# Data-driven decisions for flood risk management

**Liz Edwards*, Louise Mullagh, Ross Towe, Vatsala Nundloll, Claire Dean, Graham Dean, William Simm, Faiza Samreen, Richard Bassett and Gordon Blair**
Data Science Institute, School of Computing and Communications, InfoLab21, Lancaster University, LA1 4YW, UK

**liz.edwards@lancaster.ac.uk*

## Abstract

Ensemble is an interdisciplinary research team, working to explore the opportunity of new and emergent digital technologies in understanding, mitigating and adapting to environmental change. Using methods drawn from computer science, environmental science, social science, statistics, art, design and writing the team aim to transform the work of environmental scientists and decision makers, and the experience of communities by addressing themes related to complexity, uncertainty and abstraction of data. This paper discusses activities undertaken within a research 'sprint' directed at addressing flood risk management through data driven decision making, communication and community engagement.

***Keywords*** – flood risk management; community engagement; transparent decision making; trust; communication

## 1 Introduction

In recent years parts of the UK have experienced extreme flood events (Defra & Cabinet Office, 2016) and some communities have been repeatedly affected by flooding. The National Risk Register reported an increase in the frequency of inland flooding and the likelihood of more severe weather events (Cabinet Office, 2015). One in six homes in the UK are at risk of flooding leading to impact for 5.2 million people (National Flood Forum, 2017). The complexity that contributes to an instance of flooding means there is significant uncertainty in flood prediction and decision making about mitigation strategies, which can affect trust between local communities and decision makers. This paper discusses the work of the Ensemble research programme to support flood risk management through collaboration with environmental scientists and communities in flood risk areas.

## 2 The Ensemble Programme

Ensemble is a co-located interdisciplinary team of researchers based at Lancaster University, working to address the role of digital technologies in understanding, mitigating and adapting to environmental change. During an initial five-year period, the programme will address flood risk management, soil science, biodiversity and ecosystems services, exposing and attending to crosscutting themes. These include managing complexity and uncertainty, raising abstraction to support the work of environmental scientists and developing software architecture for the deployment of cloud technologies to manipulate large heterogeneous data sets. At the heart of this work is the desire to support decision-making at different scales, for a range of decision makers, from the scientific community, central government and national organisations to individual homeowners and small businesses. The research promotes collaborative, agile science and collaborative, agile decision-making. We address these challenges holistically which means that the strategy goes beyond the implementation of a set of technologies. It involves close and on-going collaboration with stakeholders in an iterative process, focused around a series of 'sprints' directed at particular challenges. The sprint methodology is adapted from agile software development, and consist of short, focussed research projects designed to allow the team to explore specific opportunities within a research space (Ferrario et al., 2014).

The Lancaster University-led Ensemble team comprises computer scientists, environmental scientists, statisticians, designers, artists in residence and a writer in residence, and the project partners include the UK's Environment Agency, the Centre for Ecology and Hydrology and the JBA Trust, a charity that consults on water management and flood risk. The interdisciplinarity of the team enables a broad perspective of flood risk and has the potential to bridge gaps between different groups, encouraging wider engagement and participation. Individual disciplines

communicate using different languages and styles, which means that collectively the team is able to select the most appropriate form of communication to reach particular groups of decision-makers.

# 3 Flood Sprint

In the first year of the programme the research addresses challenges related to flood management, including the development of a novel flood scenario library to support decision-making, the use of flood data walks to gather and share local knowledge about flood risk to support the development of flood risk models, and the use of diverse communication activities to gather and share information and support education about flooding. In particular the paper describes Manchester Science Festival as a focal point for reaching out to engage communities in area of flood risk. This has the potential to reach audiences beyond those who might ordinarily join flood-related community activities. It also brings children into discussion, in a deliberate response to Mort's recommendation that children be recognized as active contributors in flood response and recovery (Mort, M., et al, 2016). It is also well recognised that if you engage children you engage parents. The next section will discuss each of the activities and the motivation for exploring these approaches.

