



INSTITUTE
OF HISTORICAL
RESEARCH

University of London
School of Advanced Study

Big UK Domain Data for the Arts and Humanities

Jane Winters

(Reader in Digital Humanities, Institute of Historical
Research)

Web Archives for Arts and Humanities Research, 26 Feb. 2014



AHRC Big Data Call

- Part of the Digital Transformations theme
- Aim to address the challenges of working with big data
- A total of 21 projects funded
- A number of large-scale collaborative projects, of which BUDDAH is one

[Home](#)[About](#)[Blog](#)[Contact](#)[Documentation](#)[News and events](#)

Big UK Domain Data for the Arts and Humanities (BUDDAH)

Web archives are an increasingly important resource for arts and humanities researchers, yet we have neither the expertise nor the tools to use them effectively. Both the data itself and the process of collection are poorly understood, and it is possible only to draw the broadest of conclusions from current analytical analysis. The BUDDAH project will work with the dataset derived from the UK domain crawl from 1996 to 2013 (that is, when legal

RECENT POSTS

- [Collect, preserve, access – 10 years of archiving the online government record](#)
- [Announcing BUDDAH](#)

ARCHIVES

- [February 2014](#)



INSTITUTE
OF HISTORICAL
RESEARCH

University of London
School of Advanced Study



Project partners

- Institute of Historical Research – Jonathan Blaney and Jane Winters
- British Library – Helen Hockx-Yu, Andy Jackson and Peter Webster
- Oxford Internet Institute – Eric Meyer, Ralph Schroeder and Josh Cowls
- Aarhus University – Niels Brügger



Aims

- To highlight the value of web archives for research
- To develop a theoretical and methodological framework for the analysis of web archives
- To explore the ethical implications of big data research
- To inform collections development and access arrangements at the British Library
- To train researchers in the use of big data



What will we produce?

- To enhance the interface to the 1996-2013 data at the British Library
- To develop new tools which will support analysis of the data
- To write a history of UK web space 1996-2013
- To produce a series of case studies
- To develop an online training module

Filter by crawl_year:

- 2004 (3807)
- 2003 (2700)
- 2007 (2374)
- 2005 (1976)
- 2006 (1872)
- 2002 (1210)
- 2008 (1142)
- 2009 (793)
- 2001 (538)
- 1999 (195)
- 1997 (176)
- 2000 (167)
- 1998 (74)
- 1996 (4)

Search

Enter your search terms

Sort by

Crawl Date ▾

Order

Asc ▾

Search

Clear

1. **Title:** Services for research (University of Yor...**Crawl Date:** 1996-12-07T00:17:32Z [more]**Host:** tower.york.ac.uk (html)2. **Title:** University of York Department of History**Crawl Date:** 1996-12-07T03:11:18Z [more]**Host:** tower.york.ac.uk (html)Filter by
links_public_suffixes:

- ac.uk (10396)
- co.uk (7430)
- org.uk (7347)
- gov.uk (6338)
- net (4905)
- org (4382)
- com (3577)
- edu (1570)
- bl.uk (1307)
- ca (644)
- gov (630)
- de (516)
- nl (516)
- ie (357)



Case studies

- 10 bursaries of £2,000 each
- Range of arts and humanities disciplines
- Open to researchers in universities, libraries, archives and museums, as well as independent researchers
- Work with developers at the British Library, and produce a 2,000-word case study
- Closing date 25 April



Possible subjects

- Health and education
- Material culture and display
- Political and literary reputation
- Public engagement, and social and political movements
- Patterns of culture
- Networks of social communication
- History of institutions and organisations
- The changing structure of the web etc.



INSTITUTE
OF HISTORICAL
RESEARCH

University of London
School of Advanced Study

<http://buddah.projects.history.ac.uk>