

JISC, the DPC and the UK Web Archiving Consortium Workshop



Date: 21st July 2009

Venue: [The British Library Conference Centre](#) , London

- [Full Conference Report \[PDF 337KB\]](#)
- Presentations from the conference (see below)
- [Selection of images from the conference](#)
- [Conference report from Peter Webster and Jane Winters, Institute for Historical Research](#)
([external website](#))
- [Conference report from Alexandra Eveleigh, West Yorkshire Archive Service \(Ariadne\)](#)
- [Conference report from Jonathan Clark's Blog](#)
- [Conference report from Marieke Guy](#)

The web runs at risk. Our generation has witnessed a revolution in human communications on a trajectory with the origins of the written word and language itself. Early web pages have an historical importance with prehistoric cave paintings or proto-historic pressed clay ciphers. They are just as fragile. The ease of creation, editing and revising gives content a flexible immediacy: ensuring that sources are up to date and, with appropriate concern for interoperability, content can be folded seamlessly into any number of presentation layers. How can we carve a legacy from such complexity and volatility?

Key issues for long-term access and preservation remain unresolved. How can content creators make sure their creations are durable without impairing their flexibility? How does web-archiving relate to data-curation and traditional archiving? What constitutes an appropriate legacy from a web site? What audiences should web archives anticipate and what does this mean for selection, ingest and preservation? What will the web be like as an historical source, and what use will be made of archived web sites by future generations? How will they validate them? How will they cite them? What are our missing links? How can these be filled?

The challenges of web archiving have long been recognised and there are a number of tools and services that already offer – or purport to offer – long-term access to web content. But gaps remain in policy, expertise and implementation and the tools for web-harvesting need a clearer link between the technical needs of preservation services and the deferred needs of user communities.

Only by developing and strengthening the links between content creators, tools developers, preservation services and users can we hope to secure an enduring web.

Sponsored by the Digital Preservation Coalition (**DPC**) and the Joint Information Systems Committee (**JISC**) the six partners of the

UK Web Archiving Consortium

(British Library, National Library of Wales, JISC, Wellcome Library, The National Archives and the National Library of Scotland) are organising a joint workshop on the 21st July 2009 at the British Library Conference Centre, Euston Road, London.

This event will bring together key stakeholders – **archive managers, preservation experts, national libraries, web archivists and content providers**

- for practical and focussed discussion on shared perspectives, requirements, problems and solutions. Formal presentations and case studies will be presented with an opportunity for posters and demonstrations of tools. The day will close with a plenary discussion and a follow up Report.

Who should come?

- Web editors and web site managers
- Web archivists
- Archive managers
- Digital preservation and curation specialists
- Researchers interested in the web as an historical resource
- Librarians and policy makers with interests in electronic legal deposit

Why should you come?

- Help to shape the UK's web archiving agenda
- Learn from organisations currently preserving websites
- Share your preservation concerns and aspirations
- Understand new services and projects working in web archiving

- Anticipate the evolution of a rapidly changing field
- Contribute your web archiving expertise to the repository community

Online registration is now closed, however if you wish to be added to the reserve list please send your contact details to:

webconference@confpeople.co.uk

Places are limited. We encourage early registration and there will be no facility to register on the day. Although the event is free of charge, there will be a penalty for those who register and don't come. Preference will be given to DPC Members.

Can't come but still want to participate?

Questions can be sent in advance to the chair of the organising committee by email to info@dpconline.org

, and elements of the event will be blogged live during the day. A conference report will follow on the DPC website and sent out via the

dpc-discussion@jiscmail.ac.uk

and

digital-preservation@jiscmail.ac.uk

lists where discussion will follow. A full report will follow.

Programme

0900-09:30

Registration

0930-09:40

Welcome & introduction

William Kilbride, DPC Executive Director & Neil Grindley, Programme Manager - Digital Preservation JISC

Session 1: Setting the scene

09:40-10:20

[Keynote & questions \[PDF 2.37MB\]](#)

Adrian Brown, Assistant Clerk of the Records, Parliamentary Archives

10:20-10:40

[Web Archive and Citation Repository in One: DACHS \[PDF 1.33MB\]](#)

Hanno Lecher, Librarian, Sinological Library, Leiden University, the Netherlands.

10:40-11.00

[The future of researching the past of the Internet \[PDF 7.02MB\]](#)

Eric T. Meyer, Research Fellow, Oxford Internet Institute, Oxford University

11.00-11.30

Morning coffee break

Session 2: Creation, capture & collection

11.30-11.50

[Web Archiving Tools: An Overview \[PDF 554KB\]](#)

Helen Hockx-Yu, Web Archiving Programme Manager, The British Library

11.50-12.10

[Context and content: Delivering Coordinated UK Web Archive to User Communities \[PDF 2.03MB\]](#)

Cathy Smith, Collections Strategy Manager, The National Archives

12.10-12.30

[Capture and Continuity: Broken links and the UK Central Government Web Presence \[PDF 1.45MB\]](#)

Amanda Spencer, The National Archives.

12.30-12.45

Questions and discussion

12.45-13.30

Lunch break and exhibition

Session 3: Issues and approaches to long term preservation of web archives

13.30-13.50

[**Diamonds in the Rough: Capturing and Preserving Online Content from Blogs \[PDF 2.07MB\]**](#)

Richard Davis, Project Manager, University of London Computing Centre (ULCC)

13.50-14.10

[**Beyond Harvest: Long Term Preservation of the UK Web Archive \[PDF 577KB\]**](#)

Maureen Pennock, Web Archive Preservation Project Manager, The British Library.

14.10-14.30

[**From Web Page to Living Web Archive \[PDF 9.55MB\]**](#)

Thomas Risse, Senior researcher, L3S Research Center

14.30-14.50

[Emulating access to the web 1.0 \[PDF 3.05MB\]](#)

Jeffrey van der Hoeven, Koninklijke Bibliotheek, The Netherlands

14.50-15.00

Questions

15.00-15.15

Afternoon tea break

Session 4: Discussion and next steps

15.15 – 15.40

[What we want with web-archives; will we win? \[PDF 1.07MB\]](#)

Kevin Ashley, Head of Digital Archives, University London Computer Centre (ULCC)

15.40-16.40

Roundtable – all speakers in question time format

16.40 - 16.45

Thanks and close