

Implementation of digital repository at the Ruđer Bošković Institute: organizational and technical issues

Alen Vodopijevec
Ruđer Bošković Institute, Library
Bijenička cesta 54, Zagreb, Croatia
alen.vodopijevec@irb.hr

Bojan Macan
Ruđer Bošković Institute, Library
Bijenička cesta 54, Zagreb, Croatia
bojan.macan@irb.hr

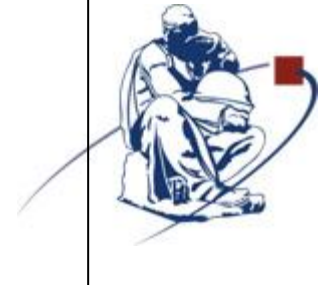


About RBI



- the largest Croatian scientific research institution in the fields of natural sciences
- 530 academic staff (375 researchers and 155 Ph.D. students)
- 136 projects in basic research (in 2008)
- 41 international projects, 67 applied and technological projects and 4 HITRA projects
- 78 undergraduate courses and 245 graduate courses
- 446 research articles published in 2008

Benefits of IR



- **Personal**

- Individual researchers gain better visibility of their papers and therefore, can be cited earlier and more often than papers which are not in OA

- **Institutional**

- institutions gain on their visibility and prestige by collecting all its scientific output in one place
- long-term data preservation

- **Community**

- Research community can find and access information more easily

Open access in science



- "Science and Technology Policy of the Republic of Croatia 2006-2010"

"Scientific output emerged from publicly funded research projects should be available to the public free of charge."

- worldwide open access movement

Challenges...



- type of archived materials
- digitization of printed materials
- copyright issues
- metadata structure
- file formats
- content gathering strategy
- choosing of appropriate software
(installation and configuration)

Repository structure



RBI DIGITAL REPOSITORY

Institutional repository
Scientific material &
Research output

Documentary
Data Archive

Press-clipping
Archive

Types of archived materials



- scientific content
 - books, book chapters, articles, poster presentations, master thesis, doctoral dissertations, datasets etc.
- documentary materials
 - photographs
 - printed materials (brochures, reports etc.)
 - digitization project
- press clipping materials
 - press cut archive
 - press releases

Copyright issues



- scientific content
 - SHERPA/RoMEO project
- documentary materials
 - RBI is copyright holder
- press clipping materials
 - informative news - free to archive in IR
 - problems with TV and radio shows

