

## Software Development Priorities

**Date:** November 12, 2020

**Status:** Draft for review meeting



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).  
**Citation guideline:** Simon Fraser University Archives (Nov 2020), *Software Development Priorities* .

**HIGH PRIORITY**

New issue

Jump to bottom

# Link AIPs to fonds / collections #1

🔓 Open

radancy opened this issue on Jul 2 · 0 comments

Labels

archivematica

high priority

👤 radancy commented on Jul 2

Contributor

Support searching and reporting on AIPs by fonds.

- Currently the Archives relies on its FileMaker AIS database for this (via the Accession and AIP tables).
- Project would build better reporting into the Archivematica dashboard.
- We currently have ca 800+ AIPs, so need to consider how a solution could be applied retroactively.

🏷

👤 radancy added

archivematica

not yet ranked

high priority

and removed

not yet ranked

labels on Jul 2

New issue

[Jump to bottom](#)

# Minor Archivemata user interface enhancements #8

Open radancy opened this issue on Jul 3 · 0 comments

Labels archivemata high priority in progress

Milestone 2019-20 dev contract

radancy commented on Jul 3

Contributor

Enhance Archivemata by making improvements to various UI features.

Backlog tab:

- Add column showing package size.
- Add column showing accession number.
- In Actions column, show status when there is an outstanding request for package deletion.

Archival storage tab:

- Add column showing number of files.
- Add column showing accession number.
- Add column showing encryption status.
- In Actions column, show status when there is an outstanding request for package deletion.

radancy added high priority archivemata in progress labels on Jul 3

radancy added this to the 2019-20 dev contract milestone on Jul 3

New issue

[Jump to bottom](#)

# Archivemata processing configuration #9

🔓 Open

 radancy opened this issue on Jul 3 · 0 comments

Labels archivemata high priority in progress

Milestone 🏁 2019-20 dev contract

👤 radancy commented on Jul 3

Contributor

You can create multiple processing configurations to support different workflows that pre-select decision point options. As of Archivemata 1.9 you can download a custom processing configuration as an xml file, include it in a transfer to use that configuration. This project would make it easier to select configurations.

Provide interface to switch between processing configurations.

- Currently you can do this only for AIP re-ingest.

🏷️ 👤 radancy added new high priority archivemata in progress labels on Jul 3

🏁 👤 radancy added this to the 2019-20 dev contract milestone on Jul 3

🏷️ 👤 radancy removed the new label on Jul 3

New issue

[Jump to bottom](#)

# Metadata-only DIP upload #10

! Open

radancy opened this issue on Jul 3 · 0 comments

Labels      archivematica-atom    high priority    in progress


Milestone      ➔ 2019-20 dev contract

 radancy commented on Jul 3

Contributor

Improve Archivematica's metadata-only DIP to AtoM so that:

- It is obvious to the AtoM user that a digital object exists, but is not currently displayed in AtoM (currently this is not clear).
- There is a clear distinction between technical metadata for the original file vs metadata for the access copy.

🏷  radancy added    archivematica-atom    high priority    in progress    labels on Jul 3

➔  radancy added this to the 2019-20 dev contract milestone on Jul 3

New issue

[Jump to bottom](#)

# Accessions module #21

**! Open** radancy opened this issue on Jul 3 · 0 comments

Labels **atom** **high priority**



 **radancy** commented on Jul 3

Contributor

Create an AtoM module to implement the recently released [Canadian Archival Accessioning Information Standard \(CAAIS\)](#).

Possible features:

- Support all CAAIS elements.
- Add multi-repository functionality.
- Allow linking AtoM archival descriptions at all levels to an accession – would allow create finding aids by accession (e.g. show all files – which may be distributed across multi series – by accession).

  **radancy** added **high priority** **atom** labels on Jul 3

New issue


[Jump to bottom](#)

# CAS integration #28

**! Open** radancy opened this issue on Jul 3 · 0 comments

Labels archivematica-atom high priority in progress

Milestone ➡ 2019-20 dev contract

 **radancy** commented on Jul 3 Contributor

Allow staff to log on to Archivematica and AtoM using SFU's Central Authentication Service (CAS).

