AHEAD

Co-funded by the European Union

Deliverable 2.3. Summary of the workshop on requirements (and subsequent national seminars)

Subtitle: Knowledge based design of an information system to collect and analyse post disaster damage data

Report prepared by Politecnico di Milano with contributions from Ahead partners as part of work package 2 of the EU | Union Civil Protection Mechanism | Knowledge for Action in Prevention & Preparedness (KAPP) funded project "AHEAD - Advanced disaster damage and loss data information system for enHancEd impAct-baseD knowledge"

Deliverable 2.3 | 9.01.2026

Authors: Scira Menoni, Jacopo Bambini, Sandro Salari, Anna Faiella, Xavier Romao

Contributions: Frederic Tatout

Index

1. Introduction

Evidence based policy and decision making in disaster risk reduction is important in order to issue policies, legislation and design measures and interventions that are at least effective in reducing and/or mitigating the damage and its secondary effects, possibly also financially efficient from a cost benefit balance.

As part of evidence damage and losses incurred as a consequence of natural hazards are essential, as only tracking them overtime it is possible to check to what extent policies and interventions have been successful (Marin Ferrer et al 2018). Still, tools and systems to manage such data are not well developed and even missing in many regions of the world, including in developing countries. A survey of the current landscape of damage and loss databases at different spatial scales from regional to international has been drawn in Deliverable 2.2. grounding on the not so wide available literature, including grey literature such as reports that account for an important share of the latter.

An important effort in analysing persistent gaps and needs in post disaster data management has been made overtime by the European Commission, in particular DG ECHO and the JRC DRMKC (De Groeve et al 2013, 2014). In the context of such intensive effort to support Member States in the development of enhanced information systems to manage such data, a number of projects have been funded by the European Commission both by DG ECHO with the Idea project (Menoni et al 2016) and the Lode project (Faiella et al 2022) and by Framework Programs, in particular the FP7 Know4drr (Concilio et al 2022) and the Horizon 2020 Educen[1] projects. The work carried out by varying consortia in those projects constitute the legacy of the current effort in the DG ECHO funded Ahead project as can be seen in figure 1. The long pathway that has been followed in over ten years by now allowed to first provide a framing to the potential solutions and to design an information system that is in its current stage operational but in need to be completed with records that are still missing and engineered. The aim of this deliverable is first to describe how stakeholders' and experts' knowledge has been leveraged upon for requirements elicitation and second to reflect on the type of collaboration between domain experts and software developers that is still ongoing to test and complete the engineering process of the piloting information system.

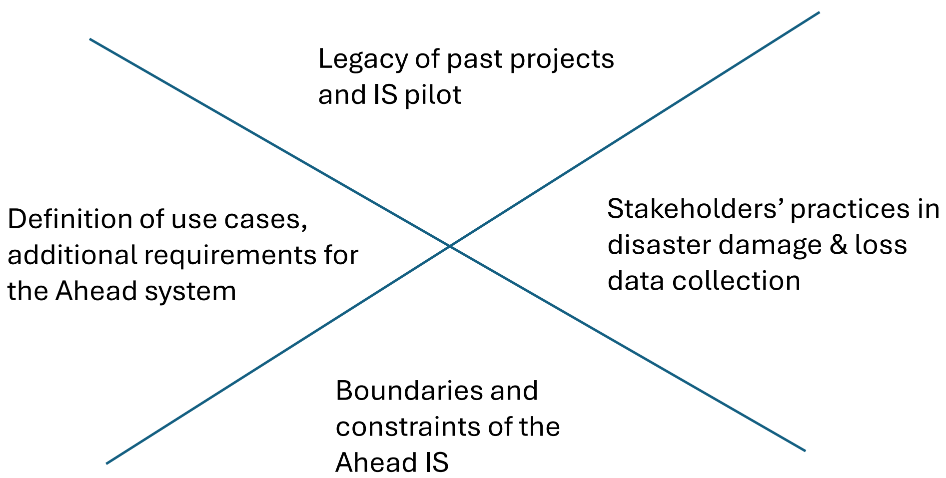


Figure 1. Main aspects of the Ahead IS development

[1] www.educenhandbook.eu

In order to do so, the other elements in figure 1 must be discussed. First the extended project priorities that have been set in the following:

- Build on prior effort instead of starting from scratch every time. The objective being to achieve a functioning information system consisting of a relational database and an interface for data input and query.
- The focus of the information system is strictly on post disaster damage assessment to assets and systems. Therefore data about affected people is not considered. Similarly, the information system is not meant to collect data about the hazardous event(s) that provoked the damage, only a generic reference to the type of hazard(s) and the date(s) of occurrence are reported.
- Beyond testing and improvement of functionalities to make it fully operational, the IS also lacks some components, i.e. some records on sectors that have not been developed yet and are necessary for its completion. In particular the records that are under development refer to: cultural heritage, some infrastructures such as defense measures, waste, water, and natural environment.

For those sectors to be developed the knowledge of stakeholders who either work in those sectors and/or are responsible for their recovery in case they are affected by natural hazards is essential. Such knowledge must be translated into requirements of what are the necessary attributes of both assets and damage that must be gathered in order to be able to forecast an expense for their repair or substitution. Stakeholders consultation though is fundamental also to understand what are key functionalities that the IS must fulfill in order to provide the added value that is sought for in comparison to existing databases and systems. As will be discussed in the following, such consultation resulted both in the specification of information that was considered relevant for each sector and in a number of use cases.

Last but not least, there are constraints that translate into boundaries within which the IS operates and will operate also at the end of the project. Some descend directly from the set priorities, so some types of damages (affected people) and some elements related to damage (events characterization) lie outside the set boundaries of the IS. Time is another important constraint. In businesses the time devoted to requirements elicitation is often shrunk due to the very limited time domain experts have to discuss with developers (Park and Kim 2018). The case is different in the context of this project, in particular researchers devote significant amount of time to study the requirements in the different sectors and to elaborate frameworks to describe the relations among the entities associated to each set of data to be input in the system. Also stakeholders consultation is an important aspect of the project that occurs through formal consultations like in workshops and seminars that have been held in the first months and informally, during exchange of opinions and discussions on the first outputs of the testing.

The effort that is described in the following pages can well be included in what Ogie and Verstaevel (2020) define as “disaster informatics”. Based on an extensive literature review, the Authors proposes a definition of informatics that connects aspects that are shared by consulted literature also across time as “the scientific discipline that deals with the study, design and application of information and communication technologies to facilitate the processing, use, storage, and management of data and information as required to support human knowledge in solving problems within different social contexts.” Disaster informatics being specifically devoted to the disaster risk reduction and management field. The Authors mention ICT such as remote satellite imagery, GIS, and those supporting emergency and crisis management, the so called EMIS (emergency management information systems) for which a quite rich literature exist. Within the latter, they provide quite large room to publications on crowdsourcing data and on data derived from social media. There may be an explanation for such focus, denoted by the relatively much larger literature in this specific domain. Crowdsourced data are big data by definition and therefore fit more easily into the capacities of current tools such as machine learning that require large datasets for training. However, there are two main challenges related to such data. On the one hand, as discussed by a thorough literature review, crowdsourced data displays pros and cons when it comes to its actually efficacy in guiding action in the field and does not always stand the promise of emergence of quality and reliability. On the other hand, authorities and public administrations involved in disaster management are proven to be reluctant to embed new technologies into their workflows for a variety of reasons, including the constraints of time that do not allow for testing new technologies under the pressure of emergency impellent needs. In addition culturally rescuers organisations have a rather high standard of evaluation the utility and effectiveness of a new technology before they decide to actually adopt it in their processes. Interestingly in their article Ogie and Verstaevel (2020) do not mention systems for collecting an querying damage and loss data, even though the latter clearly fit every aspect of the proposed definition and in principle such systems would respond to a real need that is collecting and accessing damage and loss data for a variety of purposes all relevant to disaster risk management organisations.

This deliverable builds on the outcomes of the initial workshop on requirements elicitation that was carried out at the beginning of the project (Annex 1) and a number of seminars and informal meetings with stakeholders including with the associated partners of Ahead to discuss requirements.

2. Requirements for an advanced information system of post disaster damage and loss data

Within the legacy of prior projects as depicted in figure 1, also key requirements have to be accounted for, and more precisely regarding:

- The first and perhaps most important requirement that derives from past research relates to the main required characteristics of the information system: its expected ability to enable data collection for all relevant sectors in an as much as possible coherent fashion, to allow for cross-sectoral assessments and comparison. The most advanced and complete representation of this can be found in **figure 1**.
- The type of data to be collected: related to both damaged assets, in terms of features and characteristics that may describe at least in part their vulnerability to certain hazards, and to the damage itself. The latter is intended both as physical impairment, destruction, partial or complete, to components or to the entirety of impacted assets, and functional damage, intended as lost service provided by the asset, that may or may not be directly the consequence of the damage to that asset and may instead result from interconnections among systems.
- Therefore damage data that are collected are both qualitative and quantitative. As for the former the effort already carried out in past projects relate to the collapse of narrative descriptions into standardized terms describing the damage and the assets from a pre-defined set of words. This mirrors what happens in real world post disaster damage assessment for compensation made by governmental agencies, donors and insurance.
- Some relevant relations among damages have been already modelled as critical requirements to be designed into the system, for example the possibility to track functional damage due to damage in other systems/sectors.

Based on a systematic review, Hidalgo et al (2024) categorize requirements as follows:

a. Functional Requirements: Define specific system functionalities.
b. Non-Functional Requirements: Address quality attributes like performance, security, and usability.
c. Domain-Specific Requirements: Tailored to specific industries or applications.

a. As for functional requirements, defining what the system must actually do, how data are input and extracted, prior projects identified the following requirements and shaped the current version of the IS (figure 2):

- The core ontological element of the IS is the damage, expressed both as an engineering description and an economic loss (Di Ludovico et al 2022);
- The triggering element for inserting data is an event, for example one for which a state of emergency has been declared. In some cases the initial data insertion is made to support the request of a state of emergency declaration, that once issued, requires data to be further collected and refined.
- Deriving from the latter real world condition, administrations reiterated the need for a system that is keeping memory of the various estimations made across time. The IS must therefore represent the evolution of damage estimation overtime. Such evolution may derive from two circumstances: a more refined assessment changes the initial estimation or the damage itself may deteriorate overtime (because of critical conditions that exacerbate the initial situation as may commonly happen in impacted structures but also due to additional events that may occur before repair and rehabilitation takes place).
- The process that permits to detect and estimate the damage and loss is a survey that may take place in different ways, directly by surveyors on the field, indirectly through claims by affected owners and further more formally certified. Initial damage assessment can also be carried out remotely through satellite images or interpreting pictures taken by drone or posted on different media. In all those cases an important requirement for the formalization of the process is the characterization of the survey process itself, including data regarding when and by whom (or what means) the estimation has been made.
- The asset that has been damaged can be either privately owned (like in the case of a business or a dwelling) or can be a public asset. Such distinction is key as in some procedures, for example the application to the Solidarity Fund granted by the EU in the concept of the UCPM for which only public assets are eligible. A further important distinction is between insured and non insured assets.

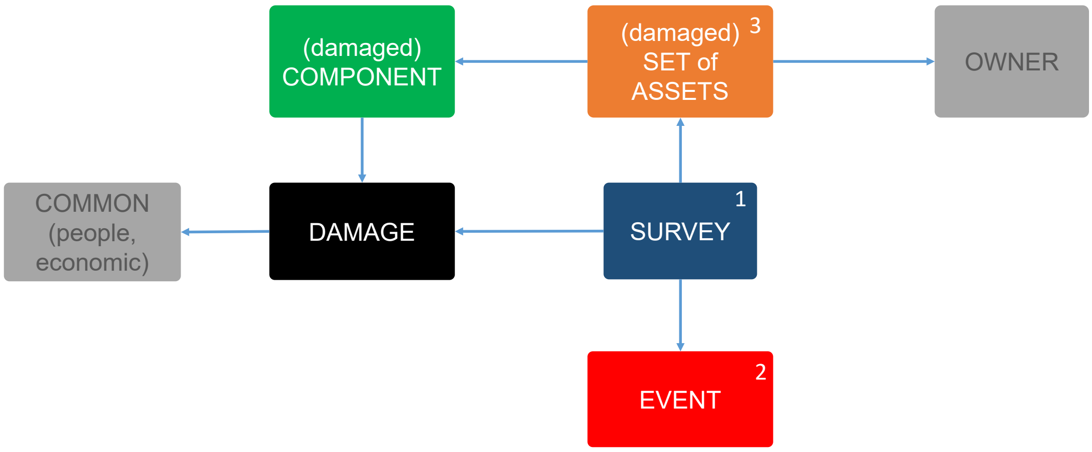


Figure 2. Abstraction of the Ahead IS physical model, the numbers represent the actual sequence of steps that are required by the system to the user.

The requirements that resulted from the historic development of the research in the field as described in the introduction led therefore to the current version of the IS that is made of three fundamental components, namely i. the elements characterizing the survey (including the triggering event), ii. The damaged asset (including positioning information that is key for representing data also geospatially), and iii. The type of damage that has been incurred.

- In fact an additional important requirement that has been highlighted by several stakeholders refer to the possibility to localize damage geographically, and possibly also aggregate to larger scale if needed. Spatial scalability is an additional more conceptual basic requirement that has been identified as crucial along with cross-sectoral and cross-temporal damage characterization (Menoni et al 2017).

