Looking Beyond the Data: an Assessment of the Emerging Data Ecosystem of Nepal's Flood Early Warning Systems

Stella van Esch
Vrije Universiteit Amsterdam
s.r.van.esch@student.vu.nl

Marc van den Homberg
510 – an initiative of The Netherlands Red Cross
mvandenhomberg@redcross.nl

Kees Boersma
Vrije Universiteit Amsterdam
f.k.boersma@vu.nl

ABSTRACT

Increasingly, data-driven instruments are used in disaster risk reduction to foster more efficient, effective, and evidence-based decision-making. This data revolution brings along opportunities and challenges, which are sometimes related to the data itself, but more often seem related to the environment in which the data is put to use. To provide insight into such an emerging data ecosystem, this paper uses a qualitative case study to assess the use of data in flood early warning systems (EWS) in Nepal. In response to the research question 'How does the data ecosystem impact the opportunities and challenges regarding data use in flood early warning systems in Nepal?', this paper discusses the importance of considering the broader context instead of regarding data as an entity unto itself. It shows how actors, policies and other contextual factors impact the effectiveness of data use by either presenting opportunities, like the establishment of a national disaster data repository, or challenges, like inadequate human resources for working with data.

Keywords

Floods, early warning systems, Nepal, data ecosystem, social shaping of technology.

INTRODUCTION

Floods are among the earth's most common and most destructive natural hazards. More than 250 million people are affected by floods every year, leading to losses of resources, incomes, homes, and lives. Due to climate change, both the amount and intensity of floods is only set to rise (Winsemius et al., 2015; Zurich Flood Resilience Alliance, 2019). Early warning systems (EWS) have been introduced as an instrument to be better prepared for floods and to reduce the losses of lives and properties (UNDP, 2018). EWS can be defined as the 'set of capacities needed to generate and disseminate timely and meaningful warning information to enable individuals, communities and organizations threatened by a hazard to prepare and to act appropriately and insufficient time to reduce the possibility of harm and loss' (UNISDR, 2012).

In the development of EWS, many countries and organizations have shifted towards using more data-driven and high-tech tools as this would foster more efficient, effective, and evidence-based decision-making, and can increase the lead time to prepare and anticipate to floods (Budimir et al., 2020; Haak et al., 2018). Both in their design and function, EWS increasingly depend on computing power (Budimir et al., 2020). This shift towards using more small and big data in disaster risk reduction is often discussed as part of the broader 'data revolution' (United Nations, 2014). Although this data revolution brings along great advantages, particularly in terms of more up-to-date, validated an reliable information, it is crucial to not regard data as an entity unto itself, but rather as a component of a broader data ecosystem in which it interacts with the technological, institutional and social context (Mulder et al., 2016; Van den Homberg & Susha, 2018).

Especially in flood-prone countries like Nepal, it is important to have a well-working EWS and make optimal use of data to establish this. To assess the opportunities and challenges regarding enhancing the use of data in EWS, it is important to have a clear overview of the data ecosystem, in which all relevant actors, infrastructures, characteristics and policies are mapped. However, despite Nepal's elaborate disaster risk reduction strategies and its increased use of data to fulfil its ambitions, the country lacks an overview of its flood EWS data ecosystem, and therewith insights in opportunities and challenges regarding the data use. This research, therefore, provides an answer to the research question: 'How does the data ecosystem impact the opportunities and challenges regarding data use in flood early warning systems in Nepal?'. By answering this research question, this paper underlines the importance of considering the broader context in which data is applied, and shows how a broad range of societal, technological and institutional factors impact the opportunities and challenges regarding data use in flood EWS in Nepal. The insights of this paper go beyond the particular case of Nepal, and are aimed to provoke further and broader research on this topic. Moreover, this research aims to provide concrete tools for practitioners in Nepal to optimize the data use regarding flood EWS.

The following section sets the framework for the analysis by discussing the relevant literature regarding the datafication of early warning systems and the importance of context when working with data. Thereafter, the methods and research design are presented, followed by a discussion of the most important outcomes, conclusions and recommendations.

BACKGROUND

The Datafication of Early Warning Systems

Early warning systems (EWS) are increasingly recognized in both international and national disaster risk reduction strategies as an instrument to become more resilient to floods and other disasters (UNDRR, 2019). EWS are systems used to reduce the impact of natural hazards by providing timely and relevant information systematically (UNDP, 2018). Although EWS vary strongly in terms of use, purpose, complexity, and scope, scientists, policymakers, and practitioners agree that complete early warning systems consist of four main elements: risk knowledge; monitoring and warning; dissemination and communication; and a response capability (e.g. Basher, 2006; UNISDR, 2012; IFRC, 2012).

Herein, risk knowledge concerns the information about the combination of hazards and vulnerabilities at a particular location. Disaster risk knowledge is based on the systematic collection of data and disaster risk assessments. Monitoring and warning refers to the system and technology for monitoring and detecting hazards in (near) real time and providing forecasts and warnings (WMO, 2018). Different kinds of information and variables are collected in order to determine risks and, if necessary, send out warnings to people exposed to the risks. Closely integrated with this monitoring and warning element, is the dissemination and communication component. This concerns the distribution of understandable warnings and preparedness information to those at risk of a hazard. Dissemination and communication include an official source of authoritative, timely, accurate, and actionable warnings and associated information on likelihood and impact. The response capability element includes the existence of management plans that contain knowledge, responsibilities, and resources needed for timely and appropriate action in the event of a flood. Preparedness at all levels (from individual households up to national level) is necessary to respond to the warnings received. The overall aim of these four components of a flood EWS is to work seamlessly end-to-end (WMO, 2018) and reach the so-called last mile. Or in other words: in an effective EWS, people to be affected by the floods (and the organizations supporting these people) should receive, understand and act upon the early warning in the lead-up to the flood event. By fostering early action, the EWS helps to protect lives and livelihoods and reduce economic and material losses (UNDP, 2018, Kumar Rai, 2020).

