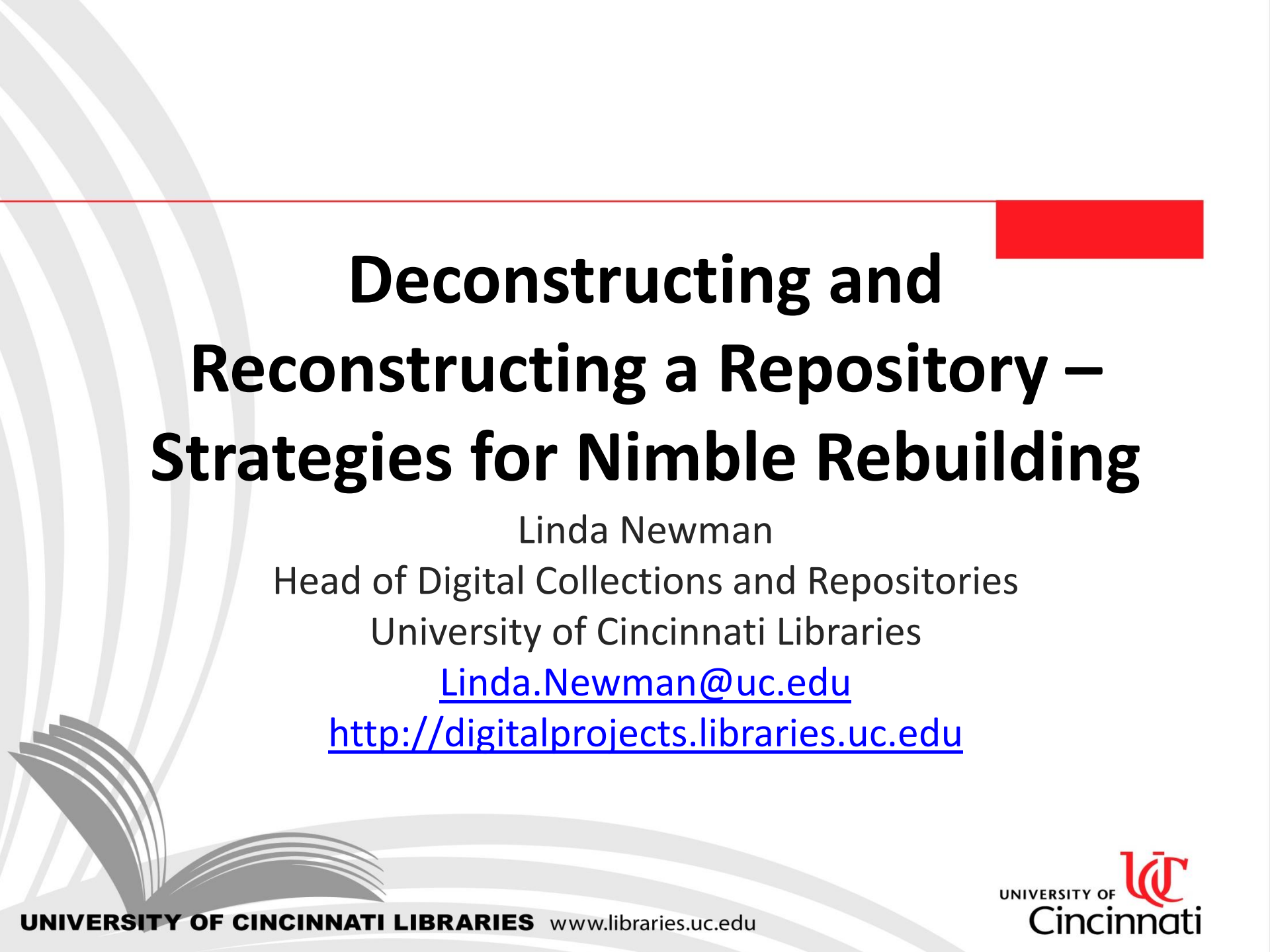




Deconstructing and Reconstructing a Repository – Strategies for Nimble Rebuilding

Linda Newman

Head of Digital Collections and Repositories

University of Cincinnati Libraries

Linda.Newman@uc.edu

<http://digitalprojects.libraries.uc.edu>

The University of Cincinnati Libraries has been relying for five years on a DSpace institutional repository (<http://drc.libraries.uc.edu>), now at over 560,000 records, and maintained by OH-Tech and OhioLINK.