🏷  **radancy** added high priority archivematica-atom in progress labels on Jul 3

➡  **radancy** added this to the 2019-20 dev contract milestone on Jul 3



**MEDIUM PRIORITY**

New issue

[Jump to bottom](#)

# Transfer analysis and validation tools #2

**! Open**

radancy opened this issue on Jul 3 · 1 comment

Labels archivematica medium priority

 **radancy** commented on Jul 3

Contributor

Build / integrate tools for better analysis, assessment, and validation of digital transfers.



**Examples\***

Compare checksums against existing AIP store to identify duplicates

- E.g. A department does not delete records after transfer and re-sends the same records at a later date.

Flag password-protected files.

- Need to know close to the moment of transfer whether there are password-protected files requiring actions on part of creator / donor.

  **radancy** added archivematica not yet ranked medium priority and removed not yet ranked labels on Jul 3

 **axfelix** commented on Jul 9

These two features seem reasonable to build into Provelt, so that's still my plan here -- I think others could easily turn into feature creep, but I like these.

New issue

[Jump to bottom](#)

# Disk images #4

 **Open**

radancy opened this issue on Jul 3 · 1 comment

Labels archivemata medium priority

 **radancy** commented on Jul 3

Contributor

Review practices for ingest of disk images.

When should disk images be used (vs copying directories / files only)?

- Should disk images be ingested or processed prior to ingest?
- Experiment with BitCurator forensics tools, other reporting tools (e.g. Brunnhilde), assess integration with Archivemata.



 **radancy** added archivemata not yet ranked medium priority and removed not yet ranked labels on Jul 3

 **axfelix** commented on Jul 9

This would be a good thing to prioritize next year, I feel like our processes around this have been rather hand-wavy for a while. I'm not convinced we need an external contract or to think about it in terms of Archivemata either, just to sit down and hammer out what we want.

[New issue](#)[Jump to bottom](#)

# Archivematica error handling #7

! Open

radancy opened this issue on Jul 3 · 0 comments

Labels

archivematica

medium priority



radancy commented on Jul 3

Contributor

Archivematica microservices sometimes fails in non-critical ways (e.g. a JHOVE partial pass is represented as a fail in Archivematica); currently fails are handled ad hoc and without documentation. This project would establish a policy for consistently handling fails and build functionality in Archivematica to support the practice.

Policy, e.g.

- Establish criteria for distinguishing critical vs non-critical fails.
- Specify actions for both cases.
- Indicate how decisions are documented and where.

Archivematica, e.g.

- Allow printing / export of identification fail reports.
- For fully automated ingest (using configuration files), generate a report on the conclusion of the ingest so archivist can review the ingest
- Set order of preference for format identification tools so Archivematica will run another tool when the preferred tool hits a fail.
- Allow using different tools on specific files without having to re-run the whole transfer.
- Provide interface for recording decisions to accept identification fails (included in METS file as PREMIS event).



radancy added medium priority archivematica labels on Jul 3

New issue

[Jump to bottom](#)

# AIP repository manager #11

🔓 Open

radancy opened this issue on Jul 3 · 0 comments

Labels archivemata medium priority

 radancy commented on Jul 3

Contributor

Develop repository management functionality for managing objects in the AIP store.

- E.g. reporting, searching, checking checksums, bulk re-normalization based on changes to the FPR, backup and recovery.

First phase could be a research project to determine the best model. Options might include:

- Enhance the Archivemata Dashboard (e.g. make some of the analysis tools from the Appraisal tab also available on the Archival storage tab).
- Enhance Storage Service (e.g. add management functionality).
- Build on existing work with Binder (make it fully open source).
- Build tools to search / report across METS files in the AIP store.
- Build integration with an existing repository management application (e.g. D-Space, Islandora)
- Build the manager as a stand-alone module separate from Archivemata pipelines.

🏷  radancy added medium priority archivemata labels on Jul 3

New issue

[Jump to bottom](#)

# DIP on demand #12

! Open

radancy opened this issue on Jul 3 · 1 comment

Labels

archivematica-atom

medium priority



radancy commented on Jul 3

Contributor

Provide interface to generate new DIPs from existing AIPs.