Decision-making based on up-to-date, validated, and reliable information is often seen as a main challenge in working with EWS, due to stress, time pressure, and uncertain information (Comes, 2016; Haak et al., 2018). Several studies have suggested that using data can help shift towards more efficient, more effective, and more evidence-based decision-making in this context (Benini, 2015; Haak et al., 2018). Therefore, small and big data are increasingly combined to improve the functioning of flood EWS.

In the risk assessment, small data, such as data collected through a national census or a survey, is increasingly supplemented by big data to enrich the information available. Big data enables the identification of infrastructure and the built environment on satellite maps, either through crowdsourcing initiatives such as volunteered geographic information, the help of experts, or artificial intelligence. This makes the exposure dimension of risks more accurate.

Regarding monitoring and warning, the detection and monitoring of hazards is facilitated by several types of ground-based monitoring stations, such as flood gauges in a river. For example, Open Source Electronics and 3D printing, cheaper monitoring stations can be developed, allowing deployment of a denser network of sensors. These can then form an Internet of Things, a network of connected embedded objects or devices, with identifiers, in which a communication using standard communication protocol can happen without human intervention.

Several technologies support dissemination and communication and these generate more and more big data streams. For example, early warning messages can be sent via Social Media (such as WhatsApp groups) or mobile services (hotlines, SMS or voice SMS (to also reach illiterate people)). Affected people can text or call specific flood-related hotlines to express their needs and send valuable locally collected data, such as water levels. More in general, digital inclusion is an important enabler for effective communication. Key EWS stakeholders can more easily extend their reach into communities through digital channels, either by directly engaging with them or by monitoring and analysing user-generated content. The smartphone allows more extensive bidirectional communication and engagement between to be affected communities and responding professionals.

Regarding the response capability, big data can for example be used for the identification of evacuation routes, demarcation of safer areas and location of temporary shelters through satellite imagery. Big data can also be used to assess whether public awareness and education campaigns are conducted and to evaluate their impact.

