HiberActive
Pro-Active Archiving of Web References from Scholarly Articles

Martin Klein, Harihar Shankar, Herbert Van de Sompel
Los Alamos National Laboratory
@mart1nkle1n

Richard Wincewicz
Edina, University of Edinburgh

http://hiberlink.org  #hiberlink
http://mementoweb.org  #memento
http://openarchives.org/rs/  #resourcesync
Link Rot
Link Rot

Conference sessions that cover cross-cutting and overarching issues, Open Repositories attempts to create opportunities to explore the challenges faced by user communities and others in today’s world.

The many repository platforms available today are changing the nature of scholarly communication. Institutions such as universities, research laboratories, publishers, libraries, and commercial organizations are creating innovative repository-based systems that address the entire lifecycle of information — from supporting the creation and management of digital content, to enabling use, re-use, and interconnection of information, to ultimately ensuring long-term preservation and archiving.

OR13: Charlottetown, Prince Edward Island July 8-12, 2013

The 2013 International Conference on Open Repositories (OR13) was held in Charlottetown, Prince Edward Island, Canada on July 8-12, 2013.

The OR2013 conference theme was Use, Reuse, Reproduce. One of the most important roles of repositories is to enable greater use and reuse of their contents—whether those contents are library collections, scholarly articles, research data, or software—and metadata. The notion of use and reuse can be extended to repository infrastructure as well. Many repositories are based on open source software that can be freely reused and adapted to serve local needs; other efforts are also emerging both in conjunction with and outside traditional repository platforms to encourage discipline or community-specific reuse and sharing of software, services, and infrastructure. In addition there is a growing interest and need to document and share the code and workflows used to produce research results — particularly in computationally intensive fields — in order to promote reproducible research.

Visit the conference website to learn more.

To get announcements about this and future Open Repositories events, you may join the OR announcements mailing list here: http://groups.google.com/group/open-repositories.

Open Repositories 2006: Sydney, Australia
1st International Conference on Open Repositories
January 31 – February, 3, 2006
Link Rot
Open Repositories 2006, 31 January - 3 February, 2006

Registrations for Open Repositories 2006 will close at 5pm AEST on Monday January 23. After that time, contact APSR.

Program now available

- Forum: The Well-Integrated Repository
- Symposium: Managing Openness in Digital Repositories
- DSpace User Group Meeting: Building Communities with DSpace
- Fedora User Group Meeting
- EPrints User Group Meeting

Guest speakers will include Sandy Payette of Cornell Computing and Information Science, Mackenzie Smith of MIT Libraries Digital Library Research Group, Robert Tansey of HP Labs, John Townsend of HarvestRoad, Les Carr and Tim Miles from EPrints and other international experts in digital repository architecture and broader e-research frameworks.

APSR is delighted to congratulate the recipients of two Hewlett Packard awards, Dr. M.G. Sreekumar from India and Ms Edna Nyika from Tanzania. More

For an overview of the order of events, take a look at the program schedule.

Open Repositories 2006 will be held at the University of Sydney. This is an interlinked set of events collaboratively convened by APSR with the ARROW project, MAMS, Macquarie University, and the University of Technology Sydney.

Those considering coming "down under" from far off places might consider extending their stay to attend the

HiberActive - Martin Klein et al.
Open Repositories, Helsinki, Finland, June 10th 2014
Scholarly Articles Increasingly Link to Web Resources

URIs extracted from PubMed Central articles
Rethinking Scholarly Communication

Building the System that Scholars Deserve

Herbert Van de Sompel
Los Alamos National Laboratory, Research Library
<herbertv@lanl.gov>

Sandy Payette
Cornell University, Computing and Information Science
<payette@cs.cornell.edu>

John Erickson
Hewlett-Packard Laboratories, Digital Media Systems Lab
<johne.erickson@hp.com>

Carl Lagoze
Cornell University, Computing and Information Science
<lagoze@cs.cornell.edu>

Simeon Warner
Cornell University, Computing and Information Science
<simeon@cs.cornell.edu>
# References

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title and Details</th>
</tr>
</thead>
</table>
Ubiquity of the problem

- Harvard Law Review
  - 75% of links are dead

- top 1% Impact Factor Journals
  - 10% of links dead just 15 months after publication

- US Supreme Court decisions
  - 29% of links dead
  - 49% of links do not point to its original target

Entrance HiberActive

• Perfect scenario:
  • All authors archive all their cited URIs e.g., with Hiberlink’s Zotero extension

