas, Reboul, Mathieu, and all e-shape partners, 'e-shape – EuroGEO Showcases: application powered by Europe contribution to EuroGEO and to the development of the EO industry', Brussels, Belgium, Jul. 2021.
[66] A. B. (Rami) Shani, S. A. Mohrman, W. A. Pasmore, B. Stymne, and N. Adler, *Handbook of Collaborative Management Research*. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications, Inc., 2008. doi: 10.4135/9781412976671.
[67] W. A. Pasmore, B. Stymne, A. B. (Rami) Shani, S. A. Mohrman, and N. Adler, 'The Promise of Collaborative Management Research', in *Handbook of Collaborative Management Research*, 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications, Inc., 2008, pp. 6–32. doi: 10.4135/9781412976671.n1.
[68] S. Cirella, M. Guerci, and A. B. Shani, 'A Process Model of Collaborative Management Research: The Study of Collective Creativity in the Luxury Industry', *Syst Pract Action Res*, vol. 25, no. 3, pp. 281–300, Jun. 2012, doi: 10.1007/s11213-011-9220-x.
[69] H. Dumez, 'How to treat the material: The templates', in *Comprehensive Research: A Methodological and Epistemological Introduction to Qualitative Research*, CBS Press, 2016, pp. 85–94.
[70] R. Barbier, P. Le Masson, and B. Weil, 'Deliverable 2.1 : Initial model for e-shape co-design. Deliverable for e-shape project.', 2019. https://e-shape.eu/index.php/resources
[71] E. B.-N. Sanders and P. J. Stappers, 'Co-creation and the new landscapes of design', *CoDesign*, vol. 4, no. 1, pp. 5–18, Mar. 2008, doi: 10.1080/15710880701875068.
[72] E. B.-N. Sanders and P. J. Stappers, 'Probes, toolkits and prototypes: three approaches to making in codesigning', *CoDesign*, vol. 10, no. 1, pp. 5–14, Jan. 2014, doi: 10.1080/15710882.2014.888183.
[73] A. Botero and S. Hyysalo, 'Ageing together: Steps towards evolutionary co-design in everyday practices', *CoDesign*, vol. 9, no. 1, pp. 37–54, Mar. 2013, doi: 10.1080/15710882.2012.760608.
[74] M. Bogers *et al.*, 'The open innovation research landscape: established perspectives and emerging themes across different levels of analysis', *Industry and Innovation*, vol. 24, no. 1, pp. 8–40, Jan. 2017, doi: 10.1080/13662716.2016.1240068.

[75] E. Enkel, M. Bogers, and H. Chesbrough, 'Exploring open innovation in the digital age: A maturity model and future research directions', *R&D Management*, vol. 50, no. 1, pp. 161–168, 2020, doi: 10.1111/radm.12397.
[76] I. Hellemans, A. J. Porter, and D. Diriker, 'Harnessing digitalization for sustainable development: Understanding how interactions on sustainability-oriented digital platforms manage tensions and paradoxes', *Business Strategy and the Environment*, vol. 31, no. 2, pp. 668–683, 2022, doi: 10.1002/bse.2943.
[77] L.-E. Dubois, P. Le Masson, B. Weil, and P. Cohendet, 'From organizing for innovation to innovating for organization: how co-design fosters change in organizations.', presented at the International Product Development Management Conference, Limerick, Ireland, 2014.
[78] J. Howells, 'Intermediation and the role of intermediaries in innovation', *Research Policy*, vol. 35, no. 5, pp. 715–728, Jun. 2006, doi: 10.1016/j.respol.2006.03.005.
[79] L. Klerkx and C. Leeuwis, 'Establishment and embedding of innovation brokers at different innovation system levels: Insights from the Dutch agricultural sector', *Technological Forecasting and Social Change*, vol. 76, no. 6, pp. 849–860, Jul. 2009, doi: 10.1016/j.techfore.2008.10.001.
[80] J. Stewart and S. Hyysalo, 'Intermediaries, users and social learning in technological innovation', *Int. J. Innov. Mgt.*, vol. 12, no. 03, pp. 295–325, Sep. 2008, doi: 10.1142/S1363919608002035.
[81] M. Agogué *et al.*, 'Explicating the role of innovation intermediaries in the "unknown": a contingency approach', *Journal of Strategy and Management*, vol. 10, no. 1, pp. 19–39, Feb. 2017, doi: 10.1108/JSMA-01-2015-0005.
[82] O. Kokshagina, P. L. Masson, B. Weil, and P. Cogez, 'Portfolio Management in Double Unknown Situations: Technological Platforms and the Role of Cross-Application Managers', *Creativity and Innovation Management*, vol. 25, no. 2, pp. 270–291, 2016, doi: https://doi.org/10.1111/caim.12121.

ACADEMIC CONFERENCES

Building the generativity of data to support the dynamics of multiple ecosystems: the case of Earth-observation data

Raphaëlle BARBIER
MINES ParisTech – PSL University
raphaelle.barbier@mines-paristech.fr

Pascal LE MASSON
MINES ParisTech – PSL University

Sylvain LENFLE
CNAM & Ecole Polytechnique

Benoit WEIL
MINES ParisTech – PSL University

Abstract

The potential of data in generating multiple usages for a large variety of actors has been largely acknowledged by practitioners and researchers. However, this so-called 'generativity' potential is not an intrinsic property of data but needs to be built and sustained through dedicated efforts. Given the high heterogeneity and unpredictability of potential data usages, it seems increasingly difficult for data producers or data users to support such efforts on their own in a long-term perspective. Based on multiple case studies in the Earth observation context, our study especially unveils the existence of specific economic actors, coined 'data generativity builders', that act as innovation intermediaries to ensure a continuous and repeated transformation of data-usage pairs, involving specific forms of brokering, configuring and facilitating actions: they act as a buffer zone between the different ecosystems related to data production and data usages, ensuring some forms of circulation without

creating direct dependences (*brokering*); they continuously design and redesign data-usage pairs, not only focusing on recontextualizing data for a new usage but also building a generic information core common to multiple usages (*configuring*); they create spaces for others to act, especially by identifying relevant capacity building actions through dedicated investigations on the specificities of each usage context (*facilitating*). This study offers interesting perspectives on how to unleash the potential of data in complex ecosystems of actors, going beyond a simple valorization logic.

I. Introduction

In recent years, the development of internet, new sensors, and computational means has dramatically increased the flow of data in almost every business, industry and research area. This phenomenon has been largely described as “digitalization” (Björkdahl and Holmén, 2019; Cappa et al., 2021), or “big data” (Bharadwaj and Noble, 2017; Blackburn et al., 2017; Chen et al., 2012). Scholars have extensively emphasized the potential of data in generating a number of usages for a large variety of actors spanning heterogeneous sectors. As highlighted by research in information systems, digital artifacts are characterized by their ‘open-endedness’, or ‘unboundedness’, insofar as data are perpetually ‘editable’ and can continuously generate new usages, different from what they were initially produced for (Aaltonen and Tempini, 2014; Kallinikos et al., 2013; Yoo et al., 2012, 2010). Data are thus associated with a certain ‘generative’ capacity, defined as “a technology’s overall capacity to produce unprompted change driven by large, varied, and uncoordinated audiences” (Zittrain, 2006). These properties result in a distributed organization logic, where both devices and knowledge associated with data are distributed across heterogeneous disciplines and communities (Kallinikos et al., 2013; Yoo et al., 2012, 2010).

In the last fifteen years, specific efforts have been made to support the circulation of data beyond the existing boundaries of organizations or sectors, especially through the development of dedicated ‘open data’ approaches. Such efforts consist in making data open and free to all, mainly involving the release of data by public actors (e.g. Berrone et al., 2016; Charalabidis et al., 2018; Zuiderwijk and Janssen, 2014), but also by private actors, e.g. in the pharmaceutical industry (Perkmann and Schildt, 2015). In this context, both researchers and practitioners have underlined a number of barriers and challenges to data ‘liquidity’ (Jetzek, 2016; Manyika et al., 2013), e.g. technical issues related to the identification, editing, formatting and sharing of data (Barry and Bannister, 2014; Janssen et al., 2012), but also issues in developing usages once data are released as users might not be aware or competent enough to use data on their own (Janssen et al., 2012; Smith and Sandberg, 2018; Zuiderwijk et al., 2012). Recent research has more specifically underlined the potential difficulties for SMEs in leveraging open data, shedding light on strong capacity requirements in terms of data acquisition, assimilation, transformation and

exploitation (Huber et al., 2020). These challenges show that generativity is not an intrinsic property of data and rather needs to be designed through dedicated efforts, that are all the more important as the potential beneficiaries of data might be highly unfamiliar with them.

This paper proposes to investigate the mechanisms involved in designing the generativity of data, i.e. their capacity to We address this question through a multiple case study approach, digging into the case of Earth observation data where specific efforts are made to support the development of multiple usages among highly heterogenous ecosystems. To further describe the mechanisms involved in generativity building, we propose an analytical framework by taking into account the specificities of data compared to other types of technology and leveraging the literature dealing with innovation intermediaries. Our study especially unveils the existence of specific economic actors, coined 'data generativity builders', that play a crucial role in supporting the continuous development of data usages in multiple and heterogeneous ecosystems.

II. Theoretical background

To better understand the efforts at stake when expanding data usages, we first rely on research in information systems that has made a specific effort in conceptualizing properties of data. We then precise our analytical framework based on an important stream of works in innovation management research describing the actions of "innovation intermediaries" that support distributed innovation processes among heterogeneous stakeholders

1) Conceptualizing the process of developing data usages

Expanding previous works on characterizing digital artifacts (Kallinikos et al., 2013; Yoo et al., 2012, 2010), a growing stream of research in information systems has recently focused on data as specific digital artifacts, shedding light on the process by which data are generated, managed and transformed into actionable and valuable objects for a certain use context (Aaltonen et al., 2021; Aaltonen and

Tempini, 2014; Alaimo et al., 2020; Alaimo and Kallinikos, 2020). Data are described as "bearers of facts that express meaning and value" (Aaltonen et al., 2021). Data are initially produced as "sign tokens used to describe, index, represent or stage (perform) reality" (Alaimo et al., 2020). Data are initially produced under particular conditions and can never be considered as completely "raw" (Gitelman, 2013). The representation of reality (or facts) conveyed by data is indeed determined by the specific formats and language used to record data, that is guided by an initial usage context: as indicated by (Aaltonen et al., 2021) "to become data, events yet to be recorded or existing records need to be imagined as data for some purpose". However, because data record facts in a certain standardized way, they become 'portable' across settings and organizations and can be repurposed by 'recontextualizing' data according to new usage contexts (Aaltonen et al., 2021; Alaimo et al., 2020). For example, (Aaltonen et al., 2021; Aaltonen and Tempini, 2014) investigate the case of a telecommunications operator that transformed data (recording of every click, call and message relayed through the network), initially produced to manage the network infrastructure (e.g. allowing optimal allocation of resources, detecting misfunctions or suspicious activity), into data commodities sold to advertisers (transforming initial data tokens in specific metrics related to the advertising audience).

