An Investigation into the Learning Goals and Requirements for the Data Wizard

Bachelor Thesis

Faculty of Engineering Technology – Civil Engineering
Rahadian Rukmana

UNIVERSITY OF TWENTE.

Interreg 2 Seas Mers Zeeën Polder2C's
European Regional Development Fund

Date: 16 July 2021

Conducted by:

Rahadian Rukmana
s2054817
Rahadianrukmana@student.utwente.nl

Commissioned by:

Polder2C's
Ir. R. Hölscher (Ron)
ron.holscher@bzim.nl

Polder2C's
Drs. W. Zomer MSc.
Wouter.zomer@bzim.nl

Supervised by:

The University of Twente
Dr. Ir. V.J. Cortes Arevalo (Juliette)
v.j.cortesarevalo@utwente.nl

The University of Twente
Dr. K. Gkiotsalitis (Konstantinos)
k.gkiotsalitis@utwente.nl

UNIVERSITY OF TWENTE.

Interreg 2 Seas Mers Zeeën Polder2C's
European Regional Development Fund

Preface

I've spent quite a fair portion of my life thinking how it would be like to finally be out there and be an adult. While I still yet to know what it is like, this bachelor thesis marks that distressing milestone of being one (huge) step closer to it.

The past three years that I've spent at the University of Twente has been nothing short of pleasant surprises. I wish to thank everyone that I've met who have made it into one of, if not the best few years of my life. In no particular order, thank you to my teachers, for educating me wholeheartedly, and being patient when I understood it different; my friends, who meant and mean the world to me, for everything - for sincerely and cheerfully putting up with me during my whole study and gave me the much-needed affirmation for the many drafts of this thesis; my best friend, Michelle, whose brace and good humour carried me through many of my days; my student psychologist, Anne, for being there and for giving me the confidence and faith to my journey; and of course, a full-hearted thanks to my family, all of them, for believing in me and for sending me endless love despite the distance and time difference.

I wish to thank Ron and Wouter, for allowing me to perform my bachelor thesis at the Polder2C's project, and for facilitating my research throughout. I also wish to thank Joanne, who have helped me to lay the groundwork early in this research. Above all, my biggest gratitude goes to my supervisor, Juliette, for always making the time for me despite her workload - and patiently, with a smile and energising enthusiasm, sharing her expertise to guide me in the right direction, without which there would be no thesis. Thank you.

Summary

Introduction

Climate change poses significant pressure on many aspects of society. The coastal regions of England, France, Belgium and the Netherlands, just to name a few are a handful of areas that have experienced the adverse effect of climate change such as rising sea levels. Facing these common challenges, Polder2C's, a European cooperation project between the aforementioned nations was initiated to learn and co-produce knowledge to enhance their adaptive capacity and resiliency to climate change. However, the knowledge that is obtained and successfully applied in one country is not necessarily effective in another due to varying contexts, conditions and strategies. Thus, taking into consideration the learning outcomes of involved parties is important to promote mutual learning and knowledge exchange. One of the ways to disseminate the collected knowledge and data from the project is through the so-called Data Wizard – a website meant for sharing data collected from the Polder2C's project.

Research goal

This research contributed to the design process of the Data Wizard. The premise is that by encompassing the learning goals and requirements of its users and stakeholders, the Data Wizard can be better implemented. To achieve the said objective, this research followed a (part of) design science methodology which consists of two phases, namely problem investigation and treatment design phase.

Problem investigation

The problem investigation phase outlined the context or problem surrounding the Data Wizard before deriving any requirements. This phase included activities such as discussions with the developers of the Data Wizard, analysis of project documents and analysis of a questionnaire that was distributed prior to this study to not only identify the problem but also the stakeholders and users of the Data Wizard.

The conversations with the developers highlighted the fact that the Data Wizard does not currently have a precise definition which resulted in different expectations and interpretations to arise – from both the developers and stakeholders as later elaborated. Distinct user groups of the Data Wizard were identified from the analysis of the questionnaire. Three of the most prominent user groups are administrative, university/academics and engineering companies – which are reflected by the stakeholders and people from the project themselves. The questionnaire also hinted at the most relevant kinds of knowledge to be included in the Data Wizard, namely new knowledge, major observations, results and cleaned experimental data – but activity from the project was not specified. Additionally, the stakeholders of the Data Wizard were found by following the information chain in the project which consisted of data production, data processing, data storage/sharing and data exploitation.

Treatment design

The treatment design phase elicited the requirements and learning goals of the stakeholders through a focus group discussion. The focus group discussion dug into their everyday practices, what they would like to learn from the project and their past experiences with platforms containing data/knowledge similar to the one produced by the Polder2C's project. It was found that activities

such as overflow test, survey, levee inspection and repair attracted the most interest – which may complement the result from the questionnaire in which the activity for the new knowledge was not defined. The stakeholders also expressed their interest in learnt lessons from the said project activities. This includes a better understanding of erosion processes, the effectiveness of levee repairs and so forth. Requirements for the Data Wizard comprises of user-friendly interface/navigation, classifications of data, and detailed descriptions of the levee conditions, just to name a few. The stakeholders also referred to researchers, students, dike engineers and levee managers as the more specific users belonging to one of the three aforementioned user groups.

Conclusion

This research provides the reader with the stakeholders, user-groups, preliminary learning goals and requirements from the stakeholders for the Data Wizard. The problem investigation phase identified the stakeholders and users of the Data Wizard whereas the treatment design gathered a minimum set of requirements and learning goals of the stakeholders for the Data Wizard. These key findings were synthesised into a design brief which formed the basis for the recommendations.

Recommendations

The recommendations from this research are not only for future research but also what this research thinks to be a necessary step for further development of the Data Wizard. The recommendations are as follows:

1. Adopt a clear definition for the Data Wizard
    a. The developers and stakeholders of the Data Wizard should construct a definition encompassing the scope and boundaries of the Data Wizard
    b. The developers and stakeholders of the Data Wizard should direct the focus of the Data Wizard towards the identified users such as administrative, levee managers, academics/university and engineering companies instead of the general public
    c. The developers and stakeholders of the Data Wizard should determine the lifespan of the Data Wizard
2. Focus the Data Wizard on areas of greatest impact
    a. Data collectors or people that are working on the experiments and data should focus future efforts on developing high-quality descriptions and materials regarding the experiments
    b. The developers and stakeholders of the Data Wizard should explore other features and needs to be incorporated into the existing architecture of the Data Wizard
3. Gather more perspectives, preferences and desires for the Data Wizard
    a. The developers and stakeholders of the Data Wizard should perform a similar workshop to gather more information regarding users' needs and desires
    b. The developers and stakeholders of the Data Wizard should create user profiles and use-case scenarios for the Data Wizard
4. Create a collaborative environment for the partners

Table of Contents

Preface 3

Summary 4

Table of Figures 8

Table of Tables 9

1. Introduction 10

1.1 Background 10

1.2 Involved parties 11

1.3 Problem definition 12

1.4 Research dimensions 13

1.5 Report outlines 13

2. Theoretical Framework 14

2.1 Knowledge Management 14

2.1.1 Knowledge creation 15

2.1.2 Knowledge retrieval/storage 16

2.1.3 Knowledge transfer 16

2.1.4 Knowledge application 17

2.2 Theory on the use of technology 17

2.2.1 Theory of Planned Behaviour (TPB) 17

2.2.2 Technology Acceptance Model (TAM) 18

2.3 Synthesis of the theoretical framework 19

3. Methodology 20

3.1 Problem investigation 21

3.1.1 Stakeholder identification 21

3.1.2 Analysis of questionnaire 21

3.2 Treatment design 22

3.2.1 Focus group discussion with the stakeholders 22

3.2.2 Design brief 24

4. Results 25

4.1 Problem Investigation 25

4.1.1 Stakeholder Identification 27

4.1.2 Questionnaire Analysis 28

4.1.3 Synthesis 34

4.2 Treatment design 35

4.2.1 Focus group discussion with the stakeholders 35

4.2.2 Design brief 41

5. Discussion 45

Recommendation 1 – Adopt a clear definition for the Data Wizard 45

Recommendation 2 – Focus the Data Wizard on areas of greatest impact 46

Recommendation 3 – Gather more perspectives, preferences and desires for the Data Wizard 47

Recommendation 4 – Create a collaborative environment for the partners 48
5.1 Limitations 50
6. Conclusion 51
Bibliography 52
Appendix A – Questionnaire Script 55
Appendix B – Screenshot of the focus group responses 58

Table of Figures

Figure 1: Partners involved in Polder2C's project (Polder2C's, 2020).....11
Figure 2: Enablers of knowledge flow (Armbrecht et al., 2001).....14
Figure 3: Schematic of proposed methodology per design phase, adapted from Wieringa (2014) 20
Figure 4: The timeline of the Data Wizard's development (adapted from the project document).....25
Figure 5: The proposed architecture of the Data Wizard.....26
Figure 6: Stages of information flow in the Polder2C's project (adapted from the project document).....27
Figure 7: Questionnaire responses based on disciplines.....29
Figure 8: Respondents' involvement in the work packages (WP) 29
Figure 9: Respondents' role on the Data Wizard based on the disciplines.....30
Figure 10: The importance of prospective content for the Data Wizard according to the respondents.....31
Figure 11: The relevance of prospective content according to engineering company discipline 32
Figure 12: The relevance of prospective content according to university/academics' discipline 32
Figure 13: The relevance of prospective content according to administration discipline.....32
Figure 14: Proportion of responses on the relevance of Data Wizard to each target audience 33
Figure 15: Focus group participants' interest in the Polder2C's activities 35
Figure 16: Screenshot of the word cloud regarding what the stakeholders think the Data Wizard is 37
Figure 17: The importance of proposed Data Wizard features 38
Figure 18: Word cloud – participants outlining other potential Data Wizard users.....40
Figure 19: The disciplines of the focus group participants 58
Figure 20: Participants' interest in project activities 58
Figure 21: Participants' interest in new knowledge from the project.....59
Figure 22: Participants' expected added value from the new knowledge 60
Figure 23: Participants' awareness regarding the Data Wizard 61
Figure 24: Word cloud - Participants' view on what they think a Data Wizard is 61
Figure 25: Data types that the participants are interested in accessing through the Data Wizard.....62
Figure 26: The importance of the proposed Data Wizard features 62
Figure 27: Participants' view regarding how the Data Wizard would complement/restrict their everyday work.....63
Figure 28: Platforms used by the participants to get data similar to the one produced by the Polder2C's project.....64
Figure 29: Participants' positive and negative past experience with the aforementioned platforms.....65
Figure 30: Participants' opinions on how the Data Wizard should be used after the project ends.....65
Figure 31: Word cloud – participants outlining other potential Data Wizard users.....66
Figure 32: Participants' opinion regarding what would encourage/discourage the aforementioned users to use the Data Wizard 67
Figure 33: Participants' opinion on whether or not a similar workshop should be organized 68

Table of Tables

Table 1: Description of questions about participants' interests 23
Table 2: Description of questions for deriving participants' requirements for the Data Wizard 24
Table 3: Stakeholders of the Data Wizard 28
Table 4: Target audience frequency table 33
Table 5: A set of initial learning goals and expected added values from it.....43
Table 6: A set of initial design requirements 44
Table 7: List of the questionnaire's respondents 57

1. Introduction

In Europe, the coastal regions of England, France, Belgium and the Netherlands that are located along the southern North Sea and the Channel are vulnerable to climate change; each of which observed a considerable sea-level rise (Secrétariat Technique Conjoint INTERREG IV A 2 Mers, n.d.). Facing these common challenges, Polder2C's, a European cooperation project between the aforementioned nations was initiated to learn and co-produce knowledge that would enhance and hasten their resiliency and adaptive capacity to the adverse effects of climate change such as sea-level rise and flooding. Resiliency is often associated with or defined as the adaptive capacity of a system (Mehvar et al., 2021; Batty, 2008). Adaptive capacity adheres to the ability or design of organizational systems, technologies, behaviours and practices (Becker & Huselid, 2006) which encourages the retention, transformation and sharing of knowledge (Loon, 2019) for adaptive management practices.

Despite the opportunity, the very nature of cross-country cooperation also presents a challenge. The knowledge/data that is obtained and successfully applied in one country is not perforce effective in another (Hanger et al., 2013). Participating countries are bound to their local context, whether it be in the disparity of their (national) adaptation strategies (Swart et al., 2009) or their vulnerabilities to climate change. But above all, extracting knowledge from data in the first place is on its own a challenge. More often than not, data are mismanaged, incomplete and carried insufficient documentation to be understandable and meaningful (Shen, 2018) which impedes the replication and reusability of the data. This is where the Data Wizard is envisioned to step in. With the growing importance of storing, organizing and re-using data sets (Shen, 2018), the Data Wizard, a publicly accessible website, is an initiative from the Polder2C's project that is envisaged to store and disseminate data/knowledge that is obtained from the project activities. Being early in its infancy presents the Data Wizard with the opportunity to embody the learning goals and requirements of the users and ensure its usability in the long run.

1.1 Background

The Polder2C's project is a part of the Interreg 2 Seas that is part of a larger programme, namely the European Territorial Cooperation, or better known as Interreg. Interreg dates back to the 1990s in its infancy as a Community Initiative encompassing cross-border collaboration. Henceforth, it has completed four successive programmes, all of which had the aim to encourage the European nations to overcome challenges in various fields such as health, transport and energy (European Commission, n.d.). Currently, the Interreg V for the period 2014-2020 is ongoing.

The 2 Seas Programme is a form of cross-border cooperation between four Member States namely England, Belgium-Flanders, France and the Netherlands. Its overarching objective is: 'to develop an innovative knowledge and research-based, sustainable and inclusive 2 Seas area where natural resources are protected and the green economy is promoted' (Interreg2Seas, 2015). This objective is further broken down into four thematic priorities, including technological and social innovation, low carbon technologies, climate change adaptation and resource efficiency.

Polder2C's addresses the climate change adaptation theme. The theme has the Specific Objective to "Improve the ecosystem-based capacity of 2 Seas stakeholders to climate change and its associated water-related effects" (Interreg2Seas, 2015). In other words, to develop a joint action plan that enhances the ability of the 2 Seas countries to face the adverse effect of climate change with regards to water such as sea-level rise, flooding, coastal erosion, acidification of marine waters, the rise of water temperature and the occurrence heavy rainfall and droughts (Interreg2Seas, 2015). To do so, Polder2C's, a pilot project, has conducted de-poldering experiments on the about-to-be replaced Hedwige-Prosperpolder levee in the Netherlands, a 6km^2 living lab, to validate and test out the flood defence for the current and future resilience of flood defences (Polder2C's, 2020; Interreg2Seas, 2019). The Polder2C's project consists of four main work packages.

- Work package 1 (WP1) – flood defence: researches on the strength of flood defences through testing such as overflow tests, wave run-up and breach growth as well as a survey of the levee and the environment around it.
- Work package 2 (WP2) – emergency response: is comprised of activities such as levee inspection via an app, emergency response exercises and levee reparation.
- Work package 3 (WP3) – knowledge infrastructure: is aimed to formalize the sharing of knowledge and to prepare young/future water engineers through lessons learnt reports, literature reviews, educational videos, winter school and the Data Wizard.
- Work package 3 (WP4) – field station: encompasses the building of a field station, which is meant to serve as a basecamp for researchers, education and visitors.

Admittedly, there are two more work packages, WP5 and WP6, but these are related to project management and communication respectively.

1.2 Involved parties

The Polder2C's is implemented by 13 partners (and 34 observing partners) from four different countries, namely the Netherlands, Belgium, France and the United Kingdom. The lead partners behind the Polder2C's project are the Dutch Foundation of Applied Water Research (STOWA) and the Belgian Department of Mobility and Public Works (MOW) (Polder2C's, 2020). This bachelor thesis is commissioned by STOWA.

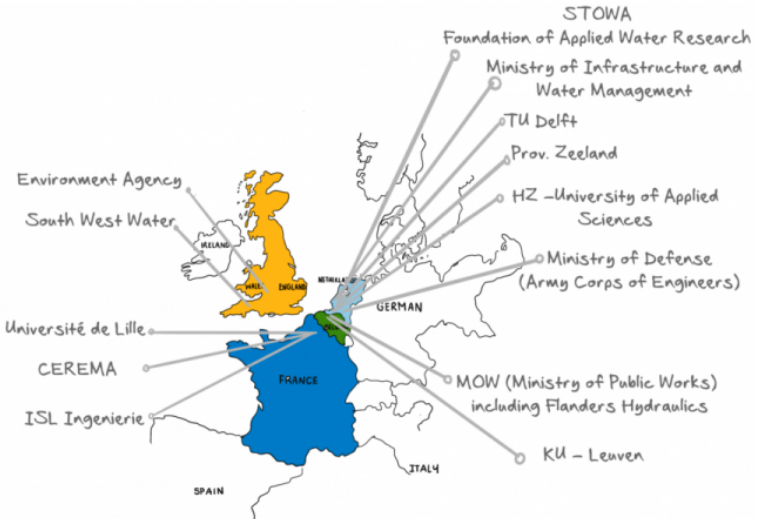

Figure 1: Partners involved in Polder2C's project (Polder2C's, 2020)

1.3 Problem definition

The cross-border cooperation nature of Polder2C's does not only aim to overcome common challenges but also act as a platform where organizations and experts from different nations can collaborate, develop and exchange knowledge for instance, on technological innovations, knowledge and experiences. For these reasons, one of the steps taken by Polder2C's is to develop the so-called Data Wizard; part of the Knowledge Infrastructure programme or Work Package 3 (WP3) (Polder2C's, 2020). The Data Wizard is envisioned as a website to distribute knowledge and data which are obtained from the project activities to the stakeholders and users. Currently, it is still early in its infancy.

This research perceives the Data Wizard as an effort taken by the Polder2C's project to improve their knowledge management and to disseminate the gathered data/information from their project activities. Following this view, this research looks at a similar domain in line with the Data Wizard, namely Knowledge Management (KM). There are various definitions of knowledge management. Webb (1998) defines it as "the identification, optimization and active management of intellectual assets to create value, increase productivity and sustain competitive advantage". Other definitions, similarly, entails the emphasis on ongoing utilization of knowledge.

The Data Wizard, through the lens of information technology (IT), can be an important enabler of knowledge management – often referred to as knowledge management systems (KMS). KMS are instruments that are outfitted to give meaning to data or knowledge sought by users or stakeholders. Therefore, this research resonates more with the interpretation of Alavi and Leidner (2001) and Armbrecht et al. (2001) – KMS is made up of the interaction of people, technology and the knowledge itself. Yet, as literature has shown, creating, identifying, finding and leveraging these pieces of knowledge optimally was found to be difficult (Armbrecht et al., 2001; Alavi & Leidner, 2001). With regards to the Data Wizard, much like in the notion of Knowledge Management, there may be barriers to its successful implementation. This includes flaws in the implementation process, misunderstanding the role of IT, not knowing the objectives and undermining the human factors for the creation and sharing of knowledge (Alavi & Leidner, 2001; Smuts et al., 2009; Martins et al., 2019).

Thus, the Data Wizard's design process must entail the overall objectives of the stakeholders, its users and be guided by the types of available and needed knowledge to contribute to a successful knowledge management system.

1.4 Research dimensions

This research aims to contribute to the design process of the Data Wizard and acts as a preliminary study that would inform the design process of the website with the following objective:

'To investigate the learning goals and requirements for the Data Wizard to be further used in the design process of the website'

To achieve the objective, the research is organized around four main research questions, namely:

1. According to the literature, which theoretical framework is appropriate to identify knowledge management processes to be supported by the Data Wizard?
2. Who are the key stakeholders of the Data Wizard?
3. Who are the end-users of the Data Wizard?
4. What are the learning goals and requirements of the Data Wizard?

The first research question concerns the use of literature to find a theoretical framework to describe the knowledge management processes that are essential to be supported by the Data Wizard. The second research question owes to identifying the key stakeholders of the Data Wizard by firstly understanding the context and circumstances that the Data Wizard is situated in; and consequently, identifying the end-users of the Data Wizard, which is the third research question. The fourth research question delves further into the perspectives of the Data Wizard users from the perspective of the stakeholders. These will be further elaborated on and discussed in the methodology section.

1.5 Report outlines

This report guides you through six successive chapters. The first chapter, as already outlined, introduced the research background and objective of this study. Chapter 2 elaborates on the theoretical framework, consisting of theory on knowledge management (KM) and theory on the use of technology. Chapter 3 expands upon the phases of the design science methodology that was applied, along with elaboration on the adopted activities/data collection methods. The methodology is made up of two phases, namely problem investigation and treatment design. Following these phases, Chapter 4 presents the result from the problem investigation and treatment design phase. The problem investigation phase provides insights into the context behind the Data Wizard – its stakeholders and users. The design treatment phase elucidates the learning goals and requirements of the stakeholders, following the focus group discussion. Key findings are then synthesised into a design brief. Chapter 5 presents the recommendations for the commissioning party based on the design brief as well as the limitations of this research. Lastly, this study is concluded with Chapter 6 – entailing the conclusion of the research.

2. Theoretical Framework

This section is divided into two parts. The first part elaborates on the theory of knowledge management (KM). The second part outlines the theories behind the use of technology, the Theory of Planned Behaviour (TPB) and the Technology Acceptance Model (TAM).

