Hiberlink joins EDINA effort to ensure continuing access to scholarly work

EDINA has begun its role of co-ordinating an international team of information scientists working on a two-year study to investigate how web links in scientific and other academic articles might fail to lead to the resources being referenced.

This is the focus of the Hiberlink project, in which a team from the University of Edinburgh and the Los Alamos National Laboratory (LANL) will assess the extent of ‘reference rot’, using a vast corpus of online scholarly work.

Increasingly, web-based scholarship includes links that point at resources needed or created in research activity, including software, datasets, websites, presentations, blogs, videos etc, as well as scientific workflows and ontologies. These referenced resources often evolve over time, unlike traditional scholarly articles. The ‘reference rot’ problem occurs whenever the original version of a linked resource is not available any more.

The problem has two aspects. First, the web link that references a resource may no longer work. Second, the content at the end of the link may have evolved and may even have become dramatically different from when originally referenced. So when eventually an online scholarly work is revisited and its references are double checked by a researcher in order to confirm evidence, to establish context, to inform policy and decision making or for any other relevant stakeholders, sharing experience, and the like, the referenced resource may have changed or even ceased to exist, hampering or preventing the research process.

The ultimate goal for the Hiberlink project is to identify practical solutions to the ‘reference rot’ problem. Technical developments at EDINA will start in June to develop approaches that can be integrated easily in the publication process. The project intends to work with academic publishers and other web-based publication venues to ensure more effective preservation of web-based resources so as to increase the prospect of continued access for future generations of researchers, students and their teachers.

Hiberlink builds directly upon a pilot study from LANL, powered by their Memento ‘Time Travel for the Web’ technology, which confirmed that as many as 30% of the web links in a selection of 400,000 arXiv.org papers did not function, and that 65% of the remaining links referred to a resource that was not archived, and hence in danger of disappearing without a trace.

Using text mining and information extracting tools developed by the Language Technology Group (LTG) at the University of Edinburgh School of Informatics, the project will examine a large body of scholarly publication to assess what links still work as intended and what web content has been successfully archived and therefore preserved for use by future researchers and students.

Related EDINA activity to ensure continuing access to scholarly content will benefit from knowledge generated from the Hiberlink project. The Keepers Registry, with an international perspective, is monitoring the extent of e-journal archiving. Nationally, the UK LOCKSS Alliance has a cooperative focus on helping HE academic libraries ensure sustainable continuing access to scholarly work. The Hiberlink project will involve these initiatives as part of engaging relevant stakeholders, sharing experience, and considering best practise to stabilise continuing access to referenced resources.

- Hiberlink project summary: edina.ac.uk/projects/hiberlink_summary.html
- Keepers Registry: thekeepers.org
- UK LOCKSS Alliance: www.lockssalliance.ac.uk
- Memento Project: www.mementoweb.org
- Language Technology Group: www.ltg.ed.ac.uk
GIS-ready background mapping data for schools

EDINA has partnered with Ordnance Survey to deliver Digimap for Schools MapStream to schools, to encourage the take-up of Geographical Information Systems (GIS) technology as a useful teaching tool in the classroom.

GIS appears in both the English and Welsh National Curriculum and the Scottish Curriculum for Excellence but is under-utilised by teachers, with pupils beginning their tertiary education unaware of the broad applicability of spatial analysis.

MapStream is a live stream of Ordnance Survey digital mapping products at a range of scales, from small-scale regional and road-style maps right down to the most detailed Ordnance Survey mapping product showing building outlines.

Subscribing to the map streaming service relieves schools of data management and storage issues, and gives access to the whole of GB contemporary background mapping.

MapStream is different from the flagship Digimap for Schools offering in having no user interface. It is a machine-to-machine web service, which requires schools to have software to access and use the service.

It is attractive to those schools that want to carry out spatial analysis more advanced than the simpler online mapping available within Digimap for Schools.

Mark Smith, Head of Science Faculty at The Grammar School at Leeds, views the service favourably:

“For many schools, the advent of stable, reliable open-source packages such as QGIS has dramatically changed the situation for teachers wishing to give their pupils free hands-on experience of real GIS, as opposed to digital mapping or Google Earth.