## 3.1 Flood scenario library

Flood data in massive volumes originates from many sources and takes many different forms including sensor data, satellite data, historical data, and community data. There is an opportunity to link the volumes of complex data in order to support deep analytics and complex queries at scale. There are great complexities in integration and interfacing to multifaceted data, however by leveraging advances in data science tools and new computing architectures, the Ensemble team are developing a scenario library which semanticly integrates structured and unstructured data from multiple sources including outputs from environmental models, at different scales and time series. This is wrapped in a query layer which aims to provide answers to questions about policy, resource allocation, assessment of long and short-term flood mitigation strategies and local interventions such as Natural Flood Management. At a general level it has the potential to empower any of the groups or individuals who might wish to query the available data for decision-making at personal, organizational, local and national scales. Additionally it will open up and increase the accessibility of flood models and their outputs to new communities of people.

The key to being able to integrate information in a reusable way is the use of semantics to enrich data and facilitate interrogation of heterogeneous data sets from various perspectives. Ontologies are an important to this process. An ontology is a semantic technology used to give meaning to data and to represent the data using a formalized notation. Ontologies provide the ability to define a domain of interest through concepts and the relationships that exist among these concepts. They can be used to describe data, metadata and systems that structure data, thereby forming flexible and dynamic links between changing data sets.

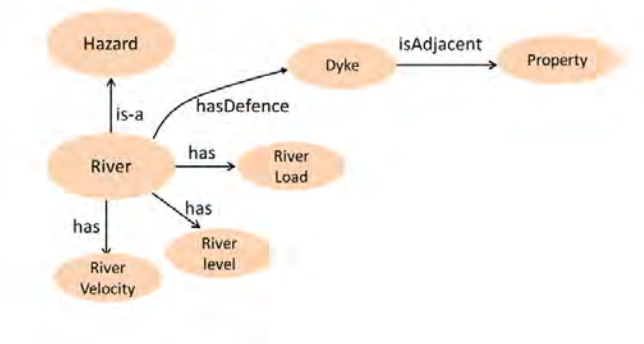


*Figure 1. An example of a flood risk ontology.*

Figure 1 illustrates a small but extendable ontology for the concept of 'river'. It shows that a river can be categorized as a hazard. It has properties such as velocity, load, and level and can have defences, of which a dyke is an example. The adaptable, extensible and fluid quality means that different data sets can be expanded, merged and recombined at will, circumventing the limitations of traditional structured databases. They can be used to specify concepts that are relevant to different audiences. This means the same linked data sets can be queried to give multiple perspectives. This means it is possible to generate ontologies to reflect the needs and interests of different user groups such as homeowners, insurance companies, individual government departments, and scientists advising on spending to mitigate against future risk.

Example of potential scenario queries include:

- If the Department of Transport has a budget to spend on flood mitigation in the next 5 years, where should it be spent?
- What is the exposure of insurers who insure properties within 100m of a named river?
- What is the risk of a flood with an Annual Exceedance Probability of 1% (a one in a hundred year flood) affecting my home within the life of

my 30 year mortgage?
- Will my business flood today?...tomorrow?...next week?

The advantage of using semantic integration in the construction of a scenario library goes beyond the ability to interrogate data from different perspectives. It also allows for the incorporation of novel data, qualitative data and localized data with existing data sources to present richer, nuanced picture of places with the potential for more refined models of risk and uncertainty. We expect this to be especially empowering to citizens living in areas at risk of flooding because the datasets can absorb local knowledge to support more effective decision-making.

At present, national flood risk models may not be sensitive enough to accurately represent the hydrology of individual catchments, leading to local mistrust of the models and the flood warning tools derived from these models. Public misgivings may extend to organizations involved in development and application of these models and the uncertainty may inhibit future planning within communities. By contrast, enabling people to contribute local knowledge in meaningful ways, for example towards the generation of new models can catalyse community activity for flood risk management. Whatmore and Landström's (2011) account of controversy around flood-risk reduction and management in the town of Pickering illustrates this point. They describe the activities of a competency group, born in the wake of disputes about how to defend the town against floods, and the group's impact on decision-making and public involvement through flood apprenticeship and interrogation of expert knowledge. Changing knowledge relationships were key to citizen empowerment. The semantically driven scenario library we describe offers another route to examine expert knowledge and introduce layers of local knowledge that complement or contradict existing data.