2.1 Knowledge Management

In this rapidly changing world, multiple organizations globally are striving to achieve the best possible way to manage their knowledge assets, not only to generate value for the marketplace but also to gain a competitive advantage over other firms. In retrospect, the focal point of this idea was around tangible assets but it was later conceded that knowledge was what sparked a competitive advantage (Armbrecht et al., 2001). Given the lack of strategies and guidelines to manage knowledge at the time, the field of Knowledge Management (KM) further gained traction and attention; as a result of Nonaka and Takeuchi's publication (1995) – *The Knowledge-Creating Company.* KM aims to provide a framework in which companies could make the most of their knowledge assets by mobilizing organizational knowledge, collaborations and knowledge gained from past experiences/practices (Carillo & Chinowsky, 2006). Henceforth, Knowledge Management has been advocated as a tool that can assist academics, practitioners and particularly businesses with benefits such as revenue growth, customer and staff satisfaction and market leadership; attested by various companies including British Petroleum, World Bank and Chevron (Carillo & Chinowsky, 2006).

Armbrecht et al. (2001) conceptualized KM as a process illustrated in Figure 2. Instead of emphasizing on 'managing', the focal point lies more so in 'enabling' knowledge flow which is influenced by culture, technology and infrastructure. These enablers can be seen as factors that would allow organizations or individuals to achieve their objectives. Culture represents values and norms that are adopted by individuals and organizations, as reflected by their visions and is arguably the most important to create the environment for knowledge creation and sharing (Armbrecht et al., 2001; Loon, 2019). In KM, technology or IT is an enabler that allows users and organizations to store, disseminate and access knowledge after a series of collecting, screening and displaying knowledge where relevant. The Role of IT in KM is diverse and often correlated with the culture of the organization. Infrastructure reflects on the organizational (i.e. hierarchical structure) and physical structures (i.e. physical layout or floorplan) of the organization.

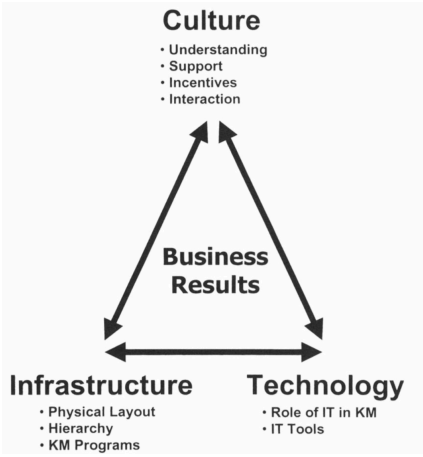

Figure 2: Enablers of knowledge flow (Armbrecht et al., 2001)

The conceptualization of knowledge management by Armbrecht (2001) emphasizes culture as a significant determinant to the success of KM efforts along with the interdependency of the so-called enablers and how if one of them were to change may have an impact on another. While this shows how fundamental the human aspect is in such endeavour, it provides little insight into how to systematically develop the IT systems and the role that they should play. Ergo, further literature research was conducted.

Still within the domain of knowledge management, Alavi and Leidner (2001) provided the framework for analysing the role of an IT system – knowledge management systems (KMS). Knowledge management systems (KMS) alludes to an information system or technology (IT) that is used to manage organizational knowledge. However, there are various roles of and technology for KMS and there is no one-size-fits-all solution (Alavi & Leidner, 2001). The implementation of KMS depends on its purpose – but at the very least has the overall objective to support the so-called knowledge processes, comprised of (1) knowledge creation, (2) storage/retrieval of knowledge, (3) knowledge transfer and (4) knowledge application (Alavi & Leidner, 2001). Other frameworks postulate a more detailed step-by-step approach to develop knowledge management systems but at the very least also alludes to the aforementioned knowledge processes. For instance, Loon (2019) postulates three mechanisms that constitute a KM practice such as learning and knowledge creation culture; organizational knowledge architecture for adaptive and exaptive capacity; and 'business model' for knowledge capitalisation and value capture. The next four subsections elaborate on each part of the knowledge process.

2.1.1 Knowledge creation

Organizational knowledge is created through the development of new knowledge that may replace the existing one (Alavi & Leidner, 2001). Nonaka (1994) created the SECI model, which believes that knowledge can change form from tacit to explicit (and vice versa) through stages of knowledge conversion (Niedderer & Imani, 2009; Nonaka et al., 2000), which is as follows:

- Socialization – tacit to tacit: This is the process where the latest tacit knowledge is created as a result of shared experiences or social contact between organizations or members; which could be in a form of discussions, apprenticeships or social interactions.
- Externalisation – tacit to explicit: Externalisation resorts to the conversion of tacit into explicit knowledge. This stage includes activities like the articulation of best practices and lessons learnt and can be in the form of concepts, metaphors or descriptions.
- Combination – explicit to explicit: Combination, is when new complex explicit knowledge is created as a result of merging, categorizing and synthesizing explicit knowledge.
- Internalisation – explicit to tacit: Internalisation refers to the creation of new tacit knowledge from explicit knowledge as if 'learning by doing' (Niedderer & Imani, 2009). This may include activities such as modelling, reading and reflection.

This model also conveys that organizational knowledge is created due to a continuous interaction of tacit and explicit knowledge dimensions between individuals, groups and organizational levels (Alavi & Leidner, 2001; Niedderer & Imani, 2009; Nonaka et al., 2000). The model is further refined and Nonaka et al. (2000) further introduced the concept of *ba. Ba* in the context of knowledge management resorts to "the shared space for emerging relationships

where knowledge is embedded" (Nonaka et al, 2000); in other words, a common space for creating knowledge (Alavi & Leidner, 2001). *Ba* resembles the aforementioned four modes of knowledge creation (Alavi & Leidner, 2001, Nonaka & Konno 1998), which is as follows:

- Originating/organizational *Ba*: Originating Ba relates to socialization. This could be a commonplace where individuals share experiences, mainly through real-life face-to-face interactions.
- Interacting *Ba*: Interacting Ba is linked with externalisation. This includes activities such as dialogue and collaboration to convert tacit into explicit knowledge and share it.
- Cyber *Ba*: Cyber Ba refers to combination, which takes place in a virtual space for interaction such as IT systems, data warehouses and document repositories.
- Exercising *Ba*: Exercising Ba corresponds to internalisation which turns explicit to tacit knowledge via continuous individual learning.

Owing to the SECI model, the concept of *Ba* thus helps to foster knowledge creation by not only informing where but also how interactions should take place to convert the knowledge. Besides, although it was previously mentioned that IT systems were in favour of the combination mode, it has the capabilities to address other modes as well. For example, interactional *ba* may take place via information systems if the system is designed for collaboration and coordination. Internalisation mode may also be promoted when systems support individual learning and learning by doing.

2.1.2 Knowledge retrieval/storage

Knowledge retrieval/storage has the aim to keep track of and retain the acquired knowledge and make it accessible to those in need. The storage and retrieval of knowledge are also known as organizational memory. It includes various knowledge types, for instance, written documentation, procedures, networks of individuals and codified human knowledge. It could be further refined into two types, namely semantic and episodic. The former is related to explicit and articulated knowledge, while the latter refers to tacit and more context-specific knowledge. KMS plays a role in storing and retrieving knowledge through retrieval techniques (e.g. query, databases) that allows users to promptly access knowledge and avoid replicating previous works by reapplying workable solutions (Alavi & Leidner, 2001). The challenge at this stage is in ensuring that knowledge could be made available, understandable and relevant for other people.

2.1.3 Knowledge transfer

Knowledge transfer is an important process of knowledge management since knowledge is dispersed throughout an organizational setting; particularly in transferring knowledge to individuals or locations where it is needed (Alavi & Leidner, 2001). It alludes to allowing knowledge to be accessible to others. An effective and successful transfer may be addressed with information systems and knowledge management practices that bridges the various disciplines of stakeholders and promote collaboration in knowledge production (Kaiser et al., 2016). Knowledge Management (KM) is also believed to increase information exchange among stakeholders (Martins et al., 2019). Yet, in reality, the transfer of knowledge is easier said than done. It is a complex process involving social factors and all levels of the organization.

2.1.4 Knowledge application

Although this stage may be out of the context and reach of this research, ultimately, the main advantage of knowledge management is in the enactment of knowledge instead of the knowledge itself (Alavi & Leidner, 2001). Similar to what was elaborated in the previous section, IT provides the opportunity to assist the process of capturing, updating and providing access to knowledge, hence accelerating the process of knowledge integration and application (Alavi & Leidner, 2001). Owing to the context of the Data Wizard, particular attention can be paid to the benefit of codified best practices, as it enables the users to embark on a faster learning curve after understanding and accessing knowledge of another with similar experiences.

2.2 Theory on the use of technology

The theory from knowledge management provided an essential framework on what ideally the Data Wizard should be able to support as a knowledge management system, such as knowledge creation, storing/retrieval, transfer and application. While it postulates the potential benefits of having such a system, which includes immediate access to knowledge, a faster learning curve and prevention from reinventing the wheel, it provides little insight into the underlying incentives and beliefs of individuals to use it in the first place; and how it may influence the system.

Therefore, in this section of the theoretical framework, two theories regarding the use of technology, namely the Theory of Planned Behaviour (TPB) and the Technology Acceptance Model (TAM) are elaborated.

2.2.1 Theory of Planned Behaviour (TPB)

The Theory of Planned Behaviour (TPB) (Ajzen, 1991), which originates from the field of behavioural sciences is a renowned and proven decision-making theory on explaining the behavioural motivation of individuals to engage in a certain activity (Wehn & Almomani, 2019). For instance, participation in data sharing and knowledge transfer between organizations (Wehn & Montalvo, 2018). The theory postulates that the human decision or willingness to engage in a certain activity, which in this case is the use of Data Wizard, is governed by the so-called beliefs. Taking the behavioural approach allows one to look closely into the near-actual behaviour of people and even gauge the relationships between two or more actors (Wehn & Montalvo, 2018). The theory classifies beliefs into three categories (Ajzen, 1991), namely behavioural beliefs – which reflects on one's attitude; normative beliefs – which alludes to the social pressures and context; and control beliefs – which relates to the perceived behavioural control over contextual aspects. These beliefs are further explained below.

- Behavioural beliefs – Attitude. Attitude alludes to one's expectations or beliefs on the favourable or unfavourable outcomes as a result of engaging or participating in a specific behaviour or activity. Aligning with this definition, therefore the attitude of individuals or organizations towards the use of Data Wizard will likely be tending towards behaviour that results in favourable outcomes. This may include benefits such as knowledge gains, business developments and innovation. Unfavourable outcomes on the other hand may lead to a negative attitude that hinders the use of the website.

- **Normative beliefs – Social pressure**. Social pressure relates to the social norms in the society on what is tolerable but can also be seen as the belief or perception that one has over the social pressure to engage in a certain behaviour or not. Another interpretation of normative belief alludes to the belief on the likelihood that individuals or organizations that one deems important to engage in certain behaviour.
- **Control beliefs – Perceived control over a certain behaviour**. The perceived control reflects on one's capability to engage in a certain behaviour in which the ease or difficulty in engaging is often influenced by circumstantial factors like resources, opportunities and relevant past experiences. Aside from having an encouraging social pressure and positive belief on the outcome, one has to possess the right technical skills, knowledge and experience which could otherwise impede their engagement. In the context of the Data Wizard, this alludes to having the required technological capabilities or integration to everyday work practices to be able to use the Data Wizard optimally.

2.2.2 Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM) originated from the theory of reasoned action (TRA) and TPB, during the times where researchers attempted to grasp the cause of the acceptance or rejection of technology (Marangunić & Granić, 2015). The TPB, which in essence is an extension of TRA, was the result of introducing *perceived behavioural control* to the already existing *behavioural intention/attitude* as elements to predict behaviours of people. These theories however were found to be unreliable to explain the adoption and rejection of technology; thus, TAM was introduced (Davis, 1986). In contrast to the other two theories, TAM, which was introduced by Davis (1986), focuses only on the attitudes to predict behaviour. It postulates that the *attitude* of a user, whether or not to accept a system, is influenced by two primary beliefs such as the *perceived ease of use* and *the perceived usefulness* – with the former influencing the latter. *Perceived usefulness* can be defined as the belief that one has over the benefit to their everyday work or performance as a result of using a certain system. *Perceived ease of use*, refers to one's belief in the perceived degree of difficulty in using a certain system (Sharp, 2007).

As of today, TAM has evolved to numerous versions. Some of which replaced *attitude* with *behavioural intention*, which came from the assumption that *attitude* is not formed when a system is perceived to be useful, but instead a strong behavioural *intention to use* is formed. Other models included *external variables* which entail aspects such as *system characteristics, user training, user participation design* and so forth. Irrespective of the changes over the years, it was nonetheless evident that the *perceived ease of use* was the determining factor to the *perceived usefulness* – and that together they influence the *attitude/intention to use* technology (Davis, 1989). In justifying this outcome, Venkatesh and Davis (2000) identified five variables that were found to be influencing the *perceived usefulness* of a technology, which included: (1) *subjective norm* – that relates to the persuasion of others on one's decision to adopt or reject a technology, (2) *image* – which refers to one's desire to hold a certain prestige over others, (3) *job relevance* – which relates to the applicability of the technology, (4) *output quality* – which alludes to how well the technology execute a certain task and (5) *result demonstrability* – which refers to giving tangible results.

2.3 Synthesis of the theoretical framework

Hence, as elaborated above, KMS has the objective to support the socially-enacted knowledge processes or framework that consists of four interdependent stages, namely: (1) Knowledge creation, (2) storage/retrieval of knowledge, (3) knowledge transfer and (4) knowledge application (Alavi & Leidner, 2001). This framework will act as a conceptual model to organize this research around and to consider the four stages that are to be supported by the Data Wizard as a knowledge management system (KMS) from the knowledge management perspective.

The knowledge processes framework shows that for IT systems to be able to support and complement knowledge management, its design must reflect on and be guided by the user's needs, and the types, scope and characteristics of knowledge to be included. This view is also supported by Armbrecht et al. (2001) that finds maladjustments of IT systems to the user's needs as one of the barriers of KMS. Yet, the framework also highlights the dynamic and continuous process of knowledge management, where given a certain point in time, individuals or groups can take part in different stages of knowledge processes, whether it is in the creation, transfer or retrieval of knowledge. This signifies the multi-faceted nature of the design process and thus the need to assess and understand the role of IT or the Data Wizard to facilitate knowledge management; the intended end-users; and the knowledge producer. For these reasons, supporting theories, namely the Theory of Planned Behaviour (TPB) and Technology Acceptance Model (TAM) was introduced to guide this research to identify the underlying motivations behind the potential users' need to use the website, their desired benefit and what may facilitate or hinders them to reach it.

3. Methodology

The intended approach to achieve the research objective is with the use of design science, particularly with the design cycle methodology proposed by Wieringa (2014). Design science aims to design and investigate the interaction of an artefact to a certain problem context that would contribute to the achievement of stakeholder's goals and resolve of a problem. It also utilizes practical knowledge or theories from science, engineering and facts for an integrated design. Artefact in this context is defined as something that has a certain objective, for example, algorithms, frameworks, techniques and are created and used by people.

The design cycle consists of several stages or design tasks which includes problem investigation, treatment design, validation and implementation. These stages together comprise an iterative process of designing and investigating, thus the term 'design cycle'. This research, however, is mainly concerned with the problem investigation and treatment design stage that would eventually contribute to or be an input for the design process of the Data Wizard. The problem investigation phase revolves around understanding the context of the Data Wizard and the identification of stakeholders and the end-users. The treatment design revolves around eliciting the requirements and learning goals of the stakeholders. Lastly, key findings from both phases are synthesised into a design brief, which acts as a starting point for the recommendations. An overview of the methodology is shown in the figure below.

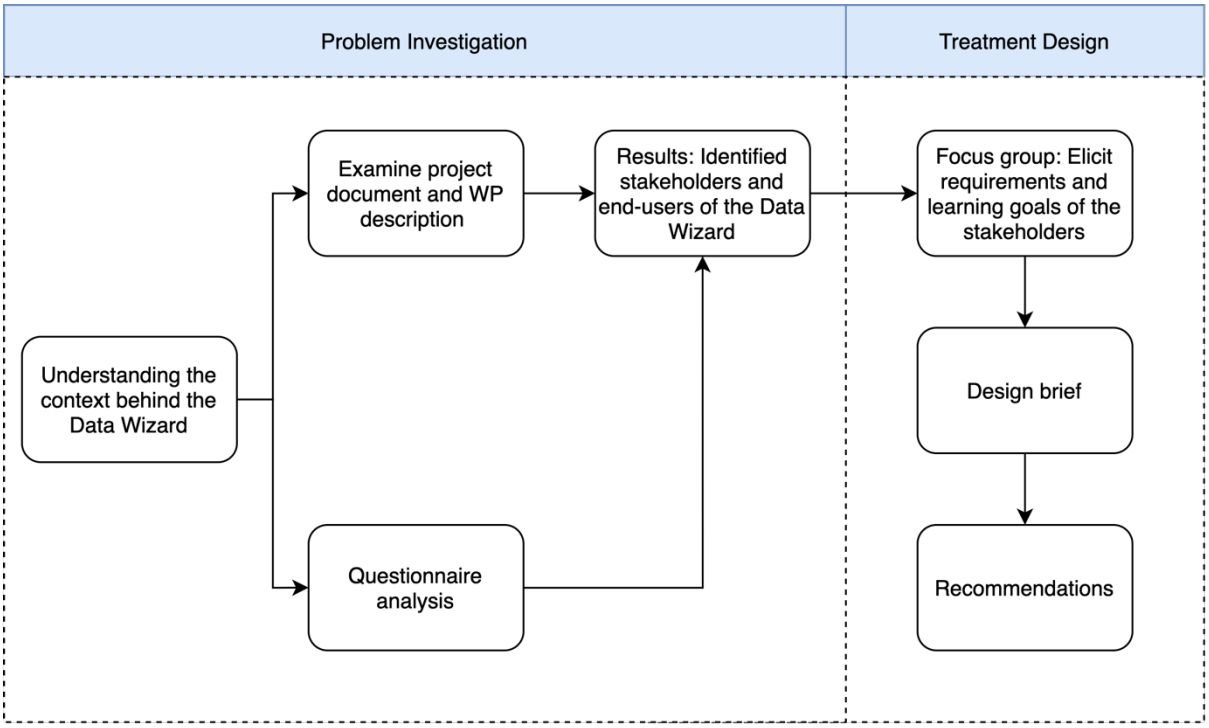

Figure 3: Schematic of proposed methodology per design phase, adapted from Wieringa (2014)

3.1 Problem investigation

The problem investigation stage aims to identify, explain and evaluate the problem to be improved before any requirements for the artefact are derived (Wieringa, 2014). In other words, to learn about the problem before the design of an artefact, which in this case is the Data Wizard. This stage of the design cycle is concerned with the second and third research question; which is to identify the key stakeholders and end-users of the Data Wizard. As such, the problem investigation consists of the following method: discussions with the developers, project document review and questionnaire analysis. Discussions with the developers, from the University of Lille and ISL Engineering, as well as document review was done to identify the information chain in the project; from data collection up to the initiation of the Data Wizard and thereby spotting the stakeholders. This study further analysed the data collected from an online questionnaire that was distributed before this study regarding the project members' view on the target audience/groups and their perspective on the content to be included by the Data Wizard; thus, giving more insight into the purpose and users of the Data Wizard.

3.1.1 Stakeholder identification

Stakeholders are generally defined as people or organizations with interests or stakes on an issue, whether that is because they are affected by or have an influence on an outcome (Wieringa, 2014; Van Velsen et al., 2013). It includes a diverse group of people, for instance: individuals, purchasers and local governments. In a design science research, these stakeholders could be categorized into but not limited to end-users; maintenance operators, that keeps the system going; and operational support, which helps end-users to use the system, just to name a few (Alexander, 2005). Other categories include beneficiaries, developers and sponsors. Anema and Pfeiffer (2017) on the other hand categorize stakeholders into three main categories based on their role in the project, such as core users, enabling environment and market forces; each of which includes their sub-categories of stakeholders. Regardless, identifying and listing stakeholders that are involved in the Data Wizard is necessary so that their needs and desires can be accounted for during the design process (Van Velsen et al., 2013).

In mapping out the stakeholders, particular attention was paid to the information chain within the Polder2C's project starting from data collection leading to the Data Wizard which not only helps to identify stakeholders but also to understand the broader picture and context of the Data Wizard. Thus, stakeholders were categorized based on the identified stages within the information chain. Project documents/descriptions related to the Work Package 3 (WP3) and the information chain were consulted and was complemented with individual discussions with representatives from the University of Lille and ISL Engineering, given their involvement in data management throughout the project and the Data Wizard.

3.1.2 Analysis of questionnaire

Prior to this study, the University of Lille had distributed a questionnaire to the project members. The questionnaire was comprised of two main parts. The first part included a series of qualitative questions to profile the respondents, including their role and preference for the content of the Data Wizard. On the other hand, the second part of the questionnaire included 5-point Likert scale questions to elicit the perspectives of the respondents quantitatively on: (1) the

target audience, (2) the content and (3) their contribution to the Data Wizard. The full list of questions is included in Appendix A.

The analysis of the questionnaire, consequently, is also divided into two parts. The first part pertains to analysing the characteristics of the respondents, whereas the second part analyses the result from the Likert scale on the target audience and content of the Data Wizard.

- Characteristics of the respondents: The questionnaire data were analysed to understand the professional background of the respondents and the relationship between the result and the characteristics of the participants.
- Target audience and content of the Data Wizard: The questions for this part of the questionnaire consists of several 5-point Likert scale questions; with 1 being the lowest and 5 being the highest. To analyse the result, these scores were divided into three categories, namely: least relevant (scores 1 and 2), undecided (score 3) and most relevant (score 4 and 5). The frequency that each category appeared per question was calculated to indicate the importance of certain attributes such as the content or potential end-users for the Data Wizard.

