Cooperation between five academic institutions for Pure

Renze Brandsma, Head Digital Services, Library UvA/AUAS  |  Wim Klerkx, Functional Application Manager, Library UvA/AUAS
Pure International Conference 2017, October 10-11
Outline Presentation

Renze Brandsma (organizational aspects)
- Why joint approach between the five institutions
- (overall) Project organization, joint tender and implementation
- Evaluation of Pure

Wim Klerkx (technical aspects)
- Cross reporting approach and solutions
- Matching, merging and deduplication process
- Limitations and issues

Cooperation between five academic institutions.
Which institutions?

In Amsterdam five organizations have jointly opted for Pure:

- University of Amsterdam (UvA)
- Academic Medical Center (AMC),
  - ‘part’ of UvA but separate Board of Management
- Vrije Universiteit Amsterdam (VU)
- VU University Medical Center Amsterdam (VUmc),
  - affiliated with the VU, but separate foundation
- Amsterdam University of Applied Sciences (AUAS)
  - cooperation with UvA

Cooperation between five academic institutions.
Cooperation

To get a new Current Research Information System (CRIS) the five started a special cooperation for:

- The tender procedure,
- Acquisition of a CRIS
- Implementation of Pure
- Migration from our former system(s)
Why a new CRIS?

- VU, VUmc used Metis (in one instance)
- UvA used Metis
- AUAS had no CRIS
- AMC had their own developed system

- Metis is developed by a university in the Netherlands
- **Future of Metis was uncertain** because several Dutch universities migrated to other systems
- Metis was regarded as a **user-unfriendly system**

Cooperation between five academic institutions.
Replacement Metis offered the chance to:

- Get a system that conforms to the **wishes of our users**

- **Minimize manual input** using bibliographic resources (Web of Science, Scopus, PubMed, etc.)

- Possibilities for **reports with citations and altmetrics** for knowledge valorization

Cooperation between five academic institutions.
For what purpose do these five use a CRIS?

- Registration of research output, activities, prizes and more of employees at the institution

- As institutional repository or alternatively interface to the repository for managing scientific full-text documents

- Reporting on research results, according to national standards: SEP & KUOZ & BKO
  - legal obligation in The Netherlands.

Cooperation between five academic institutions.
For what purpose do these five use a CRIS?

- Producing figures and overviews for the annual university report

- **Increase research impact** and academic showcasing by making research results visible and discoverable through:
  - local university web platforms or portals
  - national platforms (NARCIS, HBO Knowledge Base)
  - Google Scholar and international platforms
Why a joint approach?

- There is a lot of cooperation in research areas
  - Medical sciences
  - Dentistry
  - Natural Sciences

- There is a lot of cooperation on organizational level
  - ACTA (a collaborative institute of the Faculties of Dentistry at the UvA and VU)
  - AUC (liberal arts and sciences education, cooperation of UvA and VU)
  - Both Medical Centers talk about further integration

Cooperation between five academic institutions.
Why a joint approach?

- **Joint reporting for joint institutions:**
  - Amsterdam Movement Sciences is a collaboration between the VU Faculty of Behavioural and Movement Sciences, VU University Medical Center Amsterdam and the Academic Centre for Dentistry Amsterdam ACTA.
  - Amsterdam Neuroscience is collaboration between AMC, Vumc, UvA and VU.

- Besides separate annual reports also **combined annual reports**
  - Combined annual report UvA + AMC
  - Combined annual report VU + VUmc

Cooperation between five academic institutions.
Why a joint approach?

- The choice for a joint system is an **important money saver**
Support for joint approach

- 2013: Survey among representatives of HvA, UvA, VUmc, VU and AMC:
  - Broad support for the joint renewal of the CRIS.
- 2014: Decided to call for tenders with these 5 institutions.
Cooperation between five academic institutions.
Joint tender

- Consensus on requirements, technical issues and legal documents.

- Not only libraries: also cooperation of IT departments, Purchasing departments and lawyers, etc.

Cooperation between five academic institutions.
Joint tender solutions

- Consensus to more generic requirements, even though our previous systems differed in approach
  - Gradual process between administrators and Purchase focus group.

- Different purchasing rules and purchasing departments
  - UvA purchase was formally awarded the lead

Cooperation between five academic institutions.
Joint tender solutions

- Various ICT rules about security
  - Conform to the **strictest requirements**, namely that of the VU

- Various insights around ICT architecture landscape
  - Architects agreed to a greatly **reduced package of requirements**, as the core group and the Purchaser feared for a failed tender.

Cooperation between five academic institutions.
Joint tender

- Two suppliers submitted a tender for a contract

- Elsevier obtained the tender with the Pure system:
  - User-friendliness (test with panel) was the deciding factor

- Technical proposal 5 instances (1 per institution) + a 6th for joint reporting

Cooperation between five academic institutions.
Implementation

- Projects per institution were executed, each institution implemented their own instance.

- An umbrella project-organization for the 6th instance for joint reporting

- A lot of contact and exchange of knowledge between the 5 organizations !!!