In past projects, non-functional requirements were far less developed, as the IS was in its initial piloting phase. The need of a login system for administrations was an obvious need common to all existing systems. In addition, an important requirement was the alignment with systems already in place in various countries for damage data collection. In particular such alignment was sought in countries where usability assessment of structures following the impact of an extreme event is routinely carried out. In fact, several aspects coincide between damage assessment in general and usability: the need to carry out a survey, the characterization of damaged assets both functionally and from a physical point of view (to understand the severity and the reversibility of damage). There are also important differences as highlighted by Faiella et al (2022), in particular regarding the objectives of the two processes: whilst usability delivers a decision on the possibility to use a given asset in a safe way, the second provides an estimation of the cost of repair or replacement. The former must stay on the safe side and therefore an overestimation of damage is fully acceptable, the latter must inform instead recovery and therefore must obey to norms of fairness and coverage of the entire affected systems and assets.

Even though the domain of application may be considered obvious and simple, it is not at a more accurate examination. In fact post disaster damage and loss data collection and analysis stands at the interface between two phases of the disaster that are also managed by different organisations and administrations. Initial damage assessment for the assessment of needs is usually carried out by civil protection authorities who intervene for rescuing and re-establishing basic life conditions in affected areas. Damage assessment delivered by satellite and aerial images are part of this initial assessment, such as the service provided by the EU Copernicus system. Second level, more reliable damage and loss assessment are made generally after the first acute emergency and are meant to support the recovery and reconstruction phases, in which other administrations, including local levels of governments must take the lead. This delicate interface between damage and loss assessments across distinct phases of the disaster has been the object of some research (Walia and Menoni 2020) but has not been fully interiorized until now. Many challenges and obstacles towards effective compensation mechanisms derive from a lack of clarity of what type of damage assessment must be considered as legal reference, in many cases only initial estimations carried out during the emergency are available and they are the only guidance towards the definition of reconstruction needs. The fact that different authorities and administrations are responsible across the temporal interface, complicates the elicitation of requirements as different administrations have different views on what are the priorities and even on what has to be considered as damage (Molinari et al 2014).

2.1 New requirements elicited in the context of the Ahead project

<u>a. The following functional requirements have been set for the further implementation of the IS:</u>

- A more strict use of labels has been highlighted as a must particularly by the JRC that has a vision on the need to harmonize both among Member States and to contribute to international efforts put forward by the UN. This is an important requirement to create interoperability between the RDH of the JRC and the Ahead IS. The Ahead IS can be accessed at the following address: https://ahead-impact.app/. At the moment only authorized partners and a limited number of external users (public administrations) may input data. It has been suggested to make Ahead comply with the UN Hazard Classification (UNDRR and ISC, 2025). As for affected assets and systems it has been proposed to align to the OECD classification of assets (van de Ven 2021).

- Most of requirements that have been identified within Ahead relate to the use cases to be supported. One boundary that derives from decisions made in previous projects relate to how use cases will be supported, not creating ad hoc functionalities, but through the queries to the database. Being the database relational several complex queries can be run, thus permitting to use data for multiple purposes. The rationale of such decision lies in the very wide range of possible use cases as initially structured by De Groeve et al (2013) and further by Walia and Menoni (2020). It must be also noted that additional use cases have been proposed in the context of Ahead. Allowing a wide variety of queries would support most of use cases and therefore the focus and the investment should go in the correct and as comprehensive as possible modelling of data.

The use cases that were already set for the IS are:

- Support public administrations in their task to collect damage and loss data for compensation purposes;
- Support the evaluation of the Sendai targets;
- Support the improvement of risk modelling through empirical data on the outcome of given hazards;
- Support the forensic analysis of damage.

During the first months of the Ahead project, it has been proposed to classify use cases in three categories:

- Use cases that can be supported by data introduced each time in the system. For this purpose a set of pre-identified queries have been proposed for implementation.
- Use cases that may be available in the near future, as they require a limited time series of data
- Use cases that will be available in the longer term, once the system is routinely used and a larger set of historical data will be available.

Among the possible queries: the possibility to aggregate data from asset level to county (NUTS 3) and larger (NUTS 2 and 1).

Box 1

The following queries were considered in the testbed of the Milan area affected by the July 2023 storm.

- ✓ Total estimated loss for public entities in Milan province
- ✓ Total estimated loss for private entities in Milan province
- ✓ Buildings with structural, non-structural, and functional damages (cost division or count) in Milan province
- ✓ Total cost of repairs, replacements, vs temporary measures in Milan province for both private and public sectors
- ✓ Economic costs (or count) per province divided for roads, schools, and private buildings
- ✓ Building management type: public vs private (economic cost or count) in Milan Province
- ✓ Economic costs (or count) of different classifications of structural damage for public entities in Milan province
- ✓ Private Buildings with and without insurance in Milan province (and the total costs reimbursed by the insurance vs not reimbursed)
- ✓ Damages above a certain threshold of cost
- ✓ List of owners of the private buildings to be reimbursed by the insurance above a certain threshold of cost
- ✓ Temporal analysis of the surveying: count (or economic value) of entities recorded from the first, second, third, etc surveys
- ✓ Provinces with the highest damage on roads

The results can be seen in Annex 2.

Source: Kantaria A., An advanced information system for managing disaster damage and loss data: the 2023 Lombardy storm case study, Thesis discussed in the CERM Program, October 2025

b. Non functional requirements

- Most non functional requirements have been proposed and gradually implemented in the system in Ahead. First the need to introduce a login system to ensure that only legitimate actors will be able to input and modify data and to track those who will query the system. In the latter regard it has still to be defined what levels of access to query will be allowed, considering the variety of use cases that potentially open significantly the range of users.
- System interoperability to make access, exchange, integration and cooperation on data uses possible at least with some systems. At the very least this will have to be done for the RDH developed by the JRC. (when multiple information systems, applications, and devices can seamlessly access, exchange, integrate, and cooperatively use data). This means that certain protocols must be in place for that to happen, namely the AHEAD system should have an API or define the data exchange format so that other systems can communicate with the AHEAD system.
- Second the need to dispose of a way to navigate the large number of events for which damage and loss data are collected. A sort of coding or structuring of the events will be needed especially as the number of events for which surveys are created will grow. This was the idea behind the GLIDE number which, for some reason, was not really embraced by everyone, although still exists: https://www.glidenumber.net/glide/public/search/search.jsp (Nishikawa 2003; Shiomi 2023).
- Third some rules for FAIR treatment of data also will be needed. At the moment, as is done by most systems that we could analyse both in insurance and public administrations, private data on the damaged and affected owners, organisations, people, are collected. How to make those "invisible" to users querying the system is an important issue. According to Article 14 of the EU Regulation 2023/2854 on Harmonised rules on fair access to and use of data, emergencies and disasters represent an exception to norms to be nevertheless very carefully handled. Therefore damage and loss data that are collected in the aftermath of emergencies must be treated in such a way that they can be reused for multiple purposes and shared among different stakeholders, many of whom have contributed to their collection. However, due to sensitivity, privacy and security issues (Huang and Soete 2025) there is the need to address each concern through appropriate measure before data can be safely shared. As data are collected at the asset level, so information about the owners and their losses will be inevitably collected. What can be done is to prevent data from being queried at that level (e.g. you can only get information at the city block level) or you need to anonymize certain data if you want to allow a certain type of query at that very refined level. Still, for certain purposes you will want to have access to that fully detailed data (imagine that you are trying to validate a risk assessment model and you are replicating a given event for which you have very detailed data about the impacts. Ideally, you want to compare the model predictions with the asset-level impacts, not with the aggregated impacts). So, for certain uses, you may need to develop data-sharing protocols (Carvalho et al 2025; Vitlov et al 2025).

3. On the process of requirements elicitation in the Ahead project

The experience matured in the Ahead project and in its predecessors reinforces the understanding of requirements elicitation as a cyclic iterative process, that passes through several steps of proposing a schematization and standardization of practice and types of data to be treated, translating the latter into a logic and physical model, implementing in a partly working system and then reconsidering some choice and adding attributes and relations based on new insight and understanding, on the evolution of the practice itself and on problems that emerge once real event data have to be managed.

For example, when this work started there was less urgency in managing data from multi-hazard events, that not only have, perhaps, become more frequent, but are certainly perceived nowadays as much more central and challenging than used to be the case even a recent past.

The landscape of research and practice in disaster risk reduction and climate change adaptation is rapidly and dynamically evolving, imposing new aspects and understandings to be embedded in the system for it to be relevant to stakeholders. The landscape of damage and loss databases is changing as well, with more efforts carried out nationally and internationally, as resulted from the analysis carried out in Deliverable 2.2. Those efforts are struggling with the multisectoral damage data collection and the difficulty to coordinate among the latter under current governance arrangements; the large variety of damage and types of damage that are not always easy to monetize and require a very systematic procedure for their collection; the issue of sustainability overtime that is less acute in governmental based datasets but still dependent on the specific organization that is running them. On the one hand such difficulties that proved to be persistent overtime, defy governments from such a demanding task, and in the meantime push for the search of shortcuts such as reliance on crowdsourced data or satellite imagery. However, as discussed in literature, such methods cannot be the panacea for all hurdles of damage data collection and do not satisfy the need to dispose of good quality verified and surveyed damage for the multiple use cases that have been set. In this regard also claims of big data and artificial intelligence are defeated by the lack of data instead, by the issues in access to existing datasets that in any case have been developed according to standards and formats that make them hardly interoperable.

All those reasons explain why, despite the actual need reiterated both in policy and research debate, there is little incentive to embark in such an endeavour. Based on the experience of Ahead and predecessors projects, in the following some reflections on the process of requirements elicitation are proposed.

3.1. On knowledge co-production and sharing underpinning the development of any IS

In the Know4DRR project, specifically devoted to investigate knowledge formation, sharing and exchange across the stakeholders groups, including citizens, that must manage risk at various levels, knowledge is considered not an item, a commodity that can be transferred form one person, one group to another, but a process. In fact, knowledge is in the meantime embodied in individuals and created socially. Which means on the one hand that knowledge cannot be detached from knowing and from the knower, every time knowledge is shared, it needs to be reinterpreted and reconstructed in the mind of the learner. On the other, it means that the knowing process does not occur in a vacuum, it is intrinsically social. Because knowledge is a socially constructed process that is embodied in an individual and recognized and shared in a community,

Elsewhere (Faiella et al) the knowledge pyramid as an effective way to depict the relationship between knowledge, information and data has been discussed based on an extended literature that goes back to the Sixities. The knowledge pyramid, initially proposed by Acker in 1988, has been revisited and reinterpreted by authors who suggested to break its rigidity, recognizing that there are many feedback loops in the data-information-knowledge “pyramid” (Davenport and Prusak, 2005; Frické, 2007; Jennex, 2009).
Already in the proposal stage it was decided to group stakeholders to be considered for any knowledge management policy into four categories: scientists, public sector, private sector, and “civil society”. The first three are relevant for the sake of the present discussion.

Rescher suggests that science aims at four distinct goals, namely theoretical and cognitive, aiming at describing, explaining the world and providing the base for successful prediction, and practical and manipulative, to provide control, effective intervention over nature. Rescher also suggets correctly that science is not a collection of knowledge, but a ”system for creating it”. In the context of DRR and CCA, the construction of scientific knowledge is hence a process of reciprocal adjustment between theory and experience in an interdisciplinary space. Focusing on damage and loss data, scientific knowledge is deployed mainly for practical purposes, of organising and structuring at best data and on the other to use such data to enhance the theoretical understanding and the cognitive awareness of damage mechanisms, risk models and evaluation of measures in place to reduce/avoid damage.

Organisational knowledge is procedural, expressed as a set of routines resulting from an accumulation of experience that lead to future behavior (Levitt & March, 1996). It guides into how things should be done and processes conducted in order to provide the required type of service to collectivity. Organisational knowledge is particularly important in the domain that is discussed here: public administrations and insurance are the main organisations that carry out the task of post disaster damage and loss data collection and estimation. Therefore such organisations are the main stakeholders that is consulted on the matter of both consolidated practices and gaps in current arrangements and available tools.

Legislative knowledge refers to the procedures and the governance arrangements that are necessary to amend existing rules and laws and to issue new ones to govern processes and activities that either did not exist or were not deemed to require specific regulations and norms until when a new legislative project was initiated. In the field of DRR and CCA the dynamism that is suggested by the definition is evident, as the latter present often and more frequently in the recent past, unexpected occurrences, situations, risks that require renewed or new laws and norms. At the moment there is no specific legislation that requires post-disaster damage data collection, this is a task that is generally attributed to certain organisations and regulated each time by specific ordinances or decrees. More recently in some countries legislation on recovery and reconstruction has been issued formalizing the task of post-disaster damage data collection.

The three types of knowledge often interact, in some cases the same person is knowledgeable in more than one domain, but more often an interaction occurs between the three, that obviously are not siloed but rather interact and influences each other.

Table 1 lists a non-exhaustive list of possible influences that the three types of knowledge may exercise one on the other when they interact with specific regard to our main objective: enhancing DRR understanding and awareness.

Table 1 has been conceived as a matrix where different possible interactions between types of knowledge are considered; in the table such interactions are viewed as providing opportunities or constraints. Opportunities result from a positive interaction that produces as an outcome enhanced understanding, improved tools for decision making and action, better definitions and more mature concepts. Constraints derive from barriers to the integration of two types of knowledge that may be political, juridical, linked to the way careers are conceived in the academic world, etc.