“The increasing availability of accessible spatial data has coincided with the maturation of open-source software, making this the ideal time for schools to reconsider introducing GIS, but without some of the barriers such as cost and data acquisition that may have previously put them off.

“Launching MapStream for schools now is very timely.”

• mapstream.edina.ac.uk.

COBWEB Website

We recently launched the website of the COBWEB (Citizen Observatory Web) project, a four-year project funded by the European Union. COBWEB is creating a testbed environment enabling citizens living within or visiting UNESCO designated Biosphere Reserves to collect environmental data using mobile devices.

The website is the main hub for tracking COBWEB project progress, presentations, recent tweets and all of the essential background information on the project and the 13 project partners.

Most of the COBWEB work-packages are now underway and we are preparing to refresh the website to reflect this. Later this summer we will be switching from our current static pages to a Drupal-based website which will help us to ensure the site has the flexibility to grow with the project, that it can support pages in multiple languages, and will enable us to bring in additional features, such as a timeline of previous and forthcoming project events.

• COBWEB website: cobwebproject.eu/
• Twitter: @CobwebFP7.

Geodoc metadata now discoverable on data.gov.uk

Geodoc users can now have their dataset description discoverable on the UK Government’s data.gov.uk website.

Geodoc is the geospatial metadata creation tool available through the EDINA GoGeo service, a geodata portal freely available to the geospatial community since 2003. Geodoc users will now see their dataset description appearing not only in GoGeo but also within the collection of over 1,000 location datasets described at data.gov.uk/location.

Supported by the Jisc Managing Research Data programme, academic institutions across the UK are combining research data management tools, processes, and training to manage research data.

Geodoc usage statistics reveal considerable uptake over the past five years, with users from 49 academic institutions accessing Geodoc 3,113 times to create 2,316 metadata records. As more academic institutions adhere to an internal data management policy, and encourage the publication of their metadata on the GoGeo portal, more spatial data will be discoverable for reuse. The addition of Geodoc-created metadata to data.gov.uk widens the audience able to discover the metadata, thus increasing the potential value to researchers.

The organisation of geospatial metadata workshops has also played a key role in drawing attention to the importance of metadata for the purposes of data discovery and management. The next workshops can be found on the EDINA website at edina.ac.uk/events, or by contacting us on edina@ed.ac.uk.

• gogeo.ac.uk
New deal for JISC MediaHub

JISC MediaHub subscribers will be interested in new subscription developments recently announced by JISC Collections. From 1 Aug 2013 JISC MediaHub will be free to FE subscribing institutions.

Furthermore, HE subscribers to JISC eCollections, of which JISC MediaHub is a part, now have three subscription options:

- **JISC eCollections.** This full subscription gives full access to JISC Historic Books, JISC Journal Archives and JISC MediaHub. The price for this collection includes a built-in discount for taking all three resources and is available at www.jisc-collections.ac.uk/Catalogue/Agreements/YearlyAgreements/1012

- **JISC eCollections (MediaHub and Journals Only).** This option is a subscription to JISC MediaHub and JISC Journal Archives ONLY (you will not get JISC Historic Books). Details of the pricing at www.jisc-collections.ac.uk/Catalogue/Overview/index/1580

- **JISC eCollections (Books and Journals Only).** This option is a subscription to JISC Historic Books and JISC Journal Archives ONLY (you will not get JISC MediaHub). Details of the pricing at www.jisc-collections.ac.uk/Catalogue/Overview/index/1581

The JISC MediaHub service has also recently been enhanced. The classification of collection types has been simplified; time/date and people/organisations have been added to the Advanced Search, and commenting, bookmarking and tagging have joined the My MediaHub functions.

We have recently also included the Fitzwilliam Museum (Cambridge) collection of 1,000 still images, which was part of the bundle of collections licensed under JISC Collections’ Digital Images for Education project, and the Bioscience ImageBank collection, as requested by the former HE Academy Centre for Bioscience at Leeds University.