3.2 Treatment design

The treatment design phase alludes to the design of an artefact to treat a problem (Wieringa, 2014); the artefact being the Data Wizard. This phase correlates to the fourth research question, which is to identify the learning goals and requirements of the users. This stage of the design cycle consists of a focus group discussion with the stakeholders.

There are various interpretations of the term requirements. Van Velsen et al. (2013) outlined that there are five types of requirements, this includes: (1) Functional and modality requirements, which are generally meant for programmers, as it incorporates things like technical features and operating systems. (2) Service requirements, which resorts to services around technology such as user support and marketing that is mostly addressed for managers of such services. (3) Organizational requirements, which refers to how technology is consolidated into the organizational structures and working practices. (4) Content requirements, which resorts to the content that needs to be included in the technology. (5) Usability and user experience requirements, which correlates to the user experience, interface and design. Since this study is mainly concerned with the learning goals and requirements from the stakeholders' perspective, therefore the scope of the requirement is narrowed down into the last two types of requirements.

3.2.1 Focus group discussion with the stakeholders

Findings from the problem investigation, stakeholder identification and analysis of the questionnaire formed the basis for the focus group. The result from the questionnaire had hinted at the most important kinds of subject matter to be included in the Data Wizard, such as new knowledge and major observations just to name a few. Be that as it may, those preferences provided very little specificity to the kinds of information to be included given the vast amount of activities performed at the pilot site (e.g. new knowledge is very general and may come from all kinds of experiments that may not be of their interest). In addition, each of the three primary user groups (including those within the user group itself), which was found from the questionnaire

analysis, may have different views on why they are there and what they would like to pick up from the project. This after all relates to the problem definition of this study – where for the Data Wizard to be better implemented, the needs and learning goals of the users must be incorporated. Thus, this was addressed during the focus group session.

The focus group discussion was organized to elicit the learning goals and requirements of the stakeholders. While initially it was planned to perform three separate focus group discussions, each with one of the three main user groups recognized from the previous stages, the time constraint allowed for only one discussion. Therefore, the attention was shifted more towards the stakeholders of the Data Wizard and or people that are primarily involved in the project. The developers of the Data Wizard were also invited to the discussion. Besides, they are also better equipped with the background knowledge to interpret first-hand the suggestions and input from the participants which may guide the direction of their work. In total, 10 stakeholders were invited; four of which were involved in the information chain towards the Data Wizard. The focus group discussion lasted for about one hour.

The focus group consisted of two main parts and was designed in consultation with the external supervisor. The first part of the discussion was arranged to identify the participants' interest in the activities within the Polder2C's project, their need for information/knowledge and its benefit for them. Beforehand, a presentation regarding how the Data Wizard relates to all the work packages was given by the external supervisor to give context into how it would fit in. Subsequently, a presentation was also given concerning the activities that have been carried out in the project, to provide a point of reference to the participants; such that they have an image of what may be of their interest and can be obtained from the project, thus resulting in more discussion. Afterwards, a series of questions were asked, as summarized by the table below.

Table 1: Description of questions about participants' interests

Question	Question type
For us to analyse the results, please select which user-group do you belong to?	Multiple-choice
Which activities from the Polder2C's project are you and your organisation most interested in?	Multiple-choice
What new knowledge would you and your organisation like to learn from those activities?	Open-ended
How do you plan to use this new knowledge? How should the new knowledge benefit you and your organisation?	Open-ended

The second part was organized to derive requirements for the Data Wizard. First, a presentation regarding the Data Wizard was given, showing the timeline and the proposed architecture to give a reference point on what they could expect from the website. Furthermore, this research probed into the participants' requirements for the Data Wizard indirectly through questions relating to their line of work, such as what information would they like to get through the Data Wizard, their experience with a familiar data-gathering-platform and how the Data Wizard may complement their work, just to name a few – since no prototype of the Data Wizard could be shown. By doing so, the participants were encouraged to think more about the

requirements and functionalities that would integrate well with their everyday practices. Further questions were also asked regarding their opinion on the lifespan of the Data Wizard, the users and whether or not this workshop (focus group discussion) should be performed with other user groups. The questions asked during the discussion are summarized in the table below.

Table 2: Description of questions for deriving participants' requirements for the Data Wizard

Question	Question type
Are you aware of the Data Wizard?	Multiple-choice
What do you think is a Data Wizard?	Word cloud
What data type and information would you and your organisation like to get through the Data Wizard?	Open-ended
Which of the following is the most important feature to be included in the Data Wizard?	Scoring/rating
How do you imagine the Data Wizard to complement/restrict your work?	Open-ended
How do you generally gather new knowledge, data/information similar to the one offered by the Polder2C's for your work or interest?	Open-ended
Reflecting on that experience, what should the Data Wizard have or not have to make it easier to use?	Open-ended
How do you want the Data Wizard to be used after the project ends?	Open-ended
Outside of this group, who do you think might use the Data Wizard?	Word cloud
What would encourage/discourage them to use the Data Wizard?	Open-ended
Should we repeat this focus group with potential users of the data wizard directly related or not to the project?	Multiple-choice
Anything else you would like to add?	Open-ended

3.2.2 Design brief

The key findings from the problem investigation and treatment design phase were synthesised into a design brief – a summary of the findings, requirements and learning goals of the stakeholders – which formed the starting point for further recommendations. As a final data-gathering effort, a copy of the design brief and the preliminary recommendations was sent to the stakeholders/focus group participants to get feedback. This allowed the researcher to assess how the stakeholders perceived the recommendations and if there were any common issues – and to improve upon the recommendations.

4. Results

In this chapter, the results obtained from each stage of the design cycle are elaborated, starting with the problem investigation and followed by the design treatment.

4.1 Problem Investigation

The problem investigation stage aims to unravel the context behind the implementation of the Data Wizard, its stakeholders and eventually the intended users of the website. It began with discussions with representatives from the University of Lille and ISL Engineering to understand the current state of the Data Wizard and to identify other potential stakeholders. The project document was also consulted to elucidate the chain of information in the project and to identify stakeholders along the way.

Individual discussions were organized with the representatives from the University of Lille and ISL Engineering concerning their role in the project, as well as the purpose, content and functionalities of the Data Wizard. Regarding their roles, the Data Wizard is the primary responsibility of the University of Lille, while ISL Engineering plays more of the supporting role although it was noted that this is to be further defined as the Data Wizard continues to develop. At the moment of writing, his involvement is mainly in the internal data management system of the project.

Regarding its timeline, the Data Wizard is envisioned to be up and running in the year 2022. Currently, it sits at the second phase of the development process, namely data collection up until July after having previously completed its first phase – architecture design. Moreover, the third phase, which is set to begin in September, is a continuation of the second phase but with the addition of further software development. From the discussions and the document, it was unclear how long the website should remain available.

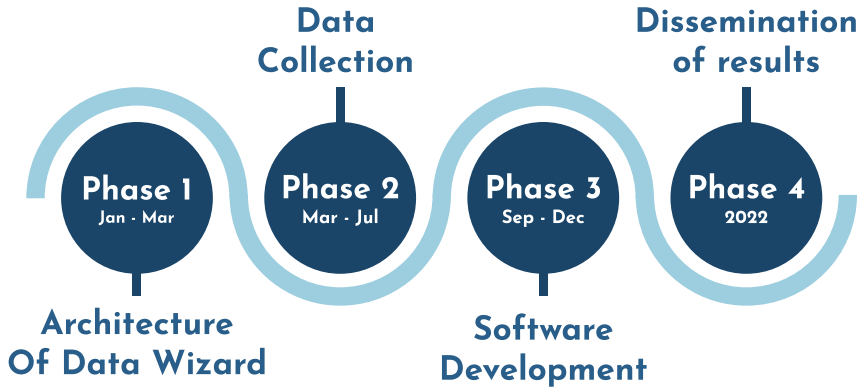

Figure 4: The timeline of the Data Wizard's development (adapted from the project document)

The conversations with the developers further highlighted that there was not one precise definition of what the Data Wizard is. However, there seems to be a convergence from the conversations, with both mentioning the fact that the Data Wizard is meant to allow data from the project to be publicly accessible with a user-friendly interface, such that it helps the users to navigate to the data or knowledge that they are looking for; with emphasis on the latter. Regarding who will access it, nevertheless, remained a grey area. One party vaguely mentioned

that it could be for everyone without much specificity whereas his counterpart suggested that it is primarily for water professionals, research institutes, private companies and other external stakeholders. He further justified his answer by saying that, to his knowledge, the internal stakeholders such as the partners of the project already have access to the data/knowledge obtained from the project via their internal data management system which includes the OneDrive and DDSC (Dike Data Service Centre). This significant difference indicates that there is a discrepancy in their interpretations on who the Data Wizard is meant for.

Being early in its inception, the Data Wizard is still very much receptive to various kinds of functionalities and content. However, it was nonetheless clear that WP1 and WP2 are the main producers of data or knowledge that will be included in the Data Wizard; as confirmed by both parties and the project document. WP1 covers the theme of climate adaptive flood defence whereas WP2 pertains to emergency responses. Regarding the former, there are two main sources of data, namely survey and monitoring. The survey encompasses the collection of data related to the initial condition of the levees and information about the site whereas monitoring entails data from tests conducted on the levees such as the overflow tests. WP2, on the other hand, have less clarity in terms of what data will be included in the DDSC or OneDrive, let alone the Data Wizard, according to the project document and the representatives. Meanwhile, there are various types of activities that have been carried out within the work package, which entails activities like levee surveillance via an app, levee reparation, emergency response exercises and BreachDefender. This was also apparent when observing the architecture of the Data Wizard, as shown below, which includes only the reparation methods from WP2.

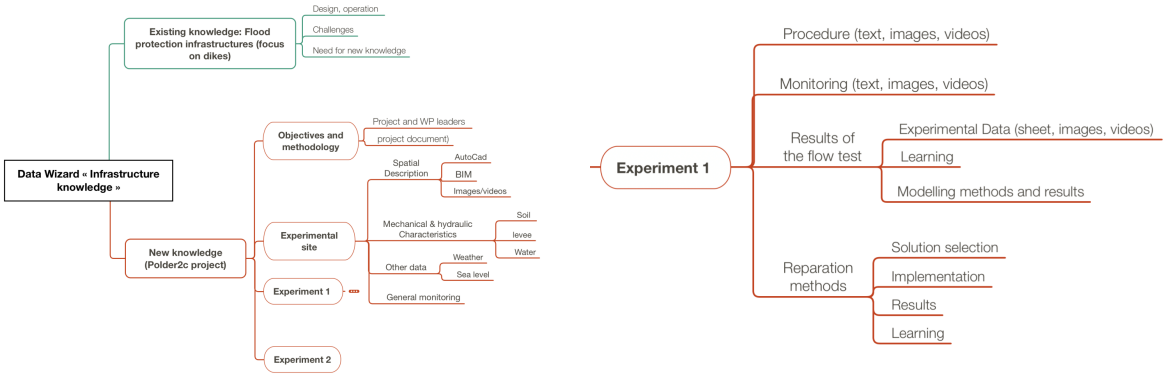

Figure 5: The proposed architecture of the Data Wizard

Concerning the functionalities of the Data Wizard, one feature that seems to be most prominent is user-friendliness, as frequently mentioned by both representatives. Being more involved in the Data Wizard, the representative from the University of Lille also suggested other possible features which entail classifications of data, digital twins of the levee (as a result of the survey from WP1) and modelling of data. Be that as it may, the external supervisor of this study unveiled that said features were still disputable in the eyes of other project members – as to whether or not it should be included in the Data Wizard – who also wondered if real-time visualisation of collected data might be necessary.

4.1.1 Stakeholder Identification

A project document encompassing the information flow in the Polder2C's project was consulted. The report specifically elaborates on the data management system (DMS) in the project but provided a solid basis that illustrates and describes the information flow in the project leading to the Data Wizard. The DMS was designed mainly for partners from WP1 and WP2 to store and share their data and output that are produced on-site that may be useful for partners from other work packages. It was found from this document that there are two separates but primary constituting components of the Data Wizard, namely OneDrive and DDSC (Dike Data Service Centre); which was also confirmed by both people from the University of Lille and ISL Engineering. The DDSC stores raster and time-series data, while other types of static data such as reports and pictures are stored in the OneDrive.

The information chain comprises four main stages, namely data production, local data processing, data storage and sharing, and data exploitation (see figure below). The data production starts with the process of data collection at the experimentation site of WP1 and WP2. Owing to the collection of data and its interaction with DMS, the WP1 is governed by a person from MOW, while for the WP2, it is still to be confirmed; whether it is from the water board (Water Board Hoogheemraadschap De Stichtse Rijnlanden) or Rijkswaterstaat.

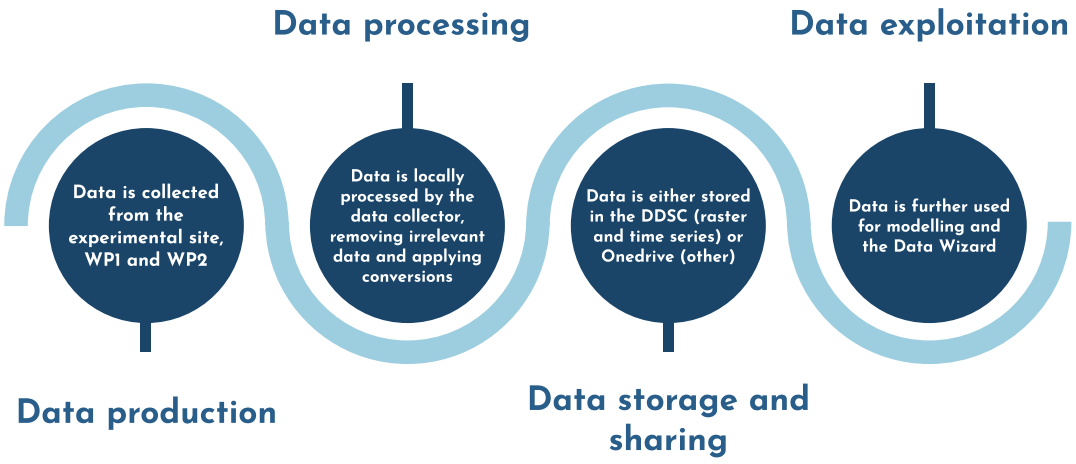

Figure 6: Stages of information flow in the Polder2C's project (adapted from the project document)

The local data processing comprises of cleaning and pre-processing of the raw data by the data collector themselves, for instance, to remove irrelevant data and to apply formulas or conversions; with the raw ones remain available to be shared. These data are then stored in the DDSC and OneDrive with the help of ISL Engineering, thus made available to other partners. The Data Wizard is part of the last stage, namely data exploitation and is the primary responsibility of the University of Lille. The fact that the Data Wizard feeds off the DDSC and OneDrive instead of creating and storing its own data thus imply that the Data Wizard could be a link between the data users to the data stored in the DDSC and OneDrive.

The identified stakeholders throughout the information chain are summarized in the table below. There is also a possibility for stakeholders outside of the information chain, such as Pieter Rauwoens who is the WP3 leader and Ludolph Wentholt, which is the project leader.

Table 3: Stakeholders of the Data Wizard

Information stages	Stakeholder	Role
Data production and processing	Davy Depreiter – MOW	Data producer and manager from WP1
	Mariaan Booltin – Water Board and/or Bart Vonk – Rijkswaterstaat	Data producer and manager from WP2
Data storage and sharing	Davy Depreiter – MOW	Data producer and manager from WP1
	Nicolas Nerincx – ISL Engineering	Facilitates storing of data into DDSC and OneDrive
Data exploitation	Ammar Aljer – University of Lille	Main developer of the Data Wizard
	Stephan Rikkert – University of Delft	Model simulations

4.1.2 Questionnaire Analysis

The questionnaire by the University of Lille was distributed to gauge the opinions of the respondents on three main aspects of the Data Wizard (see appendix A), namely the target audience, content and the respondents' contribution to the website, as will be elaborated in this section.

Characteristics of the respondents

This analysis begins with looking at the quantitative findings regarding the details of the respondents. Out of the total 28 responses gathered, the majority of people who filled out the questionnaire were project partners (16 responses, 57%), followed by the project observers which amounts to nearly a third (9 responses, 32%). The rest, or others (3 responses, 11%), distinguished themselves as either a researcher, part of the project management team or a member of the Polder2C's consortium. Furthermore, nearly half of the people, 47% (13 responses), were found to be working in the public sector whereas the private and academic sector has a share of 21% (6 responses) and 32% (9 responses) respectively. The proportions are more thoroughly distributed when looking at the respondents' disciplines, as shown by the figure below, although a similar percentage is observed for the academic sector, university/academic. The share of administration and engineering companies combined also represents nearly half of the responses. Further details of the respondents are attached in Appendix A.

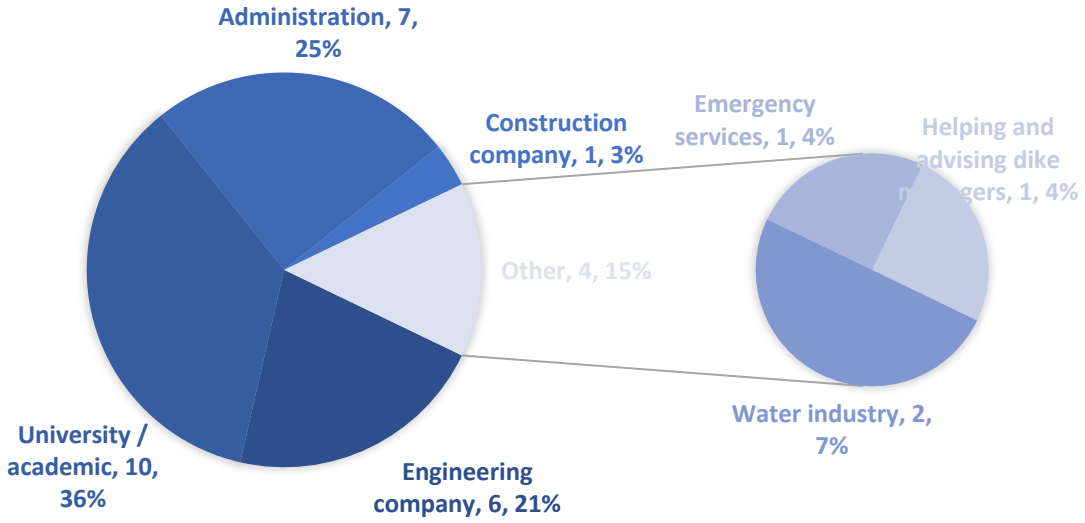

Figure 7: Questionnaire responses based on disciplines

The questionnaire also showed that the respondents came from various work packages (WP) within the Polder2C's project. The figure below shows the distribution throughout the different work packages. It is to be noted that an individual could participate in multiple work packages; thus, the sum of people from each of the work packages exceeds the number of total responses. As can be seen, the majority of people are involved in WP1, WP2 and WP3; which are closely related to the Data Wizard.

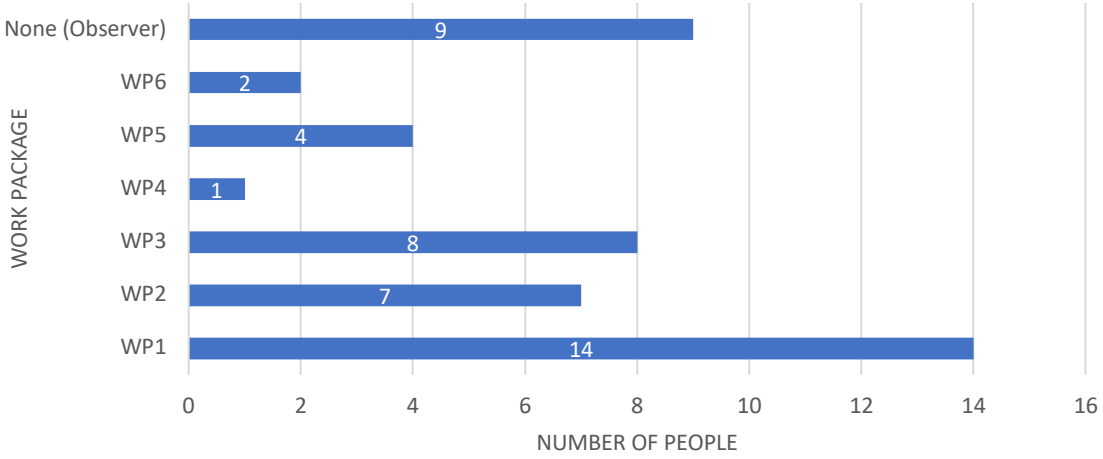

Figure 8: Respondents' involvement in the work packages (WP)

The questionnaire delves further into the roles that the respondents have in the Data Wizard; whether one is a data provider (experimental or model), data manager, data analyst or future user as shown by Figure 9. It was decided to portray the result together with the aforementioned disciplines to give a clear division or distinction of its provenance; or in other words, to show how the disciplines are distributed throughout various roles. Depicting it with the sectors proved to be too broad to give any meaning, whereas it was too difficult with the work packages because an individual can be part of multiple work packages; hence overlapping data. Furthermore, it is worth

noting that the data of the participant from the discipline of the construction company was omitted since the participant did not express their role in the Data Wizard.

As can be seen, most respondents associate themselves to have a role as either the future user of or data analyst for the Data Wizard; with each reaching a total of 14 people. Note that an individual can have multiple roles in the Data Wizard, therefore the total number of people shown in the graph exceeds the number of responses. Furthermore, as illustrated, the two prominent roles showed a similar composition in terms of the disciplines with the university/academic and administration field representing the majority. Moreover, the engineering company is evenly distributed throughout all the roles. Given the prominence of these three disciplines, as shown in Figure 7, it explains why all the roles are filled with those disciplines except for data management; where the administrative discipline is absent.

