The meaning and value of web archives for research

Deutsche Nationalbibliothek
28 November 2018

Dr Peter Webster
Webster Research and Consulting
@pj_webster
On the contemporary religious history of the Web


• ‘Lessons from cross-border religion in the Northern Irish web sphere….’ in The Historical Web and Digital Humanities. The case of national web domains (2019)
The web its own archive?

Open UK Web Archive 2004-13 comparison.
@anjacksonn http://britishlibrary.typepad.co.uk/webarchive/2014/10/what-is-still-on-the-web-after-10-years-of-archiving-.html
Who might use web archives?

• journalists
• activists
• lawyers
• anyone with a stake in a record of the recent past
Planned disappearance

Loading

North & South Tipperary are Merging - Click for updates

LATEST NEWS

- Planning Application Register Reference number 12/399
- Burst Watermain at Ballydoyle Road
- New Funding Scheme for Built Heritage
- Christmas Closing Times & Contact Details
- “No more smoke from your fire.” Delivering Cleaner Air to Clonmel
- Water Burst - Clonoulty Area
- Latest Newsletter from the Library
- Planning Application Register Reference number 13/10 Roadstone Wood Ltd
- Traveller Accommodation Programme

Follow us on Twitter - Subscribe to RSS

Office Opening Hours

South Tipperary County Council has recently revised its office opening hours as follows:-
Monday to Friday (except public holidays)
9.30 a.m. to 4.30 p.m. including lunchtime opening.

Motor Tax Office opening hours remain at 9.05 a.m. to 1.30 p.m.
Unplanned disappearance
‘Perhaps Stanley will counter that at the technical end of his period, 2000, the internet was not what it has become eighteen years later…. [but] the internet does not merit an entry in Stanley’s index, yet it has changed everything.’

Diarmaid MacCulloch, reviewing Brian Stanley, *Christianity in the Twentieth Century* (TLS, Sept 7th 2018)
The emerging discipline of Web history

Conferences: ReSAW 2015, 2017, 2019

Journal: Internet Histories (2017-)


Case studies: Web History (Peter Lang, 2010); The Web as History (UCL Press, 2017); Web25 (Peter Lang, 2017)
What is web archiving?

“deliberate and purposive collection and preservation of web material” (Brügger, 2018)

- very small- or very large-scale
- harvesting, screen capture, file delivery
- public, restricted, or no access
Who are the Web archivists?

- Internet Archive
- national libraries
- corporate archives
- research-driven: universities and individuals
- activists

National libraries

• 16 of 28 EU member states
• Iceland, Switzerland, Norway also
• Sweden the first (1996)
• some with legal deposit provision: Denmark (2005); France (2006), UK (2013)
Legal deposit web archiving: characteristics

• broad domain crawl, plus selective
• definition of the nation varies
• types of content included varies
• access restrictions
Selective harvesting

• in absence of NPLD, based on permissions
• part of the case for obtaining NPLD law
• key resources, eg. government, media
• events: elections, Olympics, Eurovision
• themes: political extremism, climate change
# Web archives in the UK

<table>
<thead>
<tr>
<th>Archive</th>
<th>Temporal scope</th>
<th>Content scope</th>
<th>Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open UKWA</td>
<td>2004-present</td>
<td>Selective</td>
<td>Online</td>
</tr>
<tr>
<td>Legal Deposit UKWA</td>
<td>2013-present</td>
<td>Comprehensive (for UK)</td>
<td>Onsite</td>
</tr>
<tr>
<td>JISC UK Domain Dataset</td>
<td>1996-2013</td>
<td>Comprehensive (for .uk)</td>
<td>Index only</td>
</tr>
<tr>
<td>UK Government Web Archive</td>
<td>1996-present</td>
<td>UK government</td>
<td>Online</td>
</tr>
<tr>
<td>Parliamentary Web Archive</td>
<td>2009-present</td>
<td>UK parliament</td>
<td>Online</td>
</tr>
<tr>
<td>Univ. of Oxford</td>
<td>2011-present</td>
<td>University sites</td>
<td>Online</td>
</tr>
</tbody>
</table>
University-based archiving

As records management

- Bodleian Libraries, Oxford

For research

- Innsbrucker Zeitungsarchiv
- Digital Archive of Chinese Studies [Leiden / Heidelberg]
DACHS - Digital Archive for Chinese Studies

Unser Ziel ist es relevante Internetseiten für die Chinaforschung zu archivieren und verfügbar zu machen. Hierbei liegt der Fokus insbesondere auf sozialen und politischen Diskursen, die durch das chinesische Internet wiedergespiegelt werden.

DACHS ist in der Entwicklung. Bitte haben Sie Verständnis, dass noch nicht alle Funktionen in vollem Umfang realisiert sind.

Schnellsuche

Titelwort / Schlagwort

Unbedingt beachten:
- Suche über englische Schlagwörter
- Wade-Giles (Pinyin - Konvertierungstabelle)
- nur Langzeichen
- Beispiel: "chinese dream" "chung-kuo-meng" "中國夢"
Five analytical strata

- element (paragraph, image, border, menu)
- page
- site
- Web sphere
- the Web as a whole

... each of which has visible and invisible aspects.