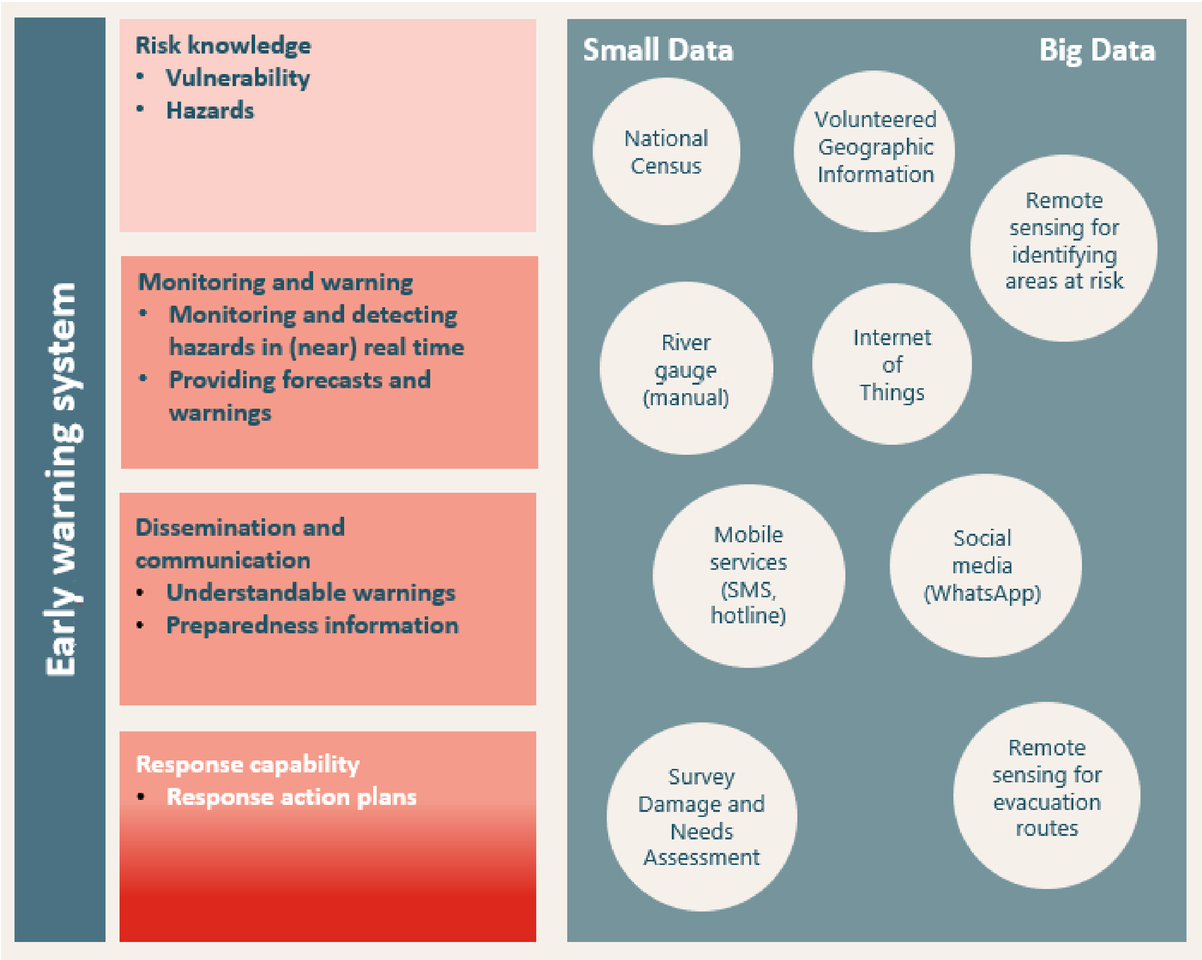

Figure 1. A visualisation of the applications of small and big data in EWS

The Importance of Context

The increased datafication of EWS thus provides many opportunities for disaster risk reduction both in terms of applications and innovations. However, although (big) data offers many possibilities for the development of flood EWS, it is important to consider the context in which the data is put to use. As Mulder et al. (2016) explain, all kinds of data are socially constructed, and can thus not be seen as fully neutral. This can have far-reaching social implications, for example, when data-driven technologies used in disaster risk reduction reinforce pre-existing patterns of socio-economic inequalities. These risks do not directly follow from data itself, but rather from how the data is put to use, and how the data interacts with the context, it is used in.

Indeed, there is by now a broad body of research that underlines the importance of approaching data and technologies not as isolated entities but to consider them in the broader context in which they are implemented. This involves, among others, the literature on the social shaping of technology (SST) (e.g. Baym, 2010, MacKenzie and Wajcman, 1999; Rip & Kemp, 1998). Rip & Kemp (1998) explain that 'technology is shaped by social, economic, and political forces, and […] in the same process, technologies and technology systems shape human relations and societies' (p. 328). Some technologists tend to see their environment only in terms of opportunities and constraints for the technology or technological system. However, it is important to acknowledge that the social environment has its own dynamics, and it has already shaped the opportunities for, as well as the ideas about the technological system (Rip & Kemp, 1998). It is, therefore, necessary to take structural aspects of the environment, existing systems and the sociotechnical landscape into account when one wants to optimize the use of data or technologies. Hence, it is important to see data as a component of a broader data ecosystem in which it interacts with the technological, institutional and societal context (Mulder et al., 2016; Van den Homberg & Susha, 2018; Bierens et al., 2020).

Defining a Data Ecosystem

A data ecosystem can be defined as 'the people and technologies collecting, handling and using the data and the interactions between them' (Parsons et al., 2011, p. 557). As any other ecosystem, data ecosystems can be understood as 'a system of people, practices, values, and technologies in a particular local environment' (Nardi & O'Day, 1999). Van den Homberg & Susha (2018) argue that it is important to consider the data ecosystem to realize the full potential of data. They created a framework, the Integrated Data Ecosystem Framework, which can be used to characterize data ecosystems. This framework is structured around five dimensions: actors and roles; data supply; data infrastructure; data demand; and data ecosystem governance. Van den Homberg & Susha (2018) used a Political Economy Analysis (PEA) to show how resources, actors, processes, and mechanisms result in output and outcome for their case study around water governance at different global, national, and local levels. Against this backdrop, they showed how the data ecosystem framework can be used for reporting on Sustainable Development Goals around water governance. However, the PEA was only loosely linked to the data ecosystem framework. In this research, we supplement the five elements by a sixth dimension, which concerns the societal context. This element is added to underline the importance of the social shaping of technology. A description of the six elements is provided in table 1. Besides enabling the characterization of data ecosystems, the data ecosystem framework also offers the possibility to identify opportunities and challenges regarding the use of data in a particular environment. Assessing the six dimensions of the framework makes it possible to get more insights into how e.g. actors interact with data and how different factors such as policies can either foster or hinder the data use.

This research aims to fill several gaps, both in theory and practice. The theoretical contribution of this paper is twofold: First, little to no research has been done on the concepts of data ecosystems and the social shaping of technology in the context of disaster risk reduction and early warning systems. Though these concepts have received attention in other context (e.g. Parsons et al., 2011; Data Pop Alliance, 2016; Van den Homberg & Susha, 2018), the application of data in early warning systems has mainly been studied from a predominantly technical and data-science perspective (e.g. Lohani et al., 2014; Sood et al., 2018). This research aims to investigate how considering the societal, economical and institutional context of data can improve both theory and practice regarding data applications in the field of disaster risk reduction and early warning systems. As EWS, and disaster risk reduction in general, increasingly become datafied, it is crucial to investigate how EWS-data interacts with the broader ecosystem and context in which it is implemented. Second, earlier studies have discussed a broad range of barriers and limits to successfully implementing big data in disaster risk reduction. Letouzé et al. discussed three barriers, big data crumbs (the access to, sharing and interoperability of data), capacities and communities (governance, ethical and political issues) for climate risk management (Data Pop Alliance, 2015). Van den Homberg and Posthumus (2014) developed a strategic framework for the socio-economic viability of community-based EWS from a design, pilot and commercialize perspective. Successful scaling up and adoption of innovations in the EWS require adequate institutes, policies and strategic alignment, business models and financing, and capabilities. This research seeks to contribute to this field of research by investigating the barriers to successful data use in the specific context of flood EWS.

The practical contribution concerns the provision of concrete insights and steps on how to better contextualize and mature the use of data and data ecosystems. This study may help practitioners and scholars to gain more insights in how to improve their data efforts and collaboration.

Table 1: A description of the six relevant dimensions in the data ecosystem. The first five dimensions are retrieved

from Van den Homberg & Susha (2018). The sixth one is added by the authors.