It thus appears that data are always related to a certain usage, that is however made more or less explicit along successive data transformations. In particular, the link with the initial usage context might become quite loose as data become recontextualized for new purposes (Aaltonen et al., 2021). However, it is also important to recall that it is ultimately this initial usage context that "sets the boundaries" of what can be later derived from data (Aaltonen and Tempini, 2014), e.g. due to the validity domain of the technology employed for data production. This has been further stressed by scholars calling for specific "data stewardship practices", that would especially keep traces of data transformations all along its circulation among organizations to guarantee the quality of data and avoid potential misinterpretations and misuses (Martin, 2015). Based on these considerations and to further underline the inseparability of data from a certain usage context, we will therefore describe *the action of developing usages of data as a transformation of an initial data-usage pair into new data-usage pairs where both data and usage can be modified compared to the initial use context.*

Although this stream of works has extensively described the complexity of the tasks underlying such a transformation process, several blind spots can be however noted. First, it has mainly considered empirical cases where the new developed usages remain within the boundaries of a specific industry, e.g. telecommunications (Aaltonen et al., 2021) or music industry (Alaimo and Kallinikos, 2020). (Aaltonen et al., 2021) also mention the relevance of such an approach for data considered as very context-specific, such as health data or hard science data, that could be repurposed in various contexts. However, these works do not yet address the question of how to recontextualize data in contexts that are not limited to a certain industry but that would span heterogeneous ecosystems spanning various industries. Second, despite describing recontextualization as an open-ended process, these studies have little addressed the question of how such an effort is managed in a longer-term perspective to ensure a continuous development of new usages. Making progress on these little-studied aspects appear as crucial in fully unleashing the generativity potential of data beyond the existing boundaries of industries. This leads us to formulate the research question investigated in this paper as follows: *how to manage a continuous and repeated transformation of data-usage pairs spanning highly heterogenous ecosystems?*

As underlined by (Aaltonen et al., 2021), repurposing data involves intricate coordination and articulation mechanisms between actors. Interestingly, an important stream of research in innovation management has more specifically investigated a specific class of actors called "innovation intermediaries" involved in facilitating exchange among a variety of actors to support innovation processes. This literature provides us with interesting insights to further describe the types of actions at stake in transforming data-usage pairs, that could indeed be seen as a specific innovation process involving multiple actors.

2) Innovation intermediaries

Innovation intermediaries are characterized as central actors in supporting innovations processes. They have been studied by various research streams, e.g. referred as knowledge brokers (Hargadon, 2002,

1998), innovation brokers (Klerkx and Leeuwis, 2009; Winch and Courtney, 2007), bridge builders (Bessant and Rush, 1995), boundary organizations (Guston, 2001; O'Mahony and Bechky, 2008). In his seminal paper, (Howells, 2006) builds upon different research steams on technology diffusion, systems of innovation, innovation management and service organizations to define an innovation intermediary as "an organization or body that acts an agent or broker in any aspect of the innovation process between two or more parties". They undertake numerous functions and activities, and can take different organizational forms, such as consultants (Bessant and Rush, 1995; Hargadon and Sutton, 1997), industry associations, chambers of commerce, innovation centers (van Lente et al., 2003), or other government-affiliated organizations (Kivimaa, 2014).

The role of such intermediaries appears particularly crucial in contexts with a high degree of unknown regarding technologies, markets or which actors to involve (Stewart and Hyysalo, 2008). Intermediaries thus act as "boundary-crossing" actors involving a large variety of disciplines, actors, interests, value systems, fields of activity and institutions (Boon et al., 2011). Scholars have highlighted various forms of intermediaries playing an important role in such contexts: intermediaries supporting demand-articulation processes (Boon et al., 2011, 2008), systemic intermediaries accompanied by an ecology of intermediaries supporting the complex and long-term changes involved in transitions to sustainable development (Kivimaa et al., 2019; van Lente et al., 2003), or intermediaries playing the role of "architects of the unknown" by stimulating and driving collective exploration and knowledge creation (Agogué et al., 2017, 2013). Among these different studies, we more specifically build upon the framework proposed by (Stewart and Hyysalo, 2008), synthesizing the activities of intermediaries in three main functions:

(1) *Facilitating*, i.e. educating, gathering and distributing resources, influencing regulations and setting local rules in order to create "spaces" for others to act.
(2) *Configuring*, i.e. shaping the technology including both technical aspects but also use-related aspects (e.g. suggesting or stimulating certain usages).
(3) *Brokering*, i.e. establishing or modifying connections between actors (potentially bringing users and suppliers together but also involving other important actors in the innovation network) .

This framework appears as particularly relevant to us in addressing the question of developing data usages in heterogeneous ecosystems. Indeed, data appear as largely unknown to the various ecosystems that could however benefit from them (e.g. farmers could benefit from yield prediction based on Earth observation data but might not have the interest or competencies in integrating data in their daily activities), thus requiring such "boundary-crossing" activities involving a large variety of expertise fields and actors. Moreover, the framework proposed by (Stewart and Hyysalo, 2008) especially describes cases where intermediation is required to bridge users and developers of a certain technology and has been more specifically developed through empirical investigations in the context of information and communication technologies. Our research question is thus investigated through the lens of innovation intermediaries and could be precised as follows: *in order to manage a continuous and repeated transformation of data-usage pairs spanning highly heterogenous ecosystems, to what extent specific actions of facilitating, brokering, configuring are required?*

III. Methodology

This research uses a qualitative methodology, relying on multiple case studies (Yin 2009). We indeed investigate the case of organizations that have been developing services based on Earth observation data for various types of actors (industries, public bodies, research communities), for at least several years, in order to study the dynamics of the ecosystems and the related actions carried out by these organizations. The selected case studies are particularly relevant for the research question under study. Indeed EO data, initially produced mainly for scientific goals to monitor the planet and its environment based on several instruments (satellites, in-situ sensors etc), are now available to every economic actor thanks to active 'open data' policies. These data appear to be a promising resource to help various sectors tackle socio-environmental grand challenges they are facing, but seem surprisingly underused in practice.

1) Empirical materials

Three authors of this paper (first, second and fourth) are involved in a large research project, funded by the European Commission under the Horizon 2020 programme. This project, called *e-shape*, brings together a team of 60 experienced partners from academia, industry, institutional entities and user communities to develop 32 pilot applications based on Earth observation data, gathered in seven showcases (food security, health, renewable energy, biodiversity, water resources, disaster resilience and climate). The authors are leading a work package dedicated to the progressive design and experimentation of a co-design approach supporting the different pilots. This setting allows the authors to conduct a multiple case study approach (Yin, 2009) in a unified context. Indeed, a distinctive feature of the portfolio of pilots lies in its variety. Each pilot corresponds to a certain context: it is managed by a focal actor - the pilot leader - that can be of different types (public research institute, private company, etc). This pilot leader interacts with various user groups on the one side and other stakeholders providing specific resources required to build the service on the other side (e.g. data sources, infrastructure, scientific models).

Thanks to our implication in the project, we have access to all the organizations involved in these pilots, as well as the users they interact with and their network of partners. Different forms of empirical materials can be exploited: questionnaires, interviews, observation notes taken during project meetings, and secondary sources of data on the different actors (mainly websites and scientific publications).

2) Data collection and analysis

The context of each pilot was analyzed following the same process:

Step 1 - Secondary sources of data were used to build a first understanding of each pilot (application forms filled-up by each pilot to participate to the project, completed with websites and research publications).

Step 2 - A framework was then used to synthetically represent each pilot as a chain linking data sources, information, usages and associated users (see a example in Annex).

Step 3 - A questionnaire was sent to each pilot to validate this framework representation and ask additional questions on unclear elements.

Step 4 - Answers were used to precise the analysis of each pilot on five main points: (1) overall understanding of the usage ecosystem (overall context, specific regulations, organization of the communities, contact points in these users' communities); (2) users' competencies (familiarity with EO data, ability to integrate the service in their own operations); (3) types of usages expected to be designed; (4) pilot-user relationship (history of the relationship, level of engagement, interaction modalities); (5) ability of the pilot to transform a first prototype in an operational service (upscaling challenges, operationalization resources, additional partners to be involved).

Step 5 - An interview of one hour and a half was carried out with each pilot to further clarify these five points.

Step 6 – Following the interview, we wrote a detailed a report detailing these five and updated the framework representing the pilot. These written documents were reviewed and validated by the pilot.

IV. Results

The outcomes of the analysis process, materialized in the completed framework and a report for each pilot, are simplified and presented in a table (see Annex) focusing on the following aspects: overall pilot's rationale (giving the general context of the pilot), type of focal actor supporting the pilot development, targeted user groups, expected usage contexts (i.e. the type of action that is expected to be taken by user groups thanks to the provided service), the type of generic information on which these various usages are based, the key data sources used to derive such information.

Based on this analysis, we can first validate the specific intermediary role played by the pilots in e-shape. Indeed, it clearly appears that they are involved in making the link between data producers and data users, transforming initial data-usage pairs (associated with contexts of data production) into new data-usage pairs (associated with the user groups' contexts). Moreover, the specificities of such intermediary actors can be further described according to their brokering, configuring and facilitating activities.

Brokering

The cases reveal an intriguing type of brokering, that does not consist in creating direct connections between data producers and data users, contrary to what is often encountered in open data initiatives (Zuiderwijk et al., 2014, 2012). This results from the high level of heterogeneity between data producers and data users and more specifically the impossibility of each side to directly respond to potential demands from the other side. To take the examples of pilots involved in providing farmers with specific EO-based services, satellite builders could not directly respond to farmers expectations making technical modifications of satellites in a way that would be satisfying for farmers, as it would involve large costs and development time (e.g. at least 10 years to launch a new satellite). In this context, the pilots serve as a kind of buffer zone between data users and data producers, managing the potential divergence in terms of dynamics and performance logics between these actors. It is also important to notice that this buffer role does not involve a complete separation between data users and producers. Indeed, when interacting with data users, the pilots represent data producers as it requires leveraging and integrating their resources and expertise. And vice versa when interacting with data producers, the pilots also represent data users by sharing the knowledge they have built on data usage contexts.

Configuring

According to the definition proposed in (Stewart and Hyysalo, 2008), configuring consists in shaping the technology including both technical aspects but also use-related aspects (e.g. suggesting or stimulating certain usages). These authors also indicate that only minor adjustments of the technology itself are involved. In the context of developing usages from data, these adjustments consist in transforming initial data-usage pairs in new ones and can take specific forms that can be more or less important depending on the initial and targeted usages. It is especially important to notice that the different data-usage pairs built by each pilot (indicated in the table by the variety of expected usage contexts) are all built on a common core, a certain type of information derived on the various data sources and compatible with various usages (see “generic information” column in the table). This suggests that the efforts of the pilots do not consist in building a new data-usage pair from scratch for each new usage but rather in building this specific kind of generic information. The form of this information is not given and fixed once and for

all but rather results from operations both related to the initial context of use and the targeted usage contexts. Operations related to the initial use context consist in considering what parts of the context need to kept, what other parts need to be hidden, involving the development of specific scientific models and documentation on the processing chain. Operations related to a targeted context of use consist in considering the parts of the context to be integrated in the generic information that should ensure that a certain usage could be developed within this specific context, but in a way that could be potentially compatible with other usage contexts.

Facilitating

As defined by (Stewart and Hyysalo, 2008), *facilitating* involves creating spaces for others to act, through educating, gathering and distributing resources, influencing regulations and setting local rules. The analysis of the pilots reveal two forms of facilitating activities, coined 'non- specific' and 'specific' facilitating activities. The former category corresponds to facilitating actions that are not specifically adapted to a certain type of users, for example help desk dedicated to answer users' questions, or training sessions on the basics of Earth observation. This 'non-specific' form of facilitating is similar to what was described by (Stewart and Hyysalo, 2008) in the case of cybercafes that provided users with training and advice related to the use of computer and software. However, in the case of EO data, this first type of facilitating is also complemented with a 'specific' form of facilitating that needs to adapt to the specificities of the users' contexts. Indeed the nature of the capacities and knowledge to be developed and shared might not be known in advance by the pilots themselves. Therefore, all pilots are involved in specific co-design activities in order to identify what kinds of spaces would be required by the actors to act on their own, and what kinds of support the pilots should bring to effectively create such spaces. This can require various efforts from the pilots such as building customized supporting tools or advice, but also pushing specific norms to make the use of EO data recognized by a certain community (e.g. defining what is a 'good' prediction of solar radiation, or promoting the potential of EO data compared to other data sources that are more commonly used).