While in the OhioLINK environment we experienced 4 migrations:

- staying in the same DSpace version but migrating to the Amazon Cloud
- migrating to OH-Tech shared infrastructure and an upgraded version of DSpace
- migrating from PostGreSQL and DSpace 1.6 to Oracle and DSpace 1.7
- merging two DSpace instances into one DSpace 1.8 installation.

The OhioLINK community determined in January of 2013 that maintaining over thirty institutional instances of DSpace was not sustainable for its future.

At the same time the University of Cincinnati Libraries determined that we had larger goals for data management and e-science, for the digital humanities and for our university born-digital records, all of which we hoped a broader repository service could encompass.

OhioLINK asked its member libraries to migrate content away from the existing DRCs by December 2013. So in February and March of 2013 the University of Cincinnati Libraries began to put together a development team and a plan, and began to evaluate our choices for a next generation open source repository.

We started asking ourselves these questions:

- Can we build, while assembling a new development team, within the OhioLINK requested timeline, a production Fedora instance with a Hydra (projecthydra.org) front end, capable of supporting all our existing use cases (images, text, audio, video, 'small data', dissertations, direct faculty submission, cultural heritage and institutional repository content) and sustaining the size of our repository?
- Can we build a DSpace instance that has the potential to develop to meet our broader repository service goals?
- Should we even settle on one platform?
- Can we maintain our domain name and handle server prefix (now registered with handle.net and doi.org) , with record identifiers maintained, so that the persistent URLs which now point to content are really persistent?
- Can we build a system to replace a successful large instance of DSpace, minimizing service disruptions for our users by maintaining most existing functionality, but at the same time design a system that is does offer a path to new functionality and service models?

What we have done in the intervening 4 months:

Formed a Digital Repository Task Force, with a platform subcommittee.

Developed a questionnaire (!) and attempted contact with repository managers or developers at all of the Hydra Partner and AP Trust sites, as well as other ARL libraries in Ohio. Most responded!

Started intensive talks with our campus computing center about a partnership; contributing not only the physical infrastructure (production servers, backup) but developer hours.

Moved a staff person into a developer position; advertised for an additional developer.

Made some decisions...

The Hydra community really impressed us – the community and the platform.

But – We concluded there was no easy entry to ‘stand up’ the repository without at least a year’s worth of development.

We decided to:

1. Adopt Hydra as our local digital repository solution.
2. Install DSpace and clone the UC DRC locally – provide our users with a system with the same URL, same functionality but now on our campus.
3. Begin a direct conversation with the AP Trust, to understand more fully the prospects, benefits, and obligations of membership.

What's now underway – Two Tracks:

DRC move

- Begin discussions with OhioLINK about a DRC cloning, and to establish a service agreement with OhioLINK clarifying the level of support they would be providing throughout this process. *(June/July 2013 – talks productive)*
- Work with Library IT to establish resources and a plan for the local installation of DSpace for the UC DRC. *(May-July 2013 – campus IT has agreed to **increase bandwidth** to library server room and to provide an Oracle database server, Library IT is bringing up 2 servers for DSpace test and production, OhioLINK will provide custom DSpace code and record/digital object migration)*
- Work with Library IT to make critical improvements in current backup solutions, to adequately secure repository content. *(July 2013, discussion with Library IT initiated)*

Repository Capacity development:

- Work with Library IT to provide a sandbox server where the digital repository development team can begin to work with Hydra. *(May/June 2013 – sandbox server in place; additional disk capacity in progress, we have built 4 hydra heads - with 3 Fedora installs – all breaking at different points.)*
- Schedule training for developers with Media Shelf, a 'Hydra Camp' -- maybe partnering with other institutions in Ohio interested in Hydra development? *(Fall 2013)*
- Fill the open Digital Repository Developer position *(June-August 2013)*
- Working with campus IT and other interested parties at the University of Cincinnati, establish agreements concerning server resources, backup, and development time. *(Summer 2013)*

It doesn't stop:

We are still actively adding content to our DSpace repository and our LUNA repository.

Our Art and Architecture library may be installing Omeka.

