

Sharing Grey Literature by using OA-x



Tools and technology
for Open Repositories

Elly Dijk

Conference **Work on Grey in Progress**
New York, 6-7 December 2004

Outline

- More about NIWI-KNAW
- What is i-Tor?
- Four areas covered by i-Tor
- DARE – Digital Academic REpositories
- OA-x project

NIWI-KNAW

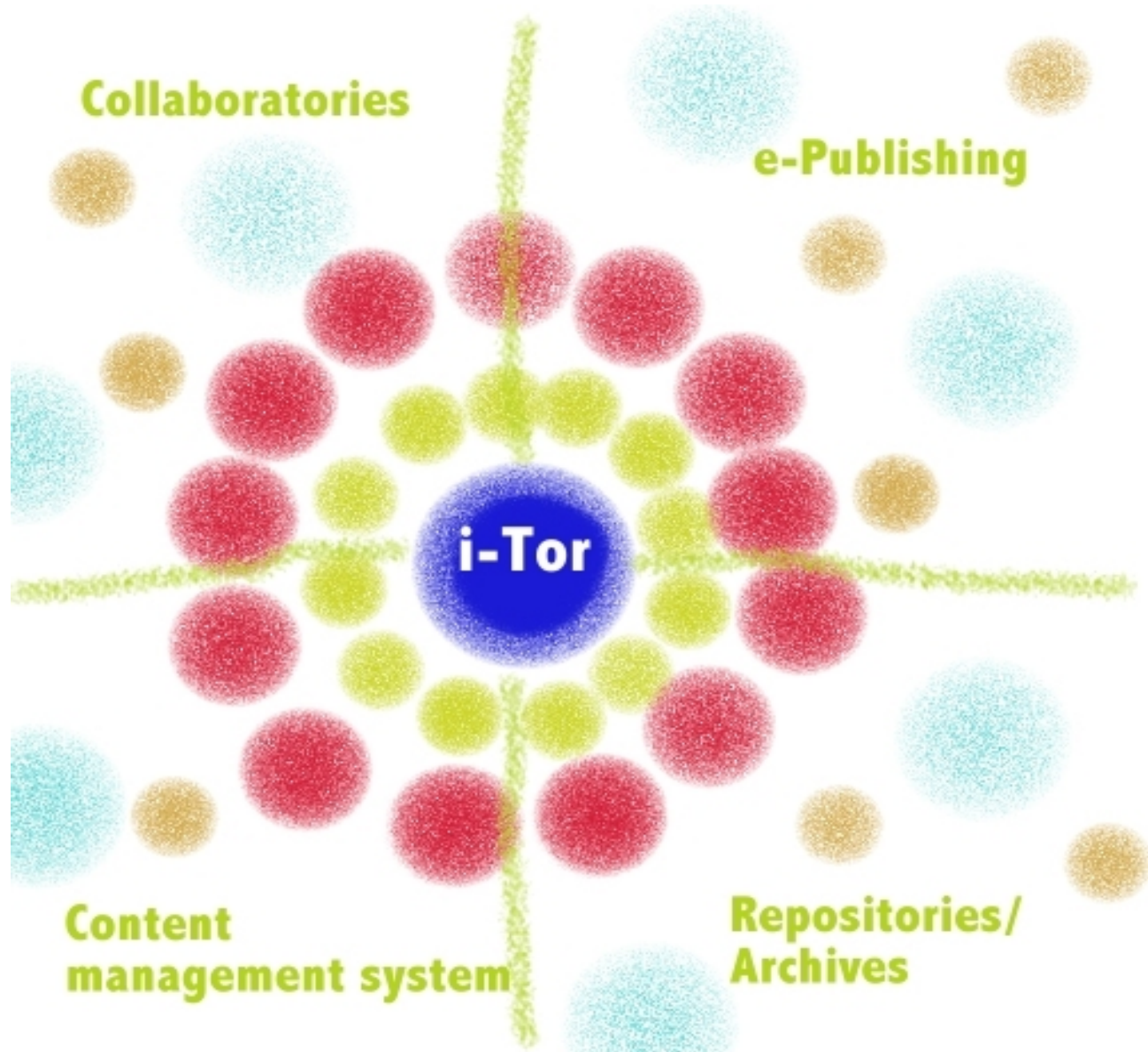
- NIWI, the Netherlands Institute for Scientific Information Services, is an institute of the Royal Netherlands Academy of Arts and Sciences



What is i-Tor?



