LODE PROJECT

Layman Report of the Lode project

–

Deliverable n° 5.3

Delivery date 15.09.2021

Reference N° 826567

Loss Data Enhancement for DRR & CCA management

Project acronym	**LODE PROJECT**
Contract Number	**826567**

Editors: Remi Harris, Silvia Torresan, Elisa Furlan, Jaro Mysiak	
Partner in charge	**CMCC**

Contributions
All project partners

Versioning			
Version	**Date**	**Name**	**Organization**
Version 01	**20/08/2021**	Remi Harris, Silvia Torresan, Elisa Furlan, Jaro Mysiak	**CMCC**
Final Version	**09/09/2021**	Remi Harris, Silvia Torresan, Elisa Furlan, Jaro Mysiak	**Umbria Region**
Revision: Scira Menoni - Polimi			

Introduction to the LODE project

The LODE project (Loss Data Enhancement for DRR & CCA management) focused on the collection, storage, organization of post-disaster damage and loss data to support a variety of applications, ranging from accounting to forensic analysis of disasters to enhancement of risk modelling capacity.

LODE aspired to define with stakeholders what are enhanced data collection tools and procedures and how the knowledge that can be extracted from such data can serve a number of useful applications to support different policies and strategies for disaster risk reduction and climate change adaptation.

Ten partners were involved from seven countries (France, Spain, Finland, Greece, Serbia, Portugal, and Italy), representing scientific research centres and universities as well as public administrations active in diverse fields of risk management and mitigation.

The baseline of LODE is a set of ten showcases across Europe where damage data collection, storage and analysis were carried out following a methodology developed for the project. In each case, data applications were carried out in order to show in practice the added value of enhanced damage and loss data, and the utilities provided by an information system constructed within the project. This information system does not duplicate already existing tools, rather it takes into account the pros and cons of existing ones. The co-development of the system aimed to rationalize what stakeholders already do in the aftermath of a disaster, and to provide more and better opportunities to capitalize on the effort of data collection beyond compensation purposes.

Why is it necessary?

Data management is increasingly important in many different fields, covering all fields of policy and decision making, ranging from economics to the environment, to public health, as has been so clearly spotlighted by the Covid-19 pandemic. However, while this data is becoming increasingly available, in terms of volume, velocity, and variety, it still remains to be seen how to most effectively make use of this data. In anyone of these fields, once of the most important questions is how can we extract value from this data? This is a task supported by different branches of the EU, not least, as in this case, in the knowledge needed to improve approaches for DRR & CCA

There are significant pressures to improve the accuracy of data available for better decision making, however finding good data, the data one needs for a specific task, or on which to ground effective action, is rarely trivial, and the realm of disaster management is no exception.

Through the experience of the LODE project partners as stakeholders, practitioners and researchers in the field of DRR, a consistent concern is that the data available following a disaster event are rarely reliable, and cannot be grounded on the detailed and laborious collection that is required after a severe disaster for the estimation of needs, financial and material, for recovery and reconstruction. When available, their production takes time and they will then come out when the interest on the event has been overshadowed in the eyes of those that can effect the greatest change. The management of such data is also fragmented among the different responsible sectoral authorities and organisations, and so a comprehensive and coordinated assessment, including a comparative analysis across sectors and at appropriate spatial and temporal scales, is missing.

For many, it is clear that disaster loss data sets are of particularly high value, however their unavailability, restricted accessibility and the many limitations that the scientific community has highlighted in Lode as well as in past projects, have limited the large potential for use and reuse across the possible applications and services to which they could be implemented.

One of the main arguments against developing and using common and better organised tools relates to perception that this may constitute an extra burden to existing obligations challenged by the very limited dedicated staff. Having different interests and needs and no legal obligation for collecting data hinders the implementation of homogeneous and systematic collecting tools and the search for a common IT system, and having the same data evaluated in a different way by different stakeholders might well often lead to data duplication and misinterpretations.

Moreover, the actual situation is characterized not only by the fragmentation of data but also and especially by the fragmentation of responsibilities in which lies the strongest influence for data quality. Disaster data reporting varies in Europe and by sector. These data are part of an administrative, cultural and sectoral tradition specific to the States. They also depend on the prerogatives devolved to the public and private sectors.