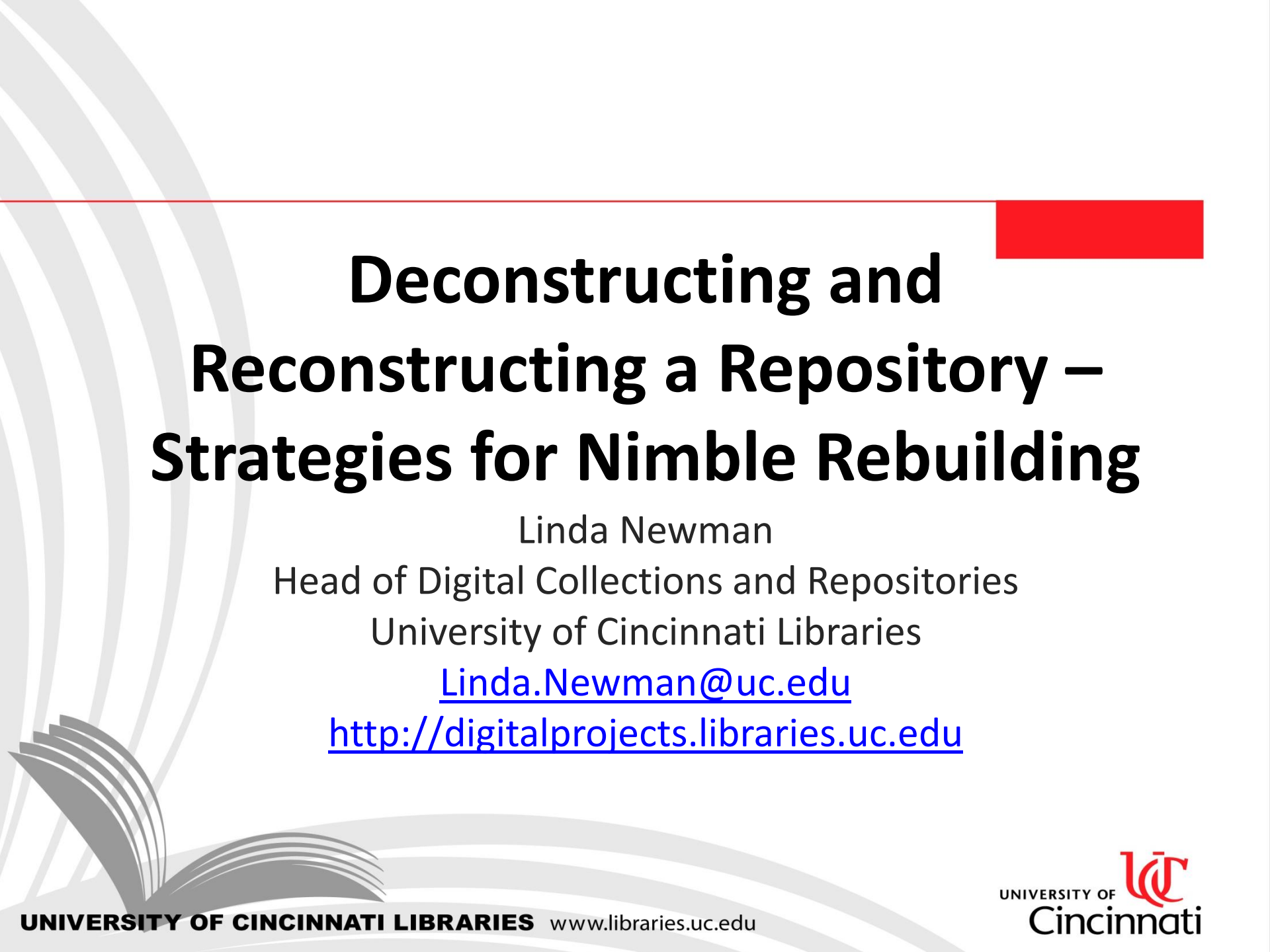
We may be installing Open Journal Systems.

Librarians are discovering more digital content than ever in senior design reports, open access journals that need better support, datasets, university born-digital but official records, the twitter archive of our university president...

We are talking to our campus IT department and the clinical and translational research folks about big data.

We hope to join the AP Trust (and are paying attention to how that may alter our repository development, to have long term preservation built in from the ground up.

We are paying attention to SHARE – and how that in turn will affect our repository development.



Deconstructing and Reconstructing a Repository – Strategies for Nimble Rebuilding

Linda Newman

Head of Digital Collections and Repositories

University of Cincinnati Libraries

Linda.Newman@uc.edu

<http://digitalprojects.libraries.uc.edu>