V. Discussion and conclusion

The specificities of the brokering, configuring and facilitating actions carried out by the pilots suggest that they can be considered as a specific form of intermediaries that cannot be fully described through existing figures reported in literature. Similar figures can be also found in other research works, although not directly referred as "innovation intermediaries". (Kokshagina et al., 2016) for example suggest a new managerial role – a 'cross-application manager' – to support the development of generic technologies. Data generativity builders might also be compared to 'platform leaders' (Gawer, 2014). These platforms would rather correspond to innovation platforms rather than transactional ones.

This paper offers several contributions. First, our study reveals the role of a specific economic actor, that we could coin 'data generativity builder', that appears as crucial in continuously expanding data usages for a large variety of actors. This question is not limited to the EO context. Indeed, unleashing the generative potential of data has been specifically acknowledged as a promising resource in facing contemporary grand environmental and societal challenges such as poverty, illness, conflict, migration, corruption, natural disasters, climate change, and pollution (Chandy et al., 2017; George et al., 2020). These authors underline the importance of stimulating data use and reuse by various actors spanning heterogeneous sectors that could not necessarily be foreseen by the data producer, as in the EO context. Our study suggests that such process could be better supported in a long-term perspective by strengthening a network of intermediary actors involved in building data generativity, although the organizational forms of such intermediaries in these different contexts would need further investigation.

This paper also contributes to information systems literature by investigating the question of 'recontextualizing' data in a repeated manner and in heterogeneous ecosystems. Our study shows that performing repeated and multiple transformations of data-usage pairs do not only involve actions related to recontextualizing data (i.e. adapting data to a new usage context), but also specific actions to build a generic information core common to different data-usage pairs.

Several limitations are worth being noticed. Our analysis is based on the analysis of actors that are currently developing usages to EO data and thus does not characterize the evolution of these actors' activities over time. Moreover, we do not specifically discuss the different organizational configurations of this 'data generativity builder' and their potential implications for the long-term success of the approach. Digging into longitudinal case studies would be interesting to further analyze these different aspects.

References

Aaltonen, A., Alaimo, C., Kallinikos, J., 2021. The Making of Data Commodities: Data Analytics as an Embedded Process. Journal of Management Information Systems 47.

Aaltonen, A., Tempini, N., 2014. Everything Counts in Large Amounts: A Critical Realist Case Study on Data-Based Production. Journal of Information Technology 29, 97–110. https://doi.org/10.1057/jit.2013.29

Agogué, M., Berthet, E., Fredberg, T., Le Masson, P., Segrestin, B., Stoetzel, M., Wiener, M., Yström, A., 2017. Explicating the role of innovation intermediaries in the "unknown": a contingency approach. Journal of Strategy and Management 10, 19–39. https://doi.org/10.1108/JSMA-01-2015-0005

Agogué, M., Yström, A., Le Masson, P., 2013. RETHINKING THE ROLE OF INTERMEDIARIES AS AN ARCHITECT OF COLLECTIVE EXPLORATION AND CREATION OF KNOWLEDGE IN OPEN INNOVATION. International Journal of Innovation Management 17, 1350007. https://doi.org/10.1142/S1363919613500072

Alaimo, C., Kallinikos, J., 2020. Managing by Data: Algorithmic Categories and Organizing. Organization Studies 017084062093406. https://doi.org/10.1177/0170840620934062

Alaimo, C., Kallinikos, J., Aaltonen, A., 2020. Data and value, in: Handbook of Digital Innovation. Edward Elgar Publishing, Cheltenham, UK.

Barry, E., Bannister, F., 2014. Barriers to open data release: A view from the top. IP 19, 129–152. https://doi.org/10.3233/IP-140327

Berrone, P., Ricart, J.E., Carrasco, C., 2016. The Open Kimono: Toward a General Framework for Open Data Initiatives in Cities. California Management Review 59, 39–70. https://doi.org/10.1177/0008125616683703

Bessant, J., Rush, H., 1995. Building bridges for innovation: the role of consultants in technology transfer. Research Policy 24, 97–114. https://doi.org/10.1016/0048-7333(93)00751-E

Bharadwaj, N., Noble, C., 2017. Finding Innovation in Data Rich Environments. Journal of Product Innovation Management 34, 560–564. https://doi.org/10.1111/jpim.12407

Björkdahl, J., Holmén, M., 2019. Exploiting the control revolution by means of digitalization: value creation, value capture, and downstream movements. Industrial and Corporate Change 28, 423–436. https://doi.org/10.1093/icc/dty022

Blackburn, M., Alexander, J., Legan, J.D., Klabjan, D., 2017. Big Data and the Future of R&D Management. Research-Technology Management 60, 43–51. https://doi.org/10.1080/08956308.2017.1348135

Boon, W.P.C., Moors, E.H.M., Kuhlmann, S., Smits, R.E.H.M., 2011. Demand articulation in emerging technologies: Intermediary user organisations as co-producers? Research Policy 40, 242–252. https://doi.org/10.1016/j.respol.2010.09.006

Boon, W.P.C., Moors, E.H.M., Kuhlmann, S., Smits, R.E.H.M., 2008. Demand articulation in intermediary organisations: The case of orphan drugs in the Netherlands. Technological Forecasting and Social Change 75, 644–671. https://doi.org/10.1016/j.techfore.2007.03.001

Cappa, F., Oriani, R., Peruffo, E., McCarthy, I., 2021. Big Data for Creating and Capturing Value in the Digitalized Environment: Unpacking the Effects of Volume, Variety, and Veracity on Firm Performance*. Journal of Product Innovation Management 38, 49–67.

Chandy, R., Hassan, M., Mukherji, P., 2017. Big Data for Good: Insights from Emerging Markets*: BIG DATA FOR GOOD. Journal of Product Innovation Management 34, 703–713. https://doi.org/10.1111/jpim.12406

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., Ferro, E., 2018. The World of Open Data: Concepts, Methods, Tools and Experiences, Public Administration and Information Technology. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-90850-2

Chen, H., Chiang, R.H.L., Storey, V.C., 2012. Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly 36, 1165–1188. https://doi.org/10.2307/41703503

Gawer, A., 2014. Bridging differing perspectives on technological platforms: Toward an integrative framework. Research Policy 43, 1239–1249. https://doi.org/10.1016/j.respol.2014.03.006

George, G., Merrill, R.K., Schillebeeckx, S.J.D., 2020. Digital Sustainability and Entrepreneurship: How Digital Innovations Are Helping Tackle Climate Change and Sustainable Development.

Entrepreneurship Theory and Practice 104225871989942. https://doi.org/10.1177/1042258719899425

Gitelman, L. (Ed.), 2013. “Raw data” is an oxymoron, Infrastructures series. The MIT Press, Cambridge, Massachusetts ; London, England.

Guston, D.H., 2001. Boundary Organizations in Environmental Policy and Science: An Introduction. Science, Technology, & Human Values 26, 399–408. https://doi.org/10.1177/016224390102600401

Hargadon, A., Sutton, R.I., 1997. Technology Brokering and Innovation in a Product Development Firm. Administrative Science Quarterly 42, 716. https://doi.org/10.2307/2393655

Hargadon, A.B., 2002. Brokering knowledge: Linking learning and innovation. Research in Organizational Behavior 24, 41–85. https://doi.org/10.1016/S0191-3085(02)24003-4

Hargadon, A.B., 1998. Firms as Knowledge Brokers: Lessons in Pursuing Continuous Innovation. California Management Review 40, 209–227. https://doi.org/10.2307/41165951

Howells, J., 2006. Intermediation and the role of intermediaries in innovation. Research Policy 35, 715–728. https://doi.org/10.1016/j.respol.2006.03.005

Huber, F., Wainwright, T., Rentocchini, F., 2020. Open data for open innovation: managing absorptive capacity in SMEs. R&D Management 50, 31–46. https://doi.org/10.1111/radm.12347

Janssen, M., Charalabidis, Y., Zuiderwijk, A., 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management 29, 258–268. https://doi.org/10.1080/10580530.2012.716740

Jetzek, T., 2016. Managing complexity across multiple dimensions of liquid open data: The case of the Danish Basic Data Program. Government Information Quarterly 33, 89–104. https://doi.org/10.1016/j.giq.2015.11.003

Kallinikos, J., Aaltonen, A., Marton, A., 2013. The Ambivalent Ontology of Digital Artifacts. MIS Quarterly 37, 357–370.

Kivimaa, P., 2014. Government-affiliated intermediary organisations as actors in system-level transitions. Research Policy 43, 1370–1380. https://doi.org/10.1016/j.respol.2014.02.007

Kivimaa, P., Boon, W., Hyysalo, S., Klerkx, L., 2019. Towards a typology of intermediaries in sustainability transitions: A systematic review and a research agenda. Research Policy, New Frontiers in Science, Technology and Innovation Research from SPRU’s 50th Anniversary Conference 48, 1062–1075. https://doi.org/10.1016/j.respol.2018.10.006

Klerkx, L., Leeuwis, C., 2009. Establishment and embedding of innovation brokers at different innovation system levels: Insights from the Dutch agricultural sector. Technological Forecasting and Social Change 76, 849–860. https://doi.org/10.1016/j.techfore.2008.10.001

Kokshagina, O., Masson, P.L., Weil, B., Cogez, P., 2016. Portfolio Management in Double Unknown Situations: Technological Platforms and the Role of Cross-Application Managers. Creativity and Innovation Management 25, 270–291. https://doi.org/10.1111/caim.12121

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., Doshi, E.A., 2013. Open data: Unlocking innovation and performance with liquid information. McKinsey Global Institute 21, 116.

Martin, K.E., 2015. Ethical issues in the big data industry. MIS Quarterly Executive 14, 2.

O'Mahony, S., Bechky, B.A., 2008. Boundary Organizations: Enabling Collaboration among Unexpected Allies. Administrative Science Quarterly 53, 422–459. https://doi.org/10.2189/asqu.53.3.422

Perkmann, M., Schildt, H., 2015. Open data partnerships between firms and universities: The role of boundary organizations. Research Policy 44, 1133–1143. https://doi.org/10.1016/j.respol.2014.12.006

Smith, G., Sandberg, J., 2018. Barriers to innovating with open government data: Exploring experiences across service phases and user types. Information Polity 23, 249–265. https://doi.org/10.3233/IP-170045

Stewart, J., Hyysalo, S., 2008. Intermediaries, users and social learning in technological innovation. Int. J. Innov. Mgt. 12, 295–325. https://doi.org/10.1142/S1363919608002035

van Lente, H., Hekkert, M., Smits, R., van Waveren, B., 2003. Roles of Systemic Intermediaries in Transition Processes. Int. J. Innov. Mgt. 07, 247–279. https://doi.org/10.1142/S1363919603000817

Winch, G.M., Courtney, R., 2007. The Organization of Innovation Brokers: An International Review. Technology Analysis & Strategic Management 19, 747–763. https://doi.org/10.1080/09537320701711223

Yin, R.K., 2009. Case Study Research: Design and Methods. SAGE.

Yoo, Y., Boland, R.J., Lyytinen, K., Majchrzak, A., 2012. Organizing for Innovation in the Digitized World. Organization Science 23, 1398–1408. https://doi.org/10.1287/orsc.1120.0771

Yoo, Y., Henfridsson, O., Lyytinen, K., 2010. The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. Information Systems Research 21, 724–735. https://doi.org/10.1287/isre.1100.0322

Zittrain, J.L., 2006. The Generative Internet. Harvard Law Review 119, 68.

Zuiderwijk, A., Janssen, M., 2014. Open data policies, their implementation and impact: A framework for comparison. Government Information Quarterly 31, 17–29. https://doi.org/10.1016/j.giq.2013.04.003

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., Alibaks, R.S., 2012. Socio-technical Impediments of Open Data. Electronic Journal of e-Government 10, 156–172.

