

Group 2:

Data collection and content of national bibliographic databases

Greta Belogurova, Lithuania

Kinga Ciereszko, Poland

Beata Koper, Poland

Salima Rehemtula, Portugal

Angelika Tsivinsakaya, Russia

Rafael Martinez-Estevez, Cuba

Goals

The main goals of the project are:

- to have a single entry point for all the research outputs (journal articles in the first stage) produced by Cuban researchers;
- to enhance the quality of research outputs by strengthening the methods of evaluation and monitoring at national, institutional and individual level;
- to increase the visibility of the national researchers and their research outputs.

Sources

Main Sources:

- ★ National Journals
- ★ Institutional Repositories

Extra Sources:

- ★ Information about journals (MIAR)
- ★ ORCID
- ★ *Dimensions*
- ★ *Crossref*
- ★ *Datacite*
- ★ *Wikidata*
- ★ *Grid*
- ★ *Sherpa/Romeo*
- ★ *FundRef*

Data collection:

- Using OAI-PMH periodically
- External Sources API

Manual input:

Institutional Level: users add publications, librarians validate

National Level: users (with ORCID, no pubs in national journal and no national affiliation) can add publications (with DOI)

Identifications of Authors and Organizations:

- Try to identify automatically from primary data source
- Use/Validate ORCID and GRID
- Include an alias dictionary for Authors and Institutions
- Manual feedback from users (merge profiles)

Validation:

- Curation of extra sources (which sources are more reliable)
- Comparison of the primary source data with extra source data
- Addressing the deduplication problem

Enrichment:

- Citations using external sources
- Connect with others outputs not present in primary source

Thank you!