- = Tools and Technology for Open Repositories
- Open Source software, using Open Standards
- Developed by NIWI's department Innovative Technology-*Applied*
- In collaboration with other institutes like University of Twente and programmes like DARE
- It is a toolbox covering four areas...



Some functionalities

- It is easy to create or add data to webpages
- All content, including databases and PDFs, can be searched with search engines like Google
- Service provider: searching and harvesting external sources of OAI information
- Data provider: making data available in Open Archive form
- Publication tool
- Version management
- Discussion lists
- Similarities



DARE: Digital Academic REpositories



DARE

Digital Academic Repositories

Een SURF-programma

- Joint initiative by the Dutch universities, the KNAW, the Netherlands Organization for Scientific Research (NWO) and the National Library of the Netherlands
- Knowledge infrastructure for the total Dutch academic output
- DAREnet contains now about 20.000 of digital files from 15 institutions

DARE project: OA-x

- Purpose: to enable researchers and managers of digital archives (including Grey Literature) to retrieve, edit, add, combine and archive both metadata **and** data (objects) from digital repositories.
- Producing a protocol for harvesting and uploading of objects
- Built as an extension of the Open Archives Initiative/Protocol for Metadata Harvesting

The commands of the OA-x protocol

- **GetObject(s)**: used to harvest objects. The plural **GetObjects** command has been created for bulk processing
- **PutObject(s)**: used to upload objects to an external repository. The plural **PutObjects** command has been created for bulk processing
- **Checksum**: a unique number, generated using a special algorithm, which can be assigned to an object

OA-x applications: using the GetObject command

- It is possible to index PDFs or other text files on external sites
- Collaboratories often make use of collections of images which are spread across several sites. OA-x makes it possible to produce a collection of thumbnails from these distributed collections for placement on one site
- OA-x can be used to upload datasets to subject-based repositories
- OA-x can also be used to integrate full-text objects
- Analytical tools – for grammatical or textual analysis, for example – can be used even when documents are distributed across several sites.

OA-x applications: using the PutObject command

- It is easy to upload articles and other academic information sources to the repository of another institution.
- An export from an existing application or database to an external repository can be generated. At present, this is only possible between i-Tor sites.
- Users can place full-text articles or other publications on their own website and at the same time easily create an export to a repository – for example, that of their institute.
- OA-x can also be used to upload publications to electronic journals or to a central address where a grey publication is produced.

OA-x applications: using CheckSum

- A checksum is a unique number, generated using a special algorithm, which can be assigned to an object. To check that an archive holds the original object – which is important when working in a collaboratory environment, for example – such a number is vital. Even the tiniest adjustment to the data will change the checksum irrevocably and thus alert the researcher.
- Naturally, such a checksum can be extremely useful in such areas as version management and the verification of electronic publications.

Current use of OA-x

- The E-laborate collaboratory, a partnership in the humanities and social sciences, is using it to update datasets into subject-based repositories as part of the X-Past subproject
- The Bibliography of Dutch Language and Literature (BTNL) is also using OA-x, to integrate full-text objects
- The upload of publications to electronic journals, as is happening at *Studies in Mycology*
- In discussion: within the DARE community
- In discussion: with the creation of a national electronic depot

Contact

- Please see for i-Tor and OA-x:

<http://www.i-tor.org>

- Or send me an email:

elly.dijk@niwi.knaw.nl