When the scientific knowledge meets with the organizational a positive result may be the reshaping of organizational procedures, strategies and priorities. However change is often constrained by existent procedures and by unwillingness and/or time that is needed to reframe the latter. Scientific knowledge may trigger new legislation or amendment to existing one. This is particularly true for domains such as risks and medicine that are heavily dependent on scientific evidence. On the other hand, political processes and governance arrangements resulting in new or amended legislation can be extremely lengthy resulting in a time mismatch between when a certain information, result is achieved and when it is translated into practical behaviors and norms.

Organisational knowledge may trigger research to deliver solutions and tools for domains, activities in which those are needed. Information systems to manage disaster data are an eminent example. On the other hand, as a constraint, there is no guarantee that even the best solution will be implemented, as there may be reluctancy to introduce the needed adjustment and to go beyond “business as usual”. This has been considered as part of non-functional requirements, that is to design the information system as close and similar as possible to tools that are already in use for managing disaster data (including usability assessment tools or incident reporting).

<table>
<tr><th></th><th></th><th>Scientific knowledge</th><th>Organisational knowledge</th><th>Legislative knowledge</th></tr>
<tr><td rowspan="3">Scientific knowledge</td><td rowspan="2">Opportunities</td><td rowspan="3"></td><td rowspan="2">Can help reshaping organisational procedures, strategies and priorities</td><td>Leads to amendment in legislation based on new facts and understandings</td></tr>
<tr><td>Science for policy is an emerging field of expertise by itself</td></tr>
<tr><td>Constraints</td><td>Change is hampered by established protocols and arrangements</td><td>Political processes for amendment and new legislation to be issued are lenghty and contorted</td></tr>
<tr><td rowspan="2">Organisational knowledge</td><td>Opportunitie</td><td>Provide practical needs and apply tools and solutions in practice, in real world contexts</td><td></td><td>Transforming practices in norms, codifying solutions and tools to become a reference or a norm</td></tr>
<tr><td>Constraints</td><td>Implementing a solution or using a new tool is a long process, that requires adjustments</td><td></td><td>Novelty is hindered by institutional structures</td></tr>
<tr><td rowspan="3">Legislative knowledge</td><td>Opportunities</td><td>Legislation may be a trigger of scientific innovation to respond to new regulations</td><td>Need to adapt organisation procedures to new norms, organisational innovation may be required</td><td></td></tr>
<tr><td rowspan="2">Constraints</td><td>Science working for policy has its own challenges and "rules"</td><td></td><td></td></tr>
<tr><td>May be hindered by processes imposing liability on scientific expertise.</td><td>New legislation can be met with resistance and reluctance to implement, delaying impact in the real world</td><td></td></tr>
</table>

Table 1. Non-exhaustive list of possible influences that the three types of knowledge may exercise one on the other (adapted from the Know4drr project, see Mejri and Pesaro 2015)

As convincingly commented by Carlile (2004), the dynamism in knowledge sharing and production derive also from the differences in knowledge among groups and subgroups.

Carlile distinguishes three types of boundary conditions:

- syntactic or informational,
- semantic or interpretative,
- political or pragmatic.

The first type, “syntactic or information processing boundary” is encountered when there exist a knowledge “quantitative” gap among actors, or the qualitative differences easily complement each other, so that transfer of knowledge from one group (sub-group) to another may easily take place. In the field of DRR this occurs between scientists and administrations, when the latter require investigations or monitoring on certain phenomena and their potential evolution.

The second type of boundary condition is more relevant, when knowledge is qualitatively different to the point that meanings and definitions need to be discussed and agreed upon, defined by Carlile as “semantic or interpretative boundary”. In this case actors depend on each others’ knowledge and need to come to an agreement about the use of terms if they want to positively collaborate. This is the boundary that must be overcome by experts in DRR and CCA and software developers. In addition it must be pointed out that also the first group of experts is very diverse as they range from different scientists searching hazards to engineers, lawyers, sociologists, etc.

The third type of boundary is the toughest one is the “political or pragmatic boundary”, where knowledge needs to be transformed through negotiation and new forms of agreement among actors must be found. In this light a number of assumptions that are common in the “disaster community” should be discussed in a more critical way. For example the idea of “interoperability” that is implicit in the Inspire Directive suggests that data can and should be shared thanks to enhanced technological solutions. However, what the application of the Directive has not fully envisaged until now is the fact that in the disaster community for example data and information regarding the apparently same topics are measured and assessed in very different ways, so that even the apparently simple exchange of data cannot take place. What is needed instead is either a new agreement on terms or the transformation of knowledge through negotiations among stakeholders, with the inevitable repercussions on the attached set of data and information. The aspect of definitions has been raised as an important requirement for the Ahead IS.
The most important question when different stakeholders, groups holding different types of knowledge interact is how such boundaries can transformed into interfaces. Different solutions have been suggested in literature. For the sake of the present discussion, two are of particular interest: knowledge brokers and boundary objects.

As for the formers some experts may become experts in collecting knowledge from other experts and stakeholders and organize it in a way that can be easily understood and used. This has been the case in the case of the Ahead IS: some experts in DRR of the research group took this role and provided the developers with the necessary understanding of the domain on damage and loss data. This was quite a challenging task, requiring on the one hand to summarize what damage is in different sectors (from agriculture, to residential buildings, to firms or lifelines) and how such damage should be represented in the system. Given the large variety of assets and sectors, each time the work of the broker was to elicit information and knowledge from the various stakeholders and synthetize in a useful way for data modelling.

Boundary objects are instead occasions, situational contexts in which knowledge sharing, translation, and transformation may occur more easily thanks to a specific mandate or to a common task. Boundary objects have been already identified in literature as key to produce “salient” knowledge (Cash et al., 2003), as experts from different fields, practitioners and scientists, decision makers and hardware and software developers are working together to find solutions rather than working separately. Creating a common understanding and a common framing of problems sets the floor for finding solutions that are more likely to be implemented. Also, at the interface between disciplines, between practice and theory, innovation is more likely to occur, as consolidated knowledge within each field, each expertise, each sector is somehow hampering the introduction of new solutions, of new views of problems, which require as suggested by Carlile (2004) “knowledge transformation” that does not come without a cost. Innovative solutions require as a pre-condition the re-framing of problems, viewing them from a different perspective, which in turn must be based experience, as a background knowledge against which innovation and learning capacity can be tested. An IS can be considered as a boundary object, requiring all experts and stakeholders to collaborate and share knowledge in their respective domain with the developers.

3.2. The interaction between domain experts and developers through requirements elicitation and beyond

Gibert et al (2018) proposed an interesting reflection on the challenges that must be met by data scientists working in the field of environmental studies that is rather new at least to our knowledge. They suggest an original perspective digging into the historical origins of data science and reflecting on how it has evolved in more recent times, requiring a readjustment of how scientific inquiry and practical applications will take place in the near future in several domains. They quote for example the work of Lauro et al providing a very similar view on social sciences. The main standpoint of the article derives from the mission attributed to the newly created International Association of Statistical Computing in 1977 "to link traditional statistical methodology, modern computer technology, and the knowledge of domain experts to convert data into information and knowledge". This article is clearly very relevant as it has been written for a field that is closely related to disaster risk reduction and management. Several of the challenges they highlight are common to the domain of disaster data, namely the need to deal with the uncertainty of data, diversity in the formats data are available, the point that "big" (data) is not necessarily synonym of sufficient or unbiased. It can be also added that in the disaster field often lack of relevant data rather than big data is the norm (UNDRR 2023). And the case of damage and loss data is an eminent example. Therefore it can be said that it is not by chance that much literature is available on crowdsourced and social media data that fits well with the possibilities offered by modern IT, whilst much less is available on aspects such as post disaster damage assessment in different sectors from official sources the availability and accessibility of which is rather scant.

Coherently with this tripartite view of the data science as the nexus between the science of data treatment and analysis as techniques of statistical modelling, computer science supporting the latter at unprecedent speed and precision, and data and information elaborated in specific knowledge domains, requirements elicitation is not just a monodirectional passive transfer of requests and needs for a practical IT solution. Instead it has to be viewed as a complex process through which initial vaguely imagined output view result into desired configurations reflecting common views between recipients and developers on complete specifications and formal representation (Pohl 1993).

Chakraborty et al (2010) tested a quite interesting framework to depict the relations that have to be developed between developers and domain experts that is structured across four stages of scoping, sensemaking, dissension and termination. They describe how the four aspects of setting the objectives, knowledge transfer, trust building and mental models work in each of the above listed phases. Without entering into the detail of how this framework has been applied in practice and the results that the Authors have discussed reaching rather interesting conclusions, it may suffice here to focus on the generic schematization they offer on the interaction between developers and domain experts. The original framework is modified to fit the Ahead project environment (figure 3).

At a minimal level, the users of the IS under development are the members of the consortium and the stakeholders participating in the effort. However, such interpretation is rather limited. The ambition of Ahead is to serve a much wider arena of public administrations sharing similar problematics on damage and loss data collection, use and reuse. Therefore the research consortium is actually engaged in a much larger effort to identify criticalities in the disaster data collection and analysis, in understanding current (and past) gaps, in eliciting on their turn specifics of what is deemed as damage in different sectors and how damage evolves overtime, meaning both the damage itself (deterioration may occur time after the affectation) and its evaluation (at re-evaluation damage may be considered differently than initially surveyed). An important role has been played by the team expert playing at the interface with the developers. In this case a certain degree of knowledge and training on heuristics used by developers, on logic and physical models abstracting data concepts and categories, had to be acquired by the team member acting at the interface. Similarly also the system analysts and developers had to acquire some understanding of the knowledge domain on damage and loss data. In some phase of the project the system analysts and the developers were distinct persons, in some phases the same developer acted also as a system analyst. This can be considered as a real world circumstance, due to personnel leaving the team, transitional periods before a new expert is hired.

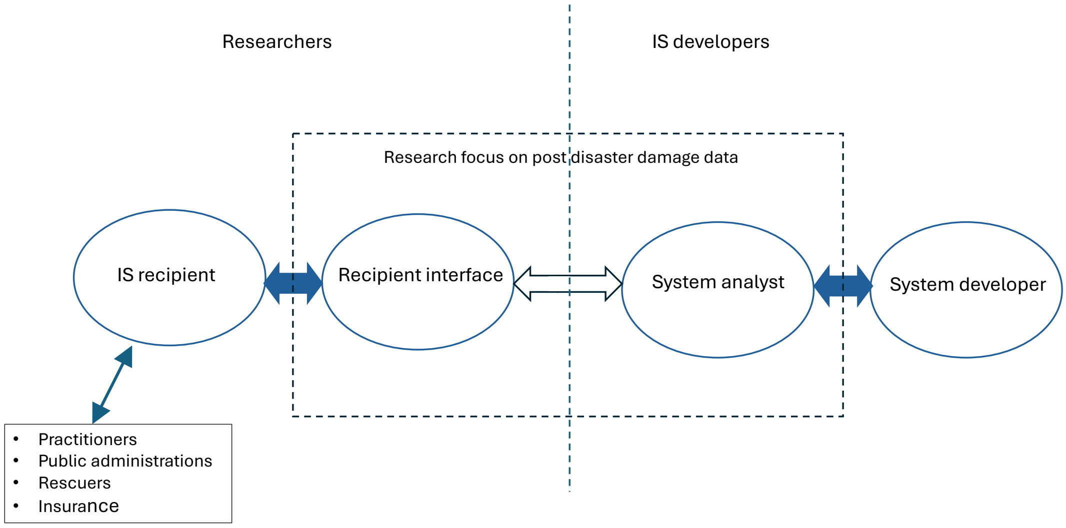


Figure 3. Interaction between researchers on damage and loss data and developersm framework by Chakraborty et al (2010) adapted to the Ahead IS

Van der Hort et al (2021) discusses the interplay between a functional view of data science as rigorous and objective and what they name "institutional" view of data science as affected by accountability, legitimacy, trustability problems as any other human endeavour. In order to reconcile the two views, they suggest to blend the data modelling process according to the functional view with the framework provided by Nonaka et al (2000) that is more in line with the institutional understanding. In the following figure a representation of Van der Hort and colleagues framework is adapted to the case of the Ahead project.