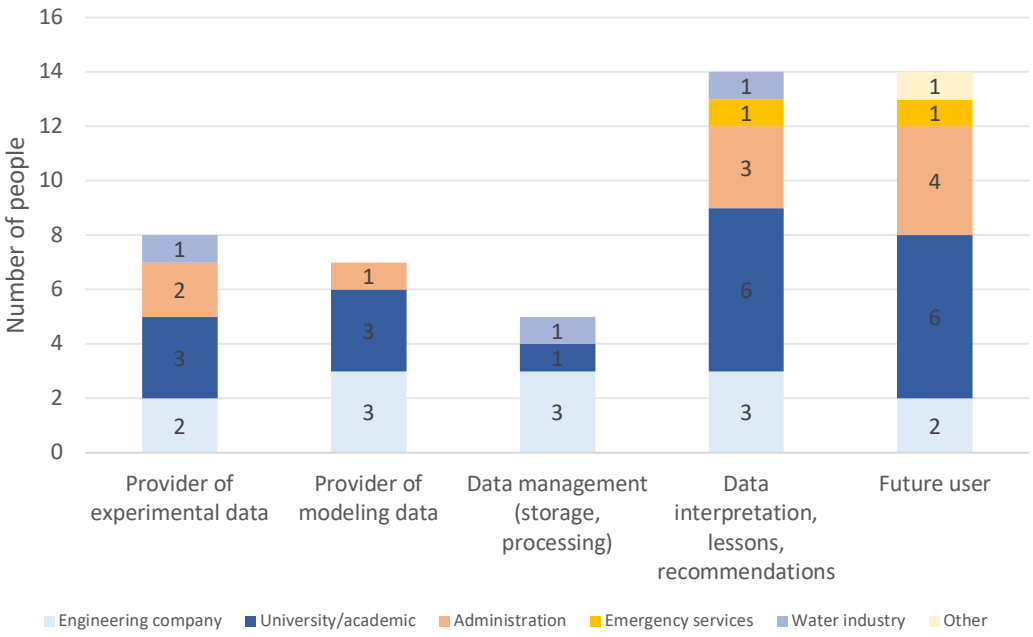

Figure 9: Respondents' role on the Data Wizard based on the disciplines

Content of the Data Wizard

The questionnaire revealed that all of the listed potential contents were relevant to the Data Wizard except for the raw experimental data as shown by the figure below. Responses for the raw experimental data was found to be quite evenly distributed, but the majority (39%) being indecisive about whether or not to include it in the Data Wizard. However, in comparison to the other types of content, it was apparent that it has the least importance/relevance to the Data Wizard as indicated by the significantly high proportion of responses giving it a score of one or two (32%). New knowledge, major observations/results and cleaned experimental data, on the other hand, were deemed most important to the Data Wizard.

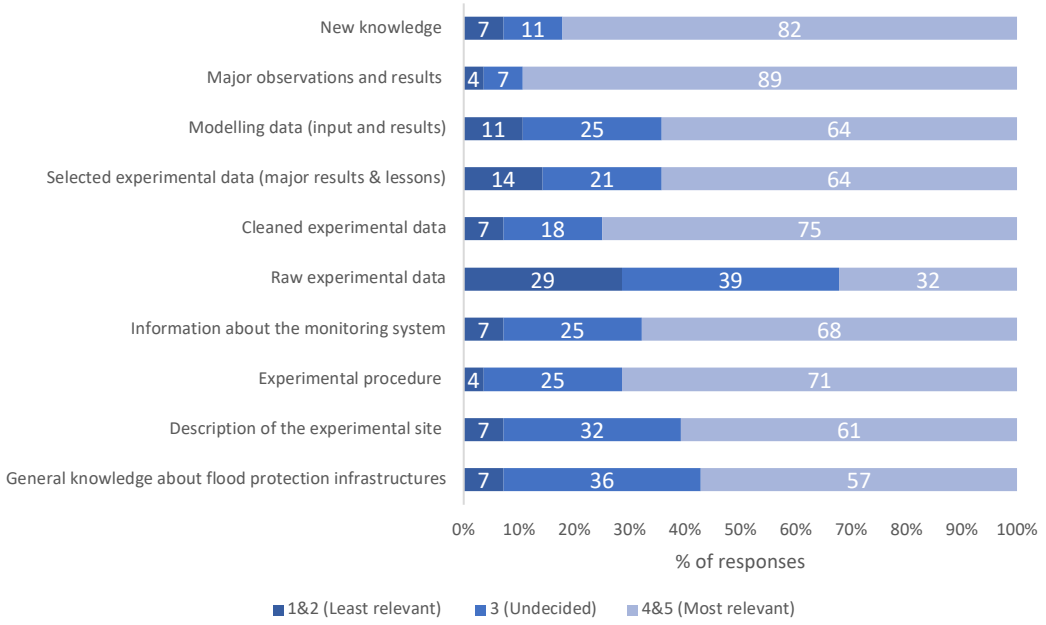

Figure 10: The importance of prospective content for the Data Wizard according to the respondents

As earlier shown in Figure 7, the majority of the responses came from three main disciplines, namely engineering company, university/academics and administration including the part where they distinguish themselves as future users. Therefore, being potentially the primary users of the Data Wizard, it was decided to further delve into their individual content preference; as this possibly also resembles their needs. From the figures below, it is apparent that new knowledge, major observations and results, selected experimental data and information about the monitoring system are of absolute importance for the engineering company discipline. Modelling data (inputs and results), cleaned experimental data and experimental procedures are of secondary importance. This shows that the engineering company values major takeaways, key lessons and methods in particular from the project instead of raw unprocessed data.

The university/academics and administration similarly are keen on new knowledge, major observations and results, with some interests in the information regarding the monitoring system, cleaned experimental data and experimental procedure. University/academics also have a significantly higher interest in the raw experimental data compared to the other disciplines.

Overall, the three disciplines have similar interests revolving around major takeaways and results from the project, with few having major interests in the raw data. One major concern in interpreting the result of this question is that the respondents were not able to indicate specifically which activities they were referring to when they indicated their preference for the content of the Data Wizard as shown below. For instance, different respondents may refer to different activities such as the overflow or breach test when choosing the same answer, experimental procedure.

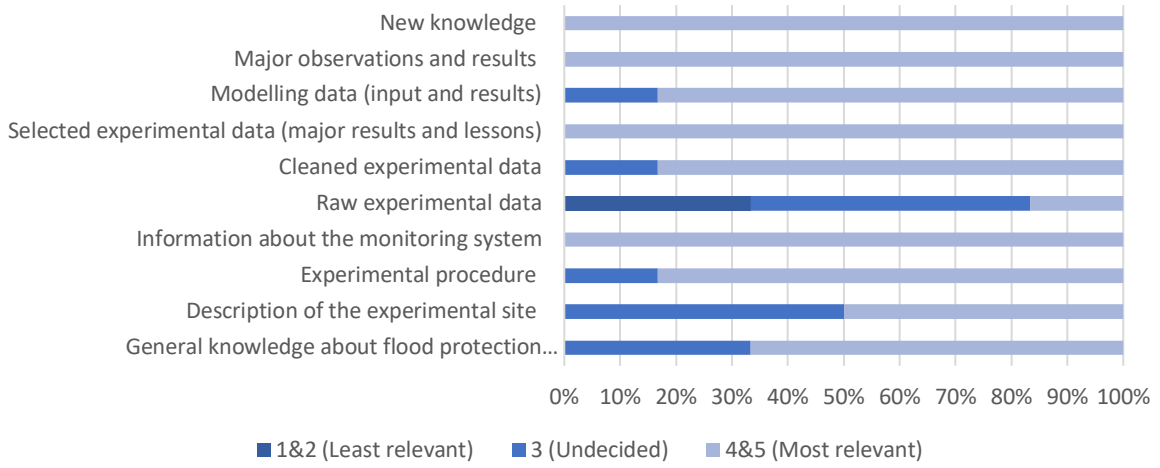

Figure 11: The relevance of prospective content according to engineering company discipline

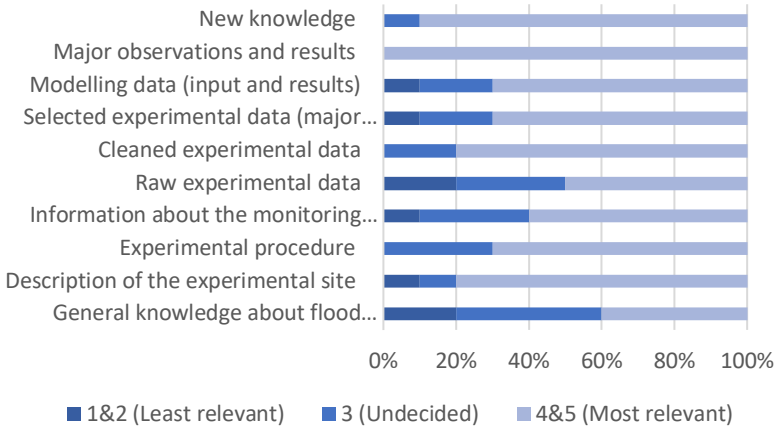

Figure 12: The relevance of prospective content according to university/academics' discipline

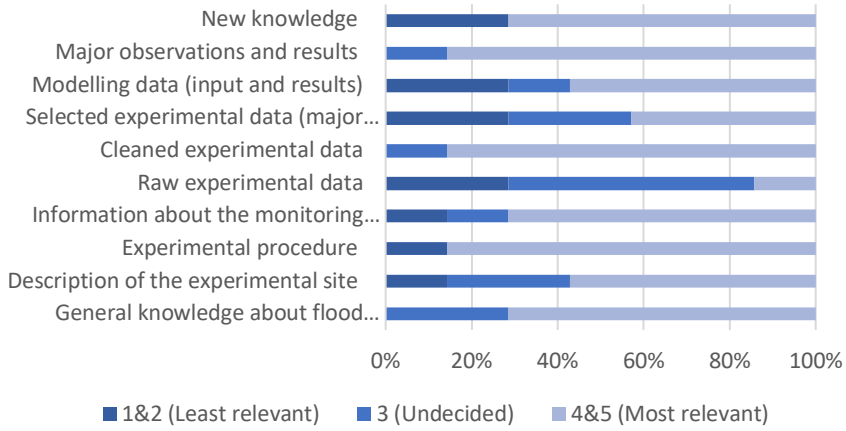

Figure 13: The relevance of prospective content according to administration discipline

Target audience

The respondents were asked to grade the relevance of Data Wizard to a list of prospective target audience, consisting of professionals, administration, emergency services, academics and students. Following the same procedure for analysing the previous Likert-scale questions, the frequency that each category appears (not relevant, undecided, relevant) per question or prospective target audience was calculated, as shown by the table below. Similarly, Figure 14, was made to better illustrate the result and proportion of each score.

The figure shows that the respondents have a general coalescence that the Data Wizard is most suitable for professionals and academics/students as indicated by the highest proportion of the scores four and five. Administration, on the other hand, has the lowest relevance according to the respondents as indicated not only by the lowest percentage of high scores (fours and fives) but also the highest percentage of low scores (ones and twos). Admittedly, as previously mentioned, the questionnaire was mainly filled by people with a background in administration, university/academics and engineering companies who also identified themselves as future users. Thus, it is no surprise that the identified future users also resemble the respondent's disciplines.

Table 4: Target audience frequency table

	Score frequency		
Potential Target Audience	1&2 (Least relevant)	3 (Undecided)	4&5 (Most relevant)
Professionals	0	5 (18%)	23 (82%)
Administration	8 (29%)	8 (29%)	12 (43%)
Emergency services	5 (18%)	8 (29%)	15 (54%)
Academics and students	1 (4%)	9 (32%)	18 (64%)

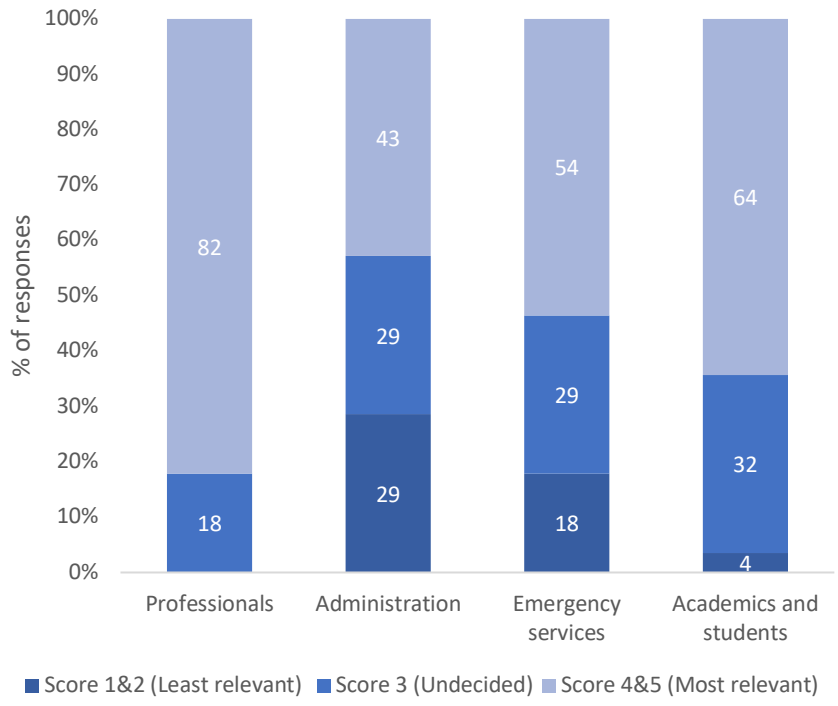

Figure 14: Proportion of responses on the relevance of Data Wizard to each target audience

Respondent's contribution to the Data Wizard

It was decided to omit the result of this section of the questionnaire due to its likeness with the first few questions asking about the respondent's role on the Data Wizard, but rather in a more unclear fashion with its 5-point Likert scale question (see Appendix A for the questionnaire).

4.1.3 Synthesis

The problem investigation has fruitfully addressed the second and third research question of this study, namely in identifying the stakeholders and end-users of the Data Wizard. The stakeholders were identified by following the information chain of the project from data collection up to the development of the Data Wizard which has four main stages: data production, processing, storage and exploitation. It can be seen that the process leading to the Data Wizard resembles a process much like the aforementioned four stages of knowledge processes by knowledge management. Data production in the project coincides with knowledge creation, data storage and sharing relate to knowledge retrieval/storage of knowledge management, while data exploitation corresponds with knowledge transfer and application. Furthermore, the primary end-users of the Data Wizard were also distinguished from the analysis of the questionnaire. These are people from administrative, university/academics and engineering companies – not the general public.

The problem investigation shows that the Data Wizard is not the sole data management system in the project. The project partners after all have access to the data obtained from the project and experiments without the existence of the Data Wizard. It is possible that the Data Wizard, with its different approach and interface, may prove to be useful to them, as shown by the high interest gauged from the questionnaire – which begs the question, what would they like to get from the Data Wizard? Furthermore, this phase also highlights that there were different expectations and understanding, or lack thereof, on what a Data Wizard is in addition to its disputable features. The discussion with the developers did not necessarily clarify what a Data Wizard is. If anything, the varying opinion shows that there was a discrepancy in their interpretations regarding the Data Wizard and nothing more than the fact that it should allow the data produced in the project to be publicly available – meanwhile, the result from the questionnaire also underlines that it is not merely made for the general public. In addition, as earlier noted during the analysis of the questionnaire, the respondents were only able to express their thoughts as to which kind of information they would like to see in the Data Wizard but without any specificity on which activity from the project it should come from. These findings were taken into consideration in the following stages of this study.

4.2 Treatment design

In this chapter, the result from the treatment design stage is elaborated. This stage of the design cycle aims to elicit the learning goals and requirements from the stakeholders/end-users which was derived from the focus group discussion.

4.2.1 Focus group discussion with the stakeholders

A focus group discussion was organized with 10 stakeholders of the Data Wizard to identify their common ambition for the Data Wizard. The discussion was divided into two main parts. The first part is concerned with the stakeholders' interests in the new knowledge from the project, whereas the second part refers to deriving requirements for the Data Wizard. Snapshots of the Mentimeter slides and responses are included in Appendix B.

Analysis of the stakeholders' interests in the Polder2C's project

Most participants of the focus group turned out to be coming off the academics/university (5 people, 50%) discipline, followed by engineering company (3 people, 30%) administrative and emergency services (with 1 person each, 10%). This distribution of participants is less evenly spread compared to the responses gathered from the questionnaire; in which 36%, 21% and 25% of the responses were from the university/academic, engineering company, administration disciplines respectively.

On the other hand, participants' interests in the activities were quite evenly distributed. Activities from WP1 represented 46% of the vote, with 25% coming from the overflow tests and 21% from survey/digital twins. Its counterpart, WP2, represented 47% of the votes, with activities such as levee repair (18%), levee inspection (18%) and BreachDefender (11%). The outcome of this question also highlights that the attendees were well invested in nearly all of the activities as shown by Figure 15, where each activity, except for the overflow test and Breachdefender, caught the interest of all disciplines. Furthermore, participants who selected *'other'* further specified that they were interested in the wave overtopping, breach growth experiment (WP1) as well as BIM (Building Information Modelling) and the Data Wizard (WP3) itself.

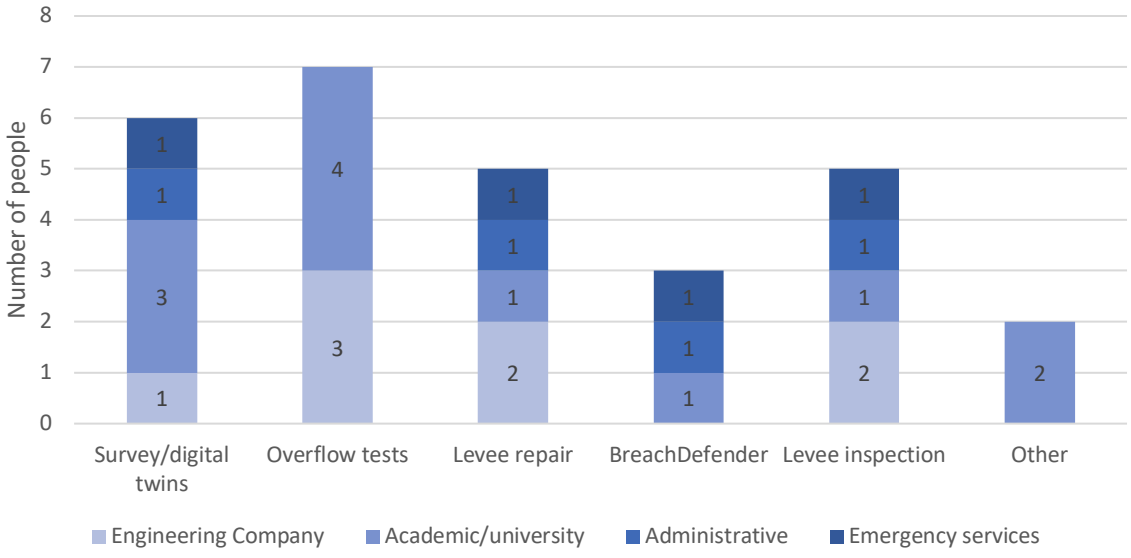

Figure 15: Focus group participants' interest in the Polder2C's activities

When asked specifically for the new knowledge that they would like to learn from those activities, some of the participants were quite like-minded. The following interests were identified:

1. Three attendees were particularly keen on erosion-related measurements, whether that is regarding the erosion process or the resistance of the levee itself towards erosion
2. Although the breach growth experiment is yet to be performed, two participants noted to not overlook the effect of foreshore vegetation on breach growth
3. A few also stressed their interest in failure processes as well as the impact of animal activity on levee strength
4. Participants who were interested in activities from WP2 addressed that they were intrigued with the efficiency/effectiveness of levee repair and flood response measures

One participant from the administrative discipline also gave a remark that the information and knowledge obtained from the project will be further diagnosed and processed into risk-based information to evaluate whether or not there is actually a risk present to the levees – and if so, how the levee design, management practices and emergency responses can be further improved. In the subsequent question, regarding how the new knowledge could benefit the represented organization, the participant believed that a better understanding of the embankments' behaviour, or the so-called "story of the dike" would allow responding to related problems more effectively. "Story of the dike" – a concept said to be introduced by the US Army Corps during the project – refers to understanding the actual historic behaviour of the dike instead of relying only on dike models/simulation for assessments. This is also valuable to validate if the models do reflect the real behaviour of the dike.

Other participants had an outlook to utilize the new knowledge to improve current practices. This includes improving models for breaches, diagnosis/levee strength analysis as well as enhancing the design of dikes and their repair to also face climate change and sea-level rise. One participant from the academics/university discipline, consistently also suggested that the new knowledge would be put to use in courses or curriculum.

Analysis of stakeholders' requirements for the Data Wizard

Before the presentation about the Data Wizard, questions were first asked to the attendees concerning their awareness and what they think the Data Wizard is through a word cloud question. Regarding the former, it is no surprise that all of the stakeholders were aware of the Data Wizard. Although I did not inquire the extent of their awareness, the result of the latter gave an indication of how they were aware of it – which seemed to aggravate throughout the discussion. Various interpretations of the Data Wizard came up. It seems as though the Data Wizard is primarily associated with the word database or a platform for data collection. A few of the participants also suggested the idea of having a user-friendly interface and being easily/publicly accessible. On the other hand, a handful of attendees specified that it is mainly a tool for levee designers and managers along with functions such as AI (artificial intelligence) and machine learning for data analysis.