Zuiderwijk, A., Janssen, M., Davis, C., 2014. Innovation with open data: Essential elements of open data ecosystems. Information Polity 19, 17–33. https://doi.org/10.3233/IP-140329

Annex (to be simplified)

Pilot's rationale	Focal actor (pilot leader)	Targeted user groups	Expected usage contexts	Generic information	Key data sources
Supporting global agricultural monitoring	Independent research institute	GEOGLAM (Group on Earth Observations Global Agricultural Monitoring Initiative): international organization gathering national, regional and global agricultural organizations and administrations	Monitoring crop calendar information, especially for two regions: Greece and Ethiopia	Crop calendars (specific field interventions such as harvest, planting or ploughing)	Satellite data (Copernicus) Satellite data from Proba V (VITO) In-situ data when available (use of a specific data sharing platform) Soil data base spectral library (iBEC) Citizen-sourced data (IIASA) Standard and Open Meteo data and services (e.g. precipitation, forecasts)
Supporting farmers for CAP (Common Agricultural Policy) compliance and farm performanc e	Public research institute	Agri-cooperatives Farmers Insurance companies Paying agencies	- Monitoring crop information expected yields coupled with meteorological risks - System for scientific purposes	Crop growth (yield assessment and forecast, damage assessments)	Satellite data (Copernicus) Soil data base spectral library (iBEC) Citizen-sourced data (IIASA) Standard and Open Meteo data and services (e.g. precipitation, forecasts) In-situ telemetric stations (provided by an IT SME)
Supporting farmers with crop insurance services in Ethiopia	Public research institute	Digital finance and payment services' provider Insurance Companies Micro-Finance Institutions Various key-government agencies	Monitoring system for droughts: based on the processing of an enriched NDVI, considering zone and season specificities.	Crop insurancce product covering impacts by droughts	Vegetation index from various satellite sources (SPOT-VGT & PROBA-V, Sentinel-3, Copernicus)

Supporting agriculture activities at farm level	Independent research institute	Agro-consultants Policy Makers Agricultural cooperatives Agro-industries Farmers	Monitoring system of crops at field level Towards decision/design support system	Crop calendars at parcel level (specific field interventions such as harvest, planting or ploughing)	Satellite data (Copernicus and VITO) In-situ data when available (use of a specific data sharing platform) Soil data base spectral library (iBEC) Citizen-sourced data (IIASA) Standard and Open Meteo data and services (e.g. precipitation, forecasts)
Providing customers with field and/or regional level information rather than simply in-situ in order to Increase the understanding of moisture across the farm	Riscognition is a private company specialized in the processing of EO imagery Elmibit is a smart agriculture software company	Riscognition: agriculture software developers Elmibit: farmers	- A specific service provision linking in-situ data with EO information to create time series and improve forecasting - A monitoring system with regularly updated information about the state of farmers' crops and fields - Decision support system	Soil and crop moisture (at 10m resolution) Vegetation stress (at 10m resolution)	Satellite data Sentinel 1 & 2 and other Copernicus services (Copernicus) In-situ data: farmers in-situ data
Providing national decision makers with high quality national sclae information based on high resolution data	Public space research institute	Governemental entities Academics Research institutes Commercial entities	3 information systems: - Country level SDG indicators information + report form - Assessment of regional crop state, crop emergency and land use + Report form - Information provision of the land use, crop area, crop state,	Information on agricultural land and crops: - SDG 15.3.1 - SDG 2.4.1 - Land cover maps - Crop maps - Agriculture productivity - Agriculture emergency - etc.	Copernicus data: Landsat-8 and Sentinel 1 & 2 Users' data: - In-situ and weather data - Crop data, crop calendars or other geospatial vector or raster layers - vector files with in-situ land cover\crop types labels and in-situ LAI\NDVI or other crop state date for validation

			land degradation and SDG Technologies and data provision: - Data and expertise for educational purpose - Land productivity and SDG assessment		
Surveilling mercury pollution	Public research institute	Health communities Conference of Parties Local and regional authorities	Decision support system for policy makers that aim at assessing the efficiency of the measures taken by policy makers to reduce mercury emissions	Mercury deposition in the environment	Satellite data (Copernicus), Multispectral high-resolution images Specific data base gathering mercury measurements
Surveilling persistent organic pollutants	Public research institute	Policy makers Conference of Parties (COP) Scientific communities	Monitoring system of POPs for policy makers: - Data collection on POPs levels in different core media; - Editing and visualization tools	Persistent Organic Pollutant concentration in core media (air, human tissues, water)	Sampling networks of polutants in air, human tissues and water
Improving information on air quality and related health issues, for use in public health assessment	Public research institute	International organizations, National authorities Cities and municipalities Insurance companies Real estate companies	Monitoring system for air quality, enriched with different layers of information (air quality, health, socio-economic, etc.), and specific visualization tools.	Air quality at global, city and municipality scales at different time horizons	- Remote sensing data from various sources - In-situ data - Open AQ Datasets - Low cost sensor networks - IIASA

and urban planning.					
Nowcasting and short-term forecasting of solar energy	Public research institute	Ministries of Electricity and Renewable Energy Power generation operators Power distribution and transmission operators	Specific service provision: - Historical solar data and aerosol load data for specific locations. - - 1-2-day solar forecasting window	- Solar surface irradiance - Maps, databases and on the fly requests	- Copernicus data: CAMS, Sentinal - AEROCOM database - Metesat (EUMETSAT) - In-situ data: ground-based aerosol optical properties and others Other data: - Pre-calculated radiative transfer model outputs
Encouraging high photovoltaics penetration in urban areas	Public research institute & national spatial agency	Energy providers Citizens Collectivities & urban planners	Monitoring system of the PV production variability at urban cale (nowcasting, forecasting) - Modelling tool for urban energy system, - Specific service provision (scientific advisory or joint inquiry)	- Variability assessment and forecast of SSI - Characterisation of existing PV and potential roofzones - Shaded surface irradiance - SSI variability - PV yield simulation - Urban classification - Meteo information	- Copernicus data: CAMS and Sentinels 1 & 2 - Other satellite data sources: airborne, building cadastral 2D plans, etc. - Other data sources: in-situ measurements and yield simulation or electric loads

Providing wind resource assessment tools for the offshore wind industry	Public research institute	Offshore wind farm developers Offshore wind farm operators Consultants Research, academia, educators	Type of system still to be discussed with users. Web-portal for on-demand information provision, and showcasing DTU's expertise: - Portal giving access to maps, explanations on the methodology, processing chains	An enriched & unified wind product - Mean wind at - Wind power density - Wind direction distributions - Weibull parameters	- Copernicus data: Sentinel 1, ENVISAT, CMEMS (EO and in-situ data) - Other data sources: Global Forecasting System model data (NOAA) and Meso-scale atmospheric model outputs (DTU Wind Energy)
Providing sustainable EO based data sets for wind energy planning	Earth Observation knowledge hub	Wind turbine manufacturers Consultancy companies Research institutes	Data package: map layers for wind ressource assessment and siting.	- Map layers - New parameters (tree height and LAI)	- Copernicus data: Sentinel 1 & 2 - Earth Observing System (EOS, NASA): ICESat-2 observations - Turbine and wind information (users' data)
Monitroing and modelling the trends and states of ecological ecosystems through the integration of remote sensing maps and in-situ data	Research institutes (public/private)	Technical staff and managers of European Protected Areas (PAs)	monitoring system of the ecosystem state and changes based on EO data.	Consistent information for indicators and variables to monitor the ecosystem state and changes Geographical scale: European scale	- Gridded datasets (e.g., EOBS - Copernicus Climate Change Services (C3S) - Remote sensing (RS) data (CMEMS and Sentinel, Copernicus) - Earth Observing System (EOS, NASA): MODIS/Aqua - In-situ measurements: eLTER network
Building an extended common registry of observation	Research institutes (public/private)	Research communities Technical and scientific staff of project groups	Monitoring system Site monitoring, providing information that	Biodiversity and ecosystem data available for a specified area of interest	- Global, documentation and sharing of site information and related datasets -European, metadata catalogue providing

and experimentation facilities of ecosystems as the basis for the integration of available information		European and national conservation agencies	leads to be more and better informed in order to take actions;		information on in-situ data provided by the eLTER RI and LTER Europe network - EO based EBV from GEOBON portal (Copernicus)
Improving the services and functionality of the Essential Biodiversity Variables data portal, providing harmonized sets of variables for biodiversity observation and conservation.	Research institutes (public/private)	Research communities Monitoring agencies that inform ministries	Biodiversity monitoring system and assessing targets Numerous tasks: - Production and delivery of data, set up indicators through a portal + reports on different topics - Set connexion between EO data production and in-situ data	Essential Biodiversity Variables (EBV) - Different scales: three levels of sites (SG1 = core sits, SG2 = integration sites, SG3 = coverage sites) + national scale	Derived EO data products EO products from ECOPOTENTIAL project (Copernicus) In-situ measurements (eLTER network)

Providing historical and near-real time information for a number of hydrological variables	Public meteorological and hydrological institute	Geological institutes Water and marine authorities Hydropower companies Emergency response Management centers Researchers	Monitoring system based on the improved hydrological information at specific locations at near-real time/forecasts	Long historical records + medium range & seasonal forecasts of different hydro-climatic variables over Sweden	EO-based data: - Copernicus Global Land Monitoring, ESA GlobSnow… - MODIS - Sentinel 3 (Copernicus) - Flood extent mapping Meteorological forcing data
EStimating flood hazard at a large-scale	Public meteorological and hydrological institute	Reasearch and meteorogical institutes R&D Companies Modelling flood risk authorities	Mainly monitoring system of water bodies over Europe: Decision/scenario-based design support system	- Water bodies and floodwater records - Surface water volumes, exceedance probabilities, return period of flood, etc.	Copernicus data: - Sentinel-1 - ENVISAT image collections (2002 - 2018)
Providing a near-real time visibility score for specific diving locations	Public laboratory	Diving centers, commercial divers	Weather report/monitoring service/ forecasting	Visibility score at near real time	CMEMS (Copernicus)

Predicting the landing areas and severity of the sargassum algae season	Private company	Local authorities Research community Tourism companies	- Monitoring system integrated in Meteo France's own workflow - Specific service provision: sargassum detection index integrated in a specific tool for operational use - With CERMES: a design support system could be considered in addition to provision of CLS algorithm's outputs	Sargassum identification and characterization Predicition several months in advance of landing areas and severity of Sargassum Time series of NFAI (Normalized Floating Algae Index)	- Earth Observing System (EOS, NASA): MODIS Aqua - Copernicus: Sentinel 1, 2 & 3 and CMEMS - Other sat-based information: ocean surface current and wind data (NOAA) - Unprecised data sources: bathymetry and coastlines
Improving monitoring and regulation of fishing activities in the Northeast Atlantic	Private company	Fishermen associations NGOs International organizations scientific communities Regional and national authorities	Monitoring system of fishing activities through a Geo-Portal	Characterization of fishing activities with visualization tools Pressure indicators Further development of CFP Data collection framework and MSFD indicators	DGRM's (user) data: - Vessel monitoring system - Fishing vessels VMS data - Vessels technical characteristics - Fishing e-logbooks data - Fishing e-logbooks data - Landing declarations - Fishing fleet landing declarations Other data: - in-situ, satellite and model EO data streams (CMEMS (Copernicus), etc.) - Automatic Identification System data - Bathymetry and sediment chart

Providing WFD ecological status products of phytoplankton biomass for management of selected water bodies	Research institute spin-off	Research center Water authorities EU level public water bodies National regulators (maybe) Dredging companies	Monitoring of phytoplankton biomass for WFD reporting based on EO data + potentially other systems to be explored (decision support system or scenario design support)	EO data to improve water quality measurements: - Better quality/accuracy and continuous spatial data Cloud-free images - Measure the concentration of Chlorophyll-a	- Sentinel 3 A & B (Copernicus) - Sentinel 2 A & B (Copernicus) - EEA, LIMNADES, GEOSS public databases
Providing aquafarmer with information about mussels in weekly bulletin, with indicators in order to prevent the impact of extreme events due to climate change	Private company (solution provider)	Aquaculture commercial companies	Rheticus® Aquaculture provides farmers with smart information in the form of analytics, trends, and alerts to boost productivity or promptly identify critical issues (weekly report)	Information to improve water quality measurements - the weekly value and estimation of growth and growth trends - 14-day forecast of mussel growth, - comparative analysis of growth rate - temperature - level of plankton	CMEMS (Copernicus): - Sentinel 2 & 3 Users' in-situ data: - Geographic coordinates of the production site or a georeferenced vector file - Seeding date; - Seed dimension; - Sorting and socking dates and sizes.