Dimension	Description / Operationalization
Actors and roles	Which actors are present in the data ecosystem? What are their roles (producers/intermediaries/consumers)? How do they relate to each other?
Data supply	What kind of data is supplied? How easy is it to get access to- and use the data? *(costs of data extraction)* Is the data up-to-date, reliable, accurate and complete? *(quality of the data)*
Data demand	Which problem does the data address? What is the expected/desired outcome of the data? What is the purpose of the data use?
Data infrastructure	What platforms are used to store, share and use data? Do these platform allow for collaboration with other actors? What knowledge, skills or resources are needed to access the data?
Data ecosystem governance	What policies or agreements are relevant for working with the data? Do actors have the right knowledge, skills and resources to participate in the ecosystem? What incentives exist for data sharing and collaboration?
Contextual factors	How do cultural, political, historical, economical and social factors impact the storing, sharing and use of data?

MATERIALS AND METHODS

Introducing the Case Study: The Data Ecosystem of Flood EWS in Nepal

Despite the increased occurrence of floods due to climate change, the amount of flood damages and causalities has worldwide dropped significantly in the past decades (Ritchie & Roser, 2019). This can be related to the increased disaster risk reduction efforts, of which flood EWS are an integral part (Perera et al., 2019). A country that has dedicated significant resources to the development of disaster risk reduction and EWS is Nepal. Due to its vulnerability to floods, the country is highly dependent on its flood EWS. Collaborate efforts of the government, NGOs, private sector parties and academia have led to the establishment of both national and more local EWS, which increasingly rely on data and technologies. However, despite the increased presence of data-driven EWS, no research has yet been done on the role of data in the Nepalese EWS and the functioning of the data ecosystem. This, combined with the country's increased efforts to improve its flood resilience, make Nepal an interesting case for this research. The case study can help to get a better understanding of a phenomenon and provides the possibility to go into depth, detail and context (Given, 2008). The focus of this case study was mainly exploratory; the goal was to explore the ways in which data interacted with its environment, and to learn more about the opportunities and challenges that arose from this.

Methodology

To better understand these interactions, relevant actors involved in Nepal's flood EWS were invited for online interviews. A total of 9 interviews were conducted, for which participants were selected through both purposive and snowball sampling. Participants were working for (I)NGOs and government institutions, such as ministries and the department of hydrology and meteorology. Performing interviews made it possible to gain more understanding about the processes, roles, and experiences and the interviewees' ideas on these processes and roles (Bryman, 2012). Following the conceptualization of data ecosystems introduced above, in the semi-structured interviews, six main themes were addressed: 'actors and roles'; 'data supply'; 'data infrastructure'; 'data demand'; 'data ecosystem governance'; and 'contextual factors'. Interviewees were asked general questions about these six themes (e.g. 'who do you consider as main data providers?'; 'how is data among actors being shared?'), and they were asked about related opportunities and challenges (e.g. 'how do you think data sharing among actors could be improved?'). After a three-step coding process (see figure 2), codes were compared with relevant literature and policy documents following an abductive approach, in which the theory used was adjusted before, during, and after the interview process. For example, the theory on the social shaping of technology and the inclusion of the societal context dimension in the data ecosystem framework was only added after conducting a few interviews, which showed the importance of the economical and political context of Nepal.

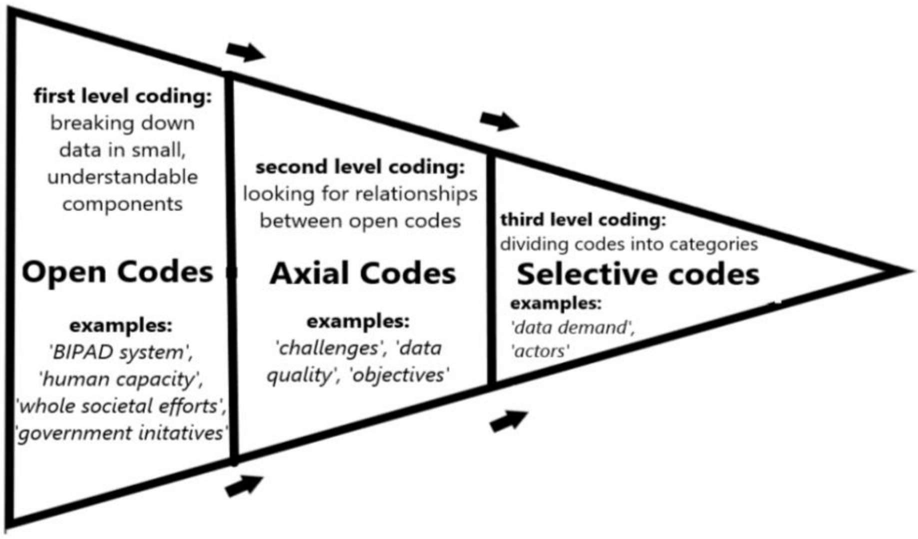

Figure 1. The three stages of coding (see Corbin & Strauss, 1990)

RESULTS

Opportunities and Challenges

The policy- and literature review and interviews provided clear and thorough insights in Nepal's data ecosystem regarding flood EWS. These insights revealed clear opportunities and challenges, of which the most noticeable ones are discussed below. Figure 3 shows how these opportunities and challenges are related to the data ecosystem.