Although building a system for disaster impact data collection, processing and archiving may be laborious and require wide co-operation, ultimately, the availability of good quality data and information services built on data offer large potential to reduce societal impacts and improve the preparedness and the resilience of the society. However, it has been observed that the high-resolution impact data required for the development of forecasts and tools are not collected or shared. It would be beneficial for several sectors to make their data available openly or at minimum, for research purposes. There is an increasing demand for policies and guidelines to promote data collection in many sectors and expand the data availability for research and development, and the potential benefits of the data collection and sharing are not yet fully understood.

What knowledge was extracted from the showcases?

The LODE project has a led a push for better systems of damage and loss data management through its series of pilot showcases covering a variety of disasters across Europe, at the same time demonstrating the wide-ranging possibilities for applying such data. Through this, there can be seen a clear added value in the accurate and precise collection of empirical data, and a plethora of opportunities for future research and improvements in data management.

In leading such a push for better systems, the showcases of the LODE project have demonstrated the benefits through two main streams of knowledge:

i. Enhanced use of data within the showcase applications, which can be invaluable for empirically based policies for risk management mitigation. Specific uses, tailored to the data available, and interests and expertise of the relevant project partners, ranged from forensic investigation of disasters, to the optimisation of risk assessment methods, the extraction of sector specific knowledge on damage mechanisms, appropriate data restructuring for disaster accounting purposes, and advanced modelling techniques that can capitalise on the increasing volume and velocity of big data that can be captured following a disaster event.

ii. An improved collaboration between public institutions, that allows for the digitisation of public administration, beyond simple tools such as the likes of Microsoft Excel, and the introduction of improved tools in this realm. These will inevitably lead to a reduction in human error, which was a major issue encountered across the showcases, as well as a more streamlined and interactive cooperation between the various stakeholders and data managers that could benefit from the effective evaluation of damage and loss data. Through practical and operational shared environments, it will become possible to use these datasets on repeat occasions and for multiple applications, such as those seen in the LODE project, with a usability across platforms and across sectors. The LODE project has accomplished a step forward by developing a tool that incorporates important and key requirements, conceptual as well as technical.

Enhanced data use

FMI: Using damage data to enhance early warning and emergency management

The Finnish showcase focused on the Tapani storm, where high winds caused severe damages and disruptions to multiple sectors over vast swathes of the country, initiating a societal pressure for action to better prevent such large scale and long-lasting disturbances to critical functions of society due to the extent of its impacts. Forensic investigation of the damages was undertaken, considering damages during the unfolding of the extreme event, in the immediate aftermath, and in a comprehensive ex-post evaluation. The data was collected into a practical online tool with maps, archives and impact graphs. The emergency data can be processed into statistical information, and information on economic loss was also extracted from physical damages. The impact database can provide a large variety of end-user solutions, for example in combination with weather data for use in impact forecasts, providing researchers with access to high quality impact information.

CMCC: Big data and new technologies

Detailed post-disaster damage data for the 2014 Secchia river flooding event was used for a GIS-based Bayesian Network (BN) approach, capable of modelling multi-sectoral flooding damages against future 'what-if' scenarios. The addition of the industrial and agricultural sectors gave a greater picture of multi flooding damages than that found in the current literature, which focuses almost exclusively on the residential sector. Access to a variety of hazard, exposure, and vulnerability indicators improved the understanding of the contributing factors to these damages (e.g. area of reported damages, land use, and flooding hazard), and their relative importance, highlighting those that should be prioritized in future data collection. A rising probability of large monetary damages was demonstrated under future scenarios of changing land use patterns, and increased flooding return periods. This Machine Learning approach allows for the integration of large, diverse heterogeneous datasets, capitalising on the availability of big data and making it possible to assimilate more information and expert knowledge within the applied risk assessment model.

What were the challenges and limitations faced?

The showcases of damage and disaster loss management applied across the case studies of the LODE project, gave a first-hand insight into the issues that are faced in terms of post disaster damage data collection across sectors, locations, and scales. The initial data collection proving to be more difficult than anticipated, with a variety of data types coming from different local, regional, or national authorities and stakeholders, there was an unexpected delay in the first project phases for most showcases, exacerbated by the negative effects of the Covid-19 pandemic.

