Institution: University of Portsmouth

Unit of Assessment: 36 Communication, Cultural and Media Studies, Library and Information Management

Title of case study: Preserving the Digital Future: The impact of the TOTEM (Trustworthy Online Technical Environment Metadata) registry on preservation professional policy and practice.

1. Summary of the impact

A global consortium of libraries has adopted the innovative TOTEM registry data model to address urgent issues surrounding the preservation of digital artefacts. The core challenge for digital archiving is to match potentially obsolete software that originally created artefacts – 'complex' objects with sound and visuals as well as data information – with later computing platforms that can thus preserve them. The TOTEM project has effected major change in the technical specifications of preservation: its technical strategy for 'emulation' enhances previous processes through which old files are 'migrated'. End-users confirm that TOTEM has had significant **cultural and technical impact on the preservation practices of national libraries** including the Netherlands, New Zealand and Australia, and US National Archives and Records Administration. Benefit to these organisations is technical, societal and economic, contributing to viable, long-term solutions in digital preservation policy.

2. Underpinning research

The prevailing digital preservation strategy is migration: as file formats or computing environments (software, operating systems, hardware) become obsolete, files are altered *en masse* to fit new formats. For libraries faced with the sheer scale of digital material, this practice is unviable in the long term, and migration alone cannot cope with increasingly complicated types of digital material. Led by Dr *Janet Delve*, TOTEM is the result of fundamental research into the metadata required to create stable, long-term solutions to preservation through a strategy of *emulation*. This approach leaves old files unchanged but recreates - through software – the functionality of the original operating hardware. Emulation requires a much greater degree of complexity in the recording of the metadata describing 'technical environments' - software, hardware and the complex web of relationships between them - than migration.

Discerning the metadata necessary to describe and locate such technical environments was a key area of research. Working within the EU FP7-funded Keeping Emulation Environments Portable (KEEP) project, EU Grant Agreement ICT 231954 [£4m], (2009 – 2012), *Delve*'s research focussed on developing the conceptual and technical framework for digital preservation in specifications of format registries, tools for format characterisation, recognition, validation and metadata extraction under the emulation strategy.

TOTEM is the resulting migration/emulation hybrid tool that matches older versions of software with compatible versions of operating systems/hardware models. Its research innovation lies in the ways in which it has modeled semantic operability protocols that can be implemented as a relational database and as an RDF (Resource Definition Framework) for linked data. It is this potential that was taken up by a global consortium of libraries that were persuaded to meet the challenge of digital preservation through the TOTEM strategy.

Research was conducted in three stages, with end-user involvement throughout:

1. The KEEP project examined the existing preservation systems of national libraries in France, Germany and the Netherlands and the Computer Games Museum in Berlin to determine the extent of current emulation practices, future plans to use emulation and the type of metadata standards needed. These standards are the foundation of digital archiving, providing metadata to describe the information resource, support its identification, location and retrieval thus facilitating content and access management. Framed by the user-group needs, the results of analysis identified the

potential for emulation-based processes and were published in EC report (**1**). This work analysed current metadata standards informing international practice such as *Preservation Metadata: Implementation Strategies* (PREMIS), *Metadata Encoding and Transmission Standards* (METS) and *Consortium of University Research Libraries Exemplars for Digital Archives* (CEDARS).

2. Using the findings, *Delve* carried out fundamental analysis of the metadata (needed to describe the technical environments necessary to render under emulation) by using five different digital objects (e.g. a radar simulation for a racing boat training package) from the German National Library. A suite of conceptual metadata models was created, together with clear explanations of its utility for a variety of stakeholders (cultural 'memory' institutions - museums, libraries archives - and computer games communities). EC Report (**2**).