The Actors and Their Common Goals

In Nepal's data ecosystem regarding flood EWS, many different actors are involved in the community-based, national and regional EWS. Both on national and international level hydrological and meteorological data is delivered by organizations such as ICIMOD, RIMES and the Nepalese Department of Hydrology and Meteorology (DHM). Whereas these organizations are seen as the important sources of flood forecast information, governmental institutions like the Ministry of Home Affairs (MoHA) are the main sources for vulnerability information, such as demographic data and infrastructure data. This data combined delivers clear insights in the flood risks. The exchange and dissemination of important flood related data is facilitated by a broad range of actors, such as (I)NGOs, (local) government institutions, and private sector parties, like telecom providers. All interviewees and policy reports stress similar objectives and desired outcomes of the flood EWS:

"The main objective of implementing this kind of system is to safe the lives of the community and the property of the people who are living in these areas."

"We can not stop the disaster. If there's going to be a disaster, we cannot stop it. This EWS also helps us to be prepared and have a plan for the new disasters"

The goal of using an early warning system is thus twofold: saving the lives, livelihoods, and properties of the people in the community, and being better prepared for future disasters. Using data-based forecasting models can, according to the interviewees, make this process more efficient and effective, and make it possible to act earlier in time. All actors seem to work towards this similar goal, which can foster collaborations and more effective decision-making. However, the interviews also make clear that actors are often not aware of other organizations' efforts, and that several organizations seem to work side by side, wanting to reinvent the wheel themselves. Data-sharing among different organizations seems not (yet) to be seen as a logical step towards more effective data-driven EWS, which is particularly striking given the actors' common goal. Hence, the recent establishment of the NDRRMA and their BIPAD-system might be very valuable for optimizing the use of data.

Data Infrastructure: The BIPAD-system for Collaborative Data Use

Also the Nepalese Disaster Risk Reduction and Management Authority (NDRRMA) recognizes that 'disaster data/information is still scattered, insufficient and not fully harmonized'. To improve this, they developed the Building Information Platform Against Disaster (BIPAD), aiming to enhance disaster preparedness and early warnings. Interviewees stress that the aim of developing this BIPAD is to create a common platform in which all disaster-related data and information can be stored, shared, and consumed:

"The objective of this BIPAD is to at the end of 2021 to only have one disaster portal in the country which is an open database."

Interviewees stress that the BIPAD is a promising development for disaster information sharing in Nepal, and that many people are enthusiastic about working together through using the system. However, they also emphasize the novelty of the system and the challenges that accompany this:

"It is very new. They have just launched in 2020 March"

"This is a kind of developing and implementing, developing and implementing."

"This BIPAD is still under progress and it has not become functional hundred percent"

Moreover, some also question the sustainability of the system, due to the short operation time of previous similar platforms:

"They have so many systems, and they are not operating for the long run. But BIPAD have some long run plan. Let's see. Because it is a baby for now. It is a baby system."

One could thus say that the BIPAD data infrastructure is promising, but not yet very mature. This is also visible in the limited amount of organizations that are actively contributing data to the BIPAD. The adoption of the infrastructure is thus still relatively low. As most interviewees expect the BIPAD to be fully operational in 1 to 2 years from now, it would be good to keep monitoring the development of the platform.

Low Costs, Low Quality

Access to Nepal's flood-related data has generally low costs. Costs should hereby not be taken literally as actual costs, but rather as the resources and effort an actor has to put into making use of certain data. The more data is difficult to find, is of different quality, hard to combine, not open or hosted at different infrastructures, the higher the costs. Though access to data of the other hydrological and meteorological data providers is less accessible for the broader public, accessing and using data from the DHM has fairly low costs. The government institution shares their information online on their website and through several other communication channels in a clear and visual way. The data can be accessed by anyone with a working internet connection:

"In that website there is a separate tab for rainfall and river. During monsoon season, on other season as well. All the data are freely available on our website. Anybody who visit our website can freely see that data and the

data of one week is freely available to the public to download or collect. After one week it will disappear and new data for next one week will be displayed."

Also the vulnerability data is easily accessible online or can be requested at the MoHA. However, most of this data is still based on the national census of 2011, which 10 years later might not be very accurate anymore. Although the use of alternative data sources and the new national census in 2021 provides a solution to this, it is important to consider the broader scope of this problem. Interviewees mention that not only the vulnerability data is lacking, but also a lot of the hydrological and meteorological data is incomplete or of considerably low quality. A lot of the rainfall- and hydrological stations are still based on manual reading of the data. This way of collecting data is highly prone to mistakes, and can therefore lead to unreliable models, unnecessary warnings or the absence of warnings while necessary. High-quality data is important for developing reliable and accurate models for predicting floods and developing policies to act upon this. If the quality of the data is bad, also the models, policies and plans may be useless. As one interviewee states it:

"garbage in garbage out, right. So this forecast has to be really accurate…"

Though there is thus still a lot of room for improvement, most of these points are also recognized by the involved actors, who are putting increased efforts into improving the data quality and models. However, these institutions face challenges in terms of funding, technology, and (particularly) human capacity, which slow down this process. These challenges in terms of resources also lead to issues regarding the alignment between policy and practice.

The Gap Between Policy and Practice

A broad range of policy documents, strategy plans, initiatives, and Standard Operating Procedures (SOPs) have been developed in the past years to optimize Nepal's flood early warning systems. Some of these address issues and plans on more abstract levels (e.g. the *National Policy for Disaster Risk Reduction 2018*), others concern practical action plans for specific scenarios (e.g. the *Standard Operating Procedure (SOP) For Flood Early Warning System In Nepal*). These documents are generally concise, elaborate and illustrate the great amounts of effort put into the disaster risk reduction in Nepal. Moreover, using data is increasingly mentioned as an important tool for strengthening EWS. However, there seems to exist a gap between the plans and strategies set out in these documents and the actual practice at this moment in time. The interviews make clear that a lack of financial and human resources, particularly regarding working with data, lead to the inability to fulfil the actions set out in the documents:

"I think everyone is super keen to and enthusiastic and want to improve the system, but it [funding and human capacity] is just very, very limited."

