EERQI Introduction
(from the technical point of view)

Slides by:

- Jenny Oltersdorf
  (Humboldt University of Berlin)

- Thomas Severiens
  (ISN Oldenburg GmbH, Oldenburg)

Technical partners

EERQI is an ICT-driven project:

- Berlin School of Library and Information Science at Humboldt-Universität (B-SLIS)

- Institute for Science Networking Oldenburg (ISN)

- Xerox Research Centre Europe, Grenoble

- Regional Computing Centre for Lower Saxony of the Leibniz University of Hanover (RRZN)

- Radboud University, ITS

- University of Lund, University Libraries

Work done

- EERQI-website has been designed and launched

A Project funded under the Socio-Economic Sciences and Humanities Theme.

Work done

An internal website has been set up as common workspace. Instruments for internal communication (e.g. a Wiki, Web DAV) have been set up and tested.

Valuable Documents

EERQI Searchengine

Internal Server

Imprint

research publications,

- propose a prototype framework for establishing such indicators and methodologies,
- make this framework operational on a multilingual basis (starting with English, German, French and Swedish),
- produce a search and query engine for resource harvesting and text analysis,

Upcoming events

ECER 2009
(1 day) (E

EERQI_SES7, EERQI open information session
(2 days) (EERQI E

Work done

- An aggregated, federated content base has been set up, including collections of several electronic resources like journal articles, monographs, and book chapters in the languages English, German, French and Swedish).

- More than 2 GByte of full text is available.

- Documents have been analysed pertaining to formats, availability, metadata, reliable log files, restrictions, etc.

Work done

- A search and query engine has been developed and already been implemented with a user interfaces in English, French, Swedish and German, giving access to:
  - Publication on the web in fulltext,
  - Metadata and two-line snipetts of publications in the content base.

Work done

- Part of the EERQI search engine has been opened to the public and is available on the project website.

- The prototype search engine has also been integrated into the German meta search engine www.metager.de.

Work in progress

- A testing methodology for indicator verification has been suggested.

- The methodological principle is to compare the validity of automatically generated statistics (like citation and usage statistics) with human assessments in a peer review process.

- We are scrutinizing, whether automatically generated statistics can be used as probabilistic indicators of quality, when they are considered in combination and in a sufficient quantity.

Details on the Evaluation Process

- Goals:
  - Configuration and optimization of the automatic qualtity checking implemenation
  - Evaluation of the whole workflow

Review of Journals

Step 1: ✓

- Build a testing corpus of 4 x 100 articles of
    - 6 – 10 pages length,
    - 2 – 10 year age (with the majority from 2006),
    - in the four project languages:
        - German
        - English
        - French
        - Swedish

Step 2: ✓

- Crosscheck by the Scientists in the project, that articles are from the field and potentially of relevance.

Review of Journals

Step 3:   **in progress**

- Find reviewers for the articles (each article needs at least 4 reviews)
- Build a matrix, who should review which article

Step 4:

- Build a questionnaire
    - Build the list of questions   **in progress**
    - Build an online tool to support the review process   ✔

Review of Journals (Questionnaire)

www.eerqi.eu

A Project funded under the
Socio-Economic Sciences
and Humanities Theme.

Review of Publications

Step 5: **can start after finishing of the intellectual review**

- Customizing and optimizing the automatic evaluation tools

Step 6:

- Run the evaluation tools on the content base (non-open access articles, books, inserts, etc.)

Step 7:

- Evaluate the results by scientists from the project.

Step 8:

- Customize and optimize the automatic evaluation tools

(Steps 7 and 8 may be recurse several times)

Step 9:

- Include the evaluation tools into the search enging harvester and run it on new media (like blogs, twitter channels etc.)

Indicators for the Algorithms

- Usage (OA vs. Non-OA!)
- References (DOARC project)
- Institutions (danger of circular reasoning)
- Review indicators (at least 100 document fulltexts for all quality levels and languages)
- Formal criteria (more usable for automatic destinguishing the genres and formats), is Readability also one?
- Originality/Significance and Rigour (do we have batches of good and bad examples?)

Next steps

- We need your input:
  - List of Reviewers
  - Matrix who should review what
  - Final version of questionnaire
  - List of relevant journals
  - List of professional institutions
  - Usage data of OA publications
  - Usage data of publications
  - Translated texts for the web server

Thank you for your attention

- For further questions:
  - Severiens@isn-oldenburg.de
  - Jenny.Oltersdorf@ibi.hu-berlin.de