Currently you can download a stored DIP (generated at AIP creation) and upload to AtoM via command line or re-ingest the entire AIP to create a new DIP.

- Ideally you should be able to upload via the Archivematica Dashboard.
- Ideally would be able to target specific files in a given AIP and create DIPs for just those files.
- Support bulk upload to existing descriptions via csv import.



radancy added medium priority archivematica-atom labels on Jul 3



axfelix commented on Jul 9

I'm starting to feel like this functionality is overdue -- would love to look at getting it in there over the next few years.

New issue

Jump to bottom


# Email #13

🔔 Open

radancy opened this issue on Jul 3 · 0 comments

Labels

archivematica-atommedium priority

 **radancy** commented on Jul 3

Contributor

Build on previous work on email to develop solutions for transfer, arrangement and description, and access to email records.

The project would involve research and policy as well as software development. Possible elements:

- Review scripts that managed export of email from Zimbra and convert to mbox – do these still work with Outlook?
- Extract descriptive metadata from mbox DIP, map to AtoM description fields.
- Integrate email viewer in AtoM.
- Configure IMAP as a transfer source in Storage Service.
- Test ePADD to manage restrictions and access.

🏷  **radancy** added medium priority archivematica-atom labels on Jul 3

New issue

Jump to bottom

# Semi-active digital transfers environment #16


🔔 Open

radancy opened this issue on Jul 3 · 0 comments

Labels

archivematica-atom

medium priority

 **radancy** commented on Jul 3 

Contributor

Provide support for semi-active transfers of born-digital records that have long-term retention requirements but may not be archival.

Create a separate Archivematica-AtoM environment for semi-active transfers.

- Build scripts to automate disposals (i.e. destructions and transfers of material from semi-active to archival environment)

🏷️  **radancy** added 

medium priority

archivematica-atom

 labels on Jul 3



[New issue](#)

[Jump to bottom](#)

# Rights management #18

[Open](#)

radancy opened this issue on Jul 3 · 0 comments

Labels

atom

medium priority



radancy commented on Jul 3

Contributor



Build on the existing PREMIS rights area in AtoM to support types of conditional access, each with their own notice screen.

Current functionality:

- You can set different access rules for different rights bases (statute, copyright etc).
- There are 3 restriction-states for each basis (allow, disallow, conditional).
- A popup notice screen for exists for one use case, where basis = "copyright" and restriction-state = "conditional".

Possible development:

- Allow each rights basis to have its own conditional pop-up notice (rather than hard-coded to copyright).
- Integrate CC license icons
- Integrate standardized rights statements (from RightsStatements.org).

  radancy added medium priority atom labels on Jul 3

New issue

[Jump to bottom](#)

# AtoM search enhancements #20

**! Open**

radancy opened this issue on Jul 3 · 0 comments

Labels **atom** medium priority



 **radancy** commented on Jul 3

Contributor

Review existing search capability in AtoM, identify and implement improvements.

Possible features:

- Add Level of description as sort option on search results
- Build admin interface to select fields for indexing (e.g. for us searching on restrictions would be very useful, add PREMIS rights fields to search index)
- Better ranking formula
- Support exclusion of records belonging to certain fonds / series from search results

  **radancy** added medium priority atom labels on Jul 3

New issue

[Jump to bottom](#)

# AtoM batch editing #23

 Open

radancy opened this issue on Jul 3 · 1 comment

Labels

atom

medium priority



radancy commented on Jul 3

Contributor

Support batch editing of a set of descriptions.

Examples:

- Add same field value (e.g. GMD) to a batch of records.
- Batch update permalink URL for a set of records (i.e. toggle from stub based on title to stub based on ref code for a given set).
- Search-and-replace on set of records.

  radancy added medium priority atom labels on Jul 3



axfelix commented on Jul 9

This has historically been really hard to fit in the regular data entry UI for repository platforms due to the number of potential edge cases to support; they'd almost be better served to add an interface where you can test and see the results of SQL queries before running them, and automatically increment migrations.

**LOW PRIORITY**

New issue

[Jump to bottom](#)

# Bulk extractor enhancements #3

🔓 Open

radancy opened this issue on Jul 3 · 0 comments

Labels

archivematica

low priority

 radancy commented on Jul 3 • edited ▾

Contributor

BulkExtractor is currently available on the Appraisal tab (when transfer is set for *Examine contents* = "Yes"); it identifies credit card and US social insurance numbers.