"It's all about funding, like if we were given unlimited funding, there's a huge amount that we could do"

"Main is educating the people not only local level, even in province, even in national level, this is the main challenge we are facing."

"I would say existing capacity of the staff, or the people engaged with this in this area, should, have some capacity building as well. Because you have to be updated on the system technology, the way how you deal with these data, […] there is room for improvement"

The lack of human and financial resources is thus seen as a barrier for optimizing the use of data in EWS, and therewith also mitigates the disaster risk reduction efforts. Moreover, also the recent changes in Nepal's federal structure lead to uncertainties regarding the use of data in Nepal's disaster risk reduction actions. This federal change, in which many responsibilities have shifted from the national government to local governments, is generally seen as an important step forward in Nepal. However, interviewees recognize that the uncertainties in terms of tasks and responsibilities negatively impact the data exchange and the availability of certain data and infrastructures. This was recognized by Vij et al. (2020) and is also mentioned by the interviewees:

"…engagement of three tiers of government, and this whole system is very important, you know, I would say like

how, who has the role for this, which role is to you know, generate the data would be analysing, would be disseminating, who will be, you know, taking care of this operational and maintenance matters, which is major challenges at some point."

"They've recently changed the federalization structure. So the municipalities now have responsibility for flood early warning within their municipalities. But they're still trying to work out how that works in practice, because they don't have, or they may not have expertise at the local level, to do this sort of work"

"The capacity at the local level or the local government is still at the newest stage you know, there still needs some recruitment of the technical people at the local level."

Though policy- and strategy documents thus show clear ambitions regarding data use and flood EWS in general, interviews reveal that aligning practices with these strategies is much more difficult due to uncertainties in terms of finances, human capital, and federal structure.

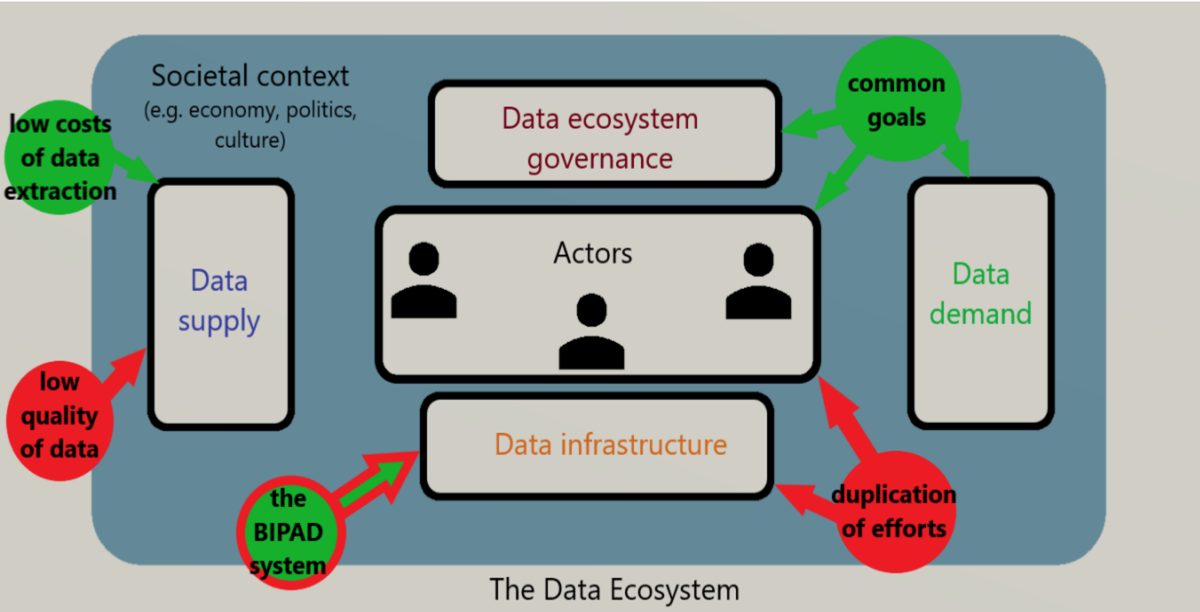

Figure 2. A visualisation of the data ecosystem regarding Nepal's flood EWS. The related opportunities are visualised in the green bubbles, the challenges are visualised in the red bubbles. The arrows show to which element of the data ecosystem these opportunities and challenges are related. As the BIPAD system provides both opportunities and challenges, this bubble is green and red.

CONCLUSION, DISCUSSION, AND RECOMMENDATIONS

Early warning systems increasingly rely on big data and technologies. This brings along opportunities in terms of more up-to-date, validated, and reliable information, leading to more efficient, effective, and evidence-based decision-making. However, we argue that data should not be seen as an entity unto itself, but rather as component of a *data ecosystem* in which it interacts with the technological, institutional, and societal context in which it is applied.