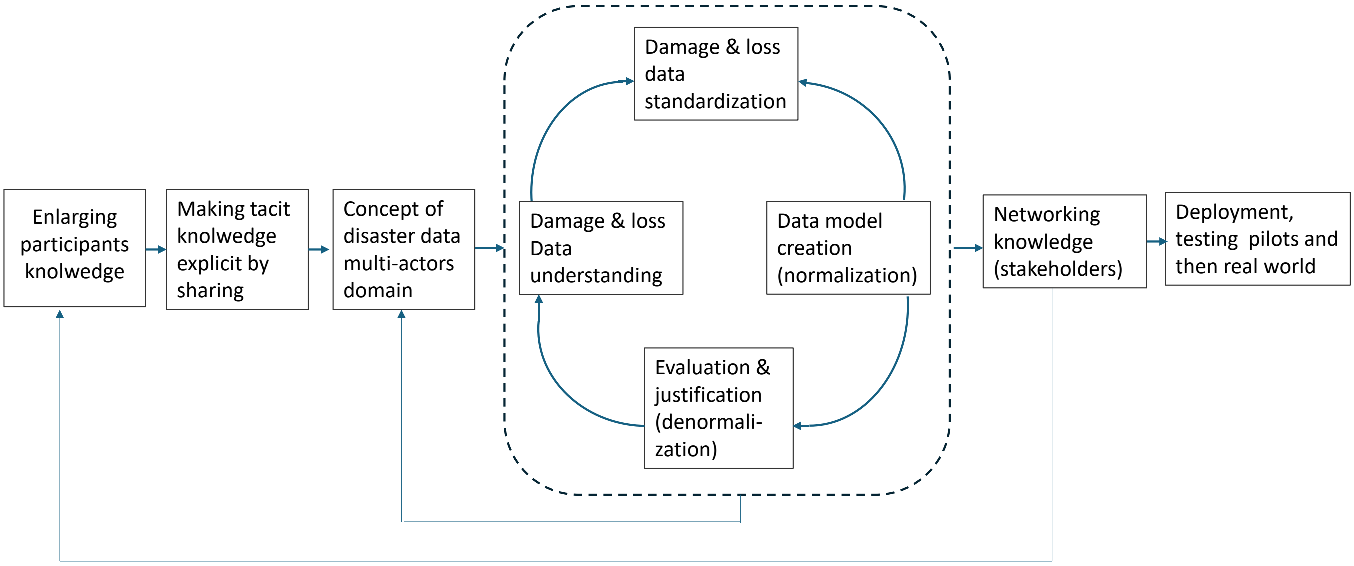


Figure 4. Process of data modelling for IS development from interaction and sharing of knowledge among developers and recipients to testing and deployment (adapted from Van der Hort et al, 2021)

3.3. The experience of knowledge sharing and requirements elicitation for developing the Ahead IS

An internal focus group was carried out by the coordinator of the Ahead project, the team member who has been in charge since the previous Lode project to communicate and discuss with developers about needs and requirements, and the developers respectively in the Lode project and in Ahead. The focus group was intended to make explicit how the process of requirements elicitation and consequent development of single stages of the IS were carried out throughout the lifetime of the two projects (Ahead still ongoing). The focus group was moderated by the project coordinator and consisted in a structured interview answering the following questions:

a. How the requirements elicitation and verification process has changed overtime since initial stages of collaboration between the research team and the developers?
b. What are the main differences that the developers identified between working in a research context and in industry (as both have extensive experience with industry and one also with public administrations)?
c. What are the best techniques for requirements elicitation according to the developers?
d. What were the advantages of an agile approach and the challenges according to the member of the team tasked with being the interface between researchers and developers.

In the following the main arguments brought by the participants are summarized, highlighting relevant aspects for the elicitation requirement process.

The developer of the IS pilot in the predecessor project Lode summarized in a quite interesting way the experience of working with domain experts. In particular he identified from his long term experience three important phases: the setting of the objectives, clarifying why the recipients were asking for a given product, a given IT solution, the definition of what the system should do, what are the functionalities that would support the activity(ies) of the recipients, and the actual development choices in terms of coding (see also Kellner et al 1999). Whilst this structuring of the process echoes literature on software development, it does it only in part. The developer's understanding and experience is helping in attaching to each phase its own specificities and challenges that were both very useful to understand in the context of the Ahead project.

- In the setting of the objectives, it was clear since the beginning the large mandate of the research team: develop an information system that could on the one hand facilitate and enhance the quality of the work conducted, generally by public administrations, to collect damage and loss data in the aftermath of a disaster. This generic request had to be nevertheless broken into the various sub-requirements that were discussed above, related to the cross-sectoral data collection, the need to consider different timelines, time and spatial scales. Being a research team and not an administration, a clear request was that once collected, data could be queried in a flexible and wide manner to provide added value to the data collection effort and to the collected data themselves.
- In defining what had to be developed in concrete terms, there are generally more tensions between developers and recipients of the product. In fact it is not always easy for recipients to translate the broad objectives into a set of functionalities, of services that the tool must provide in order to fulfill the mandate. For example the choice of a relational database is an important one that derives from the need to query data in multiple ways, keeping the relationships between the latter in terms of causality, dependency, time sequence. Consequently, the design of the logic data model is a crucial step that required the team interface to specialize in UML (unified Modeling language) to represent in the more abstract terms as possible the various entities and attributes and their relations.
- The many decisions necessary to actually develop the software that will provide the required functionalities and services relate to the "how" to actually implement the physical model, that is the translation of the logic model in specification about data and tables. The developer suggested that often recipients tend to step into the territory of developers suggesting the specific technologies and tools to use while asked to detail the "what" the software should provide in terms of functionalities and services.

According to the developer instead the how is the realm of the IT expert, even though domain experts can be brought to better understand the rationale behind certain choices.

The developer stressed that those steps are cyclic and not linear, that following implementation of design choices into a software then a re-evaluation of the adherence of the proposed software to the objectives and the expected functionalities is satisfactory or not. Such reiteration was also consequent to better understanding by the researchers of what is needed by administrations and stakeholders who actually collect the data and the imagination of further data uses and reuses that may be thought of and which, sometimes with relatively little effort, can be implemented.

The beginning of the collaboration was certainly more challenging, whilst with time collaboration has strengthened also facilitated by a more pragmatic approach that emerged also by researchers. At the beginning the presence of a highly expert data scientist with expertise in software engineering tended to impose very high standards on data quality, reliability, on the overall performance and on limitations to be imposed on data management. However, such limitations required an extensive effort and a high cost in terms of software development. A balance must be struck between correctness and reliability from an academic theoretical perspective and the pragmatic need to develop an operational working system.

The developer of the Ahead project replacing the former one who has been asked to collaborate full time with the European Commission on their risk data system (the Risk Data Hub) highlighted the effort required to shift from a pilot tool, the IS resulting from the Lode project, and a real product. This requires not only testing and debugging, but also rethinking some functionalities. In fact in the testing and evaluation process of the pilot one may discover that in two separate parts of the software contradicting functionalities exist. Those derive probably by the accumulation of requirements overtime, by the introduction of new fields and relations between data that were not part of the original design but added later upon specific needs. The transition to a final product requires actually both a vision of the entire system and its relation to the initial objectives and requests of functionalities, and the strict control on every part and component of the system. It also implies some simplifications, enhanced transparency of the structure of the system for others to be able in the future to change and adapt it as is the case of the Ahead system in the future.

An important step towards a product was the acquisition of a machine by the coordinator to guarantee longer survival and sustainability of the system. Transfer from the machine that constituted the initial environment for developing the pilot to the machine hosting a product with adequate levels of protection and differential accessibility levels imposed some important structural choices. In particular, the developer decided to move towards a configuration of the system as a set of micro-services that allow for more flexibility in the overall design and implementation.

The current system is made of five parts that are now running autonomously so that changes in one do not affect the others. Namely such parts are:

- the relational database itself as the set of established tables;
- the frontend software (the actual webpages that the user sees, with tables, buttons, menus...)
- the Webserver (an off-the-shelf software that just exposes the frontend to the public)
- the backend software that receives the calls from the frontend and reads/writes to the database.
- the Authorization Server (an off-the-shelf software that manages users, roles, permissions...)

The team member who acted as an interface between the researchers and the developers highlighted the importance of communication that was stressed also by the developers. The two ways communication consisted in long and deep discussions about what were the gaps in the current way disaster data are managed and in the best ways to fill this gap. Communication meant not only to align as for terminology and shared understanding of the problems associated with filling such gaps, but also in the shared process necessary to abstract from the real world into the logic data model though a process that is called in software development language “normalization”. “Denormalization”, meaning rollbacking relations into their original form was also practiced and useful in larger discussions with all members of the project consortium, some being unfamiliar with the technical aspects of IS design. Especially at the beginning researchers had to overcome the barrier implied in logic abstraction of data models. In fact a double barrier could be perceived: that of abstracting real world elements into objects manageable by a computer, and that of the language (UML) used to represent such abstraction. Both require exercising and getting acquainted to the underpinning methods. Such abstraction requires multiple decisions on how to standardize and synthetize best the multiple aspects of damage and loss data, starting from the need of double engineering and economic representation and using different parameters of description for different sectors.

The team member also stressed the importance of discussing also the technical choices being made among the various possible available solutions. This way researchers become more aware of the simplification that can be accepted in order to achieve a viable solution in reasonable time (and reasonable costs).

One conclusion that came from the focus group relates to the documentation of choices being made in the development of the IS: this is fundamental not only for others willing to intervene or modify the software itself, but also to make explicit the tacit knowledge that lies behind and may inform future similar efforts or improvement to the Ahead or other databases already in use by different stakeholders.

Conclusions

The deliverable has discussed the requirements elicitation of an enhanced IS for post disaster damage data collection and analysis as a complex process that actually started long ago with initial research efforts dating back to 2012. Initially the effort was more focused on understanding the challenges and the opportunities of disaster data, with time it was streamlined towards the development of an operational and functioning tool. The first observation that can be brought forward regards the time needed to fully unveil all the relevant aspects and issues implied by an apparently simple need: improving data collection and management in a certain environment. This proved to be a much more difficult and complex task than commonly though of for a variety of reasons. Here only technical aspects related to the development of the IS will be considered, referring to literature for considerations on the nature and complexity of damage assessment (Menoni et al, 2017; Molinari et al 2014). In a nutshell what emerged also through the focus group in the Ahead project, is that more than a product, the IS development proved to be an effort of knowledge co-creation. In fact the work that has been carried out stretches well beyond defining what is needed in a systematic manner, and become at the same time an effort to improve the understanding of damage to multiple sectors and their mechanisms of causality and interdependencies, and an effort to produce a logic abstraction that is not that easy in the disaster domain as stressed by some authors (Othman and Beydun 2013). The point of software development as of a knowledge development process is also highlighted in literature (Kelly 2015), even though the latter is facilitated by being conducted in a research setting rather than in a business, as time and costs considerations are more relaxed at least to a certain extent.

It has become very clear to the Ahead team that the development of the IS was much more than a technical effort to satisfy certain requirements and needs, it actually deeply involved knowledge sharing and co-production. On the one hand, researchers working on the IS design had to learn about relational databases development, about the steps needed to achieve the logic model abstraction from the real world of damage data, and the subsequent translation into the physical model to feed to the machine. On the other hand also developers had to learn about the specific problem of post disaster damage data collection and management, being involved in workshops, seminars, and through the lengthy discussion with stakeholders and the project partners. Developers also had to take a different stance on software development with respect to the same work conducted in a private company: perhaps more open to a variety of solutions and possibilities.

Such interplay between developers and domain experts, between required precision and reliability of data and pragmatic choices to be made are echoed by available literature. In a nutshell a number of aspects that have been reported were witnessed also in the Ahead IS development process, and namely: communication between developers and recipients as essential, availability to learn from each other, the effort of IS development as a cyclic one rather than a linear one through fixed steps requiring several iterations to achieve a satisfactory result. An iterative process that is not only demanding for software developer but also requiring domain experts to become more knowledgeable on data science and software development processes and techniques. It resembles more a mutual learning process rather than a traditional customer-product provider relation. This is in line with what some authors found on cross-domain collaboration that is needed for complex IS. The more knowledgeable are the two parts on each counterpart domain, the better is the collaboration the better is the result in terms of nuanced and tailored to the needs IT solutions. For IS to be deployed in complex settings and responding to articulated needs, the testing, such as the one that is carried out in Ahead, does not result in a simple evaluation of correspondence between initially set goals and provided functionalities. More often than not, while testing, new ideas arise on potential queries, on possible uses of the system that trigger new requirements and feasibility evaluation in terms of design and implementation. So that the design itself is incremental, determining even more the need of a re-evaluation of both the entire system to check for inconsistencies and contradictions and of the individual parts and components, especially those that have been added on. For the latter also the overall compatibility with the already developed system must be assessed and in case of need adjusted.

References

Carlile, P. R. (2004). Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization science*, *15*(5), 555-568.

Carvalho, T., Antunes, L., Costa Santos, C., & Moniz, N. (2025). Empowering open data sharing for social good: a privacy-aware approach. *Scientific Data*, *12*(1), 248.

Chakraborty, S., Sarker, S., & Sarker, S. (2010). An exploration into the process of requirements elicitation: A grounded approach. *Journal of the association for information systems*, *11*(4), 1.

Concilio G., Menoni, S., Norton J (2022) Innovation environments and risk management. Results from the KNOW4 European project. In De Lucia C., Borri D., Kubursi A., Khakee A. (Eds) Economics and Engineering of Unpredictable Events. Modelling, Planning and Policies, Routledge.

De Groeve, T., Poljansek, K., & Ehrlich, D. (2013). Recording Disaster Loss. Recommendations for a European approach. Luxembourg: Publications Office of the European Union. Joint Research Center. 76 pages.

De Groeve, T., K. Poljansek, D. Ehrlich, and C. Corbane (2014), Disaster loss data recording in EU member states. A comprehensive overview of current practice in EU Member States, JRC Scientific and Policy Report.

Di Ludovico, M., De Martino, G., Prota, A., Manfredi, G., & Dolce, M. (2022). Relationships between empirical damage and direct/indirect costs for the assessment of seismic loss scenarios. *Bulletin of Earthquake Engineering*, *20*(1), 229-254.

Ferrer, M. M., Do, A., Poljansek, K., & Vallés, A. C. (2018). Disaster damage and loss data for policy. *JRC Science for Policy Report*.

Faiella, A., Menoni, S., Boni, M. P., Panoutsopoulou, M., Thoma, T., Salari, S., & Rueda, N. (2022). Enabling knowledge through structured disaster damage & loss data management system. *Sustainability*, *14*(10), 6187.

Gibert, K., Horsburgh, J. S., Athanasiadis, I. N., & Holmes, G. (2018). Environmental data science. *Environmental Modelling & Software*, *106*, 4-12.

Hidalgo, M., Yanine, F., Paredes, R., Frez, J., & Solar, M. (2024). What Is the Process? A Metamodel of the Requirements Elicitation Process Derived from a Systematic Literature Review. *Processes*, *13*(1), 20.

