Institution:

KING'S COLLEGE LONDON

Unit of Assessment:

36 (Communication, Cultural and Media Studies, Library and Information Management)

Title of case study:

Freshwater Information Management and Data Sharing to Meet Environmental Standards

1. Summary of the impact (indicative maximum 100 words)

The advanced information management research of the Department of Digital Humanities (DDH) has led to a better understanding of pollution processes in inland waterways and lakes. It has also improved the standard of water quality information that is available to government and regulatory authorities. The information management framework which DDH has provided supports government-funded activities to improve environmental standards and has helped ensure that the UK Environment Agency is able to comply with the EU's Water Framework Directive, reducing the risk of financial penalties for non-compliance. Moreover, key and accurate evidence about water quality has been made freely available to beneficiaries, including governmental and non-governmental agencies, farmers and land managers, and the general public.

2. Underpinning research (indicative maximum 500 words)

The Department of Digital Humanities has engaged in long-standing and field-leading research in the areas of: digital libraries; archives; digital asset management; data integration; metadata; and research infrastructures. While DDH's work focused initially on the Humanities, the research is applicable to and has been transferred and adapted to many different domains. This case study shows the impact that the research has had on freshwater biology, freshwater hydrology, and other aspects of environmental science.

Unavailability of and lack of cohesion in environmental datasets is a major stumbling block in meeting regulatory requirements and improving environmental quality. For example, in 2009 Defra prepared an explanatory memorandum to the INSPIRE (Infrastructure for Spatial Information in the European Community) Regulations (No. 3157) stating that the five most frequent difficulties experienced in the preparation of Environmental and/or Strategic Impact Assessments related to:

- Problems with getting access to existing data (70%);
- Difficulties with finding out which data is available (56%);
- Lack of availability of necessary data (51%);
- Non-compatibility of datasets from different suppliers (47%);
- Insufficient quality of existing data (47%).

The pathway from research to impact began when the Unit was approached by the Freshwater Biological Association (FBA) as a result of the research reputation of DDH's Centre of e-Research (CeRch) in information management. The FBA and DDH collaborated on research into standards, vocabularies and infrastructure for data sharing and integration. The research outputs now support the UK freshwater science community, which includes industry, third sector, and non-HE research

organisations.

The distinctive challenges posed by water quality information are the result of its diverse and complex structure. Existing water data collections are generally hand-crafted with diverse individual characteristics that lead to an inherent 'messiness' in the way data is presented. There are over 30 criteria for assessing the status of rivers, lakes, transitional waters and coastal waters. They all include consideration of biological and chemical quality, environmental quality standards, and physical aspects such as the quantity and dynamics of water flow. Yet failure to understand or manage this complex data has serious repercussions. If part of a water body fails to meet the standards mandated by the Water Framework Directive (WFD) on any single criterion, it will be judged to have failed overall.

DDH carried out a programme of research over several years designed to lead to the more effective representation and exploitation of data related to water quality infrastructures [3.4]. Research included interviews with freshwater scientists and those with responsibility for freshwater data creation. These interviews helped elicit attitudes to sharing and reusing data and clarified the functional requirements for a freshwater data repository [3.6]. DDH's research has delivered methods for 'treating' heterogeneous collections of data [3.1]. This includes the creation of metadata (information about where each piece of data comes from) for research data archives [3.2]. In addition, DDH delivered means of carrying out semantic integration of research data sets, harmonising data stored by different groups in different formats so that any query run on the database will be able to match information appropriately. This approach was also developed in relation to scientific and experimental data sets [3.3].

The research developed innovative methods of visualising, querying, analysing and processing information in an integrated fashion that had not hitherto been possible. In the case of freshwater biology research, this results in improved knowledge and understanding of pollution processes and of their relationship to human activity, especially farming [3.6].

Key Researchers

- Mark Hedges, Senior Lecturer and Director of CeRch (at King's since 2005)
- Richard Gartner, Lecturer (at King's since 2009)
- Sheila Anderson, Professor of e-Research (at King's since 2002)
- Tobias Blanke, Senior Lecturer (at King's since 2005)

3. References to the research (indicative maximum of six references)

3.1 Mark Hedges, Tobias Blanke, Adil Hasan, Rule-based Curation and Preservation of Data, Future Generation Computer Systems, Vol. 25, No. 1, 2009. doi:10.1016/j.future.2008.10.003

3.2 Tobias Blanke, Mark Hedges, Humanities e-Science: From Systematic Investigations to Institutional Infrastructures, Proceedings of the 6th IEEE e-Science conference, Brisbane, 2010. doi:10.1109/eScience.2010.34 [winner of prize for best paper at conference]

3.3 Hedges M. & Blanke, T. (2013). Digital Libraries for Experimental Data: Capturing Process Through Sheer Curation. In Proceedings of the 17th International Conference on the Theory and Practice of Digital Libraries (TPDL 2013), Valetta, Malta, September 2013.