In the past the inclusion of heterogeneous data sources was constrained by the structure of databases and the difficulty of tracking data provenance, which was problematic because it could suggest equivalence between data despite differences in quality and levels of uncertainty. Ontologies provide a sound structure to define the metadata of data, such as the provenance of the data and versioning of the models being used in flood forecasting. Moreover, ontologies can also enable a semantically-valid integration of disparate data sets, either historical or sensor data and allow to answer richer set of queries. It may be possible that the manipulation of data through a semantic formulation can help environmental scientists to tackle uncertainties inherent in the data or reasoning based on the data. Once data streams can be tracked, traced and potentially weighted, there it becomes less controversial and more practical to combine heterogeneous data of varying quality, from a range of geographical scales.

We have described the development of a flood scenario library to support decision-making and its importance for enabling the use of regional and local flood-risk data. The next section describes an activity for knowledge sharing to inspire the collection of heterogeneous local data to populate the flood scenario library.

## 3.2 Flooding Data Walk

The Data Walk is an exploratory method that establishes meaningful practices of data collection that are socially and geographically inclusive, through the reification of data. Environmental data exists in the land, extracted via instrumental methods such as remote sensing, gauges or via satellites. The Data Walk seeks to engage people in the landscape through walking, talking and map making, making the latent data tangible and meaningful through participatory data collection. It also offers space for the integration of tacit and embodied knowledge, which is accrued by people living in a place over time but often omitted in data driven computational modelling methods.

### 3.2.1 Walking and talking

The concept of walking to understand place is not novel, it has long since been recognized as critical in the building of knowledge (Ingold, 2010). Being in a place is important in acquiring knowledge, but moving through it slowly develops the experiential and is recognized as a 'historically located practical knowledge (Edensor, 2000). Being in a place that is relevant to the problem being explored (in this case flooding) offers people the opportunity to consider data in context. Furthermore, it enables connection and engagement between the decisions regarding the extraction and re-use of data from that place, those who inhabit the place and the place itself. Moving slowly through a place on foot enables participants to explore and uncover data that are within the place, or to ask questions, prompted by visual cues or by the walk leader.

### 3.2.2 Outdoor, mobile walk-shops

Community engagement between stakeholders and agencies engaged in flood mitigation and relief has in recent years become more concerned with knowledge held by people living with flooding (Anon, 2010). However, in several interviews with local community activists prior to the development of the Data Walk, a distinct sense of apathy towards workshops was apparent, with over ten

examples cited in the period of one year after the floods of December 2015. The Data Walk offers a more innovative and exploratory approach, reflecting the interdisciplinary spirit of Ensemble and the potential for deeper engagement with different communities through sharing knowledge by walking, taking and creating maps.

### 3.2.3 Participatory Data Collection

The Ensemble Flood Data Walk was held in July 2017, at a hillside farm located in the hamlet of Warland, which lies on the border between Yorkshire and Lancashire. Seventeen participants from different backgrounds were chosen to take part in order to explore the potential of bringing together a diverse group of people to collect data together around a given problem, in this case flooding. Participant's backgrounds included local community groups (from Hebden Bridge), multi-disciplinary researchers from the Ensemble team (computer scientist, data scientist, designer), the Ensemble writer in residence, permaculturalist, specialist in natural flood management and PhD candidates from design and computer science.

Equipped with a wearable map, the group had the choice of following the owner of the farm on a pre-determined route, or following the route at their own speed. As they walked, participants were asked to consider the following questions:

- Think about water…in relation to climate change. How might we work *with* water in the future? How have we worked with it in the past?
- What data can we find in the landscape and within people?
- What data can we collect today that cannot be collected by sensors or devices?

No guidelines were given as to how they should respond and as a result a rich array of drawings, text, questions and found objects were collected.

*Figure 2. Data produced by the participants of the data walk.*

After the walk participants were asked to choose responses or objects they had collected at key points then to write on a tag what they had collected and why. The data (drawings, text and artefact) was then pinned onto a large map that corresponded to the small wearable map. This resulted in two layers of data being placed on the map, firstly the 'raw' data collected during the walk and secondly the reasoning behind its collection.

### 3.2.4 Initial findings and future work

Initial reflections from the walk demonstrated the potential of this method to engage different communities in often difficult conversations relating to flooding. Bringing together people who live with flooding with those who are working to develop models or flood management methods is not novel, however engaging participants in collecting data offers potential for unique insights and meaningful data to be gathered.