Preferred file formats



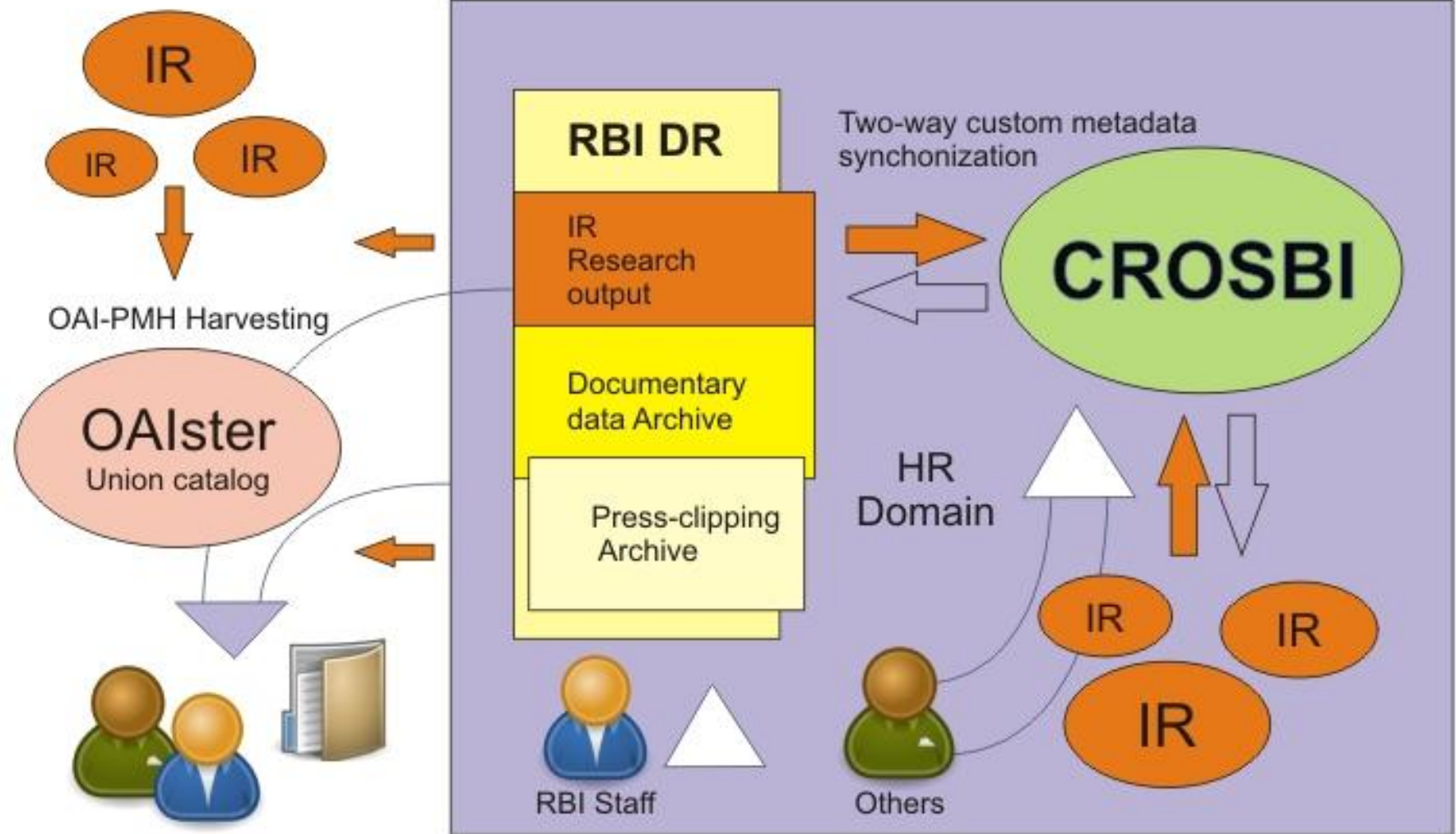
- textual materials
 - PDF, DOC, ODF, RTF, PPT
- video materials
 - OGG (Theora)
- audio materials
 - OGG (Vorbis, Flac, Speex)
- pictorial materials
 - TIFF, JPG, PNG

Metadata structure

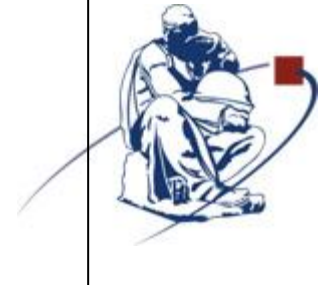


- MARC21 default
- DC, MARCXML -> data representation
- OAI-PMH compliant
- Interoperability with CROSBI

Interoperability



Content gathering strategies



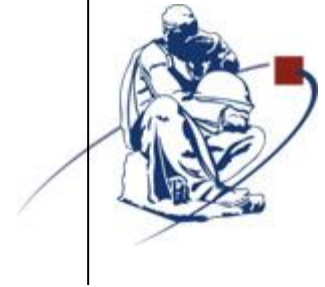
- self-archiving
- self-archiving mandate
- open access advocacy
- promotion

Software issues



FUNCTIONALITY	CDS INVENIO	DSPACE	EPRINTS
<i>Programming language</i>	Python	Java	Perl
<i>Database engine</i>	MySQL	PostgreSQL	MySQL
<i>Localization</i>	Yes – standard .po files	Yes – Language packs based on Java Standard Tag Library	Yes – XML files
<i>Customization of UI and depositing system</i>	Web interface, extensive HTML and Python programming required.	Web interface, larger changes require Java loading and recompiling application.	Majority of configuration is handled by XML files and some HTML templates. Basic knowledge of respective technologies required.
<i>Default metadata standard</i>	MARC with possible export to other Standards	DC	DC
<i>User authentication</i>	Local database, LDAP, Shibboleth	Local database, LDAP	Local database, LDAP
<i>Access control</i>	Role based	Groups as roles	3 default groups
<i>Search engine</i>	Python custom	Java Lucene	Perl custom

Future perspectives



- finishing implementation
- functional interoperability with CROSBI
- Shibboleth single sign-on implementation
- self-archiving mandate
- user training
- promotion

Thank you for your attention



- Questions and (eventually) answers