Kelly, D. (2015). Scientific software development viewed as knowledge acquisition: Towards understanding the development of risk-averse scientific software. *Journal of Systems and Software*, *109*, 50-61.

Kellner, M. I., Madachy, R. J., & Raffo, D. M. (1999). Software process simulation modeling: why? what? how?. *Journal of Systems and Software*, *46*(2-3), 91-105.

Levitt, B. and March, J.G. (1996) Organizational learning". In Cohen, M.D. and Sproull, L.S. (Eds) Organizational Learning. Thousand Oaks, CA: Sage.

Menoni, S., Molinari, D., Ballio, F., Minucci, G., Mejri, O., Atun, F., ... & Pandolfo, C. (2016). Flood damage: a model for consistent, complete and multipurpose scenarios. Natural Hazards and Earth System Sciences, 16(12), 2783-2797.

Menoni, S., Atun, F., Molinari, D., Minucci, G., & Berni, N. (2017). Defining Complete Post-Flood Scenarios to Support Risk Mitigation Strategies. In Molinari, D., Menoni S., Ballio F. (Eds) *Flood Damage Survey and Assessment: New Insights from Research and Practice*, 151-171.

Molinari, D. , Ballio, F. , Handmer, J., & Menoni, S. (2014). On the modeling of significance for flood damage assessment. International Journal of Disaster Risk Reduction, 10, 381-391.

Nishikawa, M. S. (2003). Global Unique Disaster IDEntifier Number (GLIDE): For effective disaster information sharing and management. In Proceedings of the International Conference on Total Disaster Risk Management.

Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and leadership: a unified model of dynamic knowledge creation. Long range planning, 33(1), 5-34.

Ogie, R. I., & Verstaevel, N. (2020). Disaster informatics: An overview. *Progress in Disaster Science*, *7*, 100111.

Othman, S. H., & Beydoun, G. (2013). Model-driven disaster management. *Information & Management*, *50*(5), 218-228.

Park, S., & Kim, E. J. (2018). Fostering organizational learning through leadership and knowledge sharing. *Journal of knowledge management*, *22*(6), 1408-1423.

Mejri, O., & Pesaro, G. (2015). Knowledge based strategies for Disaster Risk Reduction: A knowledge management framework to increase understanding and awareness of value of prevention and preparedness. In Safety and Reliability of Complex Engineered Systems-Proceedings of the 25th European Safety and Reliability Conference, ESREL.

Pohl, K. (1996). *Requirements engineering: An overview*. Aachen: RWTH, Fachgruppe Informatik.

Polanyi, M. 1966. The Tacit Dimension, Routledge & Kegan Paul Ltd, London.

Rescher, N. (2009). *Ignorance:(On the wider implications of deficient knowledge)*. University of Pittsburgh Pre.

Shiomi, Y. (2023). Increase Disaster Data Linkages by GLIDE. In Proceedings of the 6th Global Summit of the Global Alliance of Disaster Research Institutes, Springer.

United Nations Office for Disaster Risk Reduction, & International Science Council. (2025). UNDRR–ISC Hazard Definition & Classification Review: 2025 Update of the Technical Report. https://doi.org/10.24948/2025.04 Geneva, Switzerland, United Nations Office for Disaster Risk Reduction, Paris, France, International Science Council.

United Nations Office for Disaster Risk Reduction (2023), Closing Climate and Disaster Data Gaps: New Challenges, New Thinking, United Nations Office for Disaster Risk Reduction

van de Ven, P. (2021) Defining infrastructure. SDD/CSSP/WPNA(2021)1/REV1. Statistics and Data Directorate Committee on Statistics and Statistical Policy, Organisation for Economic Co-operation and Development. https://one.oecd.org/document/SDD/CSSP/WPNA(2021)1/REV1/en/pdf

van der Voort, H., van Bulderen, S., Cunningham, S., & Janssen, M. (2021). Data science as knowledge creation a framework for synergies between data analysts and domain professionals. *Technological Forecasting and Social Change*, *173*, 121160.

Vitlov, N., Vuković, M., Bralić, N., & Marušić, A. (2025). Waste Not, Want Not: How to Make Your Data Futureproof Through Good Data Sharing Practices. *Current Protocols*, *5*(12), e70283.

Walia, A.; Menoni, S.; Farah Dell'Aringa, M.; Bakon, M.; Rovnak, M.; Milenov, K.; Dyulgerov, A.; Radkov, R.; Posthuma, L.; Shopov, T.; et al. Methodologies for Disaster Impact Assessmen. In EU JRC Science for DRM 2020: Acting Today, Protecting To-morrow; European Commission Joint Research Centre: Luxembourg, 2020; Chapter 3.1.

Annex 1. First Workshop with stakeholders

Kick off meeting of the AHEAD project	
Project Title	**Advanced disaster damage and loss data information system for enHancEd impAct-baseD knowledge (reference nr. 101193268)**
Project Acronym:	**AHEAD**
Purpose of Meeting	**Kick off**
Date of Meeting:	**January 23-24 2024**
Location:	**Politecnico di Milano, Piazza Leonardo da Vinci 32 & On air**

23rd January afternoon: Workshop with stakeholders

Time	Item	Speaker
14.30	Welcome and brief presentation of the Ahead project to stakeholders	Scira Menoni – Polimi
14.50	Tour de table	
15.00	The background of Ahead: past projects and an information system that we wish to enhance and complete	Anna Faiella and *Sandro Salari or Jacopo Bambini (to be decided)*
15.45	Coffee break	
16.00	Stakeholders' opinion and experience with existing information systems for damage and loss data and expectations on Ahead contribution to felt gaps/needs	Moderating: Reimund Schwarze

17.15	Concluding remarks and illustration of the future steps for engaging stakeholders to the Ahead project	Frédéric Tatout - AFPCNT
17.30	End of workshop	

Participants to the First Workshop with Stakeholders and Members of the Advisory Board

External Experts		
Dandoulaki	Miranda	Visiting at the Industrial Systems Institute, Japan
Berni	Nicola	Emilia Romagna Agency for Civil Protection and Territorial Safety
Fra Paleo	Urbano	University of Extremadura, Spain
Lopez Gunn	Elena	Icatalist, Spain
Sanchez	Maria José	BC3 Basque Centre for Climate Change, Spain
Noy	Ilan	Victoria University of Wellington, New Zealand
Paupério	Esmeralda	Ministry of Culture, Portugal
Vicario	Tiziana	Italian Civil Protection Authority

Ahead Stakeholders		
Dardi	Jader	Mayor Modigliana Municipality
Conversi	Stefano	Civil Protection Directorate, Lombardia Region
Roeslin	Samuel	European Commission – JRC - DRMKC

Annex 2. Results of the queries conducted on the Lombardia Region storm 2023 testbed

Source: Kantaria A., An advanced information system for managing disaster damage and loss data: the 2023 Lombardy storm case study, Thesis discussed in the CERM Program, October 2025

In total, 12 queries were performed, each executed for two measures: total cost and record count. The results were represented through graphs and diagrams. This chapter will address each of them separately.

- ***Queries N1 and N2****: Total estimated loss for public entities in Milan province; Total estimated loss for private entities in Milan province*

By separating the public (limited to schools and roads) and private sectors and observing the damages on both in terms of the number of reports and the cumulative economic loss, we can identify which sector was more heavily impacted by the disaster and use this information for resource allocation during the recovery phase of the DRM.

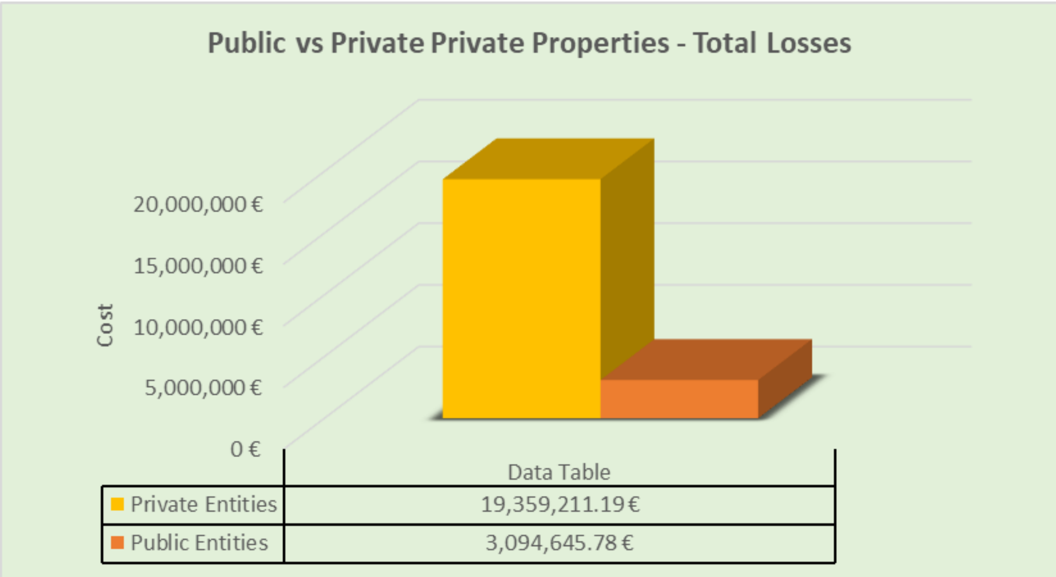


Figure 1 Public vs. private properties - Economic losses

Figure 1 shows that the total estimated monetary loss for private entities is more than six times greater than the loss from public entities. While the number of records reported to be damaged for private entities surpasses the number of recorded public assets more than three times (see Figure 2).

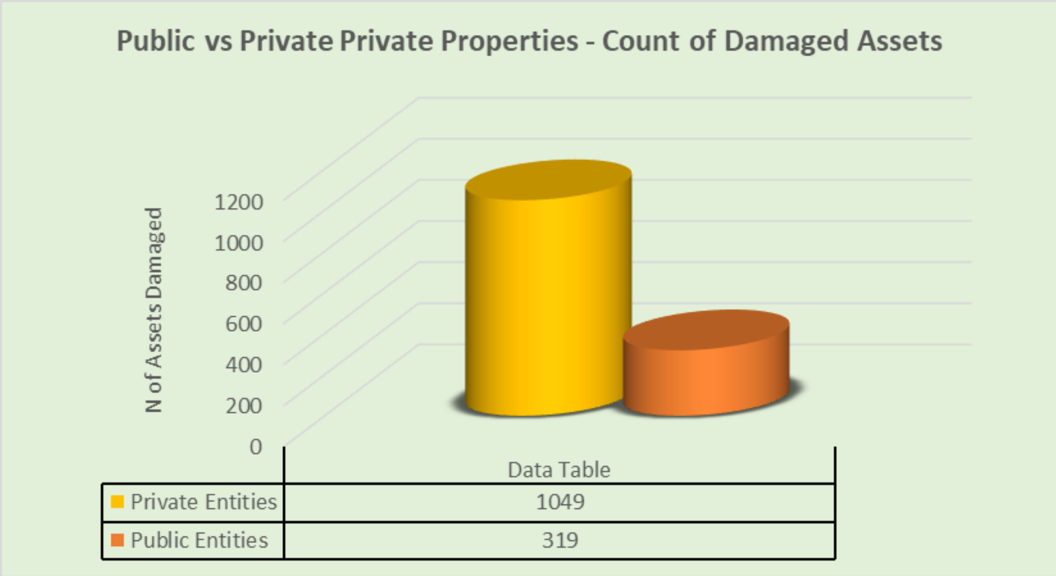


Figure 1 Public vs. private properties - count of damaged assets

- ***Query N3:** Buildings with structural, non-structural, and functional damages (cost division or count) in Milan province*

Observing the distribution between the different damage categories helps to understand the severity and the nature of the impact. Among the RASDA data regarding public assets, cases with damage to the structural elements of the buildings (mostly schools) were marked as structural damage. Damages to the transportation systems were mostly limited to non-structural damages such as debris accumulation and roadblocks due to uprooted vegetation. Non-structural damages were also present for the schools (such as clogged gutters, damaged fences, and rooms in need of cleaning). Damages in some schools were severe enough that they resulted in a long-term compromise of their functionality – these damages were marked as functional.

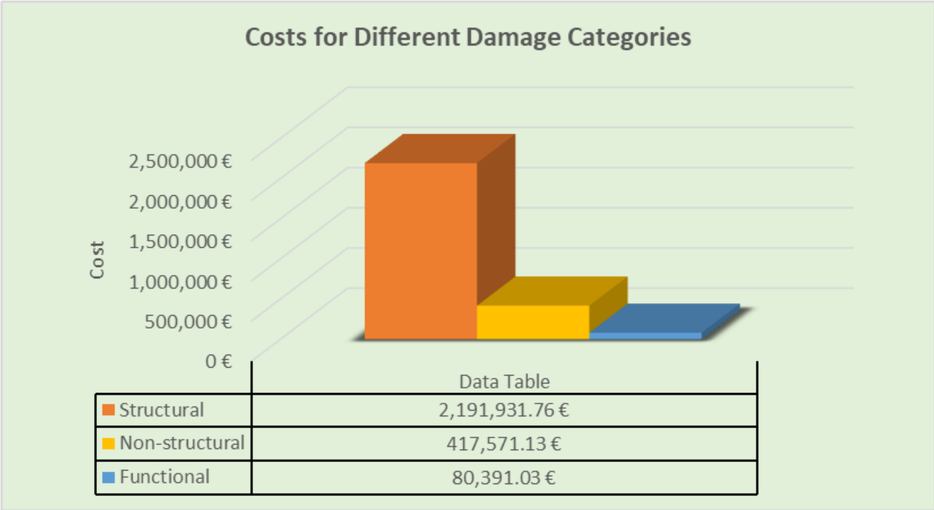


Figure 3 Cost for different damage categories

The ratio of cumulative costs for structural, non-structural, and functional damages is around 27:5:1 (see Figure 3); while the ratio for the number of reports for these damage categories is approximately 16:5:1 (see Figure 4). It is evident that within Milan Province, the structural damages had the highest occurrence.