Improving monitoring of volcanic eruptions	Public research institute	Special units in meteorologial institutes Air traffic managers Civil protection authorities Local authorities Policy makers	- Monitoring system of 4D volcanic plume distribution - Information provision	- 4D distribution of volcanic plume over Europe, - Alert and characterization of the source - Forecast of Ash dispersion	Copernicus data: - Ash plume tracking (MODIS, SLSTR –Sentinel 3) Other data: - EPOS, LIDAR and in-situ observation data - NASA/NOAA data: SO2 tracking and aerosol index ACTRIS data: Aerosol optical property profiles and Ash concentration & vertical mask
Developing a tool for fire and risk assessment and supervision scenarios	Public research institute (officially linked to the national Civil Protection Authority)	National and regional civil protection authorities Hydro-Meteorological Agencies	- Monitoring & decision support system providing scenarios for possible risks of fire and flood - Visualization tools - Specific service provision (support to improve hydro-meteorological chain)	- Enhanced meteorological information on limited area - Enhanced information on high-impact weather events - Impact assessment information - Updated Land Use Land Cover (LULC) and LULC Change Detection Maps	Copernicus data: - Sentinel-1 - Sentinel-2 L1C - Sentinel-3 L2 Personal Weather Stations Statistics on: - Number of people affected by a given natural hazard - Economic impacts/costs by a given natural hazard

Assessing geohazard vulnerability of cities and critical infrastructures	Public institute & private company	Urban managers and civil protection authorities Energy and infrastructure companies Policy makers	- On-demand service provision: INSAR data - Decision support system customized to users' competencies and daily operations and alarms on status of the network with visualization tool	- Active deformation Areas maps - Landslide and Subsidence activity Maps - Vulnerability Elements activity maps - Change detection - Pipeline Alert	- Sentinel 1 (Copernicus), Cosmo-SkyMed and TerraSAR-X - In situ ground deformation data derived from Global Navigation Satellite System observations - Urban geohazard maps: - Specific data from users
Provivind information on geo-hazard vulnerability of agriculture	Public research institute	Insurance company Farming cooperatives	- Historical information on the phenomena occurring in the area - Risk and damage assessment in real time - Monitoring system providing information to insurers - Early warning system - forecasting based on meteorological predictions. - Early warning system of disasters - Monitoring system for cotton crops' phenological state	Hazard information: - Disaster assessment and forecast - Crop type and natural ecosystems mapping Coverage: regional level	- Copernicus data: Sentinel 1, 2 & 3 raw data - EO derived parameters/indices - Partner data: in-situ datasets - Standard and Open Meteo data and services - User data: historical data

Providing territorial and ocean carbon and greenhouse gas fluxes information in order to support the Global Carbon Project	Public research institute	Global Carbon Project GHG community SOCAT community	Provision of information and models. Different components to be differentiated (different maturity levels): Atmospheric component, terrestrial component and ocean component	Improved info on terrestrial & ocean GHG fluxes: - Land carbon fluxes - ocean product that will reduce the uncertainty in ocean carbon sink estimates - Quality indices	- OCO-2 (EOS), Sentinel 5P TROPOMI (Copernicus) - ICOS (ETC & OTC), MODIS LST/NDVI/LAI/FPAR, Sentinel 3 (Copernicus), ECMWF - Statistics and historical data: ERA5 - ENVISAT SST (Copernicus), ERS2 wave/wind, SMOS,… - Global in situ EOV Inorganic Carbon surface ocean data
Providing information to municipalities on heat waves, heavy precipitations and extreme weather events	Public meteorological institute	Cities and municipalities Consultancy company for urban climatology and wind research	- Monitoring system for extreme weather events - Warning system for extreme weather events 6-week ahead - Scenario design support system - Monitoring system for extreme weather events - Monitoring system with dynamic visualization data - Scenario design support system	Weather and climate easonal forecast: - Urban Heat Index - Drought - Precipitation	Copernicus C3S Seasonal forecast HARMONIE model (NWP & Météo-France), AROME model (Météo-France), INCA model forecasts In situ-observations

Prodiving seasonal forecasting of forest harvest conditions to support forest industry organisations and companies	Public meteorological institute	Private companies in forest management support R&D company	Combination of modules going from monitoring to more decision-support oriented module Smart harvesting service: evaluation of the carbon impact of forest management practices.	- Forecast frozen soil depth and soil moisture - Improved map resolution - Forest management options - Estimate carbon emission impacts of these different harvesting scenarios - Geographical coverage: Northern European boreal zone forest areas	C3S data (Copernicus): - Copernicus C3S Seasonal forecast ensemble - C3S Climate data Store - Soil Temperature and soil moisture Other data: - 5 forest classes laser scanning analysis (Arbonaut) - Forest parcels data - In-situ data
Supporting hydropower companies by better predicting hydrological conditions	Public meteorological institute	Energy companies	Public monitoring system about snow distribution before it starts melting, assessing when the peak flow will occur and how much water it will represent to shape the hydrograph On a longer-term: a scenario-based design support system to explore ways of managing dams	- Snow information: Snow Water Equivalent, Snow Depth, Snowmelt rate - Surface runoff, soil column inter and base flow - River discharge and reservoir inflow - Reduced hydrological model uncertainties	C3S data (Copernicus): - Seasonal forecast ensemble - C3S Climate Store Other Copernicus data: Sentinel 1-2-3 (OLCI and SLSTR data). Other data - HydroSAF (Sensors + weather station observations) - Snow state verification with Webcam snow monitoring network

Helping transportation and tourism sectors to better prepare for seasonal changes	Public meteorological institute & public research institute	Tyres companies Tourism stakeholders	Optimizing/designing retail operation and logistics EO monitoring system and visualization tools and warning system	- Probalistic estimations of onset/offset dates of the winter driving conditions for Finland - Probalisitic estimations of climate-tourism indicators for Greece	Copernicus data: - European seasonal forecasts from C3S Climate data Store - Sub-seasonal model data from ECMWF - ERA5 re-analysis In-situ data Users data: - Auxiliary data from tire companies/tourism on the previous seasonal demands

Creating value from data in an ecosystem: building and expanding relationships between data and seemingly distant usages

Raphaëlle BARBIER
MINES ParisTech – PSL University
raphaelle.barbier@mines-paristech.fr

Benoit WEIL
MINES ParisTech – PSL University
benoit.weil@mines-paristech.fr

Pascal LE MASSON
MINES ParisTech – PSL University
pascal.le_masson@mines-paristech.fr

Abstract

Creating socio-economic value from data seems to be a shared concern in almost every industry, public and research area. In this paper, “value” is taken in a large meaning considering that creating value boils down to connecting data and usages. In the Earth Observation field, this concern is all the more challenging as data and usages often seem to be “distant” – that is not related at first sight. This paper explores the question of how to build relationships in such a context and how these relationships can evolve over time. Our analysis is based on the historical case study of a research lab that has progressively build services for solar energy. The importance of several elements is enhanced: (1) a new interpretation of the role of “information” in the value creation process, as playing a pivotal role between usages and data, which is thus proposed to be coined “inter-formation”; (2) the importance of how models are designed, suggesting a new way of gaining a competitive advantage that is not only focused

on the nature of data or final usages, but on the structure of models; (3) the related expansion dynamic of the ecosystem, suggesting a specific form of expansion that is related to all the elements of the data-usage relationship, and not only thanks to data pushing for new usages or usages pushing for new data. These findings contribute to bring some insights on the creation and expansion of data-based ecosystems, at both organization and ecosystem's levels.

I. Introduction

In recent years, the development of internet, new sensors, and computational means has dramatically increased the flow of data in almost every business, industry and research area. This phenomenon, commonly referred as "big data", has largely been discussed in the literature, shedding light on its definition, opportunities and challenges, especially the issue of how value can be created out of this new flow of data (Günther et al. 2017; Philip Chen et Zhang 2014; Gandomi et Haider 2015; Sheng, Amankwah-Amoah, et Wang 2017).

In this paper, we would like to bring new insights on value creation process in an ecosystem by delving into the evolution of a traditional scientific discipline - Earth Observation - that has also been affected by this "big data" phenomenon. Earth Observation (EO) refers to the gathering of information about planet Earth's physical, chemical and biological systems (for example ocean currents, solar radiation reaching the ground, composition of the atmosphere, characterization of vegetation). Different kinds of instruments are currently used for this purpose: in-situ sensors (for example floating buoys to monitor ocean currents, temperature and salinity; or land stations that record air quality and rainwater trends), airborne sensors, or satellites. In recent years, with the development of remote-sensing satellites and increasingly high-tech "in-situ" instruments, Earth observation has become more and more sophisticated and has generated an increasing amount of data, leading to concerns about a "Big Earth

Data" phenomenon (Guo 2017). Therefore concerns about handling new flows of data are raised, but also about their usages. Indeed, the European Union has significantly invested to make this scientific data a common good, freely accessible to all potential users and is now pushing for the development of new socio-economic applications that would not only benefit Earth observation scientific community but also public authorities, private companies, industry, universities, citizens. This context seems particularly interesting to study how the value creation process of data seems to be closely linked to the development of a rich ecosystem involving heterogeneous communities.

In the Earth Observation field, several issues seem to hinder the development and new usages for EO data. First, the final user is often different from the data owner and might even not be aware that satellite data can be helpful for its own activities because of the complexity of this type of data. Second, for the people in charge of processing satellite data, usually resulting in a first usage, it seems difficult to have an idea of how this first usage might be further expanded towards new usages in the future. Hence this type of situation seems to be characterized by a very large the distance between data and usages and is calling for specific approaches to cope with this distance. This large distance will help us to examine more closely the role of hidden elements in literature, especially how models are designed, and also rediscuss the role of ubiquitous elements, namely "information" usually seen as an intermediate step between data and usage.

Leveraging on this specific context, the paper aims at bringing a fresh look on the value creation process in an ecosystem. More specifically, the following elements will be enhanced:

(1) a new interpretation of the role of "information" in the value creation process, as playing a pivotal role between usage and data;
(2) the importance of how models are designed, suggesting a new way of gaining a competitive advantage that is not only focused on the nature of data sources or final usages, but on the structure of models;
(3) the related expansion dynamic of the ecosystem, suggesting a specific form of expansion that is related to all the elements of the data-usage relationship.

II. Literature review

1. Value creation as a shared issue

Literature seems to agree on the important consequences of “big data” in all scientific, industrial and business areas, stemming from its specificities (Günther et al. 2017; Philip Chen et Zhang 2014; Gandomi et Haider 2015; Sheng, Amankwah-Amoah, et Wang 2017). These specificities have been characterized with 3Vs dimensions (Chen, Chiang, et Storey 2012): *volume* (referring to the huge amount of collected data, often reported in several terabytes or petabytes), *velocity* (referring to the continuous stream of data received and the speed at which it should be analysed), *variety* (referring to the heterogeneous sources of datasets). (Gandomi et Haider 2015) mention the use of other dimensions to describe big data, such as *veracity* (importance of using reliable data and interpretations), or *variability* (continuous stream of data and that requires to take into account continuous changes in data sets).