The Web element

**Visible**
- circulation of images or memes, embedded media

**Invisible**
- Anne Helmond on trackers (*Web25*)
- Brügger on the hyperlink (*Web25*)
Information on the Irish State

President

Houses of the Oireachtas
(Parliament of Ireland)

Taoiseach (Prime Minister)

Police (Garda Síochána)

Defence Forces

About this site

Government Departments

Local Authorities

Health Boards

Links

Search  Alphabetical Index  Feedback  Help  Ask Ireland
What's New  Functional Index  Frequently Asked Questions  Information on Ireland  Home
Bienvenue à la Bibliothèque Royale Albert I

pour tout renseignement, adressez-vous à
W. De Vos

Evénements et Projets

Catalogue en ligne

Bibliographie de Belgique

European Library Automation Group

Centre National de Documentation Scientifique et Technique

Center for American Studies

Société des Bollandistes

Commission for Educational Exchange between the USA, Belgium and Luxembourg
Changing page content over time

http://www.informationr.net/ir/20-3/paper682.html
Single pages as social and political evidence

A case study of a public dispute in the UK about the place of religion in public life:

Webster, ‘Religious discourse in the archived Web: Rowan Williams, archbishop of Canterbury, and the sharia law controversy of 2008’ in Brügger and Schroeder (eds), The Web as History (2017)
Rowan Williams and sharia law

The Archbishop of Canterbury says the adoption of certain aspects of Sharia law in the UK "seems unavoidable".

Dr Rowan Williams told Radio 4's World at One that the UK has to "face up to the fact" that some of its citizens do not relate to the British legal system.

Dr Williams argues that adopting parts of Islamic Sharia law would help maintain social cohesion.

For example, Muslims could choose to have marital disputes or financial matters dealt with in a Sharia court.

He says Muslims should not have to choose between "the stark alternatives of cultural loyalty or state loyalty".

'Sensational reporting'

In an exclusive interview with BBC correspondent Christopher Landau, ahead of a lecture to lawyers in London on Monday, Dr Williams argues this relies on Sharia law being better understood.
Archbishop's disaster

Hapless prelate's Sharia views condemned by Lord Carey

By Neville Thurlbeck & Jamie Lyons

THE Archbishop of Canterbury came under devastating attack from his predecessor—who warned his Sharia law demands would be a DISASTER for Britain.

Lord George Carey—writing exclusively for the News of the World—said Dr Rowan Williams had got it WRONG.

He feared Britain accepting aspects of the strict Islamic code would be a direct challenge to British values.

WHAT DO YOU THINK? CLICK HERE TO TELL US

But despite the criticism Lord Carey said Dr Williams should NOT be forced to quit.

Lord Carey said: "His conclusion that Britain will eventually have to concede some place in law for aspects of Sharia is a view I cannot share."
George Galloway MP and the Archbishop of Canterbury's comments about Sharia law via RESPECT - THE UNITY COALITION (The Respect Supporters Blog) February 11th, 2008 at 20:10

Our brother and friend George. Last Friday, 8 February 2008, George Galloway focused on the Archbishop of Canterbury's comments about Sharia law on his talkSPORT radio show ("The Mother of All Talk Shows"). You can listen to a recording of that 8th Feb below; George's show, Friday 8 February 2008 - full show. As well as other radio shows by George, by clicking HERE.

More below: ‘Link: George’...
In george galloway, Archbishop of Canterbury, sharia law

Rowan Williams defends his remarks on Sharia Law via newsJiffy February 9th, 2008 at 16:33

The Archbishop of Canterbury has posted a statement on his website clarifying his position, as well as the Radio 4 transcript. The Archbishop made no proposals for Sharia in either the lecture or the interview, and certainly did not call for its introduction as some kind of parallel jurisdiction to the civil law. You can also hear this below on Youtube (the poster, who calls himself SirHerbertTrance, has used a single picture of Rowan Williams with teeth clenched as the visual while the audio track plays). The BBC reports that two members of the General Synod have called for the Archbishop of Canterbury to resign. Defenders Within The Faith are not only Muslims who have been defending Rowan Williams, The archbishop has been defended by the most senior woman priest in the Church of England.
In youtube, Rowan Williams, radio 4 interview with archbishop of canterbury, islam and the law, sharia law

In defence of Rowan Williams via zeroChampion February 9th, 2008 at 00:05

I've been trying for a few weeks to get the Archbishop of Canterbury to speak at this year's Think08 event. He seemed to me one of the finest thinkers we have. Unfortunately getting him to the event is an even unlikelier prospect today after the almighty ruckus he's kicked off in the last 24 hours. Yet witnessing the shouting match that his speech on Sharia law inspired on Newsnight just...
**Instant Messenger Viruses**

It seems there is an instant messenger virus at large at the moment. The virus sends messages similar to these (the URLs are deliberately obfuscated): "Hot or Not? hxpp://ymmsngallery.my funpic de/viewimage.php?youremail@someplace.com" or "this really looks like you hxpp://ymmsngallery.my funpic de/viewimage.php?youremail@someplace.com"...