This research was aimed at answering the question: 'How does the broader data ecosystem impact the opportunities and challenges regarding data use in flood early warning systems in Nepal?'. Interviews and a policy- and literature review revealed that most opportunities and challenges that impact the use of data were not directly related to the data itself but rather to the broader technological, institutional and societal context in which the data was handled. The challenge and opportunity that can directly be related to the data itself were, respectively, the low quality of the data and the low costs of data extraction. The low quality of data should definitely be recognized as a major issue, but it is important to acknowledge that the other evident challenges and opportunities mentioned by interviewees all seem more related to the broader data ecosystem. A main opportunity regarding this broader context concerns the existence of a common goal of the actors involved, fostering collaborations and thus more efficient data use. Moreover, the recent establishment of the BIPAD-system is regarded as a promising development for more active data sharing and more general access to disaster data. Lastly, the existence of elaborate policy documents and strategy plans reveal a clear ambition regarding disaster risk reduction and can help to align complex processes more smoothly. However, these opportunities can also directly be linked to main challenges. Despite the common goal of the actors involved, collaboration between actors is often still lacking, which leads to duplications of efforts and thus less efficient use of data. Moreover, both the BIPAD system and the elaborate policy plans suffer from the lack of resources in terms of finances and human capacity. This, combined with the shift of responsibilities due to the new federal structure of Nepal, lead to uncertainties in terms of investments and roles. The people responsible for handling the flood data are often not (yet) in place or lack the necessary skills or resources to work with it.

In response to the research question one could thus say that the data ecosystem does not only *impact* the opportunities and challenges regarding data use, but rather that the opportunities and challenges regarding data use are *contained within* the data ecosystem. With this, we mean that most opportunities and challenges related to the data use are not necessarily related to the quality or quantity of the data itself, but are rather related to e.g. the capacities and resources of actors, the adoption of infrastructures and the availability of resources. If one wants to optimize the data use regarding flood EWS in Nepal, it should start with looking at contextual factors, such as the financial and human resources that are available for flood EWS. These contextual factors do not only impact the success of *data use* in a flood EWS, but are also more generally important for the functioning of Nepal's flood EWS *as a whole*. The gap between the desired functioning of EWS and the application of (big) data in EWS, as presented in the introduction of this paper, and the actual practice in Nepal, as presented in the result section, is significant and requires ongoing attention, action and resources of scholars, policy makers, politicians and funding organizations.

Though this research concerned a case study, which generally provides little basis for generalization of results to the broader population or other cases, the research may provide ideas about the general interactions between data and its context. This study showed that 'optimizing the use of data'may actually mean 'optimizing the context in which the data is used'. This conclusion may also be applicable to other context, and using a data ecosystem model could also in other contexts help to provide better understanding about how to optimize the use of data. To test whether this is the case, future research could be aimed at applying a data ecosystem model to other or more cases, for example, by doing a multiple case study among different contexts. This could help to provide a better understanding of the ways in which different contextual factors impact the data use in different ways, or might give more insights in the factors that most strongly impact the use of data. Moreover, in the context of Nepal more research could be done on the ways in which data sharing among actors can be fostered and how the BIPAD system, or data repository systems in general, can add to this. Lastly, more research should be done on the ways in which higher quality flood forecasts models can be developed, based on the (limited amount of) data available in Nepal. This research could be built upon the work of Dugar (2015) and Smith, Brown & Dugar (2017).

At a practical level, we argue that a better understanding of the emerging data ecosystem and the related challenges and opportunities among all actors involved can lead to a more mature data ecosystem. The common understanding can lead to more willingness among actors to lower barriers to data sharing and may lead to a converging of data infrastructures instead of scattered data in several platforms. Also, in the other components of the data ecosystem, more awareness of opportunities and challenges helps to identify clear areas of improvement. The resulting more mature data ecosystem will help an EWS to become more effective and efficient in all its components.

ACKNOWLEDGEMENTS

We thank our interviewees for so generously making time to provide us with their valuable insights. We would like to express our gratitude to Heleen Elenbaas (510) for her contributions to this study. We are grateful as well for the support of the Nepalese Red Cross Society and the EU ECHO-III Preparedness for emergency response in the Western Region of Nepal project.

REFERENCES

Basher, R. (2006). Global early warning systems for natural hazards: systematic and people-centred. Philosophical transactions of the royal society a: mathematical, physical and engineering sciences, 364(1845), 2167-2182.

Baym, N. K. (2010). Personal connections in the digital age. Cambridge, UK ; Malden, MA: Polity.

Benini, A. (2015). The use of Data Envelopment Analysis to calculate priority scores in needs assessments. Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP), London, UK.

Bierens, S., Boersma, K., & van den Homberg, M.J.C. (2020). The legitimacy, accountability, and ownership of an impact-based forecasting model in disaster governance. Politics and Governance, 8(4), 445-455.

Bryman, A. (2012). Social research methods. Oxford: Oxford University Press.

Budimir, M., Donovan, A., Brown, S., Shakya, P., Gautam, D., Uprety, M., Cranston, M., Sneddon, A., Smith, P. & Dugar, S. (2020). Communicating complex forecasts for enhanced early warning in Nepal. Geosci. Commun. Discuss, 2019, 1-32.

Comes, T. (2016). Cognitive biases in humanitarian sensemaking and decision-making lessons from field research. In 2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA) (pp. 56-62). IEEE.

Corbin, J. M. & Strauss, A. L. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Thousand Oaks: Sage Publications.

Data-Pop Alliance (2015) Big Data for Climate Resilience. Inputs for the World Development Report 2016 “Digital Dividends”; Input 2. Available online: https://datapopalliance.org/wp-content/uploads/2017/03/WDR-Data-Pop-Alliance-input-2-Climate-Change-Resilience.pdf (last accessed: 11-04-2021)

Data Pop Alliance (2016). Opportunities and Requirements for Leveraging Big Data for Official Statistics and the Sustainable Development Goals in Latin America. Available online: http://datapopalliance.org/wp-content/uploads/2016/05/Data-Pop-Alliance-LAC-NSO-EN.pdf (last accessed: 27-01-2021).

