

CollectiveAccess

Summary

1. **Resolve the 21 collision clusters** via physical survey. Assign final identifiers to the 68 affected records.
 1. Remove 68 affected records to second sheet.
2. **Build the data model** in CA: object type list, metadata elements, entity relationship types.
3. **Build the vocabulary lists** in CA: `admc_material_category` (hierarchical) and `admc_material_terms` (flat), built from cleaned and merged DSpace subject fields.
4. **Build the storage location hierarchy** in CA.
5. **Pre-process the export**: parse `dc.description` into holdings flags, split subject strings on `||`, map rights to access values, load your new `idno` column.
6. **Write the import mapping**: two passes, entities first then objects.
7. **Test import** on a 20-record slice. Iterate.
8. **Production import** of the 310 clean records.
9. **Import the 68 collision records** after physical survey.
10. **Establish media naming convention** for future digitisation; no media ingest yet.

Data model → Lists/vocabularies → Import mapping → Test import → Production import → Media (later)

Resources

<https://camanual.whirl-i-gig.com/providence/>

<https://manual.collectiveaccess.org/providence/user/editihttps://camanual.whirl-i->

gig.com/providence/user/dataModelling/listsAuthoritiesng/lists_and_vocab.html

<https://manual.collectiveaccess.org/providence/user/dataModelling/primaryTables.html>

Primary tables

CA is structured around several primary tables, with editors that can be enabled or disabled depending on project requirements. The ones most relevant to the ADMC:

Table	What it holds
ca_objects	The physical or born-digital items themselves
ca_object_lots	Accession events grouping multiple objects acquired together
ca_entities	People and organisations (creators, donors, manufacturers)
ca_collections	Intellectual groupings of objects (series, fonds, donor collections)
ca_storage_locations	A hierarchical map of physical storage
ca_places	Geographic locations (hierarchical, linkable to GeoNames)
ca_occurrences	Flexible: events, exhibitions, publications, activities
ca_loans	Outgoing or incoming loan records
ca_movements	Optional: object movement history

An object record does not contain the entity's name as a text string. It contains a relationship to a separate entity record. This is the relational model in practice. If "Hansgrohe AG" is a manufacturer of 40 objects in the ADMC, there is one entity record for Hansgrohe, and 40 relationships from 40 objects to that record. Correct it once; it updates everywhere.

Metadata elements to create (or verify)

CA element code	Maps from	Data type	Notes
-----------------	-----------	-----------	-------

admc_idno_legacy	dc.identifier.other	Text	Preserve the old location code as a legacy field, non-searchable by default
admc_description	dc.description (substantive only)	Text (long)	Only for the ~20 records with real text
admc_physical_holdings	dc.description (parsed)	List (multi-value)	"Sample available," "Booklet available" as checkboxes
admc_manufacturer_url	dc.publisher.uri	URL	Product page; can also live on the entity record
admc_subject_classification	dc.subject.classification	List (hierarchical)	See vocabulary design below
admc_dspace_handle	dc.identifier.uri	URL	Preserve the DSpace handle for provenance
admc_intake_year	dc.date.issued	Text or Date	Record the upload year if at all; do not call it "date issued"

VRA Core fields that should already exist and need no new elements: title (`preferred_labels`), dimensions, material, technique, condition.

Sandbox -- AS

Task 2: Build the data model

Notes - 9 June

- Relationship Types:
 - Entity to collection: For ADMC, this can be the manufacturer to the objects. Related to option already included.
 - Entity to place: Includes built by, owned by, resided at, born at and worked at. (Would built by work best if place is to be included?)
 - Object to place: Includes depicts, describes, was created at and was located at. (Does created at works best for ADMC?)
- Object types:
 - Default of Works, can customize to different categories specific for ADMC like Physical Objects (ex. samples) and Other (ex. catalogs)