Posted to Tim Long (Weblog) by Tim Long on 02-09-2008

---

**God Bless Rowan Williams**

I'm not a religious person; I consider myself spiritual but I think religion does more harm than good. Nevertheless, I have much respect for Dr. Rowan Williams, Archbishop of Canterbury and head of the Church of England. He must be wondering right now what he has done to deserve calls for his resignation even from within his own ranks. The bigots...

Posted to Tim Long (Weblog) by Tim Long on 02-09-2008
Studies on whole sites

Single organisations
Allah.com (Hofheinz in *Web History*, 2010)
University of Bologna (Nanni, DHQ 11, 2017)

Platforms
Milligan on Geocities (*The Web as History*, 2017)
Paloque-Berges on Usenet (*Web25*, 2017)
The Web sphere

Definition: Web materials from more than one site with a ‘shared event, concept, theme or geographic area’ (Brügger, 2018)

Two examples: one national, one thematic
The shape of a national web sphere

Anat Ben-David (@anatbd), ‘What does the Web remember of its deleted past? An archival reconstruction of the former Yugoslav top-level domain’, *New Media and Society*, 18:7 (2016)
One island, two states

Counties of Ireland, north and south
(Wikimedia Commons)
CC-BY-SA 3.0
A unique mix of faith and politics?

Ian Paisley and Edward Carson, Stormont (1985)
(Burns Library, Boston College, CC-BY-NC-ND 2.0 via Flickr)
Cross-border religion?

- Historic Christian denominations: RC, Presbyterian (PCI), Church of Ireland, Methodist, Baptist
- all organised on an all-Ireland basis
- ... spanning two political jurisdictions
- .... and two ccTLDs - .uk and .ie
All-Ireland religion

Church of Ireland dioceses
(Col, via Wikimedia Commons)
CC-BY-SA 3.0
Research questions

Using link graph data, to ask:

• how does web estate of each church interact across the border (& between ccTLDs)?
• are there distinct web spheres for each in NI and the RoI?
Baptists in Ireland (2016)

- Association of Baptist Churches in Ireland has 117 congregations: 28 in RoI, 89 in NI
- 8.5k members, community of 20k
- Including independents, 93 in NI and 30 in RoI
- 28 congregations with domains in RoI, 77 NI
Counties of Northern Ireland

(Map by Maximilian Dörrbecker, CC-BY-SA 2.5)
## Where are the congregations?

<table>
<thead>
<tr>
<th>County</th>
<th>County Code</th>
<th>% of congregations (with domains)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antrim</td>
<td>AN</td>
<td>44</td>
</tr>
<tr>
<td>Armagh</td>
<td>AR</td>
<td>7</td>
</tr>
<tr>
<td>Down</td>
<td>DO</td>
<td>25</td>
</tr>
<tr>
<td>Londonderry</td>
<td>LD</td>
<td>10</td>
</tr>
<tr>
<td>Tyrone</td>
<td>TY</td>
<td>10</td>
</tr>
<tr>
<td>Fermanagh</td>
<td>FE</td>
<td>4</td>
</tr>
</tbody>
</table>
## Where are the domains?

<table>
<thead>
<tr>
<th>Domains</th>
<th>Coverage</th>
<th>.uk</th>
<th>.com</th>
<th>.ie</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Baptist</td>
<td>101</td>
<td>&gt; 80</td>
<td>40</td>
<td>24</td>
<td>-</td>
</tr>
<tr>
<td>Baptist (Antrim)</td>
<td>48</td>
<td>40</td>
<td>31</td>
<td></td>
<td>29</td>
</tr>
</tbody>
</table>
UK Host Link Graph (1996-2010)

• 2008 | catholic_church.co.uk | catholic_church.ie | 4

• 2001 | belfast_anglican.co.uk | derry_anglican.co.uk | 1

• 2002 | derry_anglican.org.uk | derry_catholic.co.uk | 1

Data in public domain: data.webarchive.org.uk
Coded link graph (NI-to-NI)

• 1999 | AN27 | AR07 | 11

• 2003 | DW13 | LD05 | 17

• 2010 | AN11 | AN21 | 3
### Total NI-to-NI edges

<table>
<thead>
<tr>
<th></th>
<th>Inbound</th>
<th>Outbound</th>
<th>Internal</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN</td>
<td>248</td>
<td>299</td>
<td>151</td>
</tr>
<tr>
<td>AR</td>
<td>56</td>
<td>98</td>
<td>10</td>
</tr>
<tr>
<td>DW</td>
<td>125</td>
<td>64</td>
<td>15</td>
</tr>
<tr>
<td>LD</td>
<td>52</td>
<td>20</td>
<td>2</td>
</tr>
</tbody>
</table>
Conclusions

• the Baptist web sphere very tightly localised
• ... but spread across several TLDs
• little cross-border linkage
• link analysis hard in national web archives

Webster, 'Lessons from cross-border religion in the Northern Irish web sphere' in *The Historical Web and Digital Humanities. The case of national web domains* (2019)
Questions?

Peter Webster
peter@websterresearchconsulting.com
@pj_webster / @WebsterRandC
peterwebster.me
websterresearchconsulting.com