In the academic field, a wide range of disciplines has discussed the consequences of this big data phenomenon, especially Management and Information Systems (Del Vecchio et al. 2018). Research agendas in these disciplines have particularly highlighted the crucial issue of better understanding the process of value creation from this new flow of data. In Management research, (George, Haas, et Pentland 2014) call for a better understanding “how ubiquitous data can generate new sources of value”, and how this value is distributed among the different stakeholders (companies, industries, governments) thanks to new business models and governance tools. In Information Systems research, (Abbasi, Sarker, et Chiang 2016) reports on the effect of big data on the traditional value chain leading to a specific “Big Data Information Value chain”.

Regarding the concept of “data value”, as exposed by (Furtado, Dutra, et Macedo 2017), different meanings can be found in literature. A distinction is often made between social and commercial or economic value (Chandy, Hassan, et Mukherji 2017; Günther et al. 2017), thus referring to the nature of benefits coming from the use of data. (Günther et al. 2017) define economic value as related to “the organization’s increase in profit, business growth, and competitive advantage resulting from big data adoption”, often resulting in monetary profits. For the same authors, social value refers to an improvement of social wellbeing and societal topics (in various fields such as healthcare, public safety and security, employment growth). In the following paper, we will take a more general meaning of “value” without referring to a specific type of benefits. We will consider that *there is “value” when the use of data brings benefits to the data user*. The process of creating value from data can therefore be described as *building relationships between data and these specific usages where the user benefits from*

the use of data. The question of value creation thus leads us to investigate *how this data-usage relationship is built.*

2. Two different approaches to build the data-usage relationship

In literature, two main types of situations are commonly reported. In the first type, the starting point is the access to new data and the objective is to find usages that embed value for the user. In the second type, the starting point is a target usage and the objective is to get data to generate it.

From data to usage

This scenario seems to be more frequently described in literature. The organization that has collected new data can use it to support its existing activities, such as supply chain management (Hazen et al. 2014), marketing activities (Fan, Lau, et Zhao 2015), New Product Development processes by reducing time to market and costs and improving customers' product adoption (Tan et Zhan 2017; Johnson, Friend, et Lee 2017). In the previous cases, data is used as a new source of information that brings useful insights about the existing processes (for example monitoring the efficiency of a production line) or about customers (for example needs and behaviors), thus helping to improve decision-making processes products and services delivered to customers. This approach is coined "inbound approach" by as the organization directly uses the gathered data to enhance its activities.

(Trabucchi et al. 2018) highlight the existence of a different approach coined "outbound strategy" where the collected data is sold to external parties who can see in them a greater value. The authors take the example of Strava application allowing users to track their sport activities and get useful information such as average speed, performance stats and so on. Strava sold data aggregated from all runners and bikers' activities to departments of Transportation. These data sets allowed the latter to better plan the construction of future bike lanes by identifying where they were the most needed. In this situation, the organization benefiting from the data usage is different from the one collecting data

From usage to data

As mentioned above, recent research has shed light on a second type of situation where the starting point is a target usage and the objective is to find data allowing to create it (Trabucchi et Buganza 2019). Several case studies are reported in this paper. One of them is exposed here as an illustrative example. Sage Bionetworks is a not-for-profit organization that partners with researchers, patients and healthcare

stakeholders to drive data-driven projects contributing to health improvement. In of its projects, the organization aimed at finding new ways of managing symptoms of Parkinson's disease and had to find a way to collect data on symptoms and their evolution. To do so, Sage Bionetworks launched mPower app, that could be used by people affected or not by the disease. Through this app and additional surveys, the organization was able to analyze different tasks based on the sensors embedded in the smartphone (camera, touchscreen and microphone), and therefore to analyze and monitor the symptoms over time. In this example, in order to bridge the gap between usage and data, is worth noticing that the organization had to design a whole process, including choosing instruments (smartphones and surveys) adapted to the measurements to be made (movement disorder), technical solutions to support data collection and distribution to researchers, and need for technical validity of these data (Bot et al. 2016). Based on this example, in order to better understand how the data-usage relationship is built, it seems crucial to examine *how this process is handled in practice*.

3. How the process of building data-usage relationship is handled in practice

Two complementary aspects can be found in literature. First, some authors propose frameworks detailing the different stages of the process to build a data-usage relationship. Second, given its specific features mentioned earlier (3Vs or 5Vs dimensions), handling big data and transforming it into value seems to require specific enabling factors, regarding techniques, skills and organizational aspects (Akter et al. 2016; Philip Chen et Zhang 2014; Troilo, De Luca, et Guenzi 2017).

Some frameworks are proposed in literature to clarify the important stages of building the data-usage relationship. In these processes, it seems that there is an intermediate step between data and its usage, which is the "information" or "insights" it gives (Gandomi et Haider 2015; Lim et al. 2018; Abbasi, Sarker, et Chiang 2016). Building the data-usage relationship thus includes building both a data-information relationship and an information-usage relationship.

To manage the process of building data-information-usage relationship, enabling factors are mentioned in literature, due to the specific nature of big data. (De Mauro, Greco, et Grimaldi 2016) defines big data as "the information asset characterized by such a high Volume, Velocity and Variety to require specific technology and analytical methods for its transformation into value." This definition puts forward the need for a first enabling factor related to specific technologies and methods adapted to big data characteristics. These methods can include a larger number of disciplines, such as statistics, data mining, machine learning, neural networks, social network analysis, optimization methods and visualization

approaches (Philip Chen et Zhang 2014; Wang et al. 2016). Specific tools and infrastructures are also needed to process data, for example the platform Hadoop for managing and exploiting large data sets across computer clusters (Philip Chen et Zhang 2014). In addition to technical enabling factors, personal skills mixing knowledge about domain and analytics techniques are highlighted to be important factors (Akter et al. 2016). Finally, organizational factors are mentioned in literature, for example foster collaboration between IT and marketing departments within a firm to foster service innovation (Troilo, De Luca, et Guenzi 2017). Interactions with the ecosystem are often mentioned as important factors (Tan et Zhan 2017; Del Vecchio et al. 2018).

4. Research gaps

If we compare the different situations that are commonly reported in the literature and the specific context of Earth Observation mentioned in the introduced, it seems that notion of "distance" between data and usages has little been taken into account. Indeed, in the examples given above, it is quite easy to understand that there is a link between collected data and final usage (data on consumer behaviors for new services for customers; data on internal processes to improve these processes; or even when a process needs to be built in the example of Sage BioNetworks, it seems apparent that data about the application users is related to the study of symptoms of these individuals). So in this type of situation, there is at least an apparent link between data and usage, even if this link remains to be constructed. Whereas in the case of Earth Observation data, in many cases the link is not apparent at first sight (to take an example in the energy sector, the grid operators might not be aware to what extent and how satellite data could be used for their operations). This distance idea is not completely absent in literature. Recent research has examined how big data could be used to face societal challenges and points out that in several cases new usages have been developed thanks to technologies that were not initially designed for such purposes, for example military satellites for rain forest conservation (Chandy, Hassan, et Mukherji 2017). However, it seems unclear how this distance is taken into account in the overall process of value creation, and what makes possible to expand the initial usage to new ones. Therefore this paper aims at addressing the following questions: *how to build a relationship between data and usage when they seem very distant at first sight? Once a data-usage exists, how to make it evolve overtime?*

Intuitively, it appears that specific efforts are needed to first clarify the nature of the link between data and usage, therefore questioning the nature of the intermediate information stated in literature and the mathematical models making explicit the link between data and usage. Moreover, as data and usage are

likely to belong to very different worlds and communities, it seems that the role of the ecosystem in this value creation process is central. Therefore, a specific focus will be made on these three elements: *to build data-usages relationships, what are the specific roles of information, models and ecosystem in a long-term perspective ?*

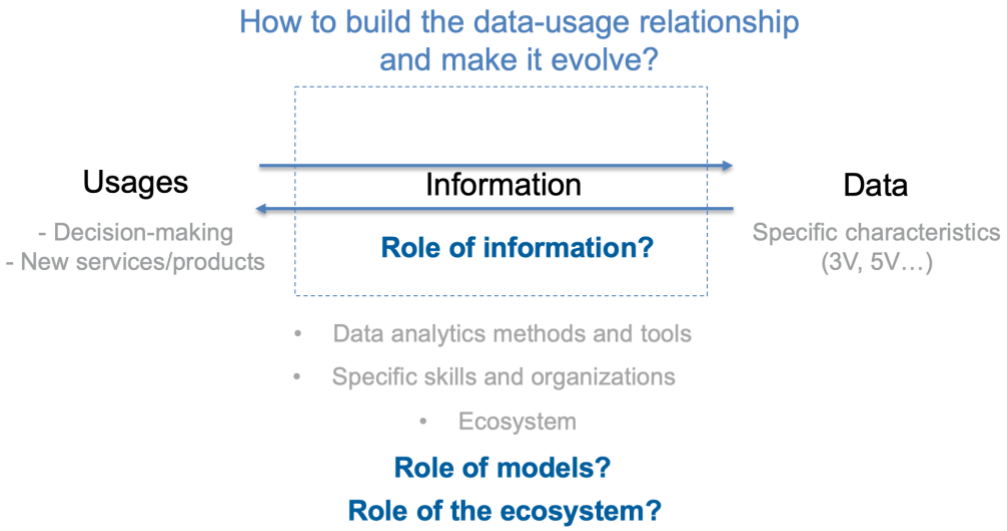

Figure 10: Summary of literature review (grey) and research questions (blue)

III. Research settings

1. Methodology

Our research work is based on a historical case study in the Earth Observation field. As suggested by (Siggelkow 2007), it is leveraged to rediscuss basic notions, "sharpening existing theory by pointing to gaps and beginning to fill them". Our case study describes a specific context that does not correspond to the two types of situations mentioned above: the starting point is neither the data nor a target usage. In this case, the organization had both access to new data and a target usage, and had to build a relationship between them. Moreover, studying a historical case allows us to analyze how this

relationship had been built progressively. This case study is analyzed by taking two complementary perspectives.

In a first stage, the organization focused on building the relationship between data and information. Moreover, because of the specific context of having both access to data and target information, the organization only had to focus on one aspect of the relationship: the model linking data and information (that is the mathematical function), and not sustainability aspects or specific processes to put in place as for the example of Sage Bionetworks. Therefore, it appears that this case is particularly adapted to examine the role of designing models, that seems to be overlooked in literature.

In a second stage, the organization had made this model evolved and had also focused on linking information with new usages. For our research questions, this case study helps us to analyse how the relationship data-usage is built overtime, especially showing the impact of the model structure and the role of the ecosystem in the long run.

2. Empirical material

Our case study is based on the history of the research work of a lab based in Sophia-Antipolis, from the 1980s up to now. This lab has been working on solar radiation estimation (information in the previous framework) from satellite data, and the development of services in the solar energy field (usages in the previous framework). Regarding the two perspectives of our analysis:

The first focus is made on the early years of the lab's research work on solar radiation, when it was involved in a project supported by the European Commission's Solar Energy R&D Programme. The project aimed at assessing solar radiation more precisely and reliably, especially by integrating new data coming from satellites (whereas at the time solar radiation was mainly derived from networks of "in-situ" solar instruments, that were installed in a limited number of locations). Within this project, three different research institutes developed models in order to link solar radiation estimates and Earth observation data including new satellite data.