SUNCAT interface previewed at UKSG

The new SUNCAT interface was demonstrated at the 36th Annual UKSG Conference held in Bournemouth in April. Some of the key features are:

- The home page brings together service information, the search box, a map showing all the Contributing Libraries and a feed of the latest news

- There are greatly improved search limiting facilities. For example, searching can be limited to:
  - a specific library or selection of libraries
  - any combination of the above
  - electronic or print serials: this is helpful if the searcher only wants to see what might be immediately available online or conversely wants to see only material likely to be available via Inter-Library Loan
  - Only the holdings for the libraries or locations specified are displayed
  - More sophisticated searching is available via the Advanced Search option
  - There is filtering of search results by library, location, subject, organisation, publisher, date etc
  - Users can choose how many records to display on the search results screen
  - Further information about each library is provided, including directions, contact details, British Library code and when it was last updated in SUNCAT

The new look SUNCAT was made available for a preview period of around six weeks and those who made visits were invited to answer a short survey to ascertain their views. The responses are currently being analysed and will be used to help improve the new look. More detailed information on the release of the new service will be provided on the service blog:

- SUNCAT Blog: sundatblog.blogspot.co.uk

The new SUNCAT interface preview
Research Data Management Training Kit for Librarians

EDINA and the Data Library, University of Edinburgh, are pleased to announce the public release of the Do-It-Yourself Research Data Management Training Kit for Librarians, under a CC-BY licence.

The training kit is designed to contain everything needed for librarians in small groups to get themselves up to speed on five key topics in research data management, with or without expert speakers.

The kit is a package of materials used by the Data Library in facilitating Research Data Management (RDM) training with a small group of librarians at the University of Edinburgh over the winter of 2012-2013. The aim was to reuse the MANTRA course developed by the Data Library for early career researchers in a blended learning approach for academic liaison librarians.

The training comprises five two-hour face-to-face sessions. These open with short talks followed by group exercises from the UK Data Archive and long discussions, in a private collegiate setting. Emphasis is placed on facilitation and individual learning rather than long lectures and passive listening. MANTRA modules are used as reading assignments, and reflective writing questions are designed to help librarians ‘put themselves in the shoes of the researcher’. Learning is reinforced and put into practice through an independent study assignment of completing and publishing an interview with a researcher using the Data Curation Profile framework developed by D2C2 at Purdue University Libraries.

• datalib.edina.ac.uk/mantra/libtraining.html

News in Brief

From the EDINA website: edina.ac.uk/news.html

Seven New Publishers to Participate in the CLOCKSS Digital Preservation Archive
The global dark archive welcomes 7 new participants.
• Full Story: edina.ac.uk/latU30z

Glasgow School of Art joins SUNCAT as new contributing library
The Glasgow School of Art brings the total of SUNCAT contributing libraries up to 91.
• Full story: edina.ac.uk/18phQopn

Two new Keepers report e-journal archiving
The Ontario Scholars Portal (Canada) and the Archaeology Data Service (UK) have joined The Keepers Registry.
• Full story: edina.ac.uk/12IjP8P

Repository Fringe 2013 registration open
Registration is now live for Repository Fringe 2013.
• Full story: edina.ac.uk/1IDRF9p

Get to know the UK Data Service
Those who rely on social and economic data have a variety of online channels to learn about the UK Data Service.
• Full story: edina.ac.uk/10Cejpz

AQMeN Phase 2 Launch
Phase 2 of the AQMeN research and training programme was launched on at the National Museum of Scotland.
• Full story: edina.ac.uk/1IDQHFV

GeoForum 2013: Programme & Booking
Booking is now open for the EDINA GeoForum Event on 20 June at the Congress Centre, London.
• Full story: edina.ac.uk/1IDQAh

Digimap Roam gets Carto features
We have introduced new print options to Digimap Roam, some of which were previously only available in Carto.
• Full story: edina.ac.uk/10C8jg4

Digimap for Schools newsletter
The latest Digimap for Schools newsletter, for the summer 2013 term, has been released.
• Full story: edina.ac.uk/1IDQd79

EDINA Strategy 2013–2016
EDINA looks ahead to working effectively in changing times.
• Full story: edina.ac.uk/1IDQ57s

New staff

Monica Warnock
I am a member of the UK federation support team at EDINA. I previously worked as an IT infrastructure engineer in an embedded software development company. Recently I set up a new software testing environment and framework for a video surveillance software development company.