The data is currently being analyzed and logged in order to gain insights into the data collected and the potential for this unstructured, heterogeneous data to be incorporated into an ontology relating to flooding in that area.

By folding in the data collected from the walk, gathered from a variety of stakeholders, a rich layer of qualitative data representing diverse communities will be featured in the scenario library. This novel approach of using data gathered from place by an interdisciplinary group, not only offers potential for rich data about a place become enacted, but there is greater potential for empowerment within local communities. There is also potential for the data collected to be stored and developed into a living community resource or physical 'data store', much like a living archive.

## 3.3 Manchester Science Festival

Manchester Science Festival presents a forum to address flood communication in which to stimulate conversations about flooding with audiences who come to the event with varying degrees of knowledge and experience.

The science festival offers a space to talk in different ways about flooding focusing on different aspects of flooding. Flood risk analysts break down flooding into component parts, source, pathway and receptor to describe the hazard function. These components refer to the river channel, flood defences and flood impact respectively. The structure proved helpful when planning activities for the science festival, to ensure that the multifaceted nature of flooding

was not lost. Some activities were designed to show the hydrological processes of river catchments, some were designed to draw attention to flood defences and others exposing impacts for people. The activities also connected to the stages of a flood; preparation, action during the flood and post-flood recovery. The festival provides an opportunity for our research partners to join us with their own interpretation and engagement activities. It allowed data to be presented in new forms. Some of the designs are described in the following sections.

### 3.3.1 Flow visualization of the River Irwell

Flow is an installation that visualizes historical flood events at locations on the River Irwell using the outputs of the Flood Foresight model, developed by JBA Consulting (Jeremy Benn Associates Ltd., 2017). Flood outlines for real events are modeled using actual observations from river gauges in conjunction with terrain models created from LiDAR radar survey data. The visualization employs an array of ShapeClips (Hardy et al., 2015) as a modular, three-dimensional shape changing display, to show terrain, and changes in river flow during a flood event at different points along the river's catchment. This information is paired with meteorological data to communicate the relationship between different kinds of weather events, and the onset of flooding. The main focus of the installation is the Boxing Day flood of 2015, which affected Manchester and the surrounding area. The located information is presented alongside historical accounts describing the power of extreme flood events emphasizes the recurrence of flooding in some place over centuries. The Flow installation sets the scene for the other flood related activities and establishes connection between the location of the Science Festival and flood risk. It adheres to the principles that communication should concentrate on "making information local, with historical context." (Fisher, 2015).

### 3.3.2 Flood Snakes and Ladders

The flood snakes and ladders game was developed as part of a project to research flood impact, recovery and resilience with children and young people. (Mort et al., 2016) It was devised in response to a quote from a young research participant who compared post flood life to a game of snakes and ladders where recovery does not occur smoothly but is intermittent and disrupted by multiple factors. The game is used to tell flood stories drawn directly from research data to underline the real consequences of flooding on children's lives. The original version of the game has been played using a person-sized board and football-sized dice with audio-visual projections to depict the negative 'snakes' and positive 'ladders' experienced after a flood. The game design deliberately foregrounds children and young people's experiential knowledge as a way of reinforcing their credentials as flood experts with valuable contributions to make in conversations about flooding, policy and future resilience. A digital version of game has been developed through the Ensemble Programme to allow the data to reach new audiences without the need of a facilitator. Both versions of the game will be played at Manchester Science Festival. The game is designed to give voice to 'flood active' publics who are 'highly aware and well informed' (National Flood Resilience Review) about flooding. Playing the game at the science festival has the potential to reach broad ranging audiences who might not otherwise engage in activities connected to flooding.

### 3.3.3 Flood Box

> "Don't just focus on the negative impacts…focus on what people can do about it." (Fisher, 2015)

The flood box is about personal empowerment and decision-making. It involves thinking about planning for a flood and the possibility of having to leave home as a result of flooding. Visitors are invited to suggest things they think they might need in flood. Afterwards they are given a box of physical items, recommended by people who have experienced flooding, to sort and rank by need at the moment of flood and at the point of evacuation. A printout of choices is given to participants to take away as a personal starting point for a making a flood box for use during a flood and an emergency bag with things needed for an extended time in temporary accommodation during flood recovery.