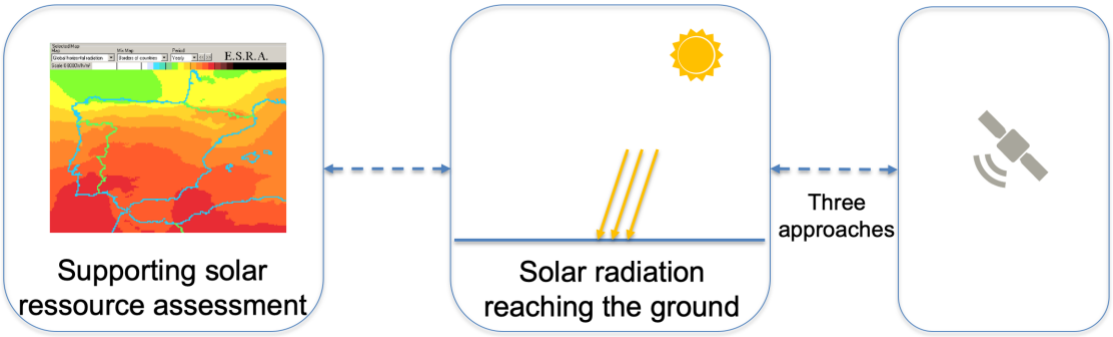

Figure 11: First focus of the analysis: comparison of three modelling approaches

The second focus of our analysis examines how the research lab has built and rebuilt over time both the relationship between data and usages, analyzing the evolution of the lab's research work and the interactions with its ecosystem.

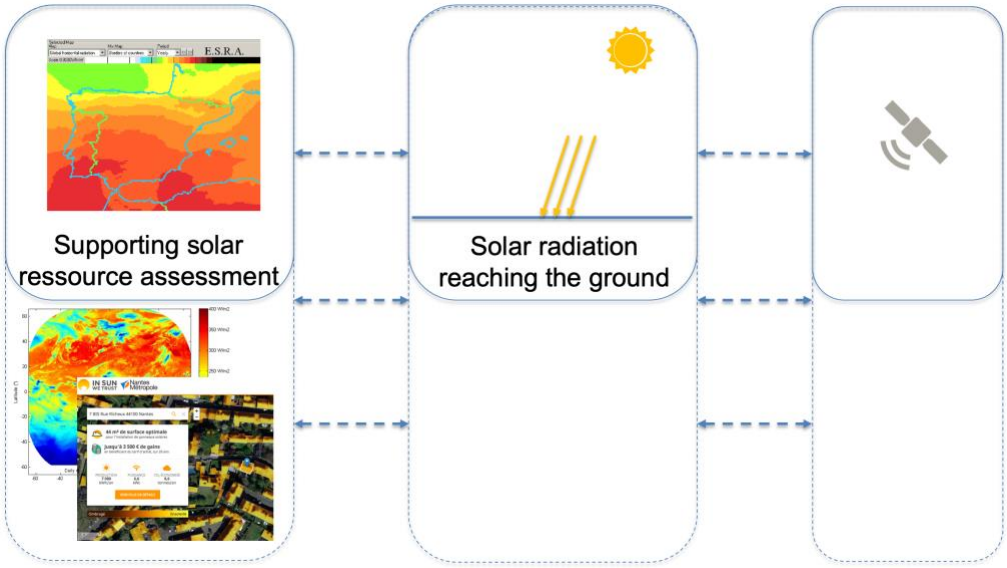

Figure 12: second focus of the analysis - expansion of data-usages relationships

3. Data collection

The first focus on our case study is based on second hand documents, composed of: (1) the final report of the European Commission describing the objectives of the project and the results of the three different approaches to derive solar radiation from satellite data (Grüter et al. 1986); (2) the PhD thesis detailing the specific modelling approach developed by the lab (Cano 1982).

The second focus of our case study is based both on interviews and second-hand documents. Semi-structured interviews (6 hours in total) were conducted with the researcher who had been working on the development of the solar radiations methods from the European project in the 80s to 2018. In addition to the interviews, we used scientific papers published by the lab allowing us to see the evolution of their research work.

IV. Case study analysis

1. First focus: building the model to obtain solar radiation estimates from satellite data

The comparison between the models developed in the 80s by the three different labs for the European project suggests the importance of the model structure. The three approaches can be described as follow:

Physical approach

One lab resorted to a physical approach: the model is based on a "radiative transfer model" that explicitly describes the physical processes (absorption, scattering etc…) occurring in the atmosphere. In this approach, the underlying idea is that satellite data are used to estimate existing parameters of the physical model. The main advantage of this type of model is first to be understandable (as it relies on physical considerations) and to be more general compared to statistical models as it does not depend on specific areas. However, this approach requires additional sources of data coming from ground stations to estimate some of the parameters, resulting in higher computational times.

Statistical approach

Another lab resorted to a statistical approach: the model is based on statistical regressions between satellite measurements and solar radiation values at the Earth's surface, measured by "in-situ" stations within the considered area. In this approach, no knowledge about physical phenomena occurring in the atmosphere. This results in relatively simple processing times. However, there is a lack of generality as the regression coefficients are determined for a specific location and have to be tuned in other areas.

Physical-statistical approach

Sophia-Antipolis lab resorted to a combination of the two previous approaches. They introduced an intermediary variable coined "cloud index", that had a physical meaning as it was describing the level of cloudiness. This intermediary variable was a way of combining the advantages of the two previous approaches:

(1) Understandability and physical-relevancy as it is based on physical considerations to link cloud index to satellite measurement, by modelling part of the radiative transfer terms

(2) Simplicity as the cloud index is then statistically linked to solar radiation with a simple linear relationship. It limits the number of regression parameters compared to fully statistical approaches, and is a way to simplify the physical approach (as parameters that are difficult to model are statistically determined).

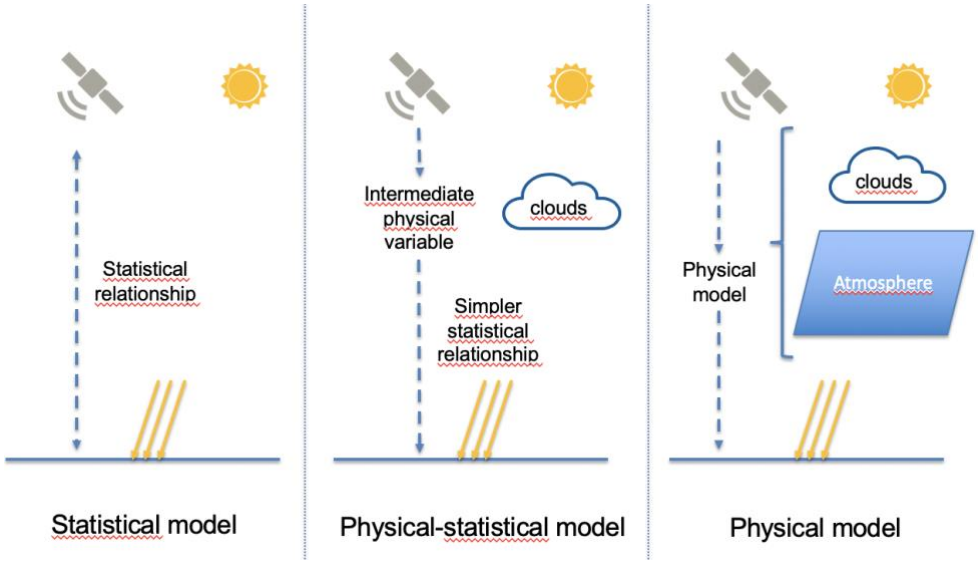

Figure 13: Comparison of the three modelling approaches

This physical-statistical approach had proved to be the most efficient one, in terms of quality of the estimation but also easiness of processing (almost ten times quicker than the physical one). Consequently, this method - called Heliosat – was the one adopted by the European Commission to make the European Solar Radiation Atlas, testifying the success of the approach.

Findings

We are still working on better characterizing the specificities of the approach and the underlying mechanisms that made it as efficient. However we can at least enhance the importance of how models are designed in the value creation process, that seemed to be overlooked in literature. Indeed, this sheds light on another differentiating factor, which is neither related to specific efforts on data sources, nor to efforts on the identification of favorable usages, but to a specific way of designing models.

2. Evolution of data-usage relationship and interaction with the ecosystem

The second focus of our case analysis investigates how the data-usage relationship has evolved overtime and what is the role of the ecosystem in this process. In our case study, the nature of the ecosystem is quite complex as it gathers both the academic community (especially the one working in the Earth Observation field and related application domains), and also private entities and public authorities.

Evolution of the model linking satellite data and solar radiation

Regarding the relationship satellite data – solar radiation estimation, the interactions with the ecosystems seems to have occurred in two ways, from the research lab towards the ecosystem and from the ecosystem towards the research lab. First, how the model was built seems to have influenced the evolution of the ecosystem. Thanks to the mixed physical-statistical approach, the model proved to be efficient but also easily understandable (thanks to the introduction of the intermediate physical variable). As a result, the same model has been adopted by a large number of research organizations that in turn have worked on improving the method, thus contributing to the creation of a community on the same research issues. Second, the ecosystem has played an important role on the evolution of the model linking data and solar radiation. Indeed, Sophia-Antipolis lab has always been modifying their model in order to capture the scientific advances on atmosphere and clouds description thanks to the development of new instrument and computation means, sometime leading to completely rebuild the structure of the method moving to a much more physical approach (notably in 2008 to take advantage

of the second generation of Meteosat instruments). To roughly summarize the overall strategy of the lab, it seems that they have always been targeting physical models but resorting to clever tricks whenever it was not possible.

From an initial usage to new usages

Moreover, the initial usage related to the European Commission's needs has also been expanded. Sophia-Antipolis lab has progressively built web services to make solar radiation information accessible and used by all kinds of communities, outside the EO academic research sphere (for example, farmers, solar energy companies, etc…). The forms of these new usages have evolved over time thanks to the interaction with the users (moving from only solar resource maps to more complex tools indicating variation in time of solar radiation at a specific location,…). A specific business model has been built for these services, based on a freemium approach: free basic solar radiation information at a certain resolution and paid sophisticated services. Furthermore, this process resulted in the creation of new actors in the ecosystem: in order to make the services operational 24h/24h and 7d/7d, the algorithms and processing chain was transferred to a commercial organization in charge of operating, maintaining and commercializing the services.

Implementation of standards

Finally, Sophia-Antipolis has become part of the Global Earth Observation community that aims at integrating observing systems and sharing data by connecting existing infrastructures. Therefore being part of this community has pushed Sophia-Antipolis lab to implement common standards impacting all the elements of the data-usage relationship. Implementing such standards has largely contributed to make Heliosat methods a reference, especially at the European level, as it is now one of the official provider of solar radiation data for the European Commission.

Findings

This case study seems to suggest that another form of ecosystem expansion can be described. In his situation, the ecosystem expansion is not only due to usage expansion resulting from new data flow, nor only to expansion of data sources resulting from new usages pushing for them. It seems that the ecosystem evolution is intertwined with the progressive evolution of the relationship between data and usages, that occurs on all the intermediary forms and processes of the data-usage chain. Moreover, this analysis further strengthens the role of designing models. In addition of being a strong differentiating factor, it seems that the capacity of redesigning them to take the best of recent advances is a key element

for supporting the expansion of the ecosystem and also has an effect on the ecosystem in a retroactive way. Although it is probably not the only factor, the specific structure of the model facilitating its understandability has certainly played a role on its future expansion and the evolution of the lab's ecosystem.

V. Conclusion and discussion

This paper aims at addressing the question of value creation from data in a ecosystem and the expansion dynamics over time.

1. Theoretical contributions

Based on the specific context of Earth Observation where data and usages appear as distant, the paper raises the question of how to build data-usages relationships and make them evolve. The specific roles of information, models and ecosystem have been more precisely examined, taking into account a long-term perspective.