Build on pattern-recognition capabilities to identify other potential pi.

- Would allow flagging materials requiring more detailed analysis (triage).




radancy added archivematica not yet ranked low priority and removed not yet ranked labels on Jul 3

New issue

[Jump to bottom](#)


# Format Policy Registry (FPR) #5

 **Open**

**radancy** opened this issue on Jul 3 · 0 comments

Labels

archivematica low priority



 **radancy** commented on Jul 3 

Contributor

Provide more user-friendly interface to let administrators make changes to the FPR.

Identity requirements

- Look into / test FPR2PAR application for sharing / querying FPR data.

  **radancy** added archivematica not yet ranked low priority and removed not yet ranked labels on Jul 3

New issue

[Jump to bottom](#)

# Submission documentation #6

🔓 Open

radancy opened this issue on Jul 3 · 0 comments

Labels

archivematica

low priority

 radancy commented on Jul 3

Contributor

Review practices for managing relationship between AIP and submission documentation (information about the transfer and its objects).

- What info would we like to capture?
- Where should it live (e.g. inside the AIP)?
- How is it related to the AIP?
- Can it incrementally change?

🏷  radancy added archivematica not yet ranked low priority and removed not yet ranked labels on Jul 3

New issue

[Jump to bottom](#)

# Object metadata extraction #14

**! Open**

radancy opened this issue on Jul 3 · 0 comments

Labels archivematica-atom low priority



radancy commented on Jul 3

Contributor

Technical metadata embedded in digital objects (e.g. image files) is captured in the AIP, but not easily accessed for descriptive purposes.

This project would develop tools to retrieve metadata from the METS file, map to AtoM / RAD fields, populate AtoM descriptive record on DIP upload.

🏷 radancy added low priority archivematica-atom labels on Jul 3



New issue

[Jump to bottom](#)

# Web archiving #15

! Open

radancy opened this issue on Jul 3 · 1 comment

Labels archivemata-atom low priority

 **radancy** commented on Jul 3


Contributor

Research and develop solutions for capturing, preserving, describing and providing access to archived websites.

- Currently Archivemata's FPR recognizes WARC files, extracts minimal metadata.

The project would include research, policy, and software development. Possible elements could include:

- Develop an acquisitions policy for web archiving.
- Determine the best methods for capturing WARC files.
- Test Archivelt software; devise a workflow to use Archivelt to create WARCs, ingest WARCs to Archivemata, link to the Internet Archive collections from AtoM as the access copy.
- Explore construction of an integrated WARC viewer for AtoM?

🏷  **radancy** added low priority archivemata-atom labels on Jul 3

 **axfelix** commented on Jul 9

Rhizome's productions are nicely standardizing right now, I think there's real possibility for an integration where there wasn't before.

New issue

[Jump to bottom](#)

# Minor AtoM user interface enhancements #17

**! Open** radancy opened this issue on Jul 3 · 0 comments



Labels **atom** low priority

 **radancy** commented on Jul 3 • edited ▾

Contributor

Identify and implement a set of relatively minor user interface enhancements. Possible features:

- Make treeview sticky: when you save a record, treeview reverts to top of tree, losing view of current record's position in the hierarchy).
- Add a "recently added" widget to the AtoM Home page to display new content; and provide an admin interface to manage content that appears.
- Support a rotating set of images for the AtoM Home page; provide an admin interface to allow easy upload / deletion.
- Restructure `Related materials`, `Associated materials`, and `Sources` fields as repeatable sets (Identifier, Note, URL); for `Associated materials` the URL can point to holdings in other repositories.
- Expose AtoM version number on the Home page (right now only an administrator can access; can be useful to know for liaising with AABC and MemoryBC).

  **radancy** added low priority atom labels on Jul 3

New issue

[Jump to bottom](#)

# Physical location module #19

! Open

radancy opened this issue on Jul 3 · 0 comments

Labels


atom

low priority

 radancy commented on Jul 3

Contributor

Identify requirements for and build out a physical locations module in AtoM to manage locations, boxes, and storage objects.