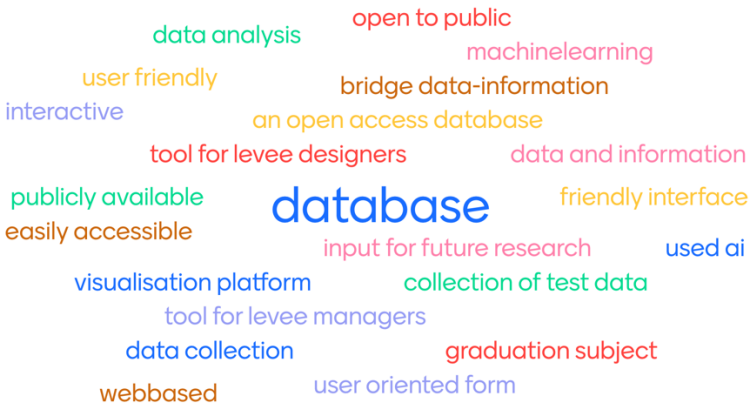

Figure 16: Screenshot of the word cloud regarding what the stakeholders think the Data Wizard is

During the presentation, the proposed architecture of the Data Wizard was shown to give an impression or reference point to the participants regarding what can be accessed through the Data Wizard, which was the succeeding question. A number of participants showed that they were mainly interested in the results, measurements, observations, output and reports from all of the experiments and work packages – much like the result from the questionnaire. A handful others were keen on test conditions and how the result can be integrated into models such as GIS and AI.

While this research was not able to inquire further about the specific activities they referred to, plenty of remarks pertaining to the architecture was given – triggering a much-needed conversation regarding the definition of the Data Wizard. A participant raised the fact that, from the proposed architecture, the Data Wizard was oriented only for the living lab project, with hardly any room for possible expansion. For example, to accommodate for information of other levee sections/systems of the four participating countries in the project. To the participant's view, the way it is now suggested that it is more of a database, comprising of collected data from the project. Although it is not necessarily a negative direction to be headed – as another attendee also said – the participant noted that having an expandable architecture will add value to the Data Wizard but particularly for knowledge development and transfer. The participant explained that the results from the experiments were situated in a specific context, on a particular levee, with certain conditions and assumptions which cannot be generalized to other levees, let alone other countries. It would be more useful instead for the Data Wizard to also include information regarding how existing levees were designed, operated and maintained in the involved countries such that new knowledge can then be developed.

As the discussion progressed, one participant affiliated with the administrative discipline also raised the question of what a Data Wizard is. The participant associated the word 'Wizard' with 'sorcery' or 'wizardry', envisioning that the Data Wizard would be able to transform the data or give an output that is different to its initial output; a contrasting view to the participants' common view for Data Wizard as a database, which typically only collects and disseminates data. While there is not one necessarily correct definition of the Data Wizard, up to this point of the discussion, it is apparent that there was a clear need for clarification on the definition and or boundaries of the Data Wizard.

Regarding the importance of the proposed Data Wizard features, classifications of data/information and friendly user interface were seen as the two most important features to be included in the Data Wizard, with both scorings above four. Furthermore, the participants seemed to have concurred that the data analysis features such as 3D visualisations and data visualisations are not intended for the Data Wizard.

One participant associated with the academic discipline justified his reason to downvote those features, saying that the Data Wizard "will not be able to replace the specialists who will work along with these data". Instead, representation of data in an understandable form through scientific publications or reports by the experts should be the aim – and others can try to reproduce this data. This view was also supported by another person with a research background, suggesting to also include reports and publications associated with the data. Within the same argument, the participant also suggested clarifying and making a distinction whether the Data Wizard is more of a knowledge-base or database. If it was the former, the participant insinuated to include a wiki-type summary that includes the aforementioned reports and publications, comprising visualisations of data instead of the raw data. Following this view, another participant also added that if the wiki were to be added, it should contain information about the knowledge that was previously known and what is learned after. In addition, alternatively, classification of data can be based on topics (of the reports) instead of by experiments – and connect it to the respective data sets.

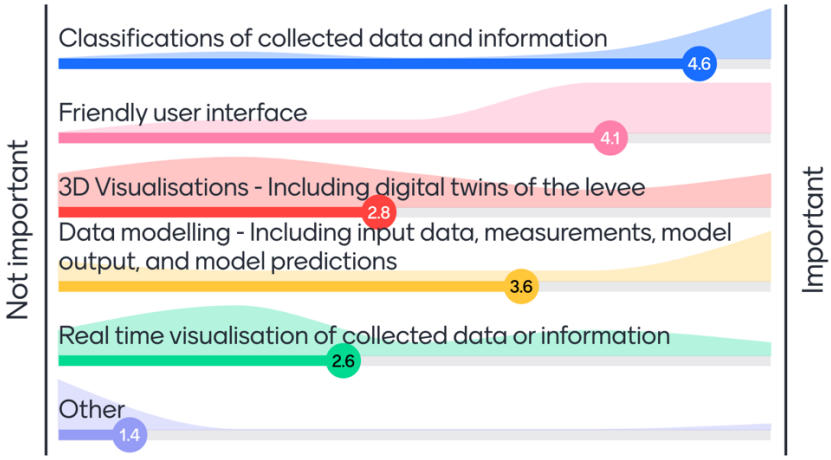

Figure 17: The importance of proposed Data Wizard features

When asked about how they imagine the Data Wizard to complement or restrict their work, the majority of the responses were mainly about the benefits instead of the restrictions. Yet, restrictions were addressed in the subsequent questions to be discussed. The participants, again, associated the Data Wizard as a facilitator that lets its users get access to data from the project. Thus, many of the responses were oriented towards features such as easy access to data or datasets along with filters to sort the data. Rather than just the data itself, a participant from the academic background further added that it should be accompanied with the context, such as how it was measured, an overview of the experiment and test conditions just to name a few. A participant of administrative background suggested that the Data Wizard – being a platform that is accessible by everyone including the project partners – could help to validate whether or not the knowledge or assumptions made by individual experts are correct.

Subsequently, participants were asked about their experience in gathering data similar to the one produced in the Polder2C's project. While it comes as no surprise that the participants have had previous experiences with other sources of data, a handful of the participants laid out that numerous data sources were needed due to how information was scattered throughout. Knowledge databases, academic papers, in-house databases, levee safety handbooks and geotechnical databases such as the Dinoloket are some of the sources that they mentioned.

During the discussion, the idea of having a more expandable architecture was conceivably made clearer by the same participant, who in this question suggested that there is no single point of reference for a platform that stores data or information similar to the one produced by the Polder2C's project. The participant elaborated that there is not one unified information platform about a specific topic – but are rather project-related, similar to the Data Wizard – therefore the need for multiple information sources. Admittedly, such project-related platforms were not that useful for the participant, who continued to express the desire to understand about the levees' prior design, the load it was designed for and how it was maintained – much like the previous comment on the Data Wizard's architecture – which cannot all be found in one place. Following these two arguments about having an expandable architecture, this research inferred that it was meant to introduce other functionalities and contents to the Data Wizard which would streamline the need for other data sources and improve its usefulness. This after all is also experienced by the other participants.

Furthermore, when asked to reflect on their experience with platforms they previously used and what can be implemented or avoided by the Data Wizard, various but similar answers to the preceding questions arise. Overlapping answers were also observed in the subsequent question, regarding how the Data Wizard should be used after the project ends. A small number of people stated that the Data Wizard should be easy to use, whether that is through a user-friendly interface or attributes that would alleviate user experience in going through the data and information. For instance, filters on data types, categorization of datasets and a detailed description supporting the data (e.g. short explanations on how it was collected, test conditions and post-processing that has been done to the data). A participant affiliated with the academic background also proposed to have a summary of all information that is available per levee section. This includes having a description of the levee condition, pre-and post-experiment as well as what happened in between; since there might be anomalies or damages purposefully introduced for the sake of the experiment – which may affect the interpretation of the respective data.

An emphasis was also given concerning the longevity of the Data Wizard. In one of the previous questions, one participant mentioned that data is often lost or inaccessible after the project ends. Therefore, a few of the participants suggested that the Data Wizard should be easy to manage and with a designated group of people to regularly update and maintain it even after the project ends; thus, prolonging the lifespan of the website. Although not a specific timeframe was mentioned, the participants expressed their desire to have the website to last even after the project ends.

In the following question, participants were asked to think of other relevant Data Wizard users. Similar to the output of the questionnaire, students and researchers came up along with professionals such as water/dike engineers and levee managers. One participant also suggested interest groups such as environmentalists. Knowing that they may or may not have any stake or interest at all in the Polder2C's project, the participants were also asked about what would persuade them to use or not use the Data Wizard. Consistent with the previous findings, a few of the participants suggested that people would likely be discouraged if the Data Wizard were to be difficult to use and only contain results from the project. On the other hand, good visualisations, user-friendly navigation, sound categorization of data and good publicity would promote its use. By publicity, the participant referred to the Data Wizard being used for scientific publications and conference presentations.

Figure 18: Word cloud – participants outlining other potential Data Wizard users

The discussion was concluded with a question as to whether or not a similar workshop should be organized with potential end-user groups. At this point of the discussion, half of the participants have left the discussion and therefore not as many inputs were obtained. The majority of those who were present, however, stated that they were unsure, supported with a similar rationale. The participants mentioned that the workshop is likely to be more useful when a prototype of the Data Wizard, including data from one or two experiments from the project, can be showed to the partners or other user groups such that a more specific and tangible input can be derived. For instance, to judge if the proposed design or interface works well for them and if it suits their desires. Hence, following this input and considering the time constraint of this study, no follow up workshops were planned.

In the next section, the result and key findings from the problem investigation and treatment design phase are synthesised into a design brief.

4.2.2 Design brief

In this section, the key findings gathered from this research are synthesized to form a preliminary design brief – and starting point for further recommendations.

Why – Data Wizard's goal

At the moment of writing, as stated on the Polder2C's website, the Data Wizard has two different aims. The first one is to "disseminate information to the wider public" and the second is "an online tool that makes all data accessible and easy to use for stakeholders such as research professionals, levee managers and students" (Polder2C's, 2020) – which is also the only two persistent visions found in this study; as said by the two developers in the problem investigation phase. These two statements, however, lack the emphasis on how it will manage to do so and the kinds of data to be shared – thus leaving some room for interpretations and different expectations to arise.

There were various interpretations regarding the Data Wizard. The word cloud from the focus group discussion shows that, at the general level, the Data Wizard is associated primarily with the word database or a platform for data collection. See Figure 16. This view, however, was challenged when considering the direction to be headed. A handful of stakeholders suggested making a distinction (or a choice) between a database and knowledge base, for which the latter is more appropriate for the collection of documentations instead of data. The suggestion for knowledge-based as an alternative arose because of its focus on reports and scientific publications with data from the project, instead of the data itself. This conforms with findings from the questionnaire, in which respondents valued new knowledge, results and major observations more than experimental data. However, data nonetheless remained important as stakeholders expressed their desires to have filters on data types and good categorization of datasets.

Furthermore, it was unclear if the Data Wizard is only confined to the information and data that are produced on-site only. Therefore, the idea to have an expandable architecture to accommodate for levee data from outside of the project, at least from the four participating member states, was introduced.

Who – Data Wizard's users

The findings from this study point at administrative, university/academics and engineering companies (according to the questionnaire) in addition to researchers, students, dike engineers and levee managers (according to the focus group) as the future users of the Data Wizard – and only once was the "wider public" (or more specifically "everyone") ever mentioned. Moreover, stakeholders or partners of the Polder2C's project also represents either one of the aforementioned disciplines which therefore gave this research little to no reason to believe that there could be other user groups or the conviction that the Data Wizard is meant for the general public. Although unlikely, it is possible that there was a misconception between the Data Wizard being publicly available and made for the public – for which the former means that it allows everyone access upon request but not necessarily for the public.

What – Initial requirements

This study aims to, at the very least, identify a preliminary set of expected added value, interests and requirements of the stakeholders/users of the Data Wizard to be further used in its design process – which may change when new information/requirement arises as it continues to develop. The focus group showed that the overflow tests, survey, levee inspection and repair attracted the most interest – complementing the result from the questionnaire in which the activities were not defined. Additionally, an initial set of added values and requirements originated from the focus group discussion are summarized in Table 5 and Table 6 respectively. In Table 6, requirements are classified into three categories for better clarity and according to the last two categories of requirements by Van Velsen et al. (2013) as earlier elaborated in the methodology.

Table 5: A set of initial learning goals and expected added values from it

#	Identified interests in lessons learnt or takeaway messages when describing the data available in the Data Wizard	Expected added values from the new knowledge
1	A better understanding of erosion processes and the resistance of levees towards it	Knowledge from the project will be further diagnosed and processed into risk-based information to evaluate whether or not there is a risk present to the levees
2	A better understanding of foreshore vegetation on breach growth	Better responses to levee-related problems as a result of understanding the "story of the dike" – an understanding of the levee's actual historic behaviour instead of relying only on models/simulations for assessments To validate if the models do reflect the real behaviour of the dike
3	A better understanding of the effects of animal activity on levee strength	To design safer dikes and dams
4	A better understanding of failure processes	To improve current practices, such as: a. Models for breaches b. Models for levee diagnosis and strength analysis c. Design and repair of dikes
5	A part may be converted into courses	Knowledge and data to be put to use in courses or curriculum
6	Getting action perspectives for the future of levee management	To get a better understanding of the levee strengths and gain more insights into action perspectives
7	To get a better understanding of how to monitor water structures with more advanced technology	Combining new knowledge to monitor water defence structures
8	Cost of interventions on the levee	
9	Remote sensing and machine learning	
10	A better understanding of the effectiveness of levee repairs and emergency measures	

Table 6: A set of initial design requirements

#	Content requirements
1	The Data Wizard should include the results, measurements, observations and output/reports from all of the work packages: physical data strength dike element, monitoring results of dike repair activities, load and resistance parameters of levees, information about flood defence structure (with GIS and AI)
2	The Data Wizard should have an expandable architecture to include information/data of levees from outside of the project (at least from the four participating member states)
3	The Data Wizard should include information regarding how existing levees were designed, operated and maintained in the involved countries
4	Specify previously known knowledge and what is learnt after – possibly with a wiki-type summary
5	Include an overview of the experiments such as how the data was measured or the methods and the test conditions alongside the data
6	Include a summary of the levee condition (per levee section), pre-and post-experiments and what happened in between (e.g. purposefully damaged) to help the interpretation of the data
7	Classify data based on topics (of the reports) instead of by experiments – therefore, the reports are connected to their respective data sets
8	Description of any post-processing that has been done or not done to the data
9	The Data Wizard should not contain information from the Polder2C's project only
10	New knowledge should be disseminated via 'standards'
	User interaction requirements
11	User-friendly interface/navigation
12	The Data Wizard should be easily and publicly accessible
13	Include features to sort and filter the data
14	Facilitate possible integration onto models such as GIS, AI and machine learning for data analysis
15	Include good visualisations
	Other
16	The Data Wizard should remain accessible after the project ends
17	The Data Wizard should be easy to manage and maintain

On top of that, the stakeholders also gave their remark on the proposed features to be incorporated in the Data Wizard; where data-processing features such as 3D visualisations and data modelling are of secondary importance to user-friendly interface and classifications of data. See Figure 17. Alternatively, visualisations of data can be in the form of scientific publications and reports instead – which again relates to the heavier focus towards knowledge instead of data.

5. Discussion

Following the key findings from the problem investigation and design treatment as summarized in the design brief, it is recommended that the design team and stakeholders of the Data Wizard focus their efforts on four recommendations:

1) Adopt a clear definition for the Data Wizard
2) Focus the Data Wizard on areas of greatest impact
3) Gather more perspectives, preferences and desires for the Data Wizard
4) Create a collaborative environment for the partners

These recommendations are directed towards the stakeholders, including the developers of the Data Wizard. The only caveat is that it may not be accurate/suitable to their roles in the project but rather to show that someone needs to take the mediating role. The recommendations presented here have incorporated the feedback from the stakeholders. In the pages that follow, each recommendation is discussed.

Recommendation 1 – Adopt a clear definition for the Data Wizard

In response to the various interpretations of the Data Wizard by the stakeholders and the developers themselves, this research believes that it is an important first step to have clarity on why the Data Wizard was initiated in the first place, its definition and what it should do before proceeding with other suggested recommendations. Having a clear statement about the Data Wizard would help to align the priorities and expectations of the developers and stakeholders going forward. The following are the recommended actions:

1.1 **Construct a definition encompassing the scope and boundaries of the Data Wizard:** In response to the various interpretations (database vs knowledge-base) and unclear boundaries (e.g. expandable architecture) of the Data Wizard, it is recommended to have a clearer statement regarding what a Data Wizard is to better communicate intends and align the expectations of the stakeholders and developers. This would also help to judge the relevance of the requirements going forward. For instance, the requirement to have information from outside of the Polder2C's project (requirement 9) may be irrelevant when the scope/boundaries for the Data Wizard is strictly set for data/knowledge from the project only. In addition, boundaries could also be in the form of money, time and skill. This is particularly important to be considered when developing certain functionalities or even an expandable architecture which could be costly and challenging.

1.2 **Direct the focus of the Data Wizard towards user groups such as administrative, academics/university and engineering companies instead of the general public:** This action is supported by the result from this study, where the general public was barely mentioned and found. Instead, the stakeholders referred to researchers, students, dike engineers and levee managers as the more specific users belonging to one of the three user groups. Emphasising this focus narrows down the scope and specifies the user groups to be considered.

1.3 **Determine the lifespan of the Data Wizard.** This action correlates with the stakeholders' interests to keep the Data Wizard even after the project ends – which signals the need for a financer and perhaps other stakeholders to maintain and update it. From the feedback, one

stakeholder suggested developing the Data Wizard with an architecture that is easily integrated or merged with an existing "Data Wizard" so that it could be managed by an external after the project ends. This recommendation therefore also alludes to exploring other systems similar to the Data Wizard that the stakeholders may be aware of or have experience with and also to consider if developing the Data Wizard is necessary in the first place if in the foreseeable future the Data Wizard will be merged with existing databases/knowledge-bases.

Recommendation 2 – Focus the Data Wizard on areas of greatest impact

In other words, this recommendation alludes to making the most of what the project has; whether that is in expertise or knowledge and data obtained. Table 5 summarizes the interests and added values that the stakeholders expect to have from the Polder2C's project and ideally through the Data Wizard. On the one hand, findings from this research showed that some of the interests of the stakeholders are not explicitly outlined in the currently proposed architecture of the Data Wizard. On the other hand, those that are proposed in the architecture such as the experiments' objectives, methods, results and site descriptions also greatly attract the interests of the stakeholders as Table 6 shows. The recommended actions are as follows:

2.1 **Data collectors or people that are working on the experiments should focus future efforts on developing high-quality descriptions regarding the data and experiments:** In response to the high interest regarding the site descriptions, methods and results of the experiments, it is important to ensure that the findings and descriptions of the experiment are preserved and communicated well to leverage this information and knowledge. Particularly, for experiments such as overflow tests, surveys, levee inspection and repair.

According to Shen (2018), the usability and trustworthiness of data are influenced by the provenance of the content, transparency and the quality of information. Researchers or data users benefit from having a well-documented relationship between the data and its context, which may include the specific steps taken, uncertainty estimations, test conditions, assumptions and calculations to arrive at the result – that are often missing or insufficient. This was also expressed by the stakeholders during the focus group. From the feedback, one stakeholder also voiced the concern that the data he had to deal with from the project are still missing detailed descriptions about the test site, experimentation goals and so forth – which is not only important for other people's understanding (particularly of those outside the project or indirectly involved) but also for the people who process the results from and conducted the experiments themselves. This aligns with one of the challenges found by Shen (2018), who consequently recommends developing a data and metadata management workflow or standards to preserve the findings and prevent data loss – similar to one of the inputs of the stakeholders in the focus group discussion to have 'standards' for knowledge dissemination.

Data loss, regardless of its form, comes with a cost. Samples of data and knowledge are obtained at a great price (funding, time and difficulty) and some, such as time-dependent events, are even considered irreplaceable/irreproducible (McNutt, 2016). Data loss also alludes to lost opportunities. On the one hand, this could refer to missing out on having

better estimates of attributes through better models, as a result of richer data sets. On the other hand, this could be translated into foregoing more cost-effective agendas/measures towards enhancing the 2 Seas countries in facing the adverse effects of climate change – as Shen (2018) suggested that agendas are often times influenced by data and information possessed at hand. Therefore, preventing data loss and being able to manage data well with its metadata is crucial. On a side note, while this cannot be directly compared to one of the benefits of KMS, namely revenue growth (Carillo & Chinowsky, 2006), having a more cost-effective agenda and worthwhile investments as a result of richer and complete data arguably achieves a similar effect.

2.2 Explore other features and needs to be incorporated in the existing architecture of the Data Wizard: This action aims to not only address the remaining needs/learning goals of the stakeholders that are not included in the proposed architecture but also feature or requirements that would improve the usability and usefulness of the Data Wizard, such as the wiki-type summaries and data filters. Table 5 and Table 6 can be used as a starting point.

Admittedly, this recommendation only reflects the desires of 10 people. However, when implemented in harmony with recommendation 3 (and 1), a complete set of requirements and learning goals/added value can be reflected by the Data Wizard and long-term utilization can be expected.

Recommendation 3 – Gather more perspectives, preferences and desires for the Data Wizard

As mentioned in the previous recommendation, only a fraction of the Data Wizard users' desires are reflected in this study – thus only enough for a starting point. Understandably, the Data Wizard is still under development and more feedback can more likely be derived when its prototype is ready as suggested by the stakeholders during the discussion. This research recommends the following actions:

3.1 Perform a similar workshop to gather more information regarding users' needs and desires: Particularly of those outside of the project such as the aforementioned levee managers and professionals from other organizations which may benefit from the Data Wizard to ensure its utility in the long term. Beforehand, it is also possible to perform the workshop using the prototype with the stakeholders and other people from the project, after implementing said requirements.