3. With user-based research, *Anderson & Delve et al*. challenged the pervading dichotomy - *either* emulation *or* migration (**3**). *Delve* reviewed the ontologies emanating from the influential PLANETS project led by the British Library, indicating where KEEP metadata work would significantly extend software and hardware OWL ontologies for emulation. A logical relational data model was developed by *Delve* and *Konstantelos* and implemented as a MySQL online database - TOTEM. TOTEM was subjected to rigorous user evaluation/testing and metadata schema created a new environment entity (*Konstantelos & Delve*) (**4** and **5**). This was incorporated into PREMIS 3.0, a metadata schema allowing environments with links to technical registries like TOTEM to be described for many preservation actions: migration, emulation, virtualization and normalization (**6**).

3. References to the research

1. Anderson, D., Delve, J., Pinchbeck, D., Alemu, G.A., & Ciuffreda, A. (2009) Preservation metadata standards for emulation access platforms. FP7 Report to the European Commission. 85pp (deliverable D3.1), click here to download

2. Delve, J., Ciuffreda, A., Anderson, D., Pinchbeck, D., Joguin, V. (2010) Documents describing meta-data for the specified range of digital objects, as well as requirements and design for the browsing system and user interface of the Emulation Framework. FP7 Report to the European Commission. 72pp. REF 2 output: 36-JD-003

3. Anderson, D., Delve, J. & Pinchbeck, D. (2010) 'Towards a workable, emulation-based preservation strategy: rationale and technical metadata'. *New Review of Information Networking*, 15 (2). pp. 110-131. ISSN 1361-4576 DOI: 10.1080/13614576.2010.530132. Journal Article.

4. Delve, J., Konstantelos, L., and Anderson, D. (2011) Requirements and design document and database implementation for the KEEP Emulation Framework GUI. Report to the European Commission. 23pp. *Available on request.*

5. Konstantelos, L., Delve, J., Anderson, D. (2012). Document recommending scalable generic metadata structures for international archiving standards, Report to the European Commission. 34pp. *Available on request.*

6. Dappert, A., Peyraud, S., Delve, J., Chou, C., (2013) 'Describing Digital Object Environments in PREMIS'. *New Review of Information Networking*, 18 (2) pp. 1-68. ISSN 1361-4576. DOI: 10.1080/13614576.2013.842494 Journal Article. REF 2 output: 36-JD-004

4. Details of the impact

The EU FP7 framework of the KEEP project located research in a European-wide range of partnerships with academic researchers, digital preservation agencies, professional archivists and librarians. The early research framework was developed by analysis of existing processes but also through workshops and conferences designed to maximise 'user' perspectives. Workshops were held to involve professional communities directly in the design, development and implementation of TOTEM at French National Library, Paris (2011/2012); Royal Dutch Library, the Hague (2011); alongside the InFuture2011 conference, Zagreb (2011); Italian Video Games Developer

Conference, Rome (2011), Novotel Cardiff (2012); German National Library, Frankfurt (2012) and the Computer Games Museum, Berlin (2012). These events were used to disseminate research to show how TOTEM contributed to preservation policy and strategy: structured tutorials, using the live online TOTEM database, were used to trial applications from which written feedback was used to improve the functionality and performance of TOTEM. Conferences were also used to advocate the use of this emulation registry, and to trial and test in applied contexts of use: at GOPORTIS (Leibniz Library Network for Research Information consisting three leading German Libraries) in 2011, iPRES 2011 and at Open Planets Foundation workshops. Those present included key figures from Portico (JSTOR), Harvard Business School, and the Library of Congress who saw for the first time that emulation was now a practical proposition due to the presence of tools such as TOTEM, having succeeded in designing suitable data for the KEEP Emulation Framework (http://emuframework. sourceforge.net/).

Utility and uptake

Delve and *Anderson's* collaboration with Chair of Digital Humanities, Professor M.Thaller at the University of Cologne and Johanna Puhl converted the created models into a form that could be used on the Web. Potentially, all data that is already 'linked' (http://linkeddata.org/) can be used to provide information about sharing and connecting in computing environments.