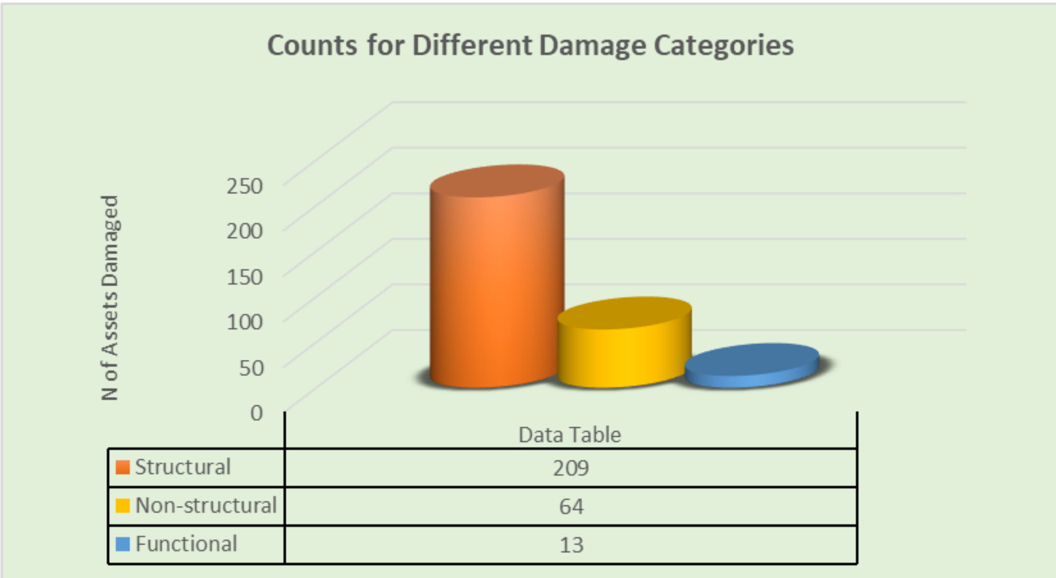


Figure 4 Count for different damage categories

- ***Query N4**: Total cost of repairs, replacements, vs temporary measures in Milan province for both private and public sectors*

To support the resource allocation during the response phase of the disaster, costs and counts of different types of interventions were evaluated. Existing data consisted of three typologies: Repairs (cost of repairs of private buildings), replacements (costs of the movable properties), and temporary measures (non-structural damages on public roads, and schools mostly encompassing debris removal and cleaning)

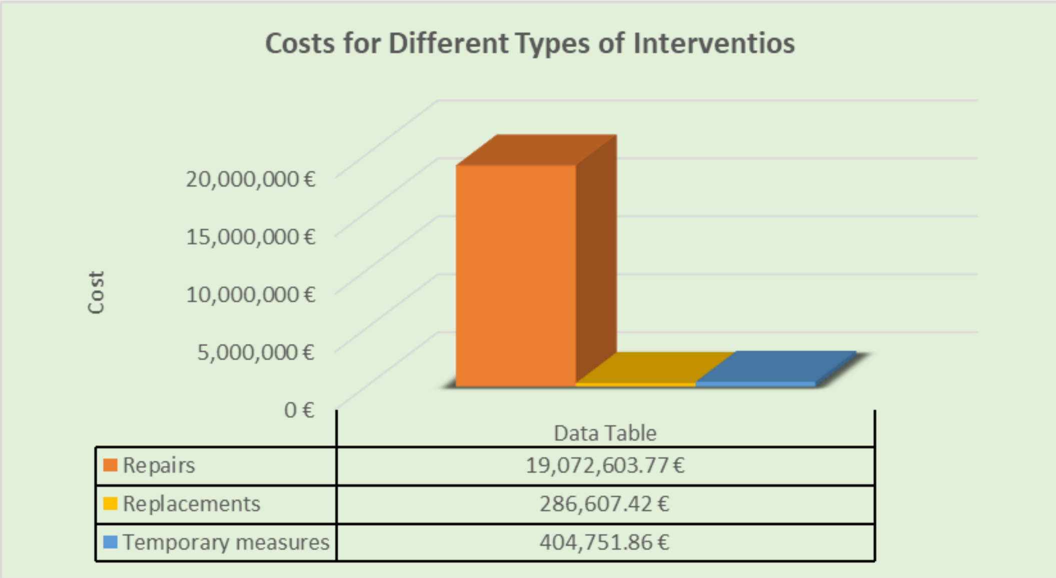


Figure 5 Costs for different types of interventions

Economic losses associated with different intervention typologies can serve as a basis for prioritizing funding and decision-making on recovery strategies. Cost ratio for repairs, replacements, and temporary measures is around 47:0.7:1 (see Figure 23). Ratio for the counts is around 32:3:1 (see Figure 5).

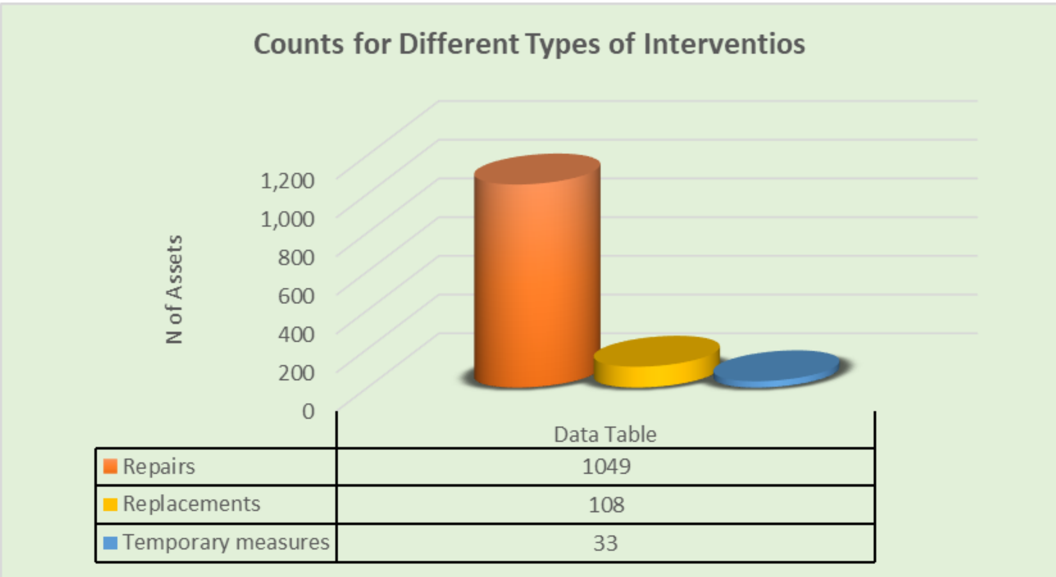
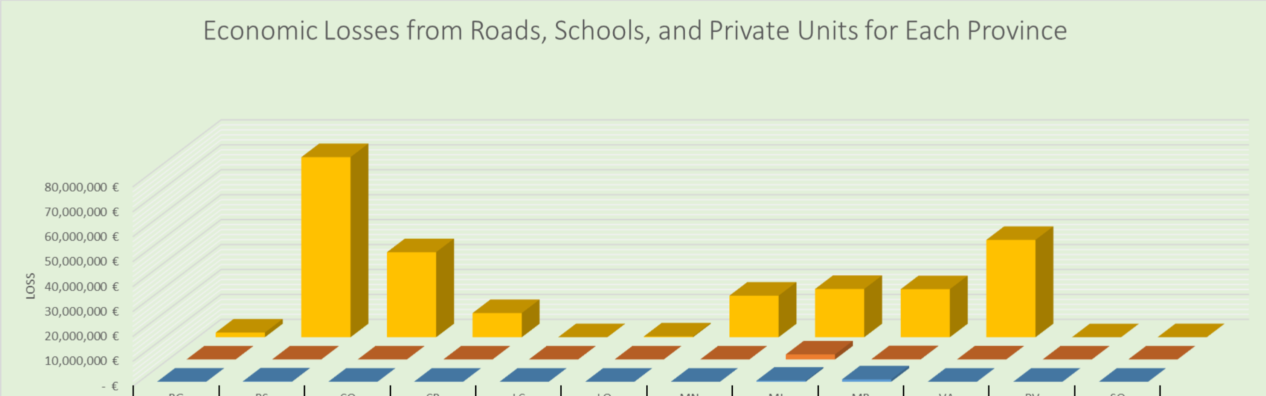


Figure 5 Counts for different types of interventions

- ***Query N5:*** *Economic costs (or count) per province divided for roads, schools, and private buildings*

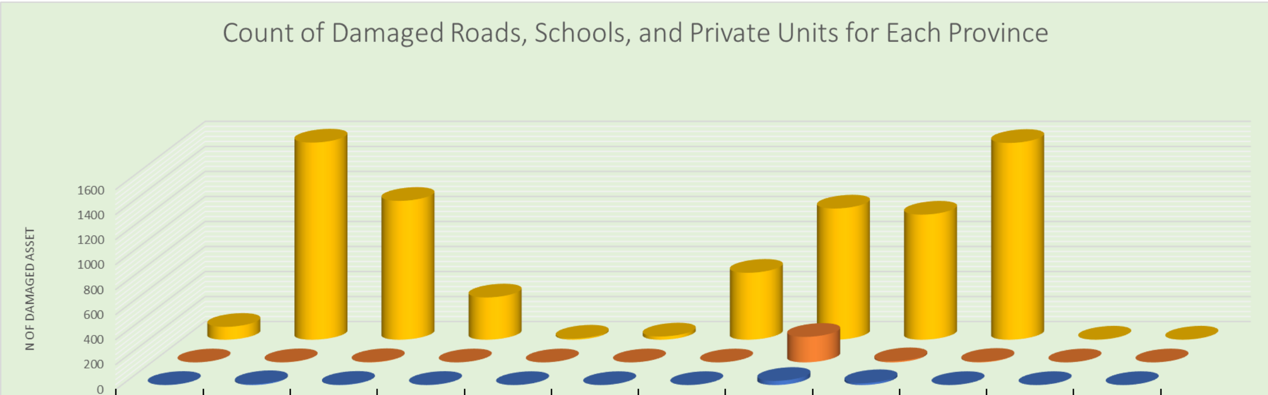
To plan for budget allocation for 12 impacted provinces and to evaluate damage distribution across three types of entities: schools, roads, and private units, the query N.5 was executed.



	BG	BS	CO	CR	LC	LO	MN	MI	MB	VA	PV	SO
Roads	23812.69	117641.4678	4879.6	3538	0	0	0	404751.8625	956437.9284	0	0	0
Schools	0	2584.8	2983.97	0	0	0	0	2016276.746	111842.34	32807.691	0	0
Private Units	1885515.84	72216301.2	34106154.08	9698298.96	26891.71	321983.68	16680949.2	19359211.19	19270884.67	39030938.27	500	13788

Figure 6 Economic losses from roads, schools, and private units for each province

Cost analysis identifies sectors with the largest financial impact, guiding funding priorities, while the number of reports for damage on each type of asset per municipality allows for better comprehension of the impact beyond economic losses, understanding of the physical damage on different asset types, and spatial distribution of the impact. The extent of damage, regardless of monetary value, is crucial for emergency management and service restoration planning.



	BG	BS	CO	CR	LC	LO	MN	MI	MB	VA	PV	SO
Roads	1	7	2	1	0	0	0	33	15	0	0	0
Schools	0	1	2	0	0	0	0	204	9	4	0	0
Private Units	103	1576	1110	338	9	26	535	1049	1000	1573	1	1

Figure 7 Counts of damaged roads, schools, and private units for each province

- ***Query N6:*** *Building management type: public vs private (economic cost or count) in Milan Province*

It is also useful to perform analysis for separate sectors; For example, the query N6 was performed to identify costs and counts of damaged assets only for the buildings sector.

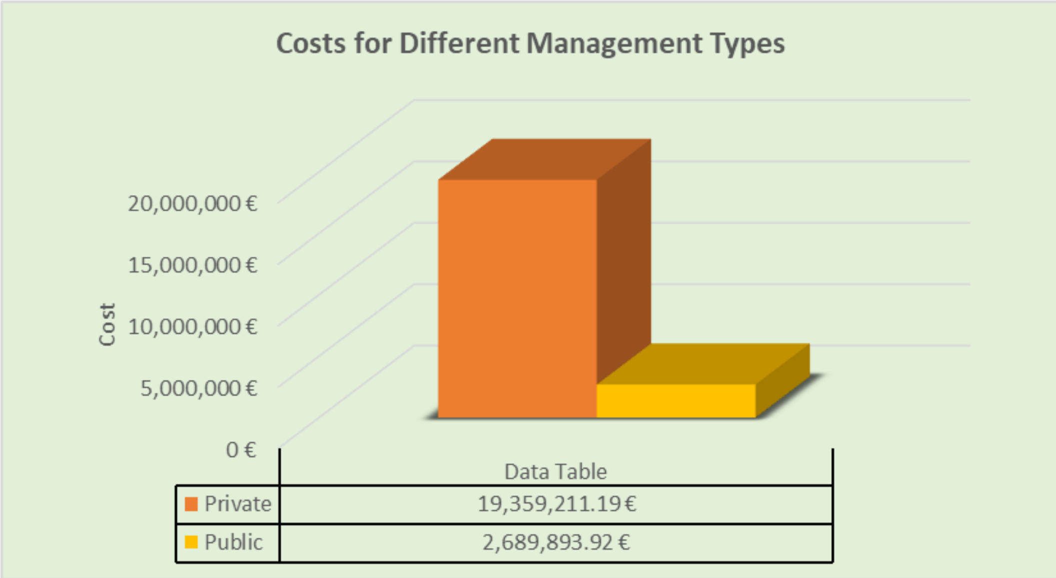


Figure 8 Costs of different management types for damaged buildings

The result shows that the economic costs from private buildings are more than seven times greater than those from public buildings (see Figure 8). In case of the amount of damaged assets, almost four times more private buildings were reported to be damaged compared to the public buildings (see Figure 9).