Dugar, Sumit. (2015). The 2014 Karnali River Floods in Western Nepal: Making Community Based Early Warning Systems Work When Data Is Lacking.

Given, L. M. (2008). The SAGE encyclopedia of qualitative research methods (Vols. 1-0). Thousand Oaks, CA: Sage Publications.

Haak, E., Ubacht, J., Van den Homberg, M., Cunningham, S., & Van den Walle, B. (2018). A framework for strengthening data ecosystems to serve humanitarian purposes. In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, 1-9.

International Federation of Red Cross and Red Crescent Societies. (2012). Early Warning Systems. Guiding Principles.

Kumar Rai, R., van den Homberg, M., Ghimire, G., & McQuistan, C. (2020). Cost-benefit analysis of flood early warning system in the Karnali River Basin of Nepal. International journal of disaster risk reduction, 47, 101534.

Lohani, A. K., Goel, N. K., & Bhatia, K. K. S. (2014). Improving real time flood forecasting using fuzzy inference system. Journal of hydrology, 509, 25-41.

MacKenzie, D. & Wajcman, J., eds. (1999) The social shaping of technology. 2nd ed., Open University Press, Buckingham, UK.

Mulder, F., Ferguson, J., Groenewegen, P., Boersma, K., & Wolbers, J. (2016). Questioning Big Data: Crowdsourcing crisis data towards an inclusive humanitarian response. Big Data & Society, 3(2), 2053951716662054.

Nardi, B., & O'Day, V. (1999). Information ecologies: Using technology with heart. Cambridge, MA: MIT Press.

Parsons, M.A., Godøy, Ø., Ledrew, E., De Bruin, T.F., Danis, B., Tomlinson, S., and Carlson, D. (2011). A Conceptual Framework for Managing Very Diverse Data for Complex, Interdisciplinary Science. J. Inf. Sci., 37, 555–569.

Perera, D., Seidou, O., Agnihotri, J., Rasmy, M., Smakhtin, V., Coulibaly, P., & Mehmood, H. (2019). Flood Early Warning Systems: A Review Of Benefits, Challenges And Prospects. United Nations University: Institute for Water, Environment and Health. Available online: https://inweh.unu.edu/wp-content/uploads/2019/08/Flood-Early-Warning-Systems-A-Review-Of-Benefits-Challenges-And-Prospects.pdf (last accessed: 27-01-2021).

Rip, A., & Kemp, R. (1998). Technological change. Human choice and climate change, 2(2), 327-399.

Ritchie, H. & Roser, M. (2019). Natural Disasters. Available online: https://ourworldindata.org/natural-disasters#number-of-deaths-by-type-of-natural-disaster (last accessed: 27-01-2021).

Smith, P. J., Brown, S., & Dugar, S. (2017). Community-based early warning systems for flood risk mitigation in Nepal. Natural Hazards and Earth System Sciences, 17(3), 423-437.

Sood, S. K., Sandhu, R., Singla, K., & Chang, V. (2018). IoT, big data and HPC based smart flood management framework. Sustainable Computing: Informatics and Systems, 20, 102-117.

United Nations. (2014). A World that Counts: Mobilising the Data Revolution for Sustainable Development. UN Data Revolution Report. Available online: https://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf (last accessed: 27-01-2021).

UNDP (2018). Five approaches to build functional early warning systems. Available online: https://www.eurasia.undp.org/content/rbec/en/home/library/environment_energy/five-approaches-to-build-functional-early-warning-systems.html (last accessed: 27-01-2021).

UNDRR (2019) Disaster risk reduction in Nepal: Status report 2019. United Nations Office for disaster risk reduction (UNDRR), Regional Office for Asia and the Pacific, Bangkok. Available online: https://reliefweb.int/sites/reliefweb.int/files/resources/68230_6nepaldrmstatusreport.pdf (last accessed: 27-01-2021).

UNISDR (2012). Terminology. UNISDR (United Nations International Strategy for Disaster Risk Reduction), Geneva. Accessed 18 October 2020 from http://www.unisdr.org/we/inform/terminology

Van den Homberg, M. J. C., & Posthumus, B. (2014). Strategic framework for socioeconomic viability of community-based early warning system. ISCRAM Asia Pacific.Van den Homberg, M., & Susha, I. (2018). Characterizing data ecosystems to support official statistics with open mapping data for reporting on sustainable development goals. ISPRS International Journal of Geo-Information, 7(12), 456.

Vij, S., Russell, C., Clark, J., Parajuli, B. P., Shakya, P., & Dewulf, A. (2020). Evolving disaster governance paradigms in Nepal. International Journal of Disaster Risk Reduction, 50, 101911.

Winsemius, H.C., Jongman, B., Veldkamp, T.I.E., Hallegatte, S., Bangalore, M. & Ward, P. J. (2018). Disaster risk, climate change, and poverty: assessing the global exposure of poor people to floods and droughts. Environment and Development Economics, 17. 1–21.

WMO (2018). Multi-hazard Early Warning Systems: A Checklist. Available online at: https://library.wmo.int/doc_num.php?explnum_id=4463 (last accessed: 08-02-2021).

Zurich Flood Resilience Alliance (2019, February 6). The Flood Resilience Portal [video]. Available online: https://www.youtube.com/watch?v=EMAkfr3lkFM&feature=emb_logo (last accessed: 27-01-2021).