3.2 Create user profiles and use-case scenarios for the Data Wizard: This action is ideally performed after recommendation 3.1, once enough information has been gathered but not too close before the Data Wizard is officially launched – otherwise, gathered inputs cannot be implemented well. With the gathered information, user profiles and use-case scenarios that reflect the fashion at which people intend to use the Data Wizard can be created – which on numerous occasions have been proved effective in designing various online tools including the RiverCare web-based knowledge platform (Cortes Arevalo et al., 2019; Van de Bildt et al., 2018) and myriads of eHealth technologies (Van Velsen et al., 2012). This study,

as summarized in the design brief has identified the main user groups to be considered for further research and development of the Data Wizard.

Recommendation 4 – Create a collaborative environment for the partners

All in all, this research believes that the challenges in developing the Data Wizard can only be overcome with collaboration across the project members, developers and stakeholders. The discussions performed highlighted that there was a dissonance between the developers and stakeholders which may prohibit the development of the Data Wizard – but can certainly be resolved with collaboration as this research shows. Project partners have demonstrated their willingness to share their opinion and potentially their resources, which this research can only imagine being crucial in the long run. There is a growing responsibility and need for the sharing of knowledge, including new knowledge from the project, prior knowledge and how experiments were carried out that can only be matched with a collective effort.

As shown by the theory of knowledge management (KM), culture plays a significant role in knowledge creation and sharing (Armbrecht et al., 2001; Loon, 2019; Alavi & Leidner, 2001). Loon (2019) postulates that the learning and knowledge creation culture is reinforced by two other mechanisms which constitute a KM; namely, knowledge architecture (such as technology and behaviours that facilitate knowledge storing and sharing) and a 'business model' for knowledge capitalisation. These three mechanisms form a cycle that reinforces the culture. For instance, the project supports the KM culture to experiment and share knowledge (e.g. overflow experiment), which in return yields new knowledge that is valuable for the partners/organizations (e.g. better understanding of failure processes to improve models for levee diagnosis) and hence reinforces the KM culture as a result of its success – in other words, people/project members see the KM culture as beneficial. Thus, ingraining the importance and purpose of the KM/Data Wizard initiative to every project member; which shapes the culture.

The aforementioned example presumes that improvement in models for levee diagnosis is the capitalized knowledge from a better understanding of failure processes. Meanwhile, it may be that others are more inclined to use said knowledge (failure processes) to improve the design of dikes. This shows that knowledge should be adaptable for use in other ways for various goals – as summarized in Table 5. As such, stages of knowledge creation and transfer of the knowledge processes by knowledge management can be considered. Assuming that the focus of the Data Wizard is towards explicit knowledge/data as the end result, stages of knowledge conversion such as *externalisation* and *combination* may be relevant. A simplified example of this conversion is as follows (Niedderer & Imani, 2009):

- Research context and goal: Often times research starts with the *internalisation* phase (explicit – tacit) such as data analysis, literature review, archival searches, experiences from practice and observations of the test site for comprehension. The focus then shifts when the acquired tacit knowledge is made available explicitly (externalisation) through descriptions and documentations for sharing and evaluation.
- Research methods: The explicit knowledge is then utilised or transformed into proposals and methodologies for a specific study or experiment (combination, explicit – explicit). *Internalisation* (explicit – tacit) then occurs when researchers internalise and use the

methods (explicit knowledge) in practice. As the chosen methods are applied, new knowledge/data is produced. In addition, *externalisation* (tacit – explicit) may happen to evaluate the methods conducted (e.g. If any adjustments had to be made).

- Research outcome: The explicit dimension of the (tacit) new knowledge is analysed and *externalised* (tacit – explicit) to be evaluated and shared. Through *combination*, the explicit knowledge is converted into a more complex explicit knowledge and means such as scientific papers and databases are used to disseminate and communicate the knowledge. Lastly, this also marks the start for the new knowledge to be applied in practice or for further research.

The example above is primarily dictated with explicit knowledge as the result of every research step because this is what the Data Wizard seems to gravitate towards. Meanwhile, the tacit dimension in each of the steps is also significant. In the research methods, it is possible that the *internalisation* of tacit knowledge may require modes such as 'learning by doing', expert training and coaching (Nonaka, 2000; Ball et al., 2004) instead of just from a paper before applying the research methods. As for the research outcome, it is equally important to transfer the tacit dimension of the new knowledge (socialization, tacit – tacit) to reach a deeper level of understanding and for further application of the new knowledge. Similar to the research methods, it requires modalities such as training and coaching to be conveyed. Furthermore, since now we know that the Data Wizard will not be a place for collaboration but rather to facilitate access to data and knowledge, it is important to consider the aforementioned modalities to take place within the project as certainly not all learning goals, as summarized in Table 5, can be achieved with reports and the Data Wizard.

5.1 Limitations

This study was posed with limitations such as having to work in an online environment and working with a limited definition of the Data Wizard. While the former is a situational problem that everyone has to adapt to, ineffective communications via email, delays from waiting for replies and scheduling meetings certainly has their repercussions to this study, in addition to the researcher's inability to perhaps ask the right questions to the right people – which relates to the latter. From the making of the proposal, even when requested, there were barely any supporting materials such as project documents that elaborate on the Data Wizard, the connection between the work packages, the people and how they relate with the Data Wizard which makes the individual meetings that much more important – when again it was difficult to know who to talk to in the first place. Furthermore, the unclear definition of the Data Wizard also made it difficult to consult particular literature, but above all to set up the focus group such that it would yield meaningful results. In other words, it was difficult to derive learning goals and requirements for the Data Wizard when the researcher cannot specifically describe what it is nor show it and that the participants have never had experience with it.

Time constraints, as a result of early difficulties in setting this research up to speed, also allowed for only one focus group session to be organized – which was informed to the stakeholders about a week prior. Thus, only a handful of people were available and the time allowed for the session was less than optimal to get the most information. In consequence, not all of the stakeholders/participants had the opportunity to elaborate their points, thus affecting the result of this research. Furthermore, the result from the focus group only entails the perspectives of the stakeholders – of people that are directly involved in the Polder2C's project. Although the preliminary sets of requirements may have provided a good starting point for the development of the Data Wizard, it certainly misses a whole array of other opinions, particularly of those outside or indirectly involved (such as the observing partners) in the project which may solidify or negate the findings of this research. In addition, this research was unable to perform the focus group with each of the user-groups individually, not only due to the limited time but also following the suggestions from the external supervisor and the stakeholders. Therefore, the recommendation to perform a similar workshop at later stages when a prototype of the Data Wizard is ready was made.

In addition, the feedback obtained for the preliminary recommendations from the stakeholders was very limited – only two responses were gathered. Therefore, this research cannot gauge how the stakeholders as a whole perceived the recommendations (e.g. whether they agree or disagree with the recommendations) and whether or not certain issues, according to their point of view, are treated or overlooked. This affects the quality and validity of the recommendations. Regardless, the premise is that the overview and recommendations that this research presents would provide discussion points that would facilitate further dialogues toward defining and developing the Data Wizard – therefore, if new information arises, the recommendations or steps towards developing the Data Wizard can be adjusted accordingly.

6. Conclusion

The goal of this research is to investigate the learning goals and requirements of the stakeholders and users of the Data Wizard to be further used in the design process of the Data Wizard.

In the problem investigation phase, the context surrounding the Data Wizard, its stakeholders and users were identified through a series of discussions with the developers, analysis of project documents and a questionnaire related to the Data Wizard.

Overall, it was found that there was not one precise definition regarding what the Data Wizard is. This has resulted in different perceptions even for the developers themselves – who suggested significantly different users of the Data Wizard. However, findings from this research showed that administrative, university/academics and engineering companies (according to the questionnaire) and more specifically researchers, students, dike engineers and levee managers (according to the focus group) are the future users of the Data Wizard – instead of just the "wider public". From the analysis of the questionnaire, it was also found that new knowledge, major observations/results and cleaned experimental data were most important to be included in the Data Wizard although the project activity/experiment it came from was not specified. Furthermore, stakeholders of the Data Wizard were found by following the information chain of the project, namely data production, data processing, data storage/sharing and data exploitation. These are summarized in Table 3. These findings were taken into consideration in organizing the focus group.

In the treatment design phase, the requirements and learning goals of the stakeholders were identified through a focus group discussion. Admittedly, the disciplines represented by the focus group participants are less evenly distributed compared to the respondents of the questionnaire.

It was evident that the effect of the imprecise definition of the Data Wizard was also carried over by the stakeholders as shown by the different interpretations and expectations – even when all of the stakeholders admitted that they are aware of the Data Wizard. Complementing the result from the questionnaire, it was found that the overflow tests, survey, levee inspection and repair were some of the project activities that attracted the most attention. Understanding that the stakeholders may or may not have any premonition regarding what a Data Wizard is, this research probed into the stakeholders' requirements through questions related to their line of work such as what they would like to get through the Data Wizard and related past experiences with a platform providing similar data/knowledge to the one produced by the Polder2C's project. These are summarized in Table 5 and Table 6. The key findings of this research from the problem investigation and design treatment phase were made into a design brief which acted as a basis for recommendations not only for future research but also what this research thinks to be a necessary step for the Data Wizard going forward.

To conclude, this research presents the stakeholders, user-groups, learning goals and requirements from the stakeholders, for the Data Wizard. While it certainly did not deliver a complete set of requirements, the premise is that these findings would stimulate discussions to not only define and set priorities for the Data Wizard but also to achieve its desired effects.

Bibliography

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211. DOI: https://doi.org/10.1016/0749-5978(91)90020-T.

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly, 25*(1), 107-136. doi:10.2307/3250961.

Alexander, I. F. (2005). A Taxonomy of Stakeholders: Human Roles in System Development. *International Journal of Technology and Human Interaction (IJTHI), 1*(1), 23-59. doi:10.4018/jthi.2005010102.

Anema, K., & Pfeiffer, E. (2017). *Deliverable D1.3: Initial Stakeholder Engagement Strategy.*

Armbrecht, F. R., Chapas, R. B., Chappelow, C. C., Farris, G. F., Friga, P. N., Hartz, C., . . . Whitwell, G. E. (2001). Knowledge management in research and development. *Research-Technology Management, 44*(4), 28-48. DOI: 10.1080/08956308.2001.11671438.

Ball, L. J., Ormerod, T. C., Morley, N. J., Cross, N. (., & Edmonds, E. (. (2004). Spontaneous analogising in engineering design: a comparative analysis of experts and novices. *Design Studies, 25*(5), 495-508. DOI: https://doi.org/10.1016/j.destud.2004.05.004.

Batty, M. (2008). The Size, Scale, and Shape of Cities. *Science, 319*(5864), 769-771. DOI: https://doi.org/10.1126/science.1151419.

Becker, B. E., & Huselid, M. A. (2006). Strategic Human Resources Management: Where Do We Go From Here? *Journal of Management, 32*(6), 898-925. DOI: https://doi.org/10.1177/0149206306293668.

Carrillo, P. M., Robinson, H. S., Al-Ghassani, A. M., & Anumba, C. J. (2004). Knowledge management in UK construction: Strategies, resources and barriers. *Proj. Manage J, 35*(1), 46-56.

Carrillo, P., & Chinowsky, P. (2006). Exploiting Knowledge Management: The Engineering and Construction Perspective. *Journal of Management in Engineering, 22*(1), DOI: 10.1061/(ASCE)0742-597X(2006)22:1(2).

Cortes Arevalo, V. J., den Haan, R.-J., Berends, K., Baart, F., van der Voort, M. C., & Hulscher, S. J. (n.d.). Drivers and barriers for the potential use of online knowledge exchange platforms for transdisciplinary river research. *To be submitted to Environmental Science and Policy*.

Cortes Arevalo, V. J., Verbrugge, L. N., den Haan, R.-J., Baart, F., van der Voort, M. C., & Hulscher, S. J. (2019). Users' Perspectives About the Potential Usefulness of Online Storylines to Communicate River Research to a Multi-disciplinary Audience. *Environmental Communication, 13*(7), 909-925. DOI: 10.1080/17524032.2018.1504098.

Davis, F. (1986). *A technology acceptance model for empirically testing new end-user information systems: theory and results.* Doctoral dissertation, MIT Sloan School of Management, Cambridge.

Davis, F. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340. doi:10.2307/249008.

European Commission. (n.d.). *Interreg : European Territorial Co-operation*. Retrieved March 2021, from Regional Policy - European Commission: https://ec.europa.eu/regional_policy/en/policy/cooperation/european-territorial/

Gupta, A., & Govindarajan, V. (2000). Knowledge Flows within Multinational Corporations. *Strategic Management Journal, 21*, 473-496.

Hegger, D., Driessen, P., Wiering, M., Van Rijswick, H., Kundzewicz, Z., Matczak, P., . . . Priest, S. (2016). Toward more flood resilience: Is a diversification of flood risk management strategies the way forward? *Ecology and Society, 21*(4).

Interreg2Seas. (2015). *Key Info*. Retrieved from https://www.interreg2seas.eu/en/content/about-programme

Interreg2Seas. (2019). Retrieved 2021 March, from Polder2C's: https://www.interreg2seas.eu/en/Polder2C%27s

Kaiser D, B., Köhler, T., & Weith, T. (2016). Knowledge management in sustainability research projects: Concepts, effective models, and examples in a multi-stakeholder environment. *Applied Environmental Education & Communication, 15*(1), 4-17.

Kelleher, D., & Levene, S. (2001). Knowledge management: A guide to good practice. *British Standards Institute*.

Klijn, F., Kreibich, H., De Moel, H., & Penning-Rowsell, E. (2015). Adaptive flood risk management planning based on a comprehensive flood risk conceptualisation. *Mitigation and Adaptation Strategies for Global Change, 20*(6), 845-864.

Loon, M. (2019). Knowledge management practice system: Theorising from an international meta-standard. *Journal of Business Research, 94*, 432-441. DOI: https://doi.org/10.1016/j.jbusres.2017.11.022.

Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society, 14*(1), 81-95. DOI:https://doi.org/10.1007/s10209-014-0348-1.

Martins, V., Rampasso, I., Anholon, R., Quelhas, O., & Filho, W. (2019). Knowledge management in the context of sustainability: Literature review and opportunities for future research. *Journal of Cleaner Production*(229), 489-500.

McGranahan, G., Balk, D., & Anderson, B. (2007). The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization, 19*(1), 17-37. DOI: https://doi.org/10.1177/0956247807076960.

McNutt, M. (2016). #IAmAResearchParasite. *Science, 351*(6277), 1005. DOI: 10.1126/science.aaf4701.

Mehvar, S., Wijnberg, K., Borsje, B., Kerle, N., Schraagen, J. M., Vinke-de Kruijf, J., . . . Hulscher, S. (2021). Review article: Towards resilient vital infrastructure systems – challenges, opportunities, and future research agenda. *Natural Hazards and Earth System Sciences, 21*(5), 1383-1407. DOI: https://doi.org/10.5194/nhess-21-1383-2021.

Mounce, S., Brewster, C., Ashley, R., & Hurley, L. (2010). Knowledge management for more sustainable water systems. *Journal of Information Technology in Construction*, 140-148.

Neumann, B., Athanasios, T., Juliane, Z., & Robert, J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding. *a global assessment: PLoS One, 10*(3), DOI: https://doi.org/10.1371/journal.pone.0118571.

Niedderer, K., & Imani, Y. (2009). Developing a framework for managing tacit knowledge in research using knowledge management models. *Undisciplined! Design Research Society Conference 2008.*

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science, 5*(1), 14-37.

Nonaka, I., & Konno, N. (1998). The Concept of “Ba”: Building a Foundation for Knowledge Creation. *California Management Review, 40*(3), 40-54. DOI: https://doi.org/10.2307/41165942.

Nonaka, I., Toyama, R., & Konno, N. (2000). ECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning, 33*, 5-34.

Parry, M., Canziani, O., Palutikof, J., van der Linden, P., & Hanson, C. (2007). *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge: Cambridge University Press.

Polder2C's. (2020). Retrieved March 2021, from Interreg 2 Seas Mers Zeeeën Polder2C's - European Regional Development Fund: https://polder2cs.eu/

Robertson, M., Swan, J., & Newell, S. (1996). The Role of Networks in the Diffusion of Technological Innovation. *Journal of Management Studies, 33*, 335-361.

Robinson, H. S., Carrillo, P. M., Anumba, C. J., & Al-Ghassani, A. M. (2001). Perceptions and barriers in implementing knowledge management strategies in large construction organisations. *Proc., RICS Foundation Construction and Building Research Conf. (COBRA)* , 451-460.

Secrétariat Technique Conjoint INTERREG IV A 2 Mers. (n.d.). Strategic Environmental Assessment of the INTERREG VA 2 Mers Seas Zeeën 2014-2020 Programme: Environmental report.

Sharp, J. (2007). Development, Extension, and Application: A Review of the Technology Acceptance Model. *Information Systems Educational Journal, 5*(9), 1-11.

Shen, Y. (2018). Data Sustainability and Reuse Pathways of Natural Resources and Environmental Scientists. *New Review of Academic Librarianship, 24*(2), 136-156. DOI: 10.1080/13614533.2018.1424642.

Smuts, H., Loock, M., Van Der Merwe, A., & Kotzé, P. (2009). Framework and methodology for knowledge management system implementation. *Proceedings of the 2009 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists* (pp. 70-79. DOI: https://doi.org/10.1145/1632149.1632160). New York: Association for Computing Machinery.

Swart, R., Biesbroek, R., Binnerup, S., Carter, T., Cowan, C., Henrichs, T., . . . Reese, M. (2009). *Europe adapts to climate change: comparing natural adaptation strategies.* Helsinski: Partnership for European Environmental Research.

Temmerman, S., Meire, P., Bouma, T. J., Herman, P. M., Ysebaert, T., & De Vriend, H. J. (2013). Ecosystem-based coastal defence in the face of global change. *Nature, 504*(7478), 79-83.

Van de Bildt, E., Cortes Arevalo, V. J., den Haan, R., & van Elzakker, C. (2018). How to create user-descriptions and scenarios to design a knowledge-base for RiverCare research? *The future river: NCR DAYS 2018 Proceedings* (pp. 72-74). Netherland Centre for River Studies publication.

Van Velsen, L., Van Gemert-Pijnen, J. E., Nijland, N., Beaujean, D. J., & Steenbergen, J. (2012). Personas: The Linking Pin in Holistic Design for eHealth. *eTELEMED 2012.*

Van Velsen, L., Wentzel, J., & Van Gemert-Pijnen, J. (2013). Designing eHealth that Matters via a Multidisciplinary Requirements Development Approach. *JMIR Research Protocols, 2*(1, e21), https://doi.org/10.2196/resprot.2547.

Venkatesh, V., & Davis, F. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science, 46*(2), 186-204.

Webb, S. (1998). *Knowledge Management: Linchpin of Change (1st ed.).* Routledge.

Wehn, U., & Almomani, A. (2019). Incentives and barriers for participation in community-based environmental monitoring and information systems: A critical analysis and integration of the literature. *Environmental Science & Policy, 101*, 341-357. Doi: https://doi.org/10.1016/j.envsci.2019.09.002.

Wehn, U., & Montalvo, C. (2018). Knowledge transfer dynamics and innovation: Behaviour, interactions and aggregated outcomes. *Journal of Cleaner Production, 171*(Supplement), 56-68.

Wieringa, R. J. (2014). *Design science methodology: For information systems and software engineering.*

Appendix A – Questionnaire Script

Your email (optional)

Your profile

Your involvement in the Polder2C's project:

- ☐ Partner
- ☐ Observer
- ☐ Other:

Your activity sectors:

- ☐ Water industry
- ☐ Construction company
- ☐ Engineering company
- ☐ University/academic
- ☐ Emergency services
- ☐ Administration
- ☐ Other:

Type of your activity sectors:

- ☐ Private sector
- ☐ Public sector
- ☐ Academic sector
- ☐ Other:

List of WP in which you participate (for the partners of Polder2C's):

- ☐ WP1
- ☐ WP2
- ☐ WP3
- ☐ WP4
- ☐ WP5
- ☐ WP6

Your role in the data wizard:

- ☐ Provider of experimental data
- ☐ Provider of modelling data
- ☐ Data management (storage, processing)
- ☐ Data interpretation, lessons, recommendations
- ☐ Future user
- ☐ Other:

The target audience of the Data Wizard

	Very low				Very high
Professionals	1	2	3	4	5
Administration	1	2	3	4	5
Emergency services	1	2	3	4	5
Academics and students	1	2	3	4	5

Do you propose to add other users?

Contents of the Data Wizard

	Very low				Very high
General knowledge about flood protection infrastructures	1	2	3	4	5
Description of the experimental site	1	2	3	4	5
Experimental procedure	1	2	3	4	5
Information about the monitoring system	1	2	3	4	5
Raw experimental data	1	2	3	4	5
Cleaned experimental data	1	2	3	4	5
Selected experimental data (major results and lessons)	1	2	3	4	5
Modelling data (input and results)	1	2	3	4	5
Major observations and results	1	2	3	4	5
New knowledge	1	2	3	4	5

Do you propose to add other content?

Your contribution to the Data Wizard

	Very low				Very high
Provide experimental data	1	2	3	4	5
Provide modelling data	1	2	3	4	5
Data management (cleaning, storage)	1	2	3	4	5
Knowledge production	1	2	3	4	5
Evaluation of the data wizard	1	2	3	4	5

Do you propose to add other contributions?