• Current scenario:
  • Mostly accidental archiving of cited URIs

• HiberActive scenario:
  • Repositories actively support archiving
  • Utilize existing pro-active web archiving infrastructure
  • Provide value-added service based on archived data
ResourceSync Notification

<?xml version="1.0" encoding="UTF-8"?>
<urlset
  xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:rs="http://www.openarchives.org/rs/terms/">
  <url>
    <loc>http://core.kmi.open.ac.uk/download/pdf/287957.pdf</loc>
    <lastmod>2014-06-08T01:13:07Z</lastmod>
    <rs:md change="created"/>
  </url>
</urlset>
ResourceSync Notification

<?xml version="1.0" encoding="UTF-8"?>
<urlset
  xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:rs="http://www.openarchives.org/rs/terms/">
  <url>
    <loc>http://core.kmi.open.ac.uk/download/pdf/287957.pdf</loc>
    <lastmod>2014-06-08T01:13:07Z</lastmod>
    <rs:md change="created"/>
  </url>
</urlset>
<?xml version="1.0" encoding="UTF-8"?>
<urlset
   xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
   xmlns:rs="http://www.openarchives.org/rs/terms/">
   <url>
      <loc>http://core.kmi.open.ac.uk/download/pdf/287957.pdf</loc>
      <lastmod>2014-06-08T01:13:07Z</lastmod>
      <rs:md change="created"/>
   </url>
</urlset>
HiberActive - Martin Klein et al.
Open Repositories, Helsinki, Finland, June 10th 2014
ResourceSync Change List

<?xml version="1.0" encoding="UTF-8"?>
<urlset
 xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
 xmlns:rs="http://www.openarchives.org/rs/terms/">
 <rs:ln rel="up" href="http://hiberactive.lanl.gov/changelist.xml"/>
 <rs:md capability="changelist" from="2014-05-01T00:00:00Z"/>
 <url>
   <loc>http://archive.today/FL2as</loc>
   <lastmod>2014-06-08T03:31:19Z</lastmod>
   <rs:md change="created" />
   <rs:ln rel="original" href="http://obofoundry.org/ro/" />
 </url>
</urlset>
<xml version="1.0" encoding="UTF-8"?>
<urlset
 xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
 xmlns:rs="http://www.openarchives.org/rs/terms/">
 <rs:ln rel="up" href="http://hiberactive.lanl.gov/changelist.xml"/>
 <rs:md capability="changelist" from="2014-05-01T00:00:00Z"/>
 <url>
  <loc>http://archive.today/FL2as</loc>
  <lastmod>2014-06-08T03:31:19Z</lastmod>
  <rs:md change="created"/>
  <rs:ln rel="original" href="http://obofoundry.org/ro/"/>
 </url>
</urlset>
<?xml version="1.0" encoding="UTF-8"?>
<urlset
    xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
    xmlns:rs="http://www.openarchives.org/rs/terms/">
    <rs:ln rel="up" href="http://hiberactive.lanl.gov/changelist.xml"/>
    <rs:md capability="changelist" from="2014-05-01T00:00:00Z"/>
    <url>
        <loc>http://archive.today/FL2as</loc>
        <lastmod>2014-06-08T03:31:19Z</lastmod>
        <rs:md change="created"/>
        <rs:ln rel="original" href="http://obofoundry.org/ro/"/>
    </url>
</urlset>
ResourceSync Change List

<?xml version="1.0" encoding="UTF-8"?>
<urlset
 xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
 xmlns:rs="http://www.openarchives.org/rs/terms/">
  <rs:ln rel="up" href="http://hiberactive.lanl.gov/changelist.xml"/>
  <rs:md capability="changelist" from="2014-05-01T00:00:00Z"/>
  <url>
    <loc>http://archive.today/FL2as</loc>
    <lastmod>2014-06-08T03:31:19Z</lastmod>
    <rs:md change="created"/>
    <rs:ln rel="original" href="http://obofoundry.org/ro/"/>
  </url>
</urlset>
Demo

https://www.dropbox.com/s/rysi5ekmozbipyq/4min_video.mp4
HiberActive
Pro-Active Archiving of
Web References from Scholarly Articles

Martin Klein, Harihar Shankar, Herbert Van de Sompel
Los Alamos National Laboratory
@mart1nkle1n

Richard Wincewicz
Edina, University of Edinburgh

http://hiberlink.org  #hiberlink
http://mementoweb.org  #memento
http://openarchives.org/rs/  #resourcesync