For the damage data applications, data quality was paramount, particularly where advanced techniques were applied. Unfortunately, it was also found that a high proportion of data were low quality. Mistakes and subjective components of the data could also be found, as in the collection of data for Finnish power outages, to the point where the data may lose all value for application.

There was often significant effort needed to transform the collected data into a usable format, as it was provided in largely unstructured manners, requiring additional time to make them usable, for example on geolocalisation of data. Even after the extra effort required for this, the available data was also largely inconsistent across the different showcases, with damage data available for different sectors, or at disparate granularities, requiring most partners to adapt their processes to best adjust for the limitations. This was highlighted in the consideration of damages to cultural heritage, where different systems for classification of conservations status were found at national and various sub-national scales.

Further, it was often impossible to harmonise the results across the different showcases. The varied sources for the collected damage data across the showcases also highlighted concerns of data sensitivity, with those showcases using insurance claims or post-disaster inspections (e.g. the Kefalonia earthquake or the Tapani storm) needing to withhold portions of the data, limit publication of the results.

Irregularites and lack of coordination between pre- and post-disaster damage data were also noted. For example, for the Lorca earthquake, it was found to be challenging to rely on pre-event risk assessments for predicting observed damages. In Serbia, forest fire data has mainly been collected for administrative and statistical purposes, with a heterogeneity in the collected data limiting the potential for coordination between prevention and disaster management.

The main thread tying these issues together was a lack of coordination. Improved coordination among the various actors, from data collection to compilation to storage, would have brought significant added value to each application. These processes varied massively, with the quality of data application for risk prevention or management at best mirroring the quality of the data collected.

What is the way forward?

Implementing a new information system for disaster loss data management involves a technological change in the way data is recorded, processed and curated. But introducing a technological change into a governmental institution is likely to present a series of challenges to the executives enforcing such a change and to the users of the new technology.

Can we lead a new wave of data collection?

Usually, in disaster management, the data collection is specifically tailored to the specific applications needed to specifically fulfil their tasks in different phases. The work only goes on until the task is over with very little critical reviews further on and/or no updates of the data.

At present, overcoming these difficulties clearly apparent in most showcases would need special attention to account for specific institutional contexts at a national or regional level. The situation is made even more complex due to the fragmentation of data among many different stakeholders, so that even in case a common information system were available, it would be very difficult to populate it. While these issues can prove challenging, they also open up room for better practice, and further development of the vital systems of information

In order to lead a successful wave of improved data management that will be taken up by policy makers, the effort of data collection must be somehow “rewarded” and recognized by stakeholders as bringing key advantages. As such, the following uses have been identified and explored:

a. Compensate damaged parties proportionately to the extent of their losses, discriminating between insured and uninsured ones.
b. Allow designers of structures to better size them in view of the intensity and frequency of the damaging events to which they may be exposed
c. Design of policies and strategies for governance and preventive management of disaster risks both before the occurrence of an event and in its aftermath
d. Increase knowledge about disasters and risks, including a better understanding of the comprehensive impact on communities, the total economic loss and the comparative overview of loss of profit among different sectors
e. The scaling of warning and preparedness in emergency and crisis situations.

In turn and considering these uses, we were able to identify five categories of data misuse presented below, which must be challenged in order to lead a successful wave of data management

a. The aggregation of data of different nature, unit and declarative size
b. The inference of a probability of occurrence on the basis of a very weak or even incomplete sampling of events
c. The characterization of an estimated amount of the cost of a disaster neglecting the cost of living in the countries, the nature of the declarative system of the listed damages as well as the distribution of the declarations of data between the public and the private sector
d. The erasure of the specificities of the damaging phenomenon and of the evolution of urbanization in the declaration of the monetary magnitude of the disaster

The hypertrophy of data declarations vocalizing mainly on accidental and acute effects to the detriment of chronic and diffuse effects of disasters

It is clear that to improve, there needs to be increased collaboration between sectors, there is already increased awareness of this from the current health crisis in terms of the digital platforms that are necessary to appropriately deal with emergency situations.

