Making Software into a Program

Columbia University Libraries’ Hyacinth User Group

(alphabetical order, * = presenting)

Columbia University Libraries
https://doi.org/10.7916/d8-2p4w-0t69
What is Hyacinth?

- Developed by CUL applications team
- Content management for a Fedora repository
- Input through web UI or .csv
- Controlled vocabulary management
- DOI registration
- Automated record creation
- Multimedia annotation

Spanish playing cards. ca 1915
Hyacinth Developers

Illustrated Map of Constantinople ("Byzantium Nunc Constantinopolis"). 1624

Eric O’Hanlon
Lead Hyacinth Developer

Benjamin Armintor
Head of Development
Infrastructure and Applications

Carla Galarza
Hyacinth Developer

Fred Duby
Hyacinth Developer
Hyacinth Users

Backend (input)
● Catalogers for digital collections and institutional repository
● Customer orders for digitization
● Marc record sync for items in Voyager ILS
● SWORD feeds for IR deposits and dissertations

Frontend (output)
● Academic Commons - institutional repository
● Digital Library Collections
● Special Collections custom sites
● DOI registration through DataCite

Notes on the language of the Tsuut’ina tribe. ca 1910
The Hyacinth User Group

Scope and Charge

A cross-divisional working group that will articulate the needs of diverse users, departments and collections is necessary to ensure collaboration, effectiveness, efficiency, transparency, and innovation in order for the system to continue to evolve.
The Hyacinth User Group

- designs, improves, and documents shared workflows.
- provides a hub for software support and issue tracking.
- facilitates communication between users and developers.
- selects controlled vocabularies.
- advises on metadata elements and schema mapping.
- manages shared projects, such as authority control, metadata cleanup, and user testing.
- recommends features for future development.

Chinese poster illustrating the story of Moses ca. 1932
Relatedness

Issue
- Linking related digital objects in Hyacinth

Stakeholders
- Cataloging/metadata staff
- Scholarly communications staff
- End users

Method
- Discussions at Hyacinth User Group
- Joint development of use cases and specifications (shared documents)

Fig.: Carnegie Corporation of New York. Report of the president, the secretary & the treasurer. 1942
Relatedness

Joint development of use cases and specifications: Challenges

- Incorporating all use cases into a MODS-based data model
- Deciding what to omit to prevent unnecessary bloat in the data model
- Determining preferred modeling when there are several possible ways to model

```xml
<relatedItem type="succeeding">
  <name>
    <namePart>Carnegie Museum</namePart>
  </name>
  <titleInfo>
    <title>Annual report of the Carnegie Museum</title>
  </titleInfo>
  <identifier type="issn">0161-5017</identifier>
</relatedItem>
```

*Fig.: Target MODS output for a related item element, with a type “succeeding”*
Relatedness

Joint development of use cases and specifications: Outcome

- A single complex relatedItem element with several subelements: type, title, name, part number, identifier, Columbia content

Fig.: draft design of a data entry form for the element
## Controlled Vocabularies

### Registered Controlled Vocabularies

<table>
<thead>
<tr>
<th>Collection</th>
<th>Action</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Form</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Frequency</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Funder</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Genre</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Geonames</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Language</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>License</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Location</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Name</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Name Role</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Rights Statement</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Subject Geographic</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
<tr>
<td>Subject Name</td>
<td>Edit</td>
<td>Manage Terms</td>
</tr>
</tbody>
</table>
Controlled Vocabularies

External
Source: id.loc.gov, FAST, AAT, Geonames, iso639-2b, Crossref Funder Registry, ISNI, etc.

Used for: Agents, Culture, Form/Genre, Subjects, Languages, Rights Statements, Funder, Roles

Local
Used for: Collection (archival), Location (repository)

Temporary
Used for: Situation when a URI is coined during data ingest. The term could not be reconciled during metadata creation or when metadata is mapped over from another data store, such as the library catalog.

Breviary. 13th century
Example: Collections

Local Term: Albert Field Collection of Playing Cