Open Repositories:

Our Infrastructure Services for Libraries and Research Institutions

Information Systems manager Esa-Pekka Keskitalo
esa-pekka.keskitalo@helsinki.fi

http://urn.fi/URN:NBN:fi-fe201208136289

Based on a presentation of Jyrki Ilva (jyrki.ilva@helsinki.fi) at Open Repositories 2012 / July 10, 2012
http://urn.fi/URN:NBN:fi-fe201207066182
FINLAND

- A country of 5 million people
- Uniform higher education sector
- Government interested in cost-efficient national infrastructures
- Among libraries, a long tradition of cooperation
Repositories in Finland

- 48 Finnish organizations have an institutional repository
  - Universities, universities of applied sciences (polytechnics), state research institutes, government agencies, scholarly societies
- 10 repository instances are only needed
  - 36 organizations utilize services by the The National Library of Finland
  - Six of the major universities have their own systems
The National Library – a major player

By organizations with a repository

Repositories using the national infrastructure (36)
Other repositories (12)

By full-text items

Repositories using the national infrastructure (62900)
Other repositories (57700)
From a little server to a national service

- The original vision: institutional repositories are - institutional
  - "Do it yourself"-oriented ideology
  - Result: a large network of repository software instances = a lot of duplication of work
  - Some of the repositories are poorly resourced, some have little content
- Consider doing it together!
History of repositories in Finland

- From late 1990s: publishing theses and serial publications online
- Early 2000s: the concept of institutional repository
- 2006: starting with Dspace, after a false start with a commercial product.
3 Types of Services provided by the National Library

A customer may

- Be part of a consortium with common policies, practices, processes
  - no kicking over the traces!

- Be a part of a multi-institutional repository
  - Freedom in metadata formats, submission processes, etc.
  - Some control over branding – logos, colours, etc

- Have its own repository hosted by the National Library
  - The technical maintenance of all repositoris is highly centralized
Division of Labour

- The curation of publications and collections is done locally

- The National Library is responsible for the development and maintenance of the technical platform
1. Theseus: all together

- http://publications.theseus.fi
- The common repository for all of the 25 universities of applied sciences
  - Growing fast, 13,000+ new publications submitted by students each year
- All organizations use the same tools, formats and processes,
- and have the same uniform appearance

NATIONAL LIBRARY NETWORK SERVICES
1. Theseus

- Essentially a big group effort, with 200+ librarians and administrators participating in 25 organizations
- Due to standard practices, requires relatively little work
- The cost of managing 25 repositories separately would be several times higher
2. Doria: all in the same place

- [http://www.doria.fi](http://www.doria.fi)
- The original idea (in 2006) was to create a neutral technical platform that any organization could easily adopt
- All of the customer organizations have their own communities
- The organizations are independent in curating their communities
2. Doria

- Many different processes,
- Different visual themes
- Different metadata formats
- There are downsides to this
  - The quality of metadata is not uniform
  - Differing organization-level user interfaces may confuse
2. Doria

- Doria also used by the National Library itself
  - Cultural heritage materials
  - Elektra = articles of the leading Finnish scientific journals (non-OA)
3. Own repository for an individual customer organization

- Some organizations prefer to have their own hosted instance
- Easy to provide, for a small extra cost
- TamPub, http://tampub.uta.fi (University of Tampere)
How to sell repository services?

- Long-term access and persistent addresses are good selling points
- Integrate the repository with the other systems and processes of the organization
- Repositories are suitable for many kinds of materials
  - Much of the discourse on repositories has concentrated on one very specific use case, green OA
  - However, there are other use cases that are just as legitimate (if done properly)
Does it cost something? Yes, it does!

- At the moment, services are fully based on fees.
  - Many other services have a centralized funding model (e.g. funded by MoEC)
- Coherent pricing scheme is a challenge
- We are trying to keep the basic services affordable
  - Smaller organizations do well with the basic functionalities, with minimal customization
- The National Library can also provide consulting and other services (metadata conversions, extensive customization, technical interfaces to other systems, etc.)
  - These cost more, but we try to come up with solutions that benefit other customers as well
Growth of the customer-base: new challenges

- Contract negotiations take time and effort
- How to balance customer projects with the development of the basic infrastructure?
  - Customer projects bring in money, but they also take up a lot of developer time
  - The infrastructure is getting more complicated with each new project
- Need for new and improved services
  - The new, customizable external ingest-system (long overdue)
  - Dark archiving on a national level? Connections to long-term preservation?
  - Standardized interfaces to other systems/processes (library catalog, CRIS, etc.)
Shared and not shared: towards common goals

- Some of the Finnish universities are going to host their own repositories even in the future
- Co-operation between repositories would benefit us all, both in technical development and policy issues
- Grassroot-level co-operation is finding its forms
- Contributions to the international community?
Measures of success for repositories?

- There is a repository
- Repository is filled with meaningful content
- The content is being disseminated (downloaded by users)

Monthly download statistics

<table>
<thead>
<tr>
<th>The entire DSpace (43605 items, 45394 bitstreams, 92577 megabytes)</th>
<th>1 / 2012</th>
<th>2 / 2012</th>
<th>3 / 2012</th>
<th>4 / 2012</th>
<th>5 / 2012</th>
<th>6 / 2012</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>607982</td>
<td>651147</td>
<td>736132</td>
<td>690223</td>
<td>686343</td>
<td>512788</td>
<td>4184616</td>
</tr>
</tbody>
</table>

Source: http://publications.theses.fi/simplestats

- The content has some kind of scientific or cultural impact
- By storing and disseminating the content the repository has an effect on the way we publish and access these materials
We are not very far yet

Types of OA full-text content in all of the Finnish repositories in June, 2012 (items):

- Self-archiving ("green OA") has been important in creating publicity and getting funding, but the actual number of submitted articles has grown only slowly
- OA publication of theses and dissertations is clearly a success story and has changed the way scholarly publishing works in Finland
9th Open Repositories Conference 
In Helsinki in June 2014 

National Library of Finland & Helsinki University Library