Table 7: List of the questionnaire's respondents

Disciplines	Organizations/people involved		
	Observer	Partner	Other
Administration	c.dast@symadrem.fr frederic.leseur@developpement-durable.gouv.fr jean.magne@omdm.fr	kristof.verelst@mow.vlaanderen.be bart.vonk@rws.nl	ammar.aljer@univ-lille.fr (researcher)
Engineering Company	Jean-jacques.fry@wanadoo.fr	alleon@isl.fr patrik.peeters@mow.vlaanderen.be	nerincx@dn-t.be wouter.zomer@bzim.nl
University/academics	ala@ou.nl j.j.warmink@utwente.nl	mohammeditair@gmail.com hanbing.bian@univ-lille.fr p.m.j.herman@tudelft.nl s.j.h.rikkert@tudelft.nl cyrille.fauchard@cerema.fr R.C.Lanzafame@tudelft.nl v.tsimopoulou@hz.nl wentholt@stowa.nl	
Water Industry		mkey@southwestwater.co.uk negassidemoz@yahoo.nl	
Other (Helping and advising dike managers)	perrine.broust@france-digues.fr		

Appendix B – Screenshot of the focus group responses

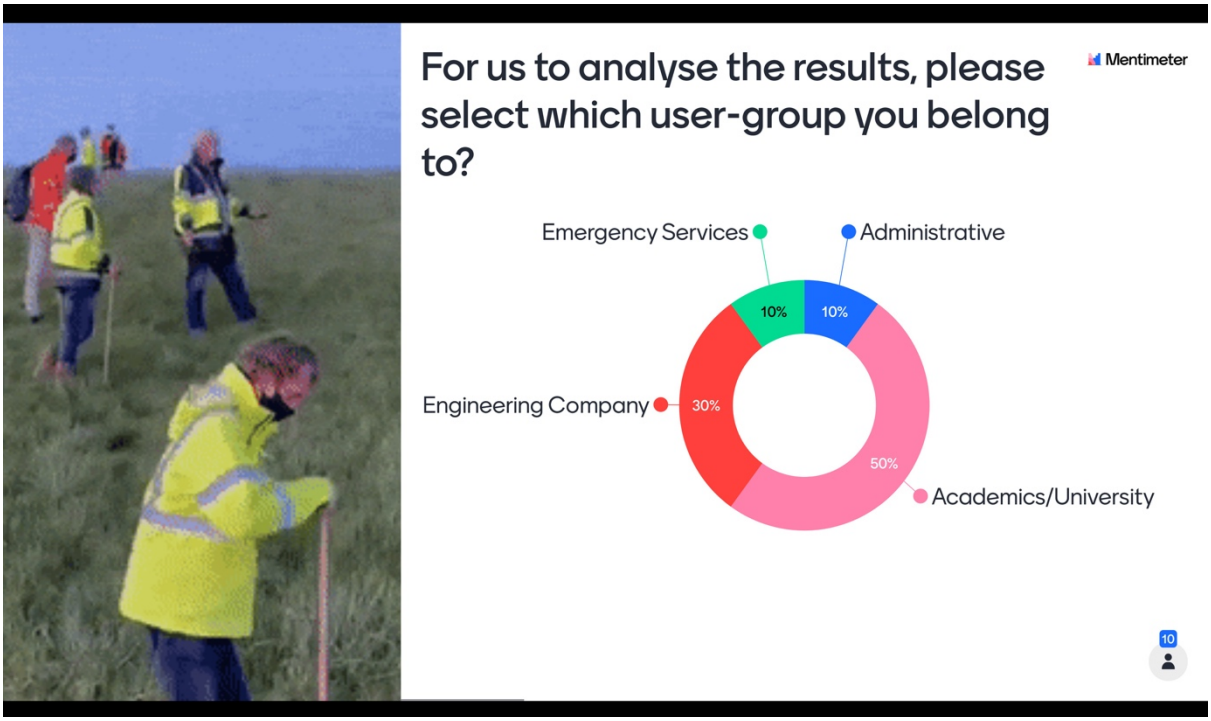

Figure 19: The disciplines of the focus group participants

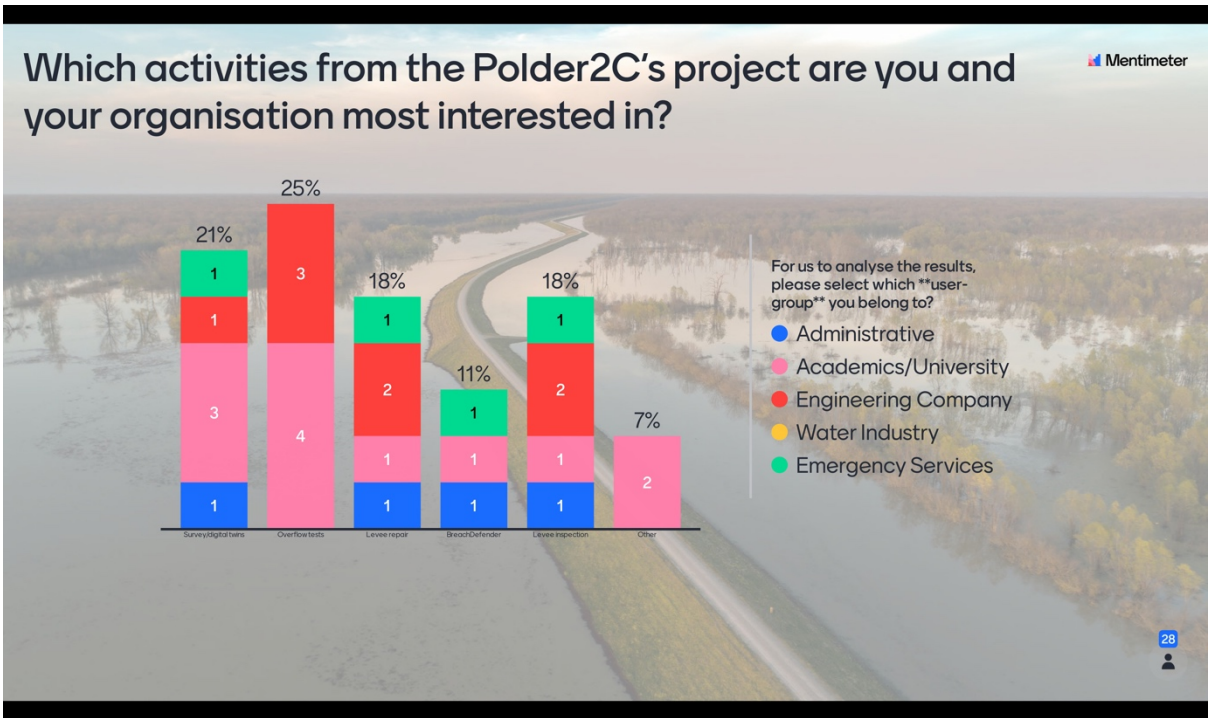

Figure 20: Participants' interest in project activities

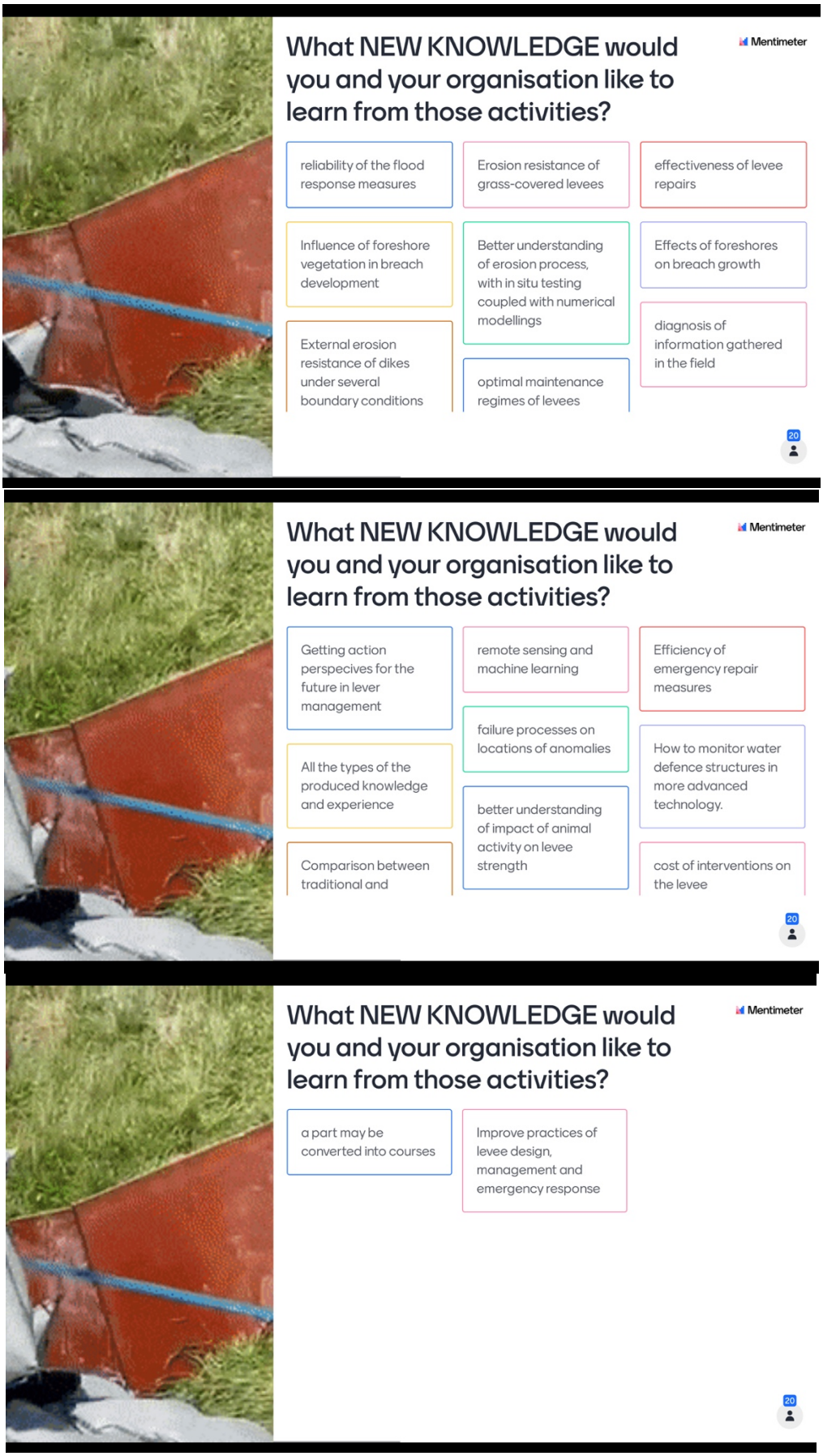

Figure 21: Participants' interest in new knowledge from the project

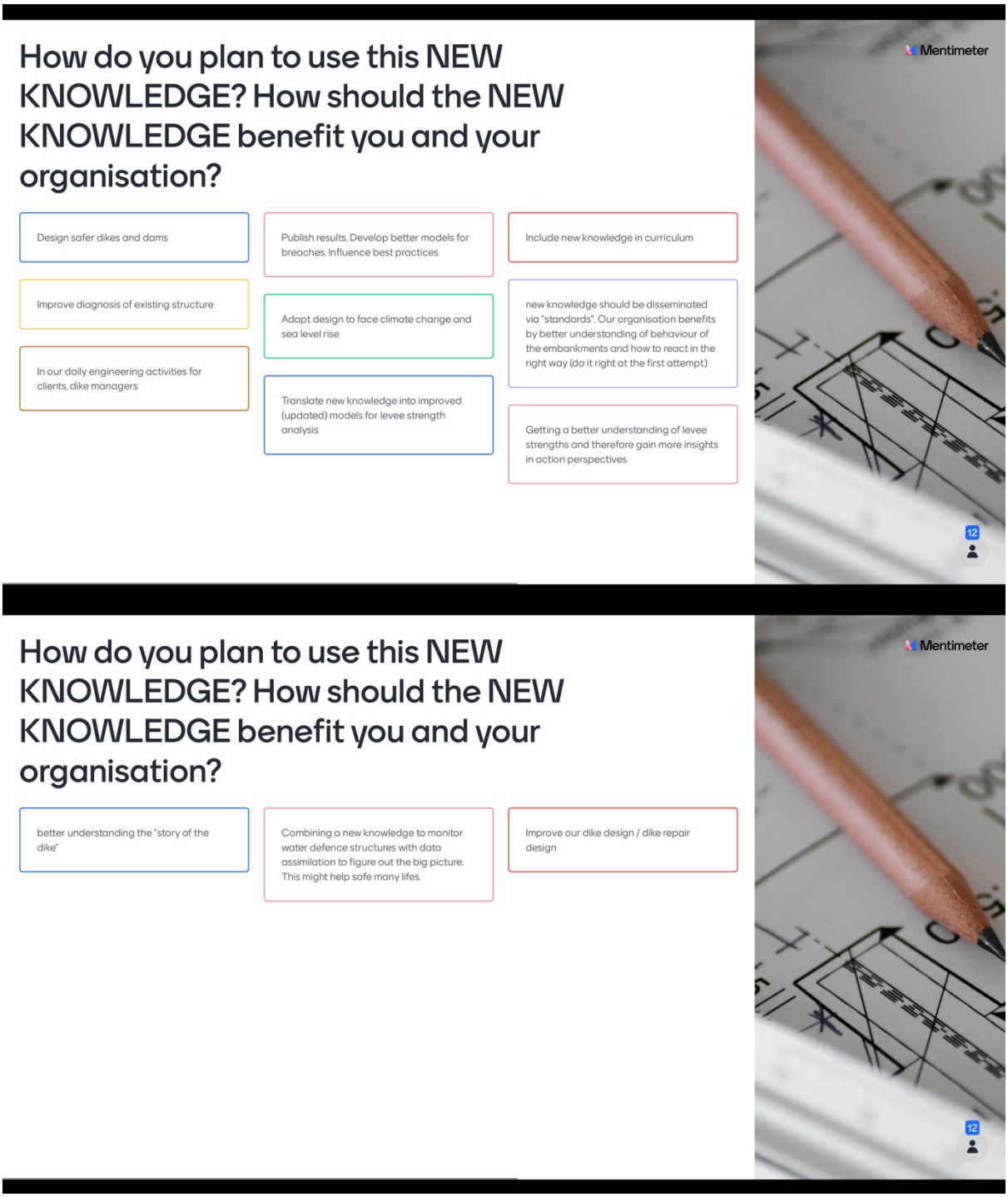

Figure 22: Participants' expected added value from the new knowledge

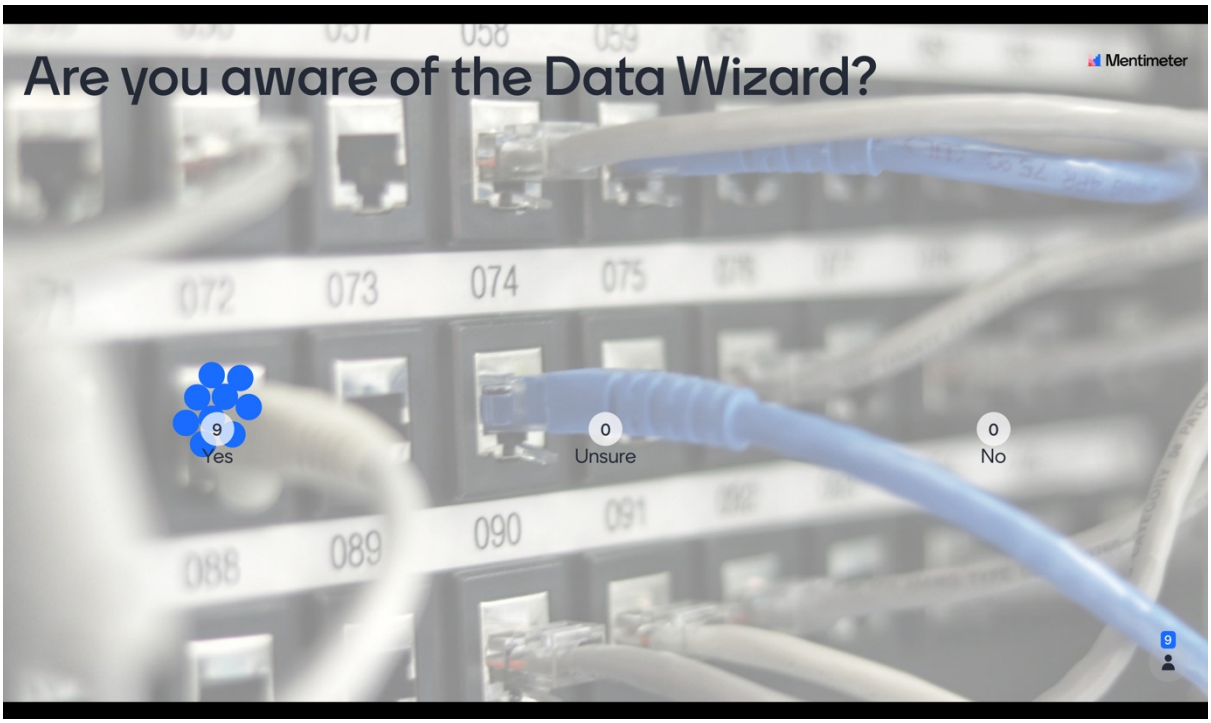

Figure 23: Participants' awareness regarding the Data Wizard

Figure 24: Word cloud - Participants' view on what they think a Data Wizard is

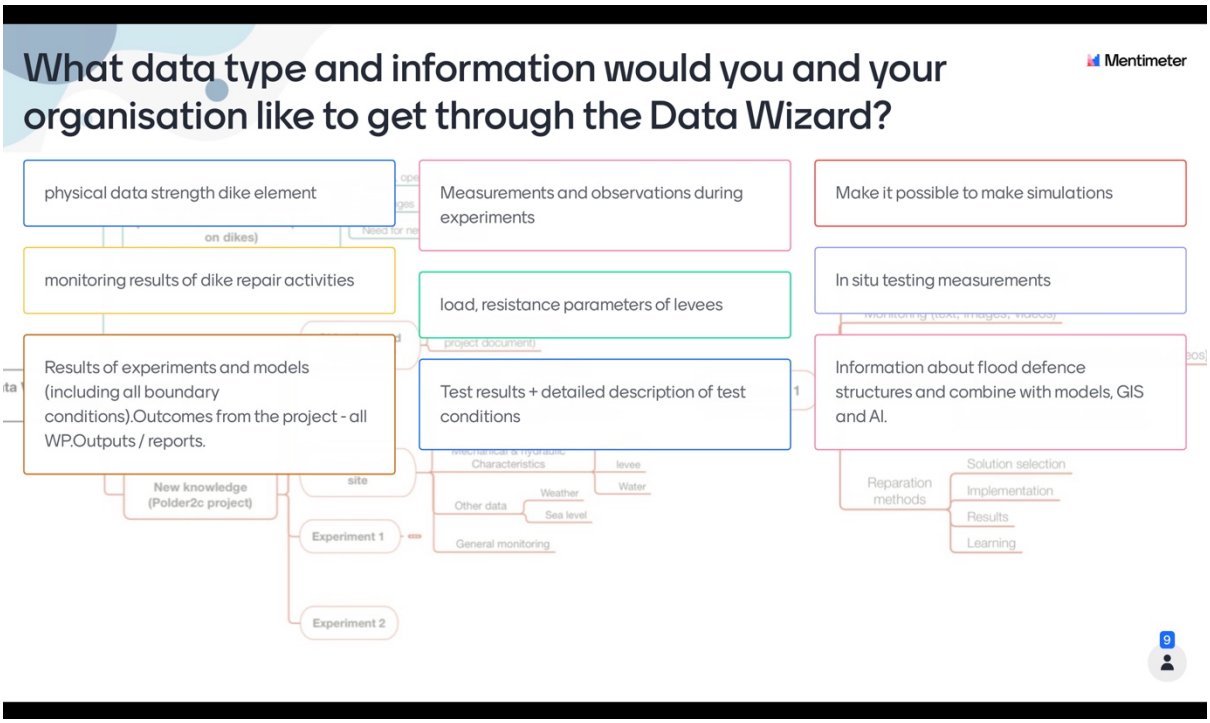

Figure 25: Data types that the participants are interested in accessing through the Data Wizard