First, a new interpretation of the concept of "information" could be drawn from this case study. In our case study, data and usages appear at first sight as very distant. Therefore, it appears that this information has a significant role to cope with this distance. It could be described as the necessary bridge or pivot between data and usage. Indeed from the user's point of view, solar radiation enables data to be understandable and manageable (unlike raw satellite data). From data owner's point of view, solar radiation is the form that allows technologists to understand user needs in a language compatible with modelling language. It is worth noticing that according the French dictionary TLFI, the verb "inform" in an ancient meaning refers to "give or receive a specific form, a signifying structure to something". We find that this meaning describes quite well the role of information proposed above, as a "form which is more signifying from both data and usage's points of view". We could even talk about "inter-formation" to enhance its role of pivot between data and usage. The framework could be represented as below:

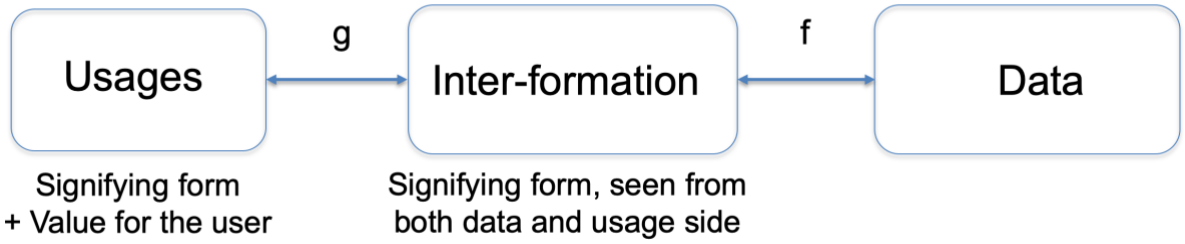

Figure 14: Proposed framework for value creation process

Second, this case study sheds light on the crucial role of designing models in the value creation process. Even when both information and data are given, there is still competition on how the model is designed. It can have both an immediate impact of the efficiency of the method, but also a long-term impact on the evolution of the ecosystem. We believe that these findings could also contribute to the current debate around data science. Indeed, issues concerning best ways of designing models are largely discussed in data-science today: do bigger samples only contribute to a better estimation or also to refine models? If new phenomena become observable, how to build the model enabling to link this new data flow with usages? How to take into account specific domain knowledge in the ways models are built? Our case study could thus give indications on (1) a way of building a model that include specific hypothesis on its possible structure based on a certain domain knowledge; and (2) the consequences on the long-run for the organization building the model and its ecosystem. To investigate this question, our findings still need to be further explored to better characterize the specificities of the physical-statistical approach and better explain the reasons of the successful evolution of this approach.

Third, this research also highlights a specific way of an ecosystem expansion: not only in a “data-push” perspective where data are expected to push for new usages, neither only in a “usage-push” perspective where usage is expected to call for new sources of data, but in both perspectives by progressively building and expanding the data-usage relationship, by playing on all its constitutive elements (data, usage, “inter-formation”, mathematical models, operationalization processes, etc…). To further investigate this question, the different actors of the ecosystem should be more precisely described, especially clarifying on which part of the data-usage relationship they specifically contribute to.

2. Practical contributions

Several practical implications could be discussed at several levels. First, it can be useful for practitioners in science, that are likely to be faced with contexts in which data and usages are not explicitly related to each other (in the case where data are produced thanks to specific instruments for an initial purpose, raising the question of how to explore other usages). Our case study suggests the importance of considering the importance of how their models are designed. It could also be worth noticing that approaches adopted in other fields or companies where the link between data and usage is clearer (for example the use of big data in marketing activities where the collected data is explicitly linked to the data) are not completely adapted and they should be careful of not making misleading parallels.

Our research work could also give some insights for practitioners in companies. Even in situations where there is an existing usage for data or a usage quite clearly related to data, looking for new distant usages could be an interesting way of differentiating themselves. This would require having people able to design models and also taking into account the long-term effect of how models are designed. Other issues will have to be further addressed: how to design this “inter-formation” in a way that encourage new usages to be developed? How could it be practically managed within the companies’ activities? Methods, infrastructures and specific organizations to support the exploration of these possibilities need to be further investigated.

The latter questions could also be considered at the ecosystem’s level. Each element of our data/inter-formation/usage framework can be supported by different actors. As stated above, these actors still need to be further analyzed, however it can be assumed that many different configurations might exist, involving heterogeneous industry, public authorities, and citizens. Our case study highlights a certain configuration, where for example the idea of solar radiation estimation did not come from the lab itself but from the European Commission’s Solar Energy R&D Programme; or the maintenance and operations were taken up by a third party. In other cases, the lab might itself form the “inter-formation” and still take care of the operations. Investigating these configurations might open up new forms of interactions between science and society.

To conclude, we hope that further research could offer a fresh look on the creation and expansion of data-centered ecosystems, at both organization and ecosystem’s levels.

References

3. Abbasi, Ahmed, Suprateek Sarker, et Roger H. L. Chiang. 2016. « Big Data Research in Information Systems: Toward an Inclusive Research Agenda ». *J. AIS* 17: 3. https://doi.org/10.17705/1jais.00423.
4. Akter, Shahriar, Samuel Fosso Wamba, Angappa Gunasekaran, Rameshwar Dubey, et Stephen J. Childe. 2016. « How to Improve Firm Performance Using Big Data Analytics Capability and Business Strategy Alignment? » *International Journal of Production Economics* 182 (décembre): 113-31. https://doi.org/10.1016/j.ijpe.2016.08.018.
5. Bot, Brian M., Christine Suver, Elias Chaibub Neto, Michael Kellen, Arno Klein, Christopher Bare, Megan Doerr, et al. 2016. « The MPower Study, Parkinson Disease Mobile Data Collected Using ResearchKit ». *Scientific Data* 3 (1). https://doi.org/10.1038/sdata.2016.11.
6. Cano, Daniel. 1982. « Etude de l'ennuagement par analyse de séquences d'images de satellite: application à l'évaluation du rayonnement solaire global au sol ». Thèse de docteur-ingénieur, France: Télécom ParisTech.
7. Chandy, Rajesh, Magda Hassan, et Prokriti Mukherji. 2017. « Big Data for Good: Insights from Emerging Markets*: BIG DATA FOR GOOD ». *Journal of Product Innovation Management* 34 (5): 703-13. https://doi.org/10.1111/jpim.12406.
8. Chen, Hsinchun, Roger H. L. Chiang, et Veda C. Storey. 2012. « Business Intelligence and Analytics: From Big Data to Big Impact ». *MIS Quarterly* 36 (4): 1165-88. https://doi.org/10.2307/41703503.
9. De Mauro, Andrea, Marco Greco, et Michele Grimaldi. 2016. « A Formal Definition of Big Data Based on Its Essential Features ». *Library Review* 65 (3): 122-35. https://doi.org/10.1108/LR-06-2015-0061.
10. Del Vecchio, Pasquale, Alberto Di Minin, Antonio Messeni Petruzzelli, Umberto Panniello, et Salvatore Pirri. 2018. « Big Data for Open Innovation in SMEs and Large Corporations: Trends, Opportunities, and Challenges ». *Creativity and Innovation Management* 27 (1): 6-22. https://doi.org/10.1111/caim.12224.
11. Fan, Shaokun, Raymond Y.K. Lau, et J. Leon Zhao. 2015. « Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix ». *Big Data Research* 2 (1): 28-32. https://doi.org/10.1016/j.bdr.2015.02.006.

12. Furtado, Leonel, Moisés Dutra, et Douglas Macedo. 2017. « Value Creation in Big Data Scenarios: A Literature Survey ». *Journal of Industrial Integration and Management* 02 (01): 1750002. https://doi.org/10.1142/S2424862217500026.
13. Gandomi, Amir, et Murtaza Haider. 2015. « Beyond the Hype: Big Data Concepts, Methods, and Analytics ». *International Journal of Information Management* 35 (2): 137-44. https://doi.org/10.1016/j.ijinfomgt.2014.10.007.
14. George, Gerard, Martine R. Haas, et Alex Pentland. 2014. « Big Data and Management ». *Academy of Management Journal* 57 (2): 321-26. https://doi.org/10.5465/amj.2014.4002.
15. Grüter, W., H. Guillard, W. Möser, J. M. Monget, Willeke Palz, E. Raschke, R. E. Reinhardt, P. Schwarzmann, et L. Wald. 1986. *Solar Radiation Data from Satellite Images: Determination of Solar Radiation at Ground Level from Images of the Earth Transmitted by Meteorological Satellites - An Assessment Study*. Solar Energy R&D in the Ec Series F: Springer Netherlands. https://www.springer.com/la/book/9789027722041.
16. Günther, Wendy Arianne, Mohammad H. Rezazade Mehrizi, Marleen Huysman, et Frans Feldberg. 2017. « Debating Big Data: A Literature Review on Realizing Value from Big Data ». *The Journal of Strategic Information Systems* 26 (3): 191-209. https://doi.org/10.1016/j.jsis.2017.07.003.
17. Guo, Huadong. 2017. « Big Earth Data: A New Frontier in Earth and Information Sciences ». *Big Earth Data* 1 (1-2): 4-20. https://doi.org/10.1080/20964471.2017.1403062.
18. Hazen, Benjamin T., Christopher A. Boone, Jeremy D. Ezell, et L. Allison Jones-Farmer. 2014. « Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications ». *International Journal of Production Economics* 154 (août): 72-80. https://doi.org/10.1016/j.ijpe.2014.04.018.
19. Johnson, Jeff S., Scott B. Friend, et Hannah S. Lee. 2017. « Big Data Facilitation, Utilization, and Monetization: Exploring the 3Vs in a New Product Development Process: BIG DATA VOLUME, VARIETY, AND VELOCITY ». *Journal of Product Innovation Management* 34 (5): 640-58. https://doi.org/10.1111/jpim.12397.
20. Lim, Chiehyeon, Ki-Hun Kim, Min-Jun Kim, Jun-Yeon Heo, Kwang-Jae Kim, et Paul P. Maglio. 2018. « From Data to Value: A Nine-Factor Framework for Data-Based Value Creation in Information-Intensive Services ». *International Journal of Information Management* 39 (avril): 121-35. https://doi.org/10.1016/j.ijinfomgt.2017.12.007.

21. Philip Chen, C. L., et Chun-Yang Zhang. 2014. « Data-intensive applications, challenges, techniques and technologies: A survey on Big Data ». *Information Sciences* 275 (août): 314-47. https://doi.org/10.1016/j.ins.2014.01.015.
22. Sheng, Jie, Joseph Amankwah-Amoah, et Xiaojun Wang. 2017. « A Multidisciplinary Perspective of Big Data in Management Research ». *International Journal of Production Economics* 191 (septembre): 97-112. https://doi.org/10.1016/j.ijpe.2017.06.006.
23. Siggelkow, Nicolaj. 2007. « PERSUASION WITH CASE STUDIES ». *Academy of Management Journal*, 5.
24. Tan, Kim Hua, et Yuanzhu Zhan. 2017. « Improving New Product Development Using Big Data: A Case Study of an Electronics Company: A Case Study of an Electronics Company ». *R&D Management* 47 (4): 570-82. https://doi.org/10.1111/radm.12242.
25. Trabucchi, Daniel, et Tommaso Buganza. 2019. « Data-Driven Innovation: Switching the Perspective on Big Data ». *European Journal of Innovation Management* 22 (1): 23-40. https://doi.org/10.1108/EJIM-01-2018-0017.
26. Trabucchi, Daniel, Tommaso Buganza, Claudio Dell'Era, et Elena Pellizzoni. 2018. « Exploring the Inbound and Outbound Strategies Enabled by User Generated Big Data: Evidence from Leading Smartphone Applications ». *Creativity and Innovation Management* 27 (1): 42-55. https://doi.org/10.1111/caim.12241.
27. Troilo, Gabriele, Luigi M. De Luca, et Paolo Guenzi. 2017. « Linking Data-Rich Environments with Service Innovation in Incumbent Firms: A Conceptual Framework and Research Propositions: SERVICE INNOVATION IN DATA-RICH ENVIRONMENTS ». *Journal of Product Innovation Management* 34 (5): 617-39. https://doi.org/10.1111/jpim.12395.
28. Wang, Gang, Angappa Gunasekaran, Eric W. T. Ngai, et Thanos Papadopoulos. 2016. « Big data analytics in logistics and supply chain management: Certain investigations for research and applications ». *International Journal of Production Economics* 176 (juin): 98-110. https://doi.org/10.1016/j.ijpe.2016.03.014.