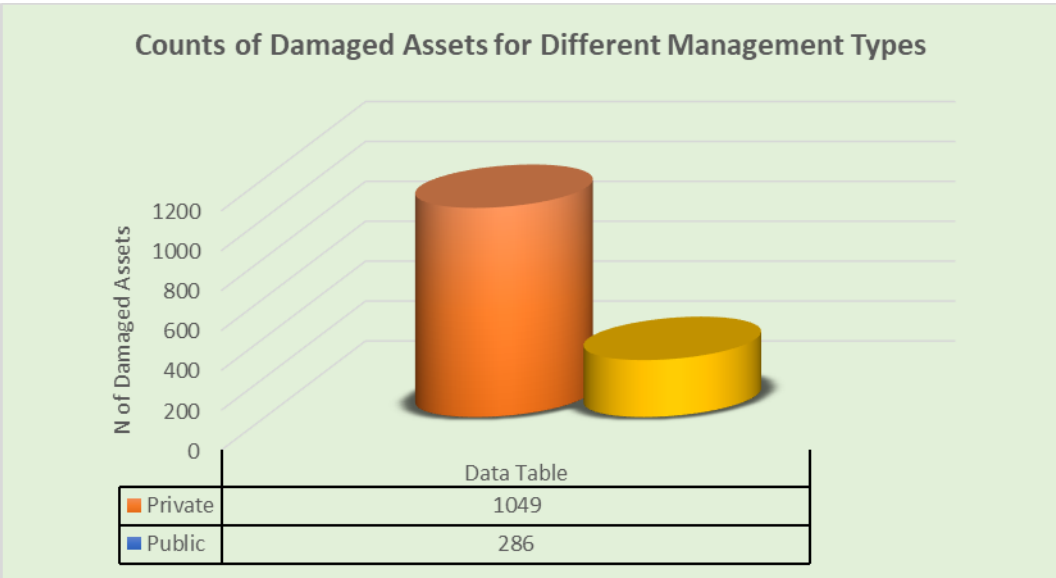


Figure 9 Counts of different management types of damaged buildings

- ***Query N7:*** *Economic costs (or count) of different classifications of structural damage for public entities in Milan province*

Damages of Public assets classified as "structural" were further subdivided into five categories: Horizontal Element (floors, slabs, beams), vertical element (columns, walls), infill walls (non-load bearing walls within the structural frame), roof, stairs, and other. The following graphs represent the breakdown of the costs and the number of reports associated with each of them.

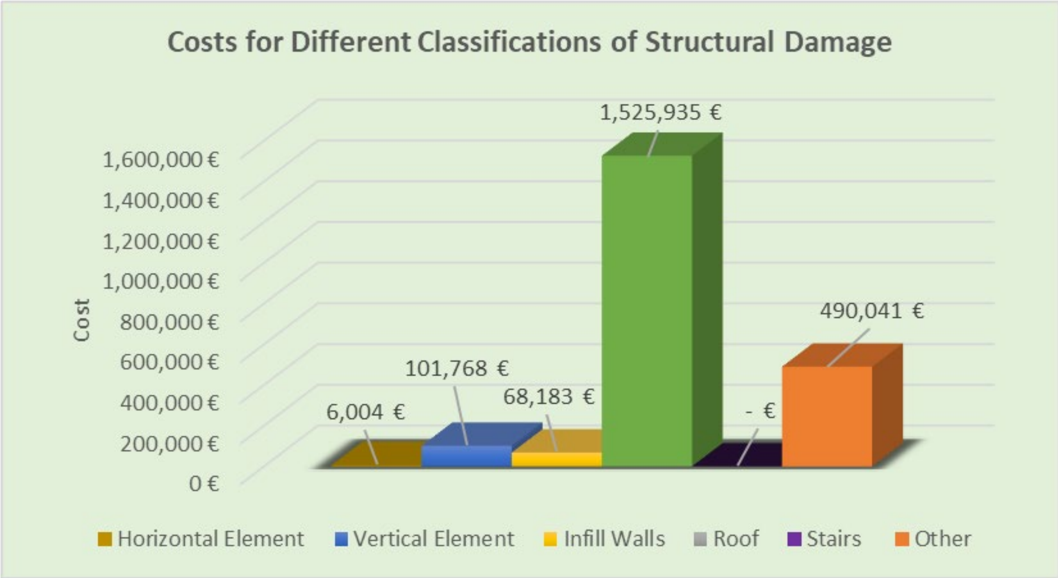


Figure 10 Costs for different classifications of structural damage

Damages to the buildings were overwhelmingly concentrated on the roofs, due to extensive rain, hailstorms, winds, or uprooted trees. There was also a significant amount of damage to windows and walls.

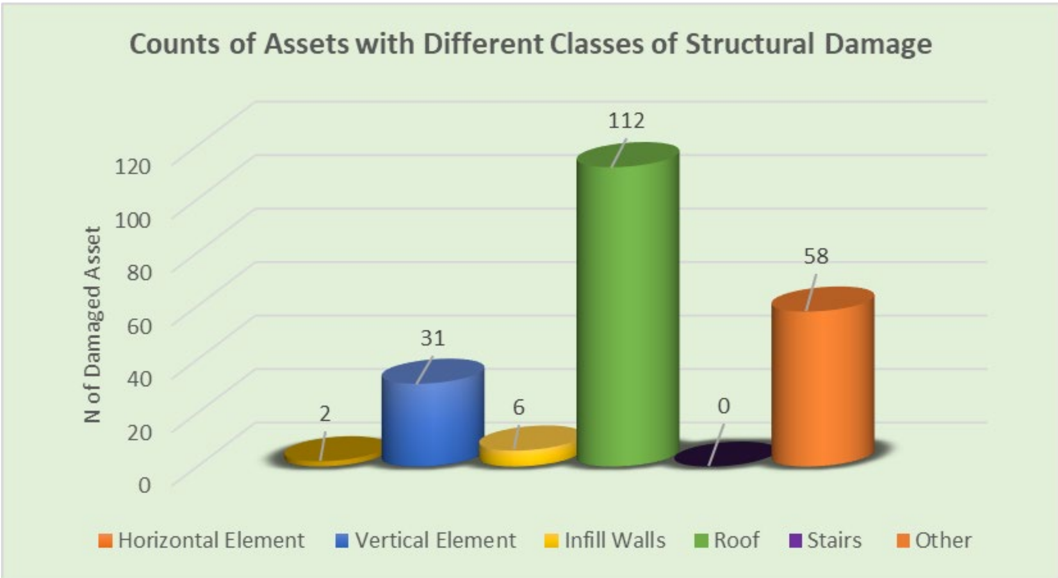


Figure 11 Counts of assets with different classes of structural damage

- ***Query N8:*** *Private Buildings with and without insurance in Milan province (and the total costs reimbursed by the insurance vs not reimbursed)*

Query N8 was performed to evaluate the costs to be reimbursed by the insurance companies and compare them with the cost of damages on uninsured private assets. Based on this analysis, we can evaluate the effectiveness of insurance in disaster recovery and identify gaps in coverage.

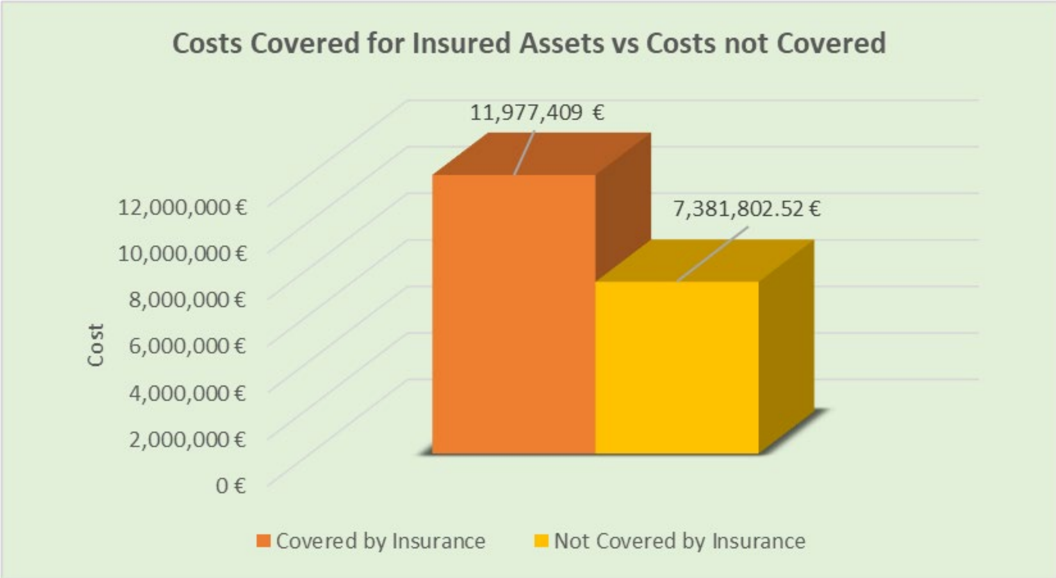


Figure 12 Insured vs. uninsured costs

Results present on Figure 12 and Figure 13 show that though most of the private properties were uninsured, the costs associated with uninsured assets are less than the insured ones.

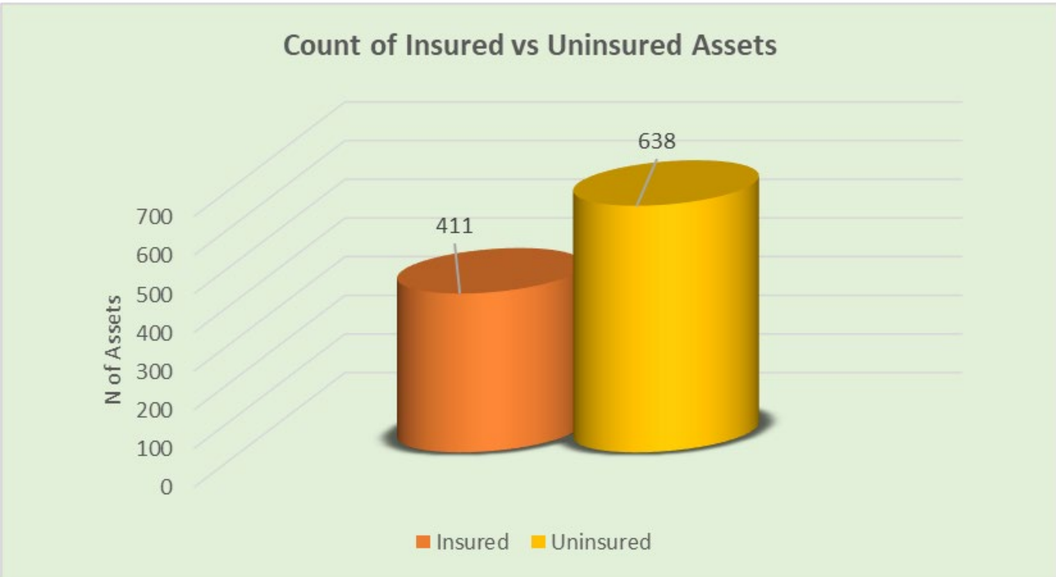


Figure 2 Counts of insured vs. uninsured assets

- ***Query N9:** Damages above a certain threshold of cost*

Data can be retrieved and visualised to display the costs and counts of damages within different cost ranges. For the testing purposes, random cost ranges were selected and the counts of data within those ranges were represented on a graph in numbers and percentages. Executing this query serves to visualize the extent of damages in recorded assets.

