



UK Web Sites Maintained only in the UKWAC Archive

Report Number 1 September 2005

About UKWAC

For many, the web has become the information source of first resort. From keeping abreast of latest news and accessing online journals and datasets, through to finding information about travel and sport, the web has become the information tool of choice.

However, despite our apparent dependence on this medium very little attention has been paid to the long-term preservation of websites. There is a danger that invaluable scholarly, cultural and scientific resources will be lost to future generations.

A number of major UK institutions have joined together to form a Consortium (UKWAC) to share the costs, expertise, and facilities required to archive selected websites for future access. The UK institutions that form the Consortium believe that there is enough consistency in the business requirements of web archiving to maintain an operational environment for the web archiving process.

UKWAC Partners

- British Library (Lead partner)
- The National Archives
- National Library of Wales
- National Library of Scotland
- JISC
- Wellcome Trust

The UK Web Archive is at www.webarchive.org.uk

UK Web Sites Maintained only in the UKWAC Archive

This report lists web sites that were no longer available at their original URL or are missing from the 'live' Internet when checked by UKWAC. It also lists web sites that were unavailable at the time their URL was checked. These sites could have been experiencing temporary technical difficulties and may now be available again.

This report uses response header codes to assess web sites availability. Only missing moved or not available error codes are reported, 'silent' and successful responses are not reported here.

Some web sites use internal 404 reporting; this means that the base URL remains available but a sub-page; sub-site or set of pages cannot be found. In this report these are classed as 404 but are marked as Internal 404.

If you are a web site owner and believe we have incorrectly reported the status of your web site please contact us.

If you have any other comments or questions about this report please contact us.

Our e-mail address is [Info @webarchive.org.uk](mailto:Info@webarchive.org.uk)

The following web sites were reported as not found (404) when checked

Title	URL	Date checked	Link to archive	Response code
Valerie Davey, Labour Parliamentary Candidate, Bristol West	http://www.valeriedavey.labour.co.uk/	09-Sep-2005	http://www.webarchive.org.uk/tep/11765.html	404
Electronic Document Delivery : the Integrated Solution (EDDIS)	http://libweb.lancs.ac.uk/eddis.htm	14-Sep-2005	http://www.webarchive.org.uk/tep/11085.html	Internal 404
Judith Blake, Labour Parliamentary Candidate, Leeds North West	http://www.20six.co.uk/leedsnwblog/	19-Sep-2005	http://www.webarchive.org.uk/tep/11895.html	Internal 404
Pensioners Manifesto	http://www.pensionersmanifesto.co.uk/	19-Sep-2005	http://www.webarchive.org.uk/tep/11718.html	404
Replicating the Voynich Manuscript	http://www.keele.ac.uk/depts/cs/staff/g.rugg/voynich/	19-Sep-2005	http://www.webarchive.org.uk/tep/12013.html	404
RIDING : Z39.50 Gateway to Yorkshire Libraries	http://www.shef.ac.uk/~riding/	19-Sep-2005	http://www.webarchive.org.uk/tep/11184.html	Internal 404
Video Medi-CAL	http://www.abdn.ac.uk/videomedical/background/	19-Sep-2005	http://www.webarchive.org.uk/tep/12146.html	Internal 404

The following sites were unavailable at their original URL when checked

Please note that some of these sites may have been only temporarily unavailable when tested and may now be available.