3.4 Gartner, Richard (2008) Metadata for digital libraries: state of the art and future directions, JISC Technology and Standards Watch Reports, [online] Available from:

http://www.jisc.ac.uk/media/documents/techwatch/tsw_0801pdf.pdf

3.5 Tobias Blanke, Leonardo Candela, Mark Hedges, Mike Priddy, Fabio Simeoni, Deploying General-Purpose Virtual Research Environments for Humanities Research, Phil. Trans. R. Soc. A 28, vol. 368 no. 1925, 2010. doi:10.1098/rsta.2010.0167

3.6 Mark Hedges, Michael Haft, Gareth Knight, Encouraging data sharing and reuse in the freshwater science community, Journal of Digital Information, Vol. 13, No. 1 (2012)

4. Details of the impact (indicative maximum 750 words)

The impact relates to three main categories: improving compliance with water quality standards; mitigating economic risks; and improving understanding through open government and information. The impact has the following key beneficiaries:

- Government agencies and non-governmental agencies with responsibility for environmental quality;
- Farmers and land-managers, who have access to information that assists them in mitigating the adverse effects on the environment while maintaining agricultural production levels;
- The general public, which benefits from the drive towards environmental improvements.

In the UK, there has been a long-standing water quality problem of diffuse pollution (pollution sources that may be small individually but which are collectively damaging). Therefore, compliance with water quality standards (such as the EU's Water Framework Directive) is challenging, particularly when the underlying causes of the pollution are not clear. DDH's work means that compliance with water quality standards has improved and pollution has been reduced as a result of better recording and managing of key information. This has had immediate benefits to the public in terms of a cleaner and safer environment and more diverse ecosystems. Michael Dobson, Former Director of the FBA (Freshwater Biological Association), has commented on DDH's impact in this area: 'The research... in the fields of digital archives, information management, data analysis and visualisation, and data modelling and description standards… has had significant impact on freshwater science... [Our collaboration] has led to a better understanding of pollution processes in inland waterways and lakes, and has improved the quality of information available to government and regulatory authorities regarding water quality, thus playing a key role in UK government-funded activities to improve environmental standards and enabling compliance with the EU's Water Framework Directive' [5.1].

Beyond the legal and ethical requirement to monitor and improve water quality, there are economic ramifications should the UK fail to meet the Water Framework Directive requirements. Foremost among these is the risk of financial penalties from the European Union, which can be as high as a €28 million fine plus €120,000 daily penalties. According to Dr Bill Brierley, Monitoring Advisor for the Environmental Agency: 'Approximately 64% (by length) of English rivers currently fail the current Water Framework Directive phosphorus standards and diffuse agricultural pollution is one of the significant reasons for failure of ecological status in England and Wales – the task of reducing nutrient levels is a very significant one for both the Environment Agency and UK plc' [5.2].

The EU is prepared to implement steep penalties for river pollution and came close to doing so in 2007 when France failed to improve water quality. As Brierley has said: 'The archive [created by the DDH team] allows us to comply with the Water Framework Directive – the penalties for not doing this are financial as well as environmental and reputational. Without the work of the King's

team it would be very difficult for us to monitor, understand and demonstrate compliance or identify the sources of diffuse pollution within agricultural catchments. It is only when we have this detailed information that we can, together with other stakeholders, identify, undertake cost benefit analyses of and prioritise both mitigation and management actions' [5.2].

Access to data and freely available evidence is a key priority for the UK government. The knowledge gained through this information management research is being communicated to farmers, leading to improved practices in relation to fertilisation of agricultural land so that pollution is reduced while land productivity is maintained. This integrated data sharing supports the work of the Environment Agency and facilitates stakeholder engagement. As Dr Brierley of the Environment Agency says: 'The archive has, for the first time, collated and stored information, monitoring data and evidence and makes this available to the Environment Agency, Defra, academics and members of the public via web portals. This ensures that the data is available for anyone to access, use, analyse and scrutinise. ... This is one of the underlying principles of the River Basin Management Plans in the Water Framework Directive' [5.2]. Moreover, according to the Science Programme Manager at the Department for Environment, Food and Rural Affairs (Defra): 'The implementation of an archive consistent with open data standards, allowing linked data approaches to analysis, is important to achieve the vision... KCL has taken a leading role in developing appropriate access restrictions to ensure data can be shared effectively in accordance with the INSPIRE directive and open data strategy but within the limits of the data protection act. Their work will ensure the investment in data collection continues to have impact without disclosing sensitive information about stakeholders' [5.3].

Beyond the Water Frameworks Directive, DDH has collaborated on other Defra-funded activities, including meeting the INSPIRE (geographical data recording) Directive and further developing the freshwater data model to support another research platform, namely the Agricultural Greenhouse Gas Platform. This platform supports activities to mitigate climate change. As Dr Dobson of the Freshwater Biological Association states, this research informs the UK government 'about the practicalities of implementing the EU's INSPIRE Directive... and contains a wide range of rules for managing geo-spatial information. Defra is responsible for overseeing the implementation of INSPIRE in the UK, and the collaborations are a showcase for implementing the Directive in data archives, and for pioneering the open publication of Defra's scientific data in accordance with the Directive. The research carried out by CeRch [DDH] was a key factor in enabling these results, and we expect the impact that it has had to date to continue to grow further over the coming years' [5.1].

5. Sources to corroborate the impact (indicative maximum of 10 references)

5.1 Former Director of the Freshwater Biological Association between 2007 and 2013. Now Principal Freshwater Consultant at APEM Ltd, a leading environmental consultancy specialising in freshwater and marine ecology and aerial surveys (statement uploaded).

5.2 Monitoring Advisor within the Evidence Directorate of the Environment Agency (statement uploaded).

5.3 Science Programme Manager, Sustainable Land and Soils, Defra.