

# PoWR: Explaining Web Preservation

Kevin Ashley, ULCC

# What might be kept ?

- Information content
- Information appearance
- Information behaviour
- Information relationships
- Change history
- Usage history

# Content preservation

An 'at the event' report on the first JISC PoWR workshop held at Senate House Library, London on Friday 27th June 2008 has been published in the recent Ariadne Web Magazine (issue 56, July 2008). The piece, written by Stephen Emmott, concluded:

The challenges are significant, especially in terms of how to preserve Web resources. No doubt the institutional repository will play a role. Arguably, the absence of a solution to the preservation of Web resources leads to either retention or deletion, both of which carry risks. The workshop's core message to practitioners was therefore to start building an internal network amongst relevant practitioners as advice and guidance emerge.

My thinking about this matter was certainly stimulated and I look forward to the next two workshops, and the handbook that will...

# Preserving Appearance

## JISC-PoWR

Preservation of Web Resources: a JISC-sponsored project

[HOME](#) [ABOUT](#) [CASE STUDIES](#) [CONTACT](#) [WORKSHOPS](#) [TEAM AREA](#)

### PoWR Themes

[Case Studies](#)

[Challenges](#)

[Digital Preservation](#)

[Events](#)

[Policies](#)

[Preservation](#)

[Project News](#)

[Records Management](#)

[Resources](#)

[Selection](#)

[Technologies](#)

[Web 1.0](#)

[Web 2.0](#)

### Ariadne Report on the JISC PoWR workshop

Posted by [Marieke Guy](#) on September 2nd, 2008

An 'at the event' report on the first JISC PoWR workshop held at Senate House Library, London on Friday 27<sup>th</sup> June 2008 has been published in the recent [Ariadne Web Magazine](#) (issue 56, July 2008). The piece, written by Stephen Emmott, concluded:

The challenges are significant, especially in terms of how to preserve Web resources. No doubt the institutional repository will play a role. Arguably, the absence of a solution to the preservation of Web resources leads to either retention or deletion, both of which carry risks. The workshop's core message to practitioners was therefore to start building an internal network amongst relevant practitioners as advice and guidance emerge.

My thinking about this matter was certainly stimulated and I look forward to the next two workshops, and the handbook that will result. Web preservation is an issue which was always important

Search

### Calendar

September 2008

M	T	W	T	F	S	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

[« Aug](#)

### Partner blogs

[Opencontentlawyer](#)  
[UK Web Focus](#)  
[ULCC DA Blog](#)

### Partners

[JISC IIE](#)  
[UKOLN](#)  
[ULCC](#)

# Preserving Behaviour

## Recent Comments


\* Christian van der Ven on Web 2.0 and Archives: Something like a Phenomenon?

\* Maureen on Web 2.0 and Archives: Something like a Phenomenon?

 \* Christian van der Ven on Web 2.0 and Archives: Something like a Phenomenon? ("da.blog (ulcc digital archives blog)")

 Sunderland, Sunderland arrived from google.co.uk on "da.blog (ulcc digital archives blog)" » Blog Archive » Digitising William Morris's Lantern Slides (Part 2)"

 New Zealand arrived from google.co.uk on "da.blog (ulcc digital archives blog)"

 Bern, Bern arrived on "da.blog (ulcc digital archives blog)" » About the blog"

working  
to promote their collections. Does  
on in the archival setting? I don't  
or  
in thoughtful consideration of the value  
ur work. I'd also love some discussion

enefits have they brought? What kind of  
ese audiences stick around? Is the  
th it? Do these audiences engage with  
ur databases? come in and use our  
they don't? How is our adoption of  
rocess and promote? With  
documentation skills, cataloguing  
lise our research collections? Any  
Yoxall, Archives Manager, Registrati  
00 2011/04/04 11:00 AM GMT+01:00

- + JISC DigiRepWiki
- + JISC Repositories Research Team
- + TASI

## Project feeds

- + JISC-PoWR » Ed Pinsent
- + JISC-PoWR » Richard M. Davis

## Archives

Select Month

# Meta



## the rise o

Richard M. Davis

learn today from Rhodri  
c Blog that the Number 10  
Wordpress over a previous  
m.

of Wordpress for a while now. I  
at it circa 2004, for an internal  
stem, when we needed an  
foul of Movable Type's new  
ts. (I think MT's now reverted to  
sir paid-for interim probably did  
vour.)

it successfully for DA Blog and several personal projects - not just for blogging, but  
CMS. We've also seen the rise of Wordpress.com, Edublogs, and, closer to home,  
routed by many web illuminati, such as Brian Kelly and Jill Walker. Wordpress is

certainly looks fresher than of old. Earlier news entries have all been imported - not  
ad the rest of this entry »



DETIS in Bolton

Watch in Real-Time

vacancies at ULCC

# Other things to preserve

- Relationships:
  - links behave
  - associated metadata survives
  - Styles and content stay related
- Usage/change logs: obvious what they are, but not whether they are needed

# Techniques

- Save within the authoring system or server
- Save appearance at the browser
- Harvest content with crawlers



# Capturing on the server

- Easy (?) if it's your server
- Captures raw information, not presentation
- May be too dependent on authoring infrastructure or CMS
- Works in short to medium term, for internal purposes
- Not good for external access



# Capture post-rendering

- You get what you see: but you don't know why
- It's relatively simple for well-contained sites
- Commercial tools exist
- Treats web content like a publication: frozen
- Loses behaviour and other attributes

# Harvesting

- Most widely-used
- Presents many problems for capture – often don't get everything (or too much)
- Defers some access issues:
  - Link re-writing
  - Embedded external content: from archive or live ?
- Lots of work, tools and experience

# When?

- What triggers things ?
- A regular schedule (yearly, monthly, termly....)
- When stuff changes (regular crawls, but throw away unchanged content)
- Manual initiation
- Intelligent agents
- Transactions