Why do we do it? Understanding the Research Value of Web Archives

Pamela Graham
Director, Humanities & Global Studies
Director, Center for Human Rights Documentation & Research
Columbia University Libraries
Challenges of user access to web archives

Diversity of Web Archiving purposes and relevance to use

Researcher initiatives
Challenges for Users of Web Archives

- Discovery Curve
- Cultural Understanding of the Web
- Functional limitations of Web Archives
- Lack of scholarly standards and practices
<table>
<thead>
<tr>
<th>Web as ....</th>
<th>Content</th>
<th>Use</th>
<th>Discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repository</td>
<td>Discrete objects; stand alone value; analogous to pre-web forms;</td>
<td>close reading/viewing and analysis; used in tandem with non-web</td>
<td>Item / series-level cataloging and description; connection to related non-web items</td>
</tr>
<tr>
<td>Platform / field site</td>
<td>Websites, “web spheres”; Context is important</td>
<td>Analysis of language use, networking and linking patterns, “utilitarian” functions</td>
<td>Collection level / website level description, web search as research, data-mining, analytical and visualization tools</td>
</tr>
<tr>
<td>Technology</td>
<td>Code, CSS, scripts, design elements</td>
<td>History of the web; programming; experience of the web</td>
<td>Help!</td>
</tr>
</tbody>
</table>
BUDDAH

• Better methods, frameworks for web archive research

• Build skills and awareness of potential value

• Inform collection development

• Demonstrate overall research value of web archives

https://buddah.projects.history.ac.uk/
Online reactions to institutional crises: BBC Online and the aftermath of Jimmy Savile by Rowan Aust. Rowan examines the way the BBC has responded to the scandal, given that Savile is present in so many archived programmes and articles.

’all writing is in fact cut ups’: the UK Web Archive and Beat literature by Rona Cran. Rona’s study is on the reception of Beat literature in the UK, in both academia and among the general public, at the end of the 20th century and the beginning of the 21st.

Revealing British Eurosceptism in the UK Web Domain and Archive Case Study by Richard Deswarte. Richard reflects on studying his topic via web archives, with the advantages and pitfalls of this methodology.

Digital barriers and the accessible web: disabled people, information and the internet, by Gareth Millward. Gareth explains how his plans for using the web archive to study what was available to the disabled had to be changed because of the particular challenges of using web archives for research.

Do online networks exist for the poetry community? by Helen Taylor. Helen used the UK web archive to investigate poetry communities and how they interact with each other, with a particular focus on The Poetry Forum and the online presence of the Oxford University Poetry Society.
RESAW, a Research Infrastructure for the Study of Archived Web Materials, is a network established in 2012, aiming at promoting a collaborative European research infrastructure for the study of archived web materials.

Tweets from https://twitter.com/resaw_eu/lists/resaw

Stories and News

Peter Webster, Web Historian, UK

Bernstein

Ian Milligan, Web Historian, Canada

Web Arch shop

Web Archives for Historians

Multidisciplinary Event Documentation: “The From Da

WORD Adventure” on our D
AUT PROJECT

Our goal is to make petabytes of historical internet content accessible to scholars and others interested in researching the recent past.
The collection

- Starting with idea of working with London 2012 and Rio 2016 Olympic collaborative crawls
  - Dataset was too large
- Ended with 2016 Rio National Olympic and Paralympic Committees collection:
  https://archive-it.org/collections/72227

AU 4.0 The British Invasion, June 2014
Welcome to the Archives Unleashed Project

Welcome

Archives Unleashed aims to make petabytes of historical internet content accessible to scholars and others interested in researching the recent past. Supported by a grant from the Andrew W. Mellon Foundation, we will be developing web archive search and data analysis tools to enable scholars and librarians to access, share, and investigate recent history since the early days of the World Wide Web.

Interested in the project? Subscribe to our mailing list! Or you can follow the links at left for information about the project, the Archives Unleashed Cloud, Archives Unleashed Toolkit, or our events.

We’re always looking for ways to engage archivists, librarians, researchers, developers, or any others interested in born-digital heritage!

Contact Us

Questions? Comments? Please contact us, either by leaving an issue on one of our GitHub projects or by sending us an e-mail. Are you a SIG user? Join our Slack team!

CC-BY 2018 | Site built with Hugo using the Material theme

http://archivesunleashed.org/