Implementing Symplectic Elements at Macquarie University as a hosted publication sourcing solution

Josephine Morton
Rhiannon Rasins
Vladimir Bubalo
Implementing Symplectic Elements at Macquarie University as a hosted publication sourcing solution

Josephine Morton
Rhiannon Rasins
Vladimir Bubalo
Research output collection and reporting process at Macquarie University

- Jointly managed by researchers and staff from the Library and Faculties

- Research Office has ultimate responsibility for the audit and submission of the final report to the Government
Process for entering publications:

- Up until 2011, academics were responsible for manually entering their publications into the University's research management system (RMS), IRIS, directly.

- Manual process = poor data quality

- Mid 2011: Library took over responsibility for the data collection and entry for research reporting (HERDC and ERA)

- Entered into IRIS, via the University's institutional repository, ResearchOnline, http://www.researchonline.mq.edu.au/
Data quality significantly improved with Library involvement, but still many manual processes involved

- Prior to Symplectic implementation, the data for ResearchOnline sourced via searching citation databases, and either:
  - manually creating records
  - bulk uploading records

- Library staff would then enrich the metadata of each record, add required information/documentation and load the record into ResearchOnline

- New records added into ResearchOnline are extracted daily and uploaded into IRIS

- Tedious, repetitive and mostly manual process for which automation could be beneficial - freeing up staff resources
• Library staff would then enrich the metadata of each record, add required information/documentation and load the record into ResearchOnline

• New records added into ResearchOnline are extracted daily and uploaded into IRIS

• Tedious, repetitive and mostly manual process for which automation could be beneficial - freeing up staff resources
A solution was sought to help automate the pulling in of publication records from citation databases, as well as de-duplication of these records

- University at that time preferred hosted IT solutions for all new systems
- Additionally, we didn't have capacity for IT support internally

We approached Symplectic about hosting our Elements instance and providing support

- New contract was negotiated as we were their first hosted customer
- The platform is hosted by Symplectic in Amazon Web Services (AWS)
We approached Symplectic about hosting our Elements instance and providing support

- New contract was negotiated as we were their first hosted customer
- The platform is hosted by Symplectic in Amazon Web Services (AWS)
Implementation

Phase 1
Phase 2
Getting our infrastructure built in the cloud

Understanding import process and data structure and preparing historical publication data for import
Implementation

- Phase 1
- Phase 2
Ensuring our systems can communicate with Symplectic:

- Worked with our IT to develop an application to extract publications and upload into our system
- Worked with our IT and Symplectic to generate a user feed
Getting our infrastructure built in the cloud

• Suggested by Symplectic to build infrastructure in AWS because of local presence in Australia

• Quick deployment enabled by AWS's flexible infrastructure options

• Hardware and software managed and maintained by Symplectic:
  • always up to date hardware/software
  • clearly defined in the SLA
Understanding import process and data structure and preparing historical publication data for import

- Learning the process of importing
- Where to source the data
- Cleaning the data
- Matching points between publications and people (only could use staff/student number from source system, duplications in that)
- Mapping data from RMS to Symplectic
Worked with our IT to develop an application to extract publications and upload into our system

- Symplectic didn't previously have any connections to any of our systems: we worked with Symplectic and our Library IT to develop an app utilising the API to export publication metadata of claimed publications into our repository workflow system

- Data mapping from Symplectic API to our workflow system
Worked with our IT and Symplectic to generate a user feed

- Historical users extracted from Databank initially and loaded to assist with historical publication matching

- Current users extracted from Databank daily

- Ports being opened

- FTP server set up
Publication metadata export via script running daily

Import of certain metadata
Challenges:

• System connections:
  • Symplectic
  • Repository
  • Research Management System (RMS)

• Dirty data for migration from RMS:
  • had to clean the data
  • map to the correct fields so the data could validate

• Had to go through central IT for firewall changes, which could take time

• Time difference/turnaround delays prolonged the process
Challenges:

- Duplication (publications and people) and cleaning up pending relationships
- Learning to divide time:
  - staff working to incorporate another system in workflow
- Learning curve: first time for both us and Symplectic to be involved in this kind of process
- Had to register Symplectic under MQ IP to access Scopus and WoS metadata
Benefits:

- Staff have positive experiences using the system
  - Very stable and responsive
  - User-friendly and fairly quick to learn

- Dynamic system: most of the data is frequently updated in an automated way

- Opportunities to meet with the Symplectic team:
  - User Conference
  - training
Benefits:

- Working in Symplectic has helped us identify issues and seek corrections of source data in:
  - Internal data sources (HR, RMS, etc)
  - External data sources (Scopus, WoS, PubMed, etc)

- Working with our developer was positive experience on both sides - quick and agile app development:
  - 4 week initial rollout
  - 1 subsequent release shortly thereafter

- Continuous improvement focus on the app: frequently log tickets for new improvements
Benefits:

- First time our developer was able to create and deploy the app via deployment packaging, so our app gave him the opportunity to test this.

- App gives us flexibility of what we can export to our workflow system:
  - can pick and choose what fields we want
  - can pick and choose from which data sources

- Further developed and expanded our cross-campus collaborations (HR, Faculty staff, DVC-R office)
Going forward:

- We're not currently opening Symplectic up for academic use:
  - currently a closed system used by the repository team in the library
  - may open it up for academics to use in the future

- One-way connection for data currently: hope for it to flow back (publication identifiers could be used as a match point)
Going forward:

- Look at how deletions in RMS can be dealt with in Symplectic

- There are some gaps in the data Symplectic can get (missing fields and snippet descriptions):
  - raised feature requests
  - plan to work with Scopus on the best API for Symplectic to use
Questions?

Contacts:

Josephine Morton: josephine.morton@mq.edu.au
Rhiannon Rasins: rhiannon.rasins@mq.edu.au
Vladimir Bubalo: vladimir.bubalo@mq.edu.au