- They are all listed as physical objects in DSpace, but are there are differences between a pack of fabric samples and a pricelist catalog when it comes to specifics in the metadata?
- Term and Vocabulary:
 - CollectiveAccess already includes Getty's Art & Architecture vocab list
 - Or, customize using the terms from the subjects columns in DSpace
- Linking different subjects to a single object (ex. Walls, Metal): terms taken from dc.subject column, CollectiveAccess has the ability to use Library of Congress Subject Headings (Probably not using if taking from present subject terms? Cannot find LCSH in list/vocabulary page)
 - In List/Vocabulary List:
 - Subject (term) for Topics: conceptTopic or descriptiveTopic included, or customized with ADMC terms
 - Object label types: cannot find much related to labels on CollectiveAccess website, many sub-categories already present but does not seem to be related to subjects?
 - In Relationship List:
 - Object to vocabulary term: includes depicts or is described by sub-categories (Would these work for subjects? Since some objects use the subject to describe what the object is made of? Would that be depicting?)
 - How does the process of creating these relationships work? Would this process include first creating a list of vocab related to the ADMC, then putting the terms in the record/object metadata? Would this process automatically link the object to these terms, or would a relationship have to be manually created?
- Metadata elements to possible include: Object representations (can just be a media file, would this have to be included when creating records with images available?)
- Additional questions: Will dc.subject.other be included in the metadata as well? (dc.subject.classification already listed as to be created). Some records don't include dc.subject, but do include subject.classification or subject.other, does the level of subject matter depending on the record?

Notes - 10 June

- Creating entities:

- Corporate entity - before any editing
 - The only requirements is an entity identifier (should this follow the same format as the object identifiers?)
 - Preferred labels include Forename, Surname, etc. (should this be edited to better fit the need of the entities, since they will mostly be corporations, not individual people?)
 - Vocab, refid, Attribution, Culture, Dates and Role Set
- Corporate entity - possible customization
 - List and Vocabulary: entity types, Corporate (org)
 - Entity class (can chose between Individual person, Organization and Individual person without additional forenames). Would choosing Organization work best?
 - <https://support.collectiveaccess.org/d/301869-entity-types>
- Enabling Getty Art & Architecture Thesaurus
 - <https://support.collectiveaccess.org/d/301886-enabling-gettys-art-architecture-thesaurus-aat-vocabulary>
 - Add metadata element, "informationService" as the type, AAT as the service, "aat_vocabulary" as element code, follow same restriction as Library of Congress Subject Headings - done.
 - Questions: How to enable Getty to UI usage after adding metadata element? Will this then allow for fields to be added to records (object, entity, etc.)? Does a field for this need to be added (like how a metadata field for the Library of Congress Subject Headings is included in an object record)?
 - Seems to be available to search under subjects when editing an object record, but searches for terms that are on the Getty website (<https://www.getty.edu/research/tools/vocabularies/aat/>) yield no results in CollectiveAccess. Still marked as disabled under Lists & Vocabulary.
- Other issues
 - Uploading images to media object records
 - Able to upload a test image (downloaded from a different object record from DSpace as a test image) to the object record. Successfully uploaded and able to view through this interface and marked as "accessible to public". However, the image is not viewable through the public interface. When downloading the file, it says that the format is unsupported or corrupted.

Notes - 6 July

- Batch editing - must create a Set first
 - <https://docs.collectiveaccess.org/providence/user/editing/batch>
 - <https://docs.collectiveaccess.org/providence/user/workflow/sets>

Notes - 8 July

- Making customized metadata fields present in Pawtucket
 - Have to first customize the theme for the Pawtucket site:
 - <https://pawtucket2.readthedocs.io/en/latest/pawtucket/themes/defining.html>
 - Seems to be specifically editing the Configuration Files that will allow for customizing specific metadata fields in object records:
<https://pawtucket2.readthedocs.io/en/latest/pawtucket/themes/conf.html>
 - app.conf:
<https://pawtucket2.readthedocs.io/en/latest/pawtucket/general/app.html>
 - Add the element code for the metadata fields in the
ca_objects_detail_display_attributes:
<https://support.collectiveaccess.org/d/471-471>
 - Maybe also in detail pages - detail.conf:
<https://pawtucket2.readthedocs.io/en/latest/pawtucket/detail/detail.html>
 - From Pawtucket help page forum:
 - <https://support.collectiveaccess.org/d/294944-294944>
 - <https://support.collectiveaccess.org/d/299450-299450>
 - <https://support.collectiveaccess.org/d/110-110>
 - <https://support.collectiveaccess.org/d/299483-299483>
 - <https://support.collectiveaccess.org/d/299092-299092>

Revision #9

Created 2026-06-09 07:52:36 UTC by Librarian

Updated 2026-07-08 11:35:06 UTC by AS