### 3.3.4 Flood Story Dice

> "Stop talking about probability and risk in mathematical language as it means very little to a lot of people." (Fisher, 2015)

Flood Story Dice were developed as a way to explain the probability of the one in a thousand year flood, but the activity has expanded to become a way to talk about risk and extreme events using stories. Visitors to the installation can use the pre-made Flood Story Dice, or make their own dice, drawing images on each side to represent places that flood, things and people that help in a flood, people and animals who are affected by flooding, and the causes of flooding. When sets of dice are thrown together they can be used to create flood stories and talk about flood risk.

#### 3.3.5 Flood Resilient Home makeover

> "Flood-affected children also understand the need for adaptation. Rather than promoting 'back to normal', please support them when they say 'we need a NEW NORMAL'." (Mort et al., 2016)

The resilient home focuses on the new normal in which people living in areas that have flooded repeatedly may need to live differently, making changes to their homes to live with water. The activity involves identifying features in a flood adapted home and has the potential to be used in discussions about insurance and future planning.

# 4 Discussion

The paper has outlined an integrated series of interdisciplinary activities to address flood risk management through data driven decision making, communication and community engagement. The activities undertaken are diverse, recognising the strength of a multi-faceted approach to a complex domain. They encompass the development of software architecture to support data interrogation and transparency, methods for collection of heterogeneous situated data and design of artefacts for provocation and communication

Semantically enriched data facilitates interrogation from various perspectives and widens audience engagement for decision making. Data walks stimulate the development of new ontologies embracing community knowledge and communication activities include different publics in discussions about flood risk management.

The work is intended to improve communication before, during and after a flood event so people are better prepared and have an improved understanding of risk. The strategies seek to combine open data to facilitate transparent decision making at all levels. As a result, this will promote trust between communities, environmental scientists and decision makers to help the decision making process, which recognise and account for the complex nature of flooding.

# Acknowledgements

We would like to thank EPSRC grants EP/P002285/1 (The Role of Digital Technology in Understanding, Mitigating and Adapting to Environmental Change) and EP/N027736/1 Models in the Cloud: Generative Software Frameworks to Support the Execution of Environmental Models in the Cloud and the Digital Economy programme (RCUK Grant EP/G037582/1), which supports the HighWire Centre for Doctoral Training (highwire.lancs.ac.uk). Our special thanks go to Rob Lamb and the JBA Trust, Jacqui Cotton and the Environment Agency, and Maggie Mort and Marion Walker at Lancaster University.

# References

Anon, 2010. Doing flood risk science differently: an experiment in radical scientific method. pp.1–22.

Cabinet Office (2015) National Risk Register of Civil Emergencies, London: Cabinet Office

Defra and Cabinet Office (2016) National Flood Resilience Review, Policy Paper, London: Defra

Edensor, T., 2000. Walking in the British countryside: reflexivity, embodied practices and ways to escape. *Body & Society*, 6(3-4), pp.81–106.

Ferrario, M. A., Simm, W., Newman, P., Forshaw, S., & Whittle, J. (2014, May). Software engineering for 'social good': integrating action research, participatory design, and agile development. In *Companion Proceedings of the 36th International Conference on Software Engineering* (pp. 520-523). ACM.

Fisher, H., 2015 Public dialogues on flood risk communication *Enivronment Agency* Final Report - SC120010/R1

Hardy, J., Weichel, C., Taher, F., Vidler, J., & Alexander, J. (2015). ShapeClip: Towards Rapid Prototyping with Shape-Changing Displays for Designers. *Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems*, *1*, 19–28.

Ingold, T., 2010. Footprints through the weather□world: walking, breathing, knowing. *Journal of the Royal Anthropological Institute*, 16(s1), pp.S121–S139.

Jeremy Benn Associates Ltd., 2017 *Flood Foresight* https://www.floodforesight.com

National Flood Forum, 2017, https://nationalfloodforum.org.uk

Mort, M., Walker, M., Lloyd Williams, A., Bingley, A., & Howells, V. (2016) Final project report for 'Children, Young People and Flooding: Recovery and Resilience', Lancaster University, Lancaster, UK

Whatmore, S.J., & Landström, C. (2011) Flood apprentices: an exercise in making things public, Economy and Society, 40:4, 582-610