Title	URL	Date checked	Link to archive	Response code
Bill Griffiths	http://www.bgriffiths7.freemove.co.uk/	09-Sep-2005	http://www.webarchive.org.uk/tep/11259.html	
British Organ Donor Society : BODY	http://users.argonet.co.uk/body	09-Sep-2005	http://www.webarchive.org.uk/tep/10514.html	11004 - Host not found
Crafts Study Centre Digitisation Project	http://www.arc.ucreative.ac.uk/	14-Sep-2005	http://www.webarchive.org.uk/tep/10725.html	11004 - Host not found
Cynghorau Cyllido Cymru - Welsh Funding Councils	http://www.wfc.ac.uk/	14-Sep-2005	http://www.webarchive.org.uk/tep/12261.html	10061 - Connection refused
Jeffrey Donaldson	http://www.jeffreydonaldson.org/	19-Sep-2005	http://www.webarchive.org.uk/tep/11779.html	11002 - Host not found
Mathematics Portal Project	http://www.mathgate.bham.ac.uk/mathportal/	19-Sep-2005	http://www.webarchive.org.uk/tep/11053.html	401
Our Healthier Nation	http://www.ohn.gov.uk/	19-Sep-2005	http://www.webarchive.org.uk/tep/11052.html	10060 - Connection timeout
Penglais School	http://www.penglais.ceredigion.sch.uk/	19-Sep-2005	http://www.webarchive.org.uk/tep/10502.html	11002 - Host not found
Two Books by Judith Habgood-Everett	http://www.habgood-everett.freemove.co.uk/	19-Sep-2005	http://www.webarchive.org.uk/tep/10681.html	11004 - Host not found
In Control	http://www.selfdirectedsupport.org/	19-Sep-2005	http://www.webarchive.org.uk/tep/12307.html	Site has moved, automatic re-direction from old URL

Title	URL	Date checked	Link to archive	Response code
MIDRIB : Medical Images Digitised Reference Information Bank	http://www.midrib.ac.uk/	19-Sep-2005	http://www.webarchive.org.uk/tep/11618.html	URL produces 'fatal error' and/or 'No such file or directory' message
National Radiological Protection Board	http://www.nrpb.org/	19-Sep-2005	http://www.webarchive.org.uk/tep/11358.html	Site has moved, automatic re-direction from old URL
Richard Porter, Liberal Democrat Parliamentary Candidate, Camberwell and Peckham	http://www.richardporter.org.uk/unfinished.html	19-Sep-2005	http://www.webarchive.org.uk/tep/11701.html	Error message 'This site is in development and is not yet ready. Please come back later'

Common host response codes

This list is not exhaustive. See <http://www.w3.org/Protocols/HTTP/HTRESP.html>

Response code range	Basic meaning	Response code	Common meaning
400-499	Client Error - Request is incomplete for some reason.		
		400	The request had bad syntax or was inherently impossible to be satisfied
		401	The parameter to this message gives a specification of authorization schemes which are acceptable
		403	The request is for something forbidden. Authorization will not help.
		404	The server has not found anything matching the URI given
		407	The request must be authorized before it can take place
		408	The request timed out
		409	Conflict. Too many people wanted the same file at the same time
		410	The page used to be there, but now it's gone
500-599	Server Error - An error has occurred in the server		
		500	The server encountered an unexpected condition which prevented it from fulfilling the request.
		502	The server cannot process the request due to a high load (whether HTTP servicing or other requests). The implication is that this is a temporary condition which maybe alleviated at other times. Also known as "Bad Gateway" error
		503	This is equivalent to Internal Error 500, but in the case of a server which is in turn accessing some other service, this indicates that the response from the other service did not return within a time that the gateway was prepared to wait. As from the point of view of the client and the HTTP transaction the other service is hidden within the server, this maybe treated identically to Internal error 500, but has more diagnostic value.
		504	The gateway as timed out
		505	The HTTP protocol requested is not supported.

Glossary

Date checked	The date at which the URL was tested and the response code noted
Internal 404	Occurs when a sub-site is unavailable and the main URL remains live and the server provides a customised error message, eg when a university project finishes and removes its web site but the main university URL remains live
Link to archive	The URL that provides a link to the archived instance of a website in the UKWAC archive
Response code	The code as returned by the host server when the URL was tested
Title	The title of a web site as it is recorded in the UKWAC web archive. Archived web sites can be retrieved by their title from the archive
URL	The URL from which a web site was originally archived