To ensure wide compatibility with existing initiatives, TOTEM includes a link in the file format data to the PRONOM registry http://www.nationalarchives.gov.uk/PRONOM/Default.aspx. At GOPORTIS, TOTEM was presented as part of a global 'eco—registry' initiative alongside the UK National Archives PRONOM format registry. This practical application of TOTEM's functionality was identified as an exemplar of best practice by S. Knight, Program Director of the Preservation Research & Consultancy division of the National Library of New Zealand (http://www.digitalpreservation.gov/series/edge/newzealand.html, http://natlib.govt.nz). Knight recognised TOTEM as a significant tool upon which to build a future meta registry for use in 'memory' institutions worldwide. This led to global strategic collaboration for its adoption with the National Library of New Zealand (NLNZ), the National Library of Australia (NLA), their registry of media (http://www.nla.gov.au/mediapedia) and the US National Archives and Records Administration (NARA) (http://www.archives.gov/), conducted under the auspices of the National and State Libraries Australasia (NSLA) from 2011- present. *Delve* and *Anderson* contributed further to this development in Wellington, New Zealand in March 2013, together with key personnel from NLA, NLNZ and NARA, produce the model behind the meta-registry.

From discussions and communications with industry professionals, key utilities of TOTEM have been defined as:

- Preservation planning
- Digital content review to consider preservation strategies
- Checking the compatibility of given formats with software
- Gathering information on platforms
- Investigational work around support/dependencies on software and data
- Finding formats; software and operation system combinations for emulation paths
- Finding tools for working with files
- Discovering software/hardware tools necessary for handling particular files
- Identifying software/hardware configurations for testing and research
- Identification of problem file types from legacy collections
- Working with legacy digital material not yet investigated
- Cross-checking metadata terms to ensure consistency
- Gleaning information about hardware
- Appraisal of electronic records
- Documenting the AIP (Archival Information Package)
- In the context of formats, migration to identify file formats readers
- Documenting reference environments for data assets

By collaborating with organisations such as the Digital Preservation Coalition

http://www.dpconline.org

and Open Planets Foundation

http://www.openplanetsfoundation.org/blogs/2013-04-01-software-archiving-eaas),

TOTEM is being validated and implemented by the digital preservation community, with robust data entry methods being employed. Current projects such as bwFLA at the University of Freiburg (http://bw-fla.uni-freiburg.de/wordpress/?page_id=7) are using TOTEM as the basis of future research.

£3 billion of public finance is invested annually in research in the UK alone yet the full economic and cultural value of research in all disciplines is lost where it cannot be preserved. TOTEM directly contributes to the sustainability of global digital architectures, lowering barriers to effective conservation and curatorial management of collections, and preserving cultural value through technical application of research to preservation policy and practice.

5. Sources to corroborate the impact

1. Digital Preservation Coalition - TOTEM was Finalist for Award in 2012 for 'an outstanding contribution to research and innovation in digital preservation in the last 2 years'.http://www.dpconline.org/advocacy/awards/2012-digital-preservation-awards/928-finalists-2012-research-and-innovation

2. Letter of support from Executive Director, Digital Preservation Coalition cites TOTEM as 'not just clever, it is timely. By understanding technical environments it makes virtualisation a working possibility for long term access, anticipating the elasticity in provision that on-demand preservation-as-a-service is certain to need in the next decade'.

3. Digital Preservation Coalition Editorial Board review, citing Delve's lead role on the editorial board as expert in community-led preservation research. This letter includes data on downloads for Technology Watch Reports (ISSN: 2048-7916) showing demand and reach within the international digital preservation community. Also cited by US Library of Congress as one of the 'Top Ten Digital Preservation Developments of 2012'.

4. Letter of support from Program Director of the Preservation Research & Consultancy division of the National Library of New Zealand, cites adoption of TOTEM collateral as component in construction of NSLA Digital Preservation Technical Registry

5. Confidential tender from NLNZ to NSLA to build technical registry through TOTEM in 'Re-Imagining Libraries' project (2012)

6. Letter of support from Manager of Web Archiving and Digital Preservation Branch, National Library of Australia.

7. Letter of support from Head of Digital Scholarship at The British Library, cites TOTEM's role in development of technical environments (rather than simple properties of digital objects) as making an 'essential shift' in digital preservation.