- Special attention to the cooperation process and the user involvement

Cooperation between five academic institutions.
Implementation

- In production at the five institutions end 2016 and begin 2017

- Subsequently, joint reporting via the sixth instance was partly arranged during 2017 and is still ongoing!
Organization after the project

- Cooperation agreement (document)
- Coordination group (steering committee) with 5 representatives
- Administrators meetings with 5 representatives

Cooperation between five academic institutions.
Evaluation and conclusion

- We are *pleased* with the product Pure and Elsevier as supplier.
- Cooperation between the five institutions is successful.
- With regard to the basic module Pure:
  - The *users find Pure generally fine*.
  - The implementation was quite a job.
  - The default interface options offered are limited.
  - The default reporting function is very extensive.
Evaluation and conclusion

- With regard to the repository
  - Harvesting not yet completely funder-compliant (in the harvested metadata the code of the research project is not correctly included)

- Regarding the umbrella reporting function:
  - Promising
  - Strongly dependent on the underlying data quality and mutual agreement between the delivery settings

Cooperation between five academic institutions.
Evaluation and conclusion

- As regards to Elsevier as a supplier:
  - The maturity level of the Elsevier organization still needs to be improved
  - Coping with the wishes of customers needs to be better
  - Testing new releases should be significantly improved

Cooperation between five academic institutions.
Maybe we should have waited until the 2019 conference!

- 6th instance issues still ongoing
- Production environment with current data not yet available
- Still testing and refining deduplication, considering workflow and procedures
- No reports yet – but the first deadlines coming in sight

Cooperation between five academic institutions.
7 faculties

2016: 31K students
2016: 2798 academic staff

Cooperation between five academic institutions.
9 Faculties

2015: 2,196 academic staff
2015: 22,897 students

Cooperation between five academic institutions.
7 Faculties (Domains)
current: 2,596 academic staff
2016 46,140 students

Cooperation between five academic institutions.
Universiteit van Amsterdam
Faculty of Medicine
2016:
2300 medical students, 1500 PhD candidates
2016: 1500 staff in medical research

Cooperation between five academic institutions.
Vrije Universiteit Faculty of Medicine
2.200 students
1.500 academic staff

Cooperation between five academic institutions.
Stuff we share …

Cooperation between five academic institutions.
Use cases for joint reporting:

- UvA + AMC annual report on research output (formerly done by importing RIS files in UvA CRIS)
- MOVE annual report on research output
- Other collaborations reporting (formerly done by manually combining data)

VU + VUMC annual report on research output
formerly easy as they shared their CRIS!

Cooperation between five academic institutions.
Cross reporting: solution 1

One shared CRIS

Filtering & customizing layer (using API)

Web pages, portals reports

UvA

VU

AUAS

AMC

VUMC

+ All content entered once!
+ Enriching from different sources
+ Optimize sharing of data

- Profiling of local content
- Licences differences
- Workflow differences
- Handling differences in definitions

Cooperation between five academic institutions.
Cross reporting: solution 2

Exporting & aggregation layer

“6th instance” reporting platform
aggregated & deduplicated content

CRIS

CRIS + portal

VU

CRIS + portal

CRIS

+ customization to local needs
+ licences
+ local workflow choices
- same data entered on overlapping persons, output
- 1-way sync, no automated enrichment
- handling differences in definitions

Cooperation between five academic institutions.
Challenges - reaching agreement on:

**definitions**
- classifications
- templates, types, categories
- reports

**procedures**
- Meetings, decision making, define strategy
- Master data sync: journals
- System upgrades control
- Update classification schemes

**workflows**
- Define steps for data improvement
- Assign tasks between institutions

Cooperation between five academic institutions.
Elsevier’s choice (after discussions): feeding the 6th instance

- 5x harvest of local systems
- Transform to importable XML files
- Sequential executing cronjob: import sequence
  - 7 sync jobs & import wizard
  - Organisations/persons/external organisations/external persons/publishers/journals/events/research output

- Managed by admin in 6th instance
- Logfiles available

- Deduplication sequence
  - Publications, journals

Cooperation between five academic institutions.
Synchronisation

Synced but not merged!
Duplicates remain
Complex filters in reports
No grouping

Cooperation between five academic institutions.
Deduplication process

MATCHING
- Find matching source ID, else ...
- Find matching DOI, else...
- Find records with matching metadata fields (12)

Still discussing details!

MERGING
- Target record preserved in single valued fields (e.g. title, workflow)
- Most fields added if missing

Rule of thumb: multi-valued fields (i.e. persons, organisations, ISBNs, Publication Statuses, keywords) are merged

BEFORE / AFTER DEDUPLICATION (result of 1st run):

**Publications**
- 430466 / 398697 (-7%)
- Duplicate titles: 38847 / 14059 (-36%)
- Duplicate isbn: 497 / 391 (-22%)
- Duplicate doi: 21044 / 77 (-99.6%)

**Journals**
- 42606 / 30435 (-29%)
- Duplicate ISSN: 10119 / 2830 (-72%)
- Duplicate title: 15433 / 4441 (-71%)

Cooperation between five academic institutions.
Duplicate journals and research output

- Deduplication logfile as source
- Avoid manual merges

- Instead correct in local systems:
  - *add DOI*
  - *correct typos, etc.*
- Merge by running deduplication process again

Cooperation between five academic institutions.
Working on corrections and progress

Hmmm, looks like a duplicate

Why wasn’t it merged into the other one?

Cooperation between five academic institutions.
Issues big and small

**Staff capacity for necessary corrections**
- start publication year 2017, then backward (… to 1980??)

**Corrections to be coordinated, by report owners?**
- strategy to be chosen …

**Solutions to be found for**
- Article number in PubMed > start page = match criterium
- Chapters with identical DOI??
- ……

**Planning to add activities, prizes, press / media?**

Cooperation between five academic institutions.
Lessons learned

Sharing:

- Everything
- Most data
- Within privacy limits
- Limited data
- Nothing

Cooperation between five academic institutions.
Questions?

Cooperation between five academic institutions.