The next steps of the LODE project are to further engineer the prototype database, capitalising on its functionalities and potential for wider scale implementation within Disaster Risk Reduction strategies, with the intention of putting the successful work into real world practice, potentially in collaboration with other platforms for shared data. Further developments will be necessary to deal with the implications of multiple risks from different sources that can act in parallel and in tandem, and also from recurrent and lingering disasters. The LODE project has accomplished a step forward by developing a tool that incorporates important and key requirements, conceptual as well as technical. Yet to make it a tool used in practical situations, some further steps must be accomplished, including more longitudinal disasters studies and pilots to dig deeper into the association of available data with disaster impacts throughout the entire disaster cycle. The analysis of disaster impacts (and related information and data) should not be confined within the limits of "the physical outcomes", but should involve relief and recovery policies, and how these alter the pre-disaster conditions and causing an array of social and economic effects on communities and spatial settings.

In order to develop adequate strategies that will facilitate the required transition to a modernised approach, as well as to ensure the long-term sustainability and use of the proposed information system, challenges at both an institutional and a technological scale must be envisaged and acted upon. Institutionally, as highlighted by the work of the LODE project, there remain issues in terms of knowledge gaps and misconceptions on the benefits of such a system, coupled with a resistance from users to adapt to the changes necessary, to the perceived perplexity of the system, or to accept an additional workload that it may imply. Other concerns with the data, from reliability of the system to privacy and sensitivity concerns, must also be addressed. On a technical level, software related challenges come about through the likes of georeferencing methods, baseline data, and compatibility with existing datasets. These issues are exacerbated when considering the technical skills that may be necessary to operate high functioning systems, particularly when it comes to more advanced techniques and models. These challenges must therefore be tackled head on in any future developments, as only by doing so will it be possible to fully implement the most meaningful change in disaster and loss damage data practices.

The LODE information system

The LODE information system is a Database Management System designed to be multi-purpose, dynamic and flexible by comprising standardization, a result of many years of interactions between researchers, stakeholders and developers culminated in the tight development process within the Lode project lifetime.

The tool has been developed to respond to the need of an improved damage and loss data collection to account for damages according to a precise analysis of current practices, taxonomy and terminologies in use, and proper consideration of temporal and spatial scales. It has been designed in order to allow for an accurate, consistent and transparent collection of damage & loss data. The proposed tool takes into consideration all the phases of damage and loss collection recording, storing, managing, maintaining an up-to-date documentation of the database, and performing queries to retrieve information.

The model has been designed according to three main requirements:

i. Flexibility and standardization: adaptable to different societal sectors and their specific characteristics but creating a structure that allows to collect information in a standardized fashion
ii. Spatial and temporal characteristics: information is collected at the asset level, however it can be queried and filtered at different scales
iii. Damage causality and dependency: model features embracing the inter- and intra-connection of damages, within the same sector or due to the dependency of others, tracking the causality

In turn the system functionality comprises three main features:

i. Data collection: through the LODE system, data are collected precisely eliminating redundancy and inconsistency, and delineated sector by sector with the objective to serve different purposes. Those data are organized through a relational database which consists in a collection of tables that store interrelated data.
ii. Data storage: A relational database management system allows to store and retrieve data represented in tables through different types of queries elaborated in advance to support a range of different purposes and multiple objectives, such as risk assessment curation, the understanding of damage mechanisms, the delineation of real and robust trends, identification of priorities etc.
iii. Data management: while most damage assessments are carried out manually filling pre-compiled forms with data that lose their interrelation, or through datasets that function only as a static archive of data collected from different and heterogeneous sources; a relational database system anticipates the use of data for multiple objectives allowing to properly collect them and to store large amount of data, permitting efficient search performance through prefigurated queries. The use of a well-designed database management system allows to store data directly into the archive and offers the opportunity to have timely available data and remote accessibility to the information.

The LODE information system provides significant added value over the classic paper-forms, by allowing the creation of a comprehensive inventory of georeferenced information, that stores large amounts of data in an integrated and coordinated manner, due to the predefined relationships contained within the data, which guarantees consistency in data collection, and permits efficient search performance through a wide range of prefigurated queries. In addition, it allows rapid and more accurate filling of documentation and sharing of big damage and loss data, through the potential of such an information system to facilitate a fast flow of information and instantaneous connection of numerous organizations across wide geographic areas.

To be informed and updated about **LODE**'s activities access:

www.lodeproject.polimi.it