🏷  radancy added low priority atom labels on Jul 3

New issue

[Jump to bottom](#)

# Export digital objects #22

🔓 Open

radancy opened this issue on Jul 3 · 0 comments

Labels

atom


low priority

 radancy commented on Jul 3

Contributor

Provide an interface to allow users to download all digital objects for a given level of description (e.g. from a fonds, all child objects).

- Use case: researchers want to run visualization or data analytics on set of objects.

🏷  radancy added low priority atom labels on Jul 3

New issue

[Jump to bottom](#)

# User contributed content #24

! Open

 radancy opened this issue on Jul 3 · 0 comments

Labels


atom low priority

 radancy commented on Jul 3

Contributor

Provide support for AtoM public users to contribute descriptive content.

The first phase would likely be a research project to identify and scope out requirements.

🏷  radancy added low priority atom labels on Jul 3

New issue

[Jump to bottom](#)

# Authority records and access points #25


**Open**
radancy opened this issue on Jul 3 · 1 comment

Labels
atom
low priority




**radancy** commented on Jul 3

Contributor

Review issues relating to shared authority records across multiple repositories.

Possible enhancements could include:

- Allow multiple admin histories / bio sketches for one authority (e.g. contributed by different institutions).
- Allow linking of authority records to URIs maintained by other agencies (e.g. link "Simon Fraser" record to entries in SNAC, VIAF, LC Name Authorities, Wikipedia etc).
- Add controls on data entry for personal names (e.g. LastName, FirstName) to support consistent data entry.
- Support the use of previous forms of name as an access point.
- Load different thesauri of subject terms (e.g. Special Collections use different set of terms than Archives).
- Select which subject terms (by thesaurus) to include in export (e.g. don't want to send SFU-specific terms to MemoryBC)



**radancy** added low priority atom labels on Jul 3


**axfelix** commented on Jul 9

I'm doing some work during my upcoming leave on integrating Linked Data authorities which will hopefully come in handy here...

**NOT YET RANKED**

New issue

[Jump to bottom](#)

# CSV AtoM import utility #26

 **Open**

radancy opened this issue on Jul 3 · 1 comment

Labels

atom not yet ranked



 **radancy** commented on Jul 3

Contributor

The Archives has an Excel template to capture minimal file / item list data for use when processing records. It needs to be transformed into the standard AtoM csv template (with columns for every AtoM field). Currently RD does this with a FileMaker utility. Ideally we would have a Python-based drag-and-drop app that would do the transformation.

Map Archives' template fields to AtoM template.

- Develop script.
- Create drag-and-drop interface.

  **radancy** added atom not yet ranked labels on Jul 3

 **axfelix** commented on Jul 9

Should be doable, though at some point having all these apps becomes a maintenance burden, and if it's just for data entry rather than working on the filesystem (unlike Movelt and Provelt), might be better to look at ways of getting this into AtoM.



[New issue](#)[Jump to bottom](#)

# Selective de-indexing of digital objects #27

! Open

radancy opened this issue on Jul 3 · 3 comments

Labels

atom

not yet ranked



radancy commented on Jul 3 • edited ▾

Contributor

Support ability to remove indexing from selected digital objects for privacy reasons.

Use case: individual requests removal of name from public document (e.g. Convocation Book). Record would remain available in AtoM, but individual's name would not be picked up in searches.

radancy added atom not yet ranked labels on Jul 3



axfelix commented on Jul 9

I'm guessing this would require setting a database flag on the objects that's editable from the interface and gets ignored by all the indexing calls -- should be very doable within the existing codebase with a small-to-medium amount of effort.



axfelix commented on Jul 9

Would be easier if we don't need to worry about pre-setting this permission, as then we would only have to worry about handling it on objects that already exist in AtoM, rather than on ingest or communication from Archivematica.



axfelix commented on Aug 31

Adding this after the `page.goto();` block seems to work:

```
await page.evaluate((sel) => {
  var elements = document.querySelectorAll(sel);
  for(var i=0; i< elements.length; i++){
```

```
        elements[i].parentNode.removeChild(elements[i]);  
    }  
    }, "signup-prompt")
```