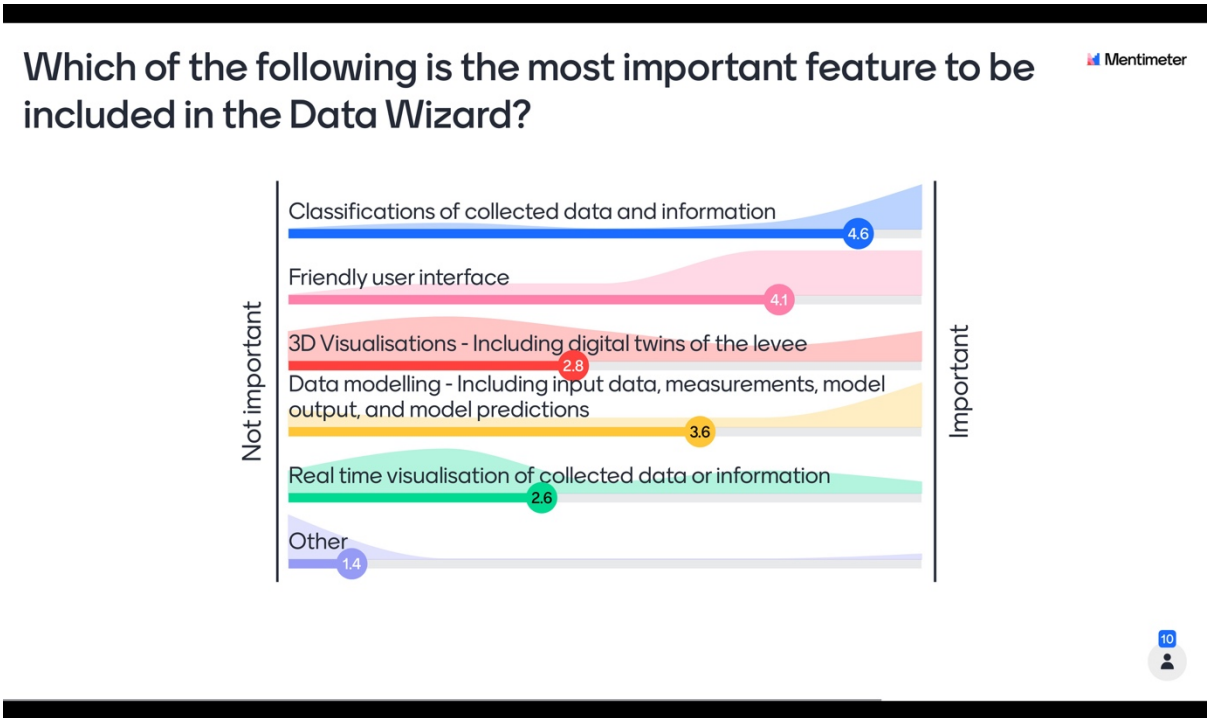

Figure 26: The importance of the proposed Data Wizard features

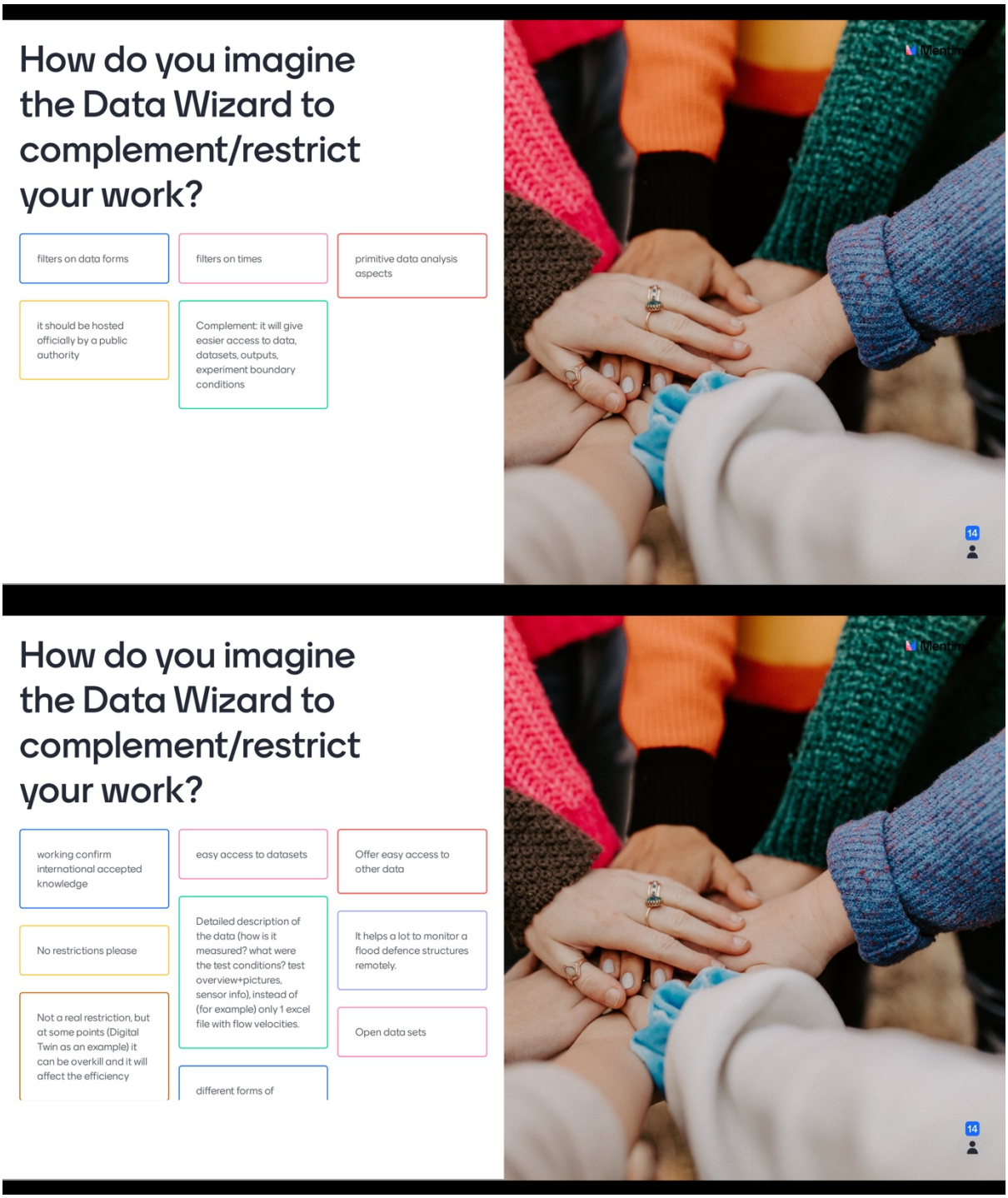

Figure 27: Participants' view regarding how the Data Wizard would complement/restrict their everyday work

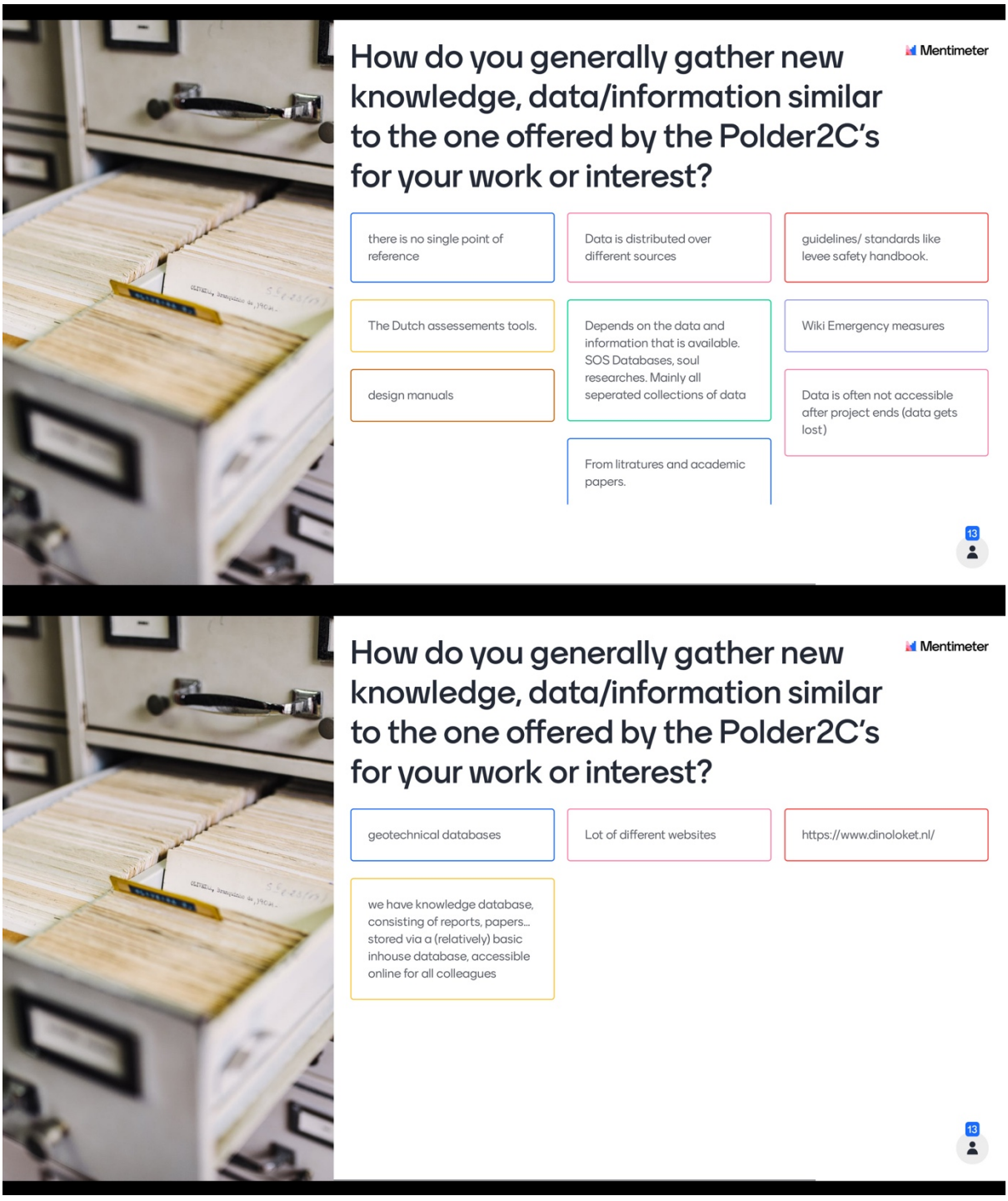

Figure 28: Platforms used by the participants to get data similar to the one produced by the Polder2C's project

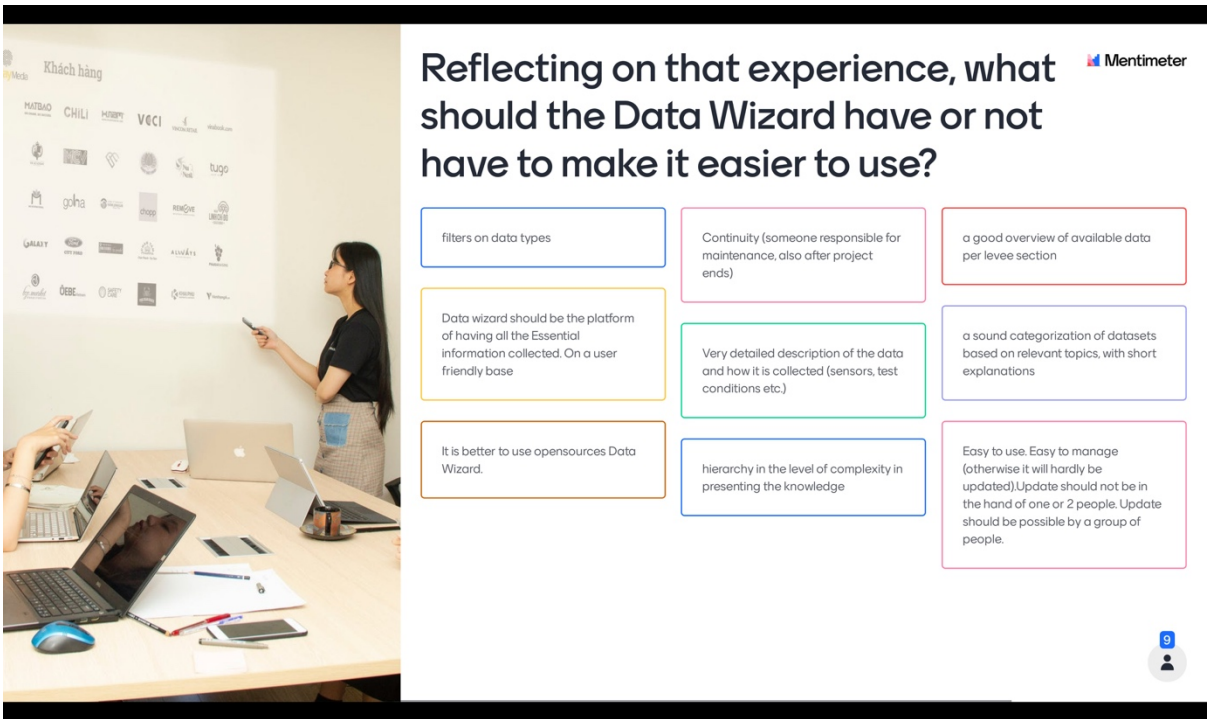

Figure 29: Participants' positive and negative past experience with the aforementioned platforms

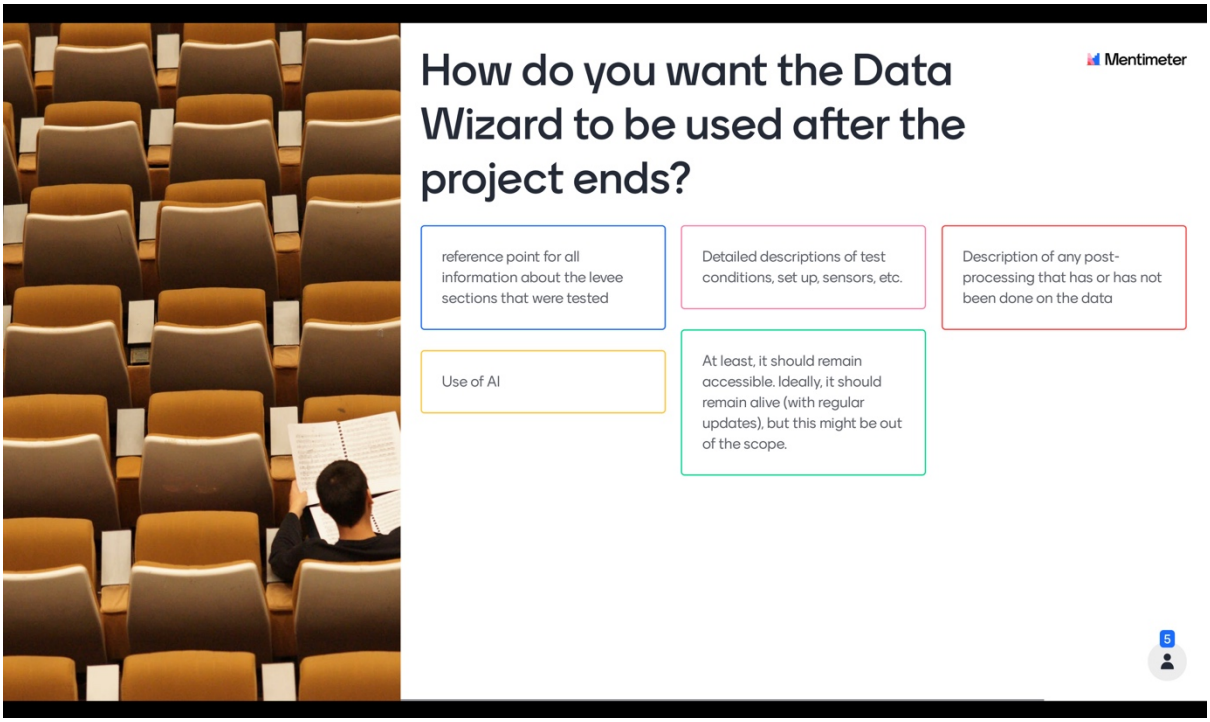

Figure 30: Participants' opinions on how the Data Wizard should be used after the project ends

Figure 31: Word cloud – participants outlining other potential Data Wizard users

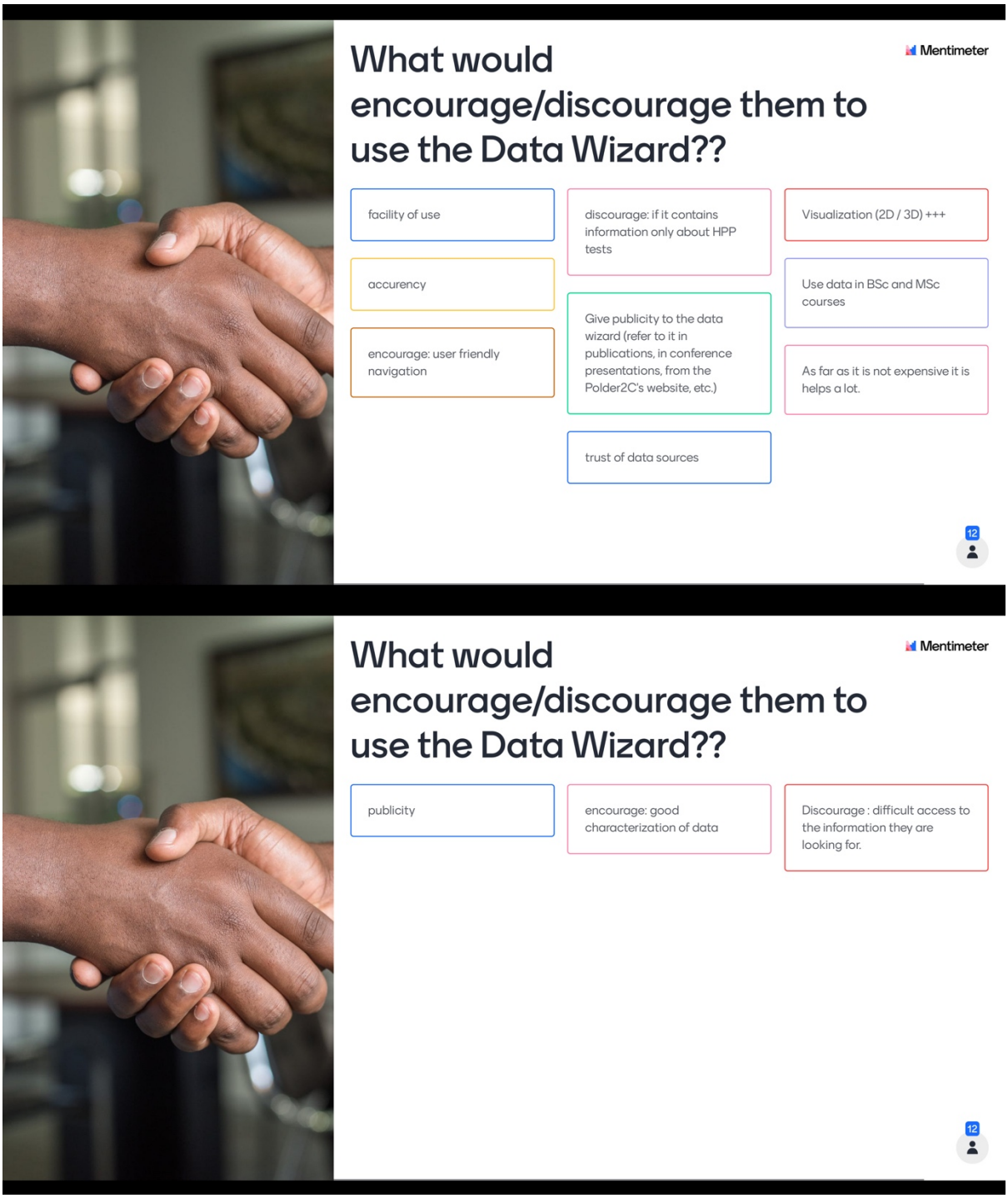

Figure 32: Participants' opinion regarding what would encourage/discourage the aforementioned users to use the Data Wizard

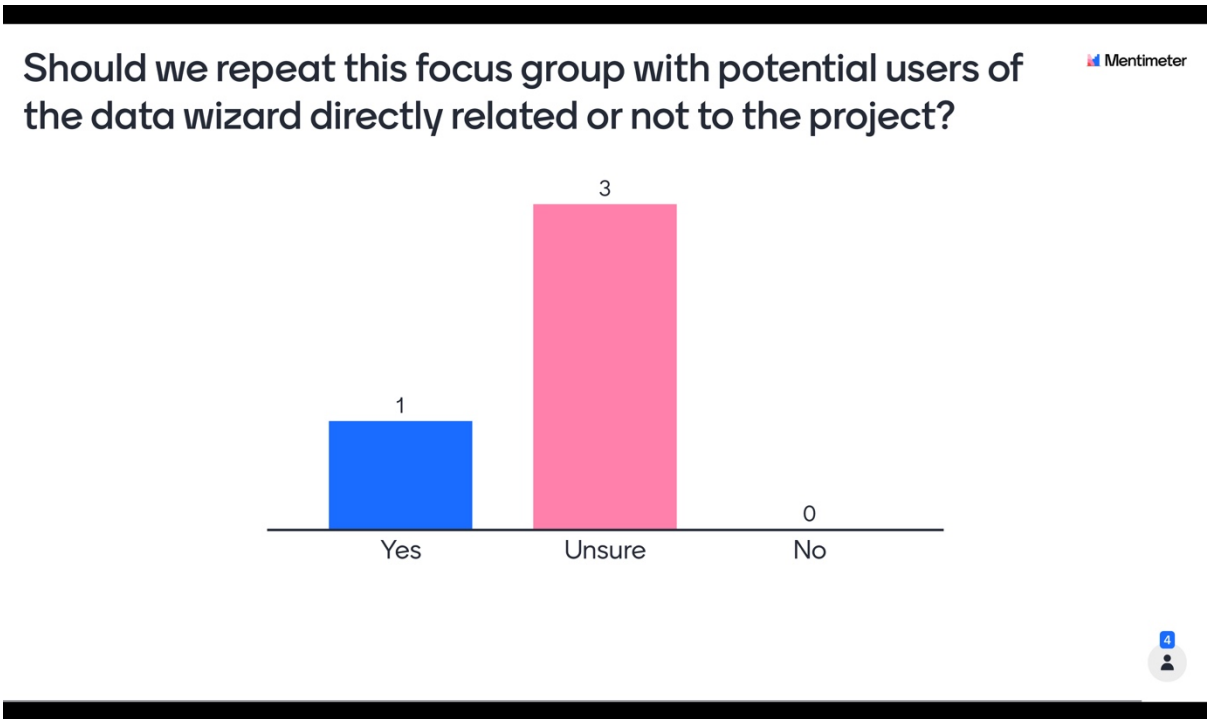

Figure 33: Participants' opinion on whether or not a similar workshop should be organized