
Helena Byrne
helena.byrne@bl.uk
@HBee2015
London 2012 Publications
British National Bibliography (BNB)

http://bnb.bl.uk/
British National Bibliography (BNB)


http://bnb.bl.uk/
British National Bibliography (BNB)


List reduced to 36 publications that potentially had URL references.

http://bnb.bl.uk/
British National Bibliography (BNB)


List reduced to 36 publications that potentially had URL references.

Sampling method used to select publications.

http://bnb.bl.uk/
British National Bibliography (BNB)

http://bnb.bl.uk/
British National Bibliography (BNB)

http://bnb.bl.uk/
British National Bibliography (BNB)

http://bnb.bl.uk/
What Was Included/Excluded
What Was Included/Excluded

Included:
• All URL links transcribed the exact way it was printed in the text.
• Only a few URL’s referenced in the Notes sections were repeated in the Bibliography/References section and were counted twice.
What Was Included/Excluded

Included:
• All URL links transcribed the exact way it was printed in the text.
• Only a few URL’s referenced in the Notes sections were repeated in the Bibliography/References section and were counted twice.

Excluded:
• DOI’s
• Ibid’s
• Duplications in the one reference
Challenges
Challenges
Challenges


Challenges


Challenges

Hard to tell the difference between:
the number 1 or lower case L,
the number 0 or the letter O
upper case I or lower case L
Lower case c and e
Challenges

Hard to tell the difference between:
the number 1 or lower case L,
the number 0 or the letter O
upper case I or lower case L
Lower case c and e

www.bctreaty.net.
Challenges

12. For a more complete account of these comments see: http://www.theyworkforyou.com/debates/?id=2005–07–06a.404.0.


E-books also put some spaces in the URL that were hard to read, these spaces were ignored and assumed to be an error in width between letters.
Total URL References
Total URL References

685 URLs in the Notes & References Section out of a total of 2,924
HTTP Checker Results
HTTP Checker Results

Total URLs = 685

- 200 Code - 430
- 301 Code - 1
- 302 Code - 1
- 403 Code - 7
- 404 Code - 173
- 410 Code - 3
- 500 Code - 11
- Unknown - 59
What Does The 202 Status Mean?
202 Status Review

TOTAL REVIEWED
30/430
Web Archives With Copies

archive.org
nationalarchives.gov.uk
bibalex.org
vefsan.is
webarchive.org
loc.gov
amazonaws.com
webharvest.gov
archive-it.org
arquivo.pt
Most Popular Domain
Most Popular Domain

- http://www.guardian.co.uk
- http://www.london2012.com
- http://www.olympic.org
- http://www.telegraph.co.uk
- http://www.bbc.co.uk
- http://www.dailymail.co.uk
- http://news.bbc.co.uk
- http://www.dailymail.co.uk
- http://www.sportengland.org
- http://www.newham.gov.uk
- http://getset.london2012.com
Where Was The Content From?
Where Was The Content From?

295 URLs

- UK 190
- USA 53
- Canada 13
- Unknown 11
- Australia 6
- Switzerland 6
- France 3
- Brazil 2
- Greece 2
- Netherlands 2
- Russia 2
- China 1
- Germany 1
- Kenya 1
- Spain 1
- UAE 1
IIPC 2016 Summer Olympics and Paralympics

Total: 125

https://archive-it.org/collections/7235

http://netpreserve.org/about-us/working-groups/content-development-working-group/
URL Issues
URL Issues

URL Issues


URL Issues


Text of soliloquy reprinted with permission of Richard Howlett. To see a video of the performance, see “BP or Not BP? The Debut Performance of the Reclaim Shakespeare Company,” http://www.youtube.com/watch?v=0MA1mPJHqHk&feature=player_embedded.
Types Of Content Referenced
Types Of Content Referenced

- http://www
- PDF
- WordPress
- YouTube
- Online Survey
Current Literature
Current Literature

Link Rot/URL Decay
Current Literature

Link Rot/URL Decay

Caroline Nyvang, Thomas Hvid Kromann and Eld Zierau  (The Royal Library of Denmark)

Persistent Web References

Issues With Citation Guidelines


The Chicago manual of style. 16th Edition 2010
Issues With Citation Guidelines

MLA – No http/https or retrieval date

APA – Includes http/https but no retrieval date

Chicago – Includes http/https and either last modified or retrieval date
Suggestion
Suggestion

All citations in the notes or reference/bibliography section should include the date the URL was last accessed, a link to the URL in the browser and if necessary a canonical URL.
Suggestion

All citations in the notes or reference/bibliography section should include the date the URL was last accessed, a link to the URL in the browser and if necessary a canonical URL.

Example:

Suggestion

All citations in the notes or reference/bibliography section should include the date the URL was last accessed, a link to the URL in the browser and if necessary a canonical URL.

Example:

What Can You Do?
What Can You Do?

Welcome to the UK Web Archive
Thousands of UK websites have been collected since 2004 and the Archive is growing fast.

Here you can see how sites have changed over time, locate information no longer available on the live Web and observe the unfolding history of a spectrum of UK activities represented online. Sites that no longer exist elsewhere are found here and those yet to be archived can be saved for the future by nominating them.

The Archive contains sites that reflect the rich diversity of lives and interests throughout the UK. Search is by Title of Website, Full Text or URL, or browse by Subject, Special Collection or Alphabetical List.

Cookies on the UK Web Archive website
We use cookies to ensure that we give you the best experience on our website. If you continue without changing your settings, we'll assume that you are happy to receive all cookies on the UK Web Archive website. Click here for more information.

Explore the Special Collections
Special Collections are groups of websites brought together on a particular theme by librarians, curators and other specialists, often working in collaboration with key organisations in the field. They can be events-based (e.g. The Olympic & Paralympic Games 2012), topical (e.g. The Credit Crunch Collection) or subject-oriented (e.g. The British Countryside Collections).
What Can You Do?

Nominate a site
Fields marked with * are mandatory
Personal details you provide on this form are protected by UK data protection law. Please view our Privacy Statement.

* Full name
Please enter your full name

* Title of website
Please enter the title of the website

* URL of website
Please enter the URL of the website

* Email address
Please enter your email address

Telephone number
Please enter your telephone number

Address

* Are you the copyright holder or owner of the website?
Yes ☐ No ☐

* Your justification: this will aid selection e.g. 'a typical business blog' or a 'prize winning site' or 'representative of Internet culture' or even 'humorous'

Notes about your special request e.g. 'please archive urgently as this site will disappear next month' or 'this site relies on a database. Is this a problem?'
What Can You Do?

INTERNET ARCHIVE
WayBackMachine

http://
BROWSE HISTORY

DONATE

Explore more than 294 billion web pages saved over time

Save Page Now

http://
SAVE PAGE

Capture a web page as it appears now for use as a trusted citation in the future.

Only available for sites that allow crawlers.

FAQ | Contact Us | Terms of Service (Dec 31, 2016)
International Internet Preservation Consortium

http://netpreserve.org/about-us/members/
Conclusions
Conclusions

- Need to update citation rules to improve traceability
Conclusions

- Need to update citation rules to improve traceability
- More awareness of web archiving limitations
Conclusions

- Need to update citation rules to improve traceability
- More awareness of web archiving limitations
- Need to improve font and publishing formats to make it easier to read URL references
Conclusions

- Need to update citation rules to improve traceability
- More awareness of web archiving limitations
- Need to improve font and publishing formats to make it easier to read URL references
- Researchers need to take more responsibility to nominate content
Contact Details

webarchive.org.uk/ukwa

webarchive.org.uk/blog

helena.byrne@bl.uk

@HBe2e2015