

EERQI

“European Educational Research Quality Indicators”

WP 2 Technical Corpus Analysis

Kick-off meeting
3-5 April 2008
Hamburg, Germany

Duration:

3 months (months: 1 – 3)

Starting Date: 01 April 2008

Partners involved:

- ***ISN: Institute for Science Networking Oldenburg***
- ***IBI: Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin***
- ***XRCE: Xerox Research Centre Europe, Grenoble***

Objectives:

- ✓ *Technical analysis of the document*
 - *formats,*
 - *availability,*
 - *metadata,*
 - *reliable log files,*
 - *restrictions,*
 - *etc.*
- ✓ *Needed for the preparation of the search engine (WP 5)*
- ✓ *Needed for Conversion implementation and automated metadata extraction for research analysis (WP 6, 8, 9)*

Tasks:

- ✓ *Relevant Document Space needs to be known first
(Collection in our internal Wiki)*
- ✓ *Analysis and Statistics on*
 - *file- and collection-formats,*
 - *data-availability,*
 - *metadata application profiles used,*
 - *reliable log files established,*
 - *restrictions implemented,*
 - *gateways preferred (http, ftp, oai-pmh etc.),*
 - *URN or other persistent identifiers available (DOI?),*
 - *granularity of the document space*
- ✓ *Metadata application profile and crosswalks...*

Deliverables:

- ✓ *Officially: One Report after PM 3*
- ✓ *Effectively: Report + Application Profile + Configuration for Harvester...*

Thank you for your attention!

R. Kaspera & T. Severiens

kaspera@isn-oldenburg.de - severiens@isn-oldenburg.de

**ISN Oldenburg GmbH
(Germany)**

Partner no. 4