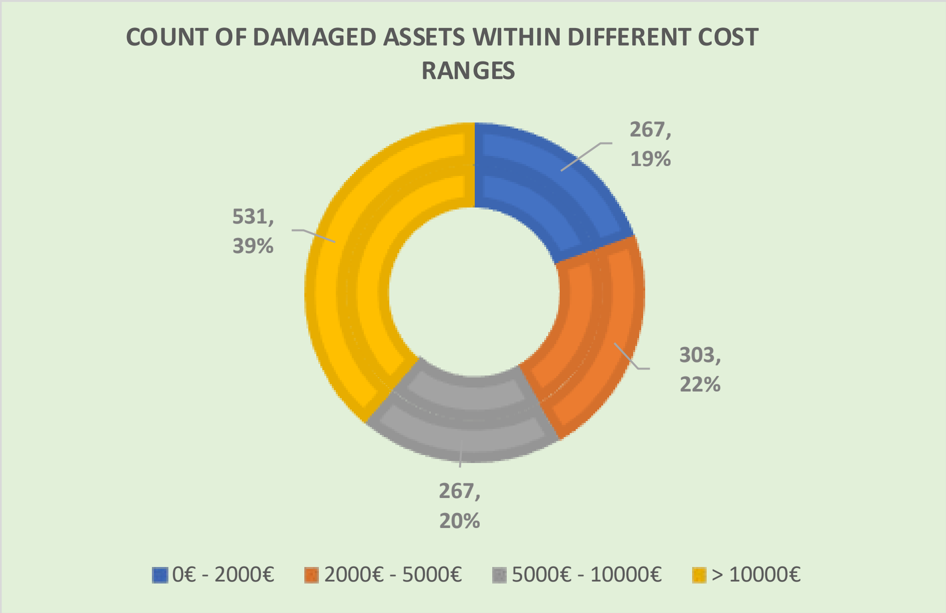


Figure 14 Count of damaged assets within different cost ranges

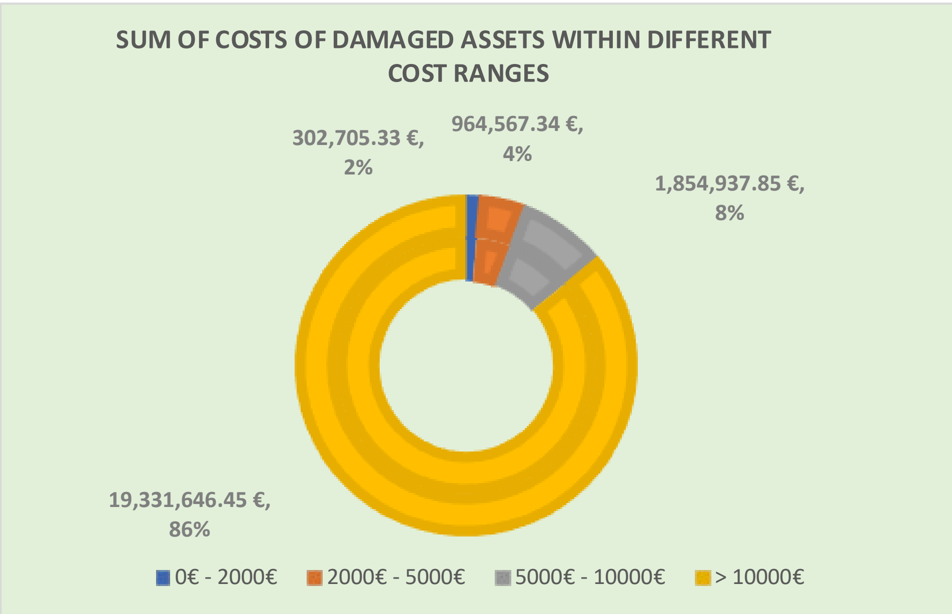


Figure 15 Sum of costs of damaged assets within different cost ranges

- ***Query N10:*** *List of owners of the private buildings to be reimbursed by the insurance above a certain threshold of cost*

Another useful query application can be to identify the owners (and associated contact information) that fit a defined criterion, in this case damage cost above 10000 euros. Please note that all names and private information (including VAT) are invented, they are not the real ones for FAIR treatment of data.

Codice Fiscale !Fabricated!	Name !Fabricated!	Surname !Fabricated!	Phone Number !Fabricated!	Email !Fabricated!	Partita IVA !Fabricated!
RSSMRA85A01H501Z	Marco	Rossi	+39 320 456 7890	marco.rossi@testmail.com	PIV5154801PRIV
VRDLGU93C15F205S	Giulia	Verdi	+39 328 987 6543	giulia.verdi@testmail.com	PIV5111315PRIV
BNCLRA75L09G273T	Lara	Bianchi	+39 331 245 6789	lara.bianchi@testmail.com	PIV5136621PRIV
FNTMRC88E22L219U	Marco	Fontana	+39 347 998 1234	marco.fontana@testmail.com	PIV5087127PRIV
GLLFNC90B12D325V	Francesco	Galli	+39 339 654 3210	francesco.galli@testmail.com	PIV5159512PRIV
MRNLGU92D30H703X	Luisa	Marino	+39 348 210 9876	luisa.marino@testmail.com	PIV5211015PRIV
RNLMLR94H11L781N	Mario	Ranalli	+39 340 123 7894	mario.ranalli@testmail.com	PIV5193816PRIV
BNZSLV81F03G902J	Silvia	Benazzi	+39 333 222 1110	silvia.benazzi@testmail.com	PIV5162544PRIV
VNTPLA89T21D969E	Paolo	Venturi	+39 327 100 2003	paolo.venturi@testmail.com	PIV5184172PRIV
MRRGNN86S18C351M	Ginevra	Marrone	+39 345 112 4578	ginevra.marrone@testmail.com	PIV5184532PRIV
FBRCLD91A07E472Y	Claudio	Fabbri	+39 329 873 4562	claudio.fabbri@testmail.com	PIV5189187PRIV

Figure 16 List of owners of Assets with damage above 10000 euros

- ***Query N11:*** *Temporal analysis of the surveying: count (or economic value) of entities recorded from the first, second, third, etc surveys*

As mentioned, many times in this paper, temporal dimension of disasters is one of the key areas of study that can improve the contemporary damage management scene. Data of damages on public assets provided by the Regione Lombardia consisted of records produced on three different dates: July 21st, 24th, and 25th, of 2023. Figure 17 and Figure 18 shows that most of them were recorded on the last survey date, while number and records collected on the first day was very low.

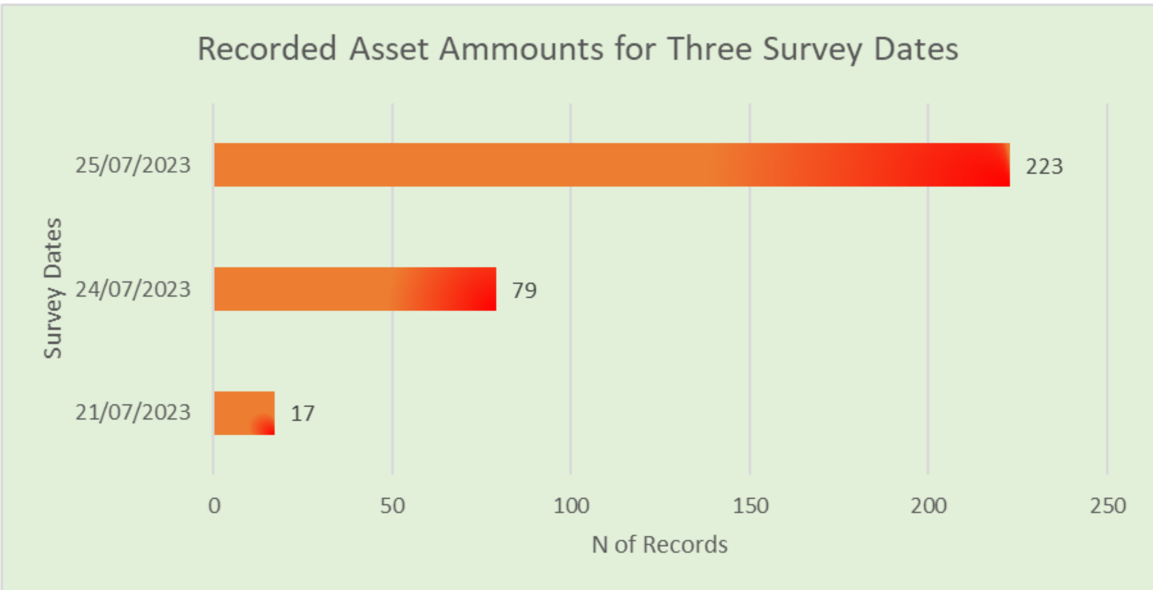


Figure 17 Recorded asset amounts for three survey dates

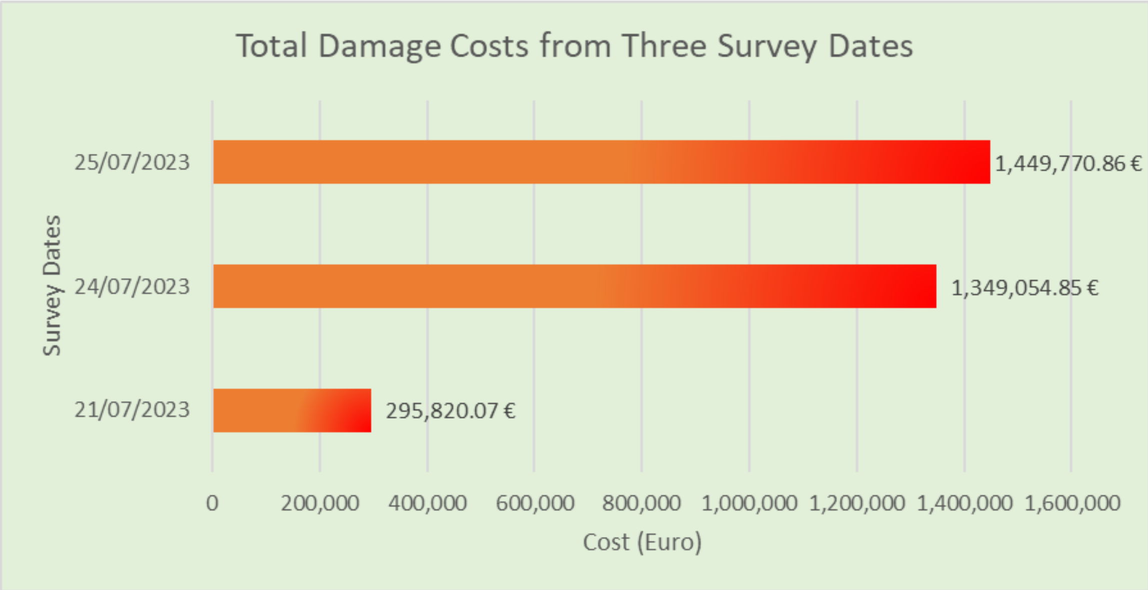
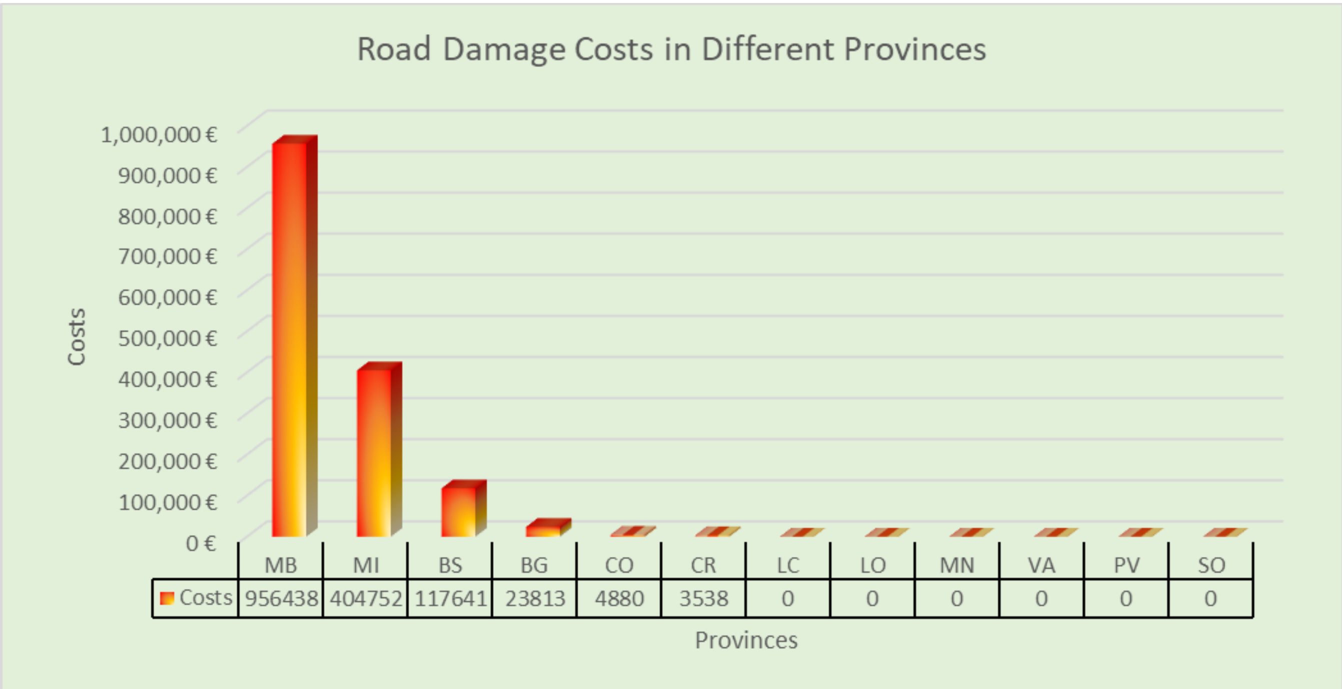


Figure 18 Total damage costs from three survey dates

- ***Query N12:*** *Provinces with the highest damage on roads*

Example of a query to identify economic losses associated with a certain sector (in this case roads) for each of 12 provinces. Figure 19 shows that the provinces with the highest economic damage on roads is Monza e Brianza, followed by Milan, Brescia, Bergamo, Como, and Cremona. There were no damaged roads recorded in the rest of the provinces, including Lecco, Lodi, Mantua, Varese, Pavia, and Sorano.



	MB	MI	BS	BG	CO	CR	LC	LO	MN	VA	PV	SO
Costs	956438	404752	117641	23813	4880	3538	0	0	0	0	0	0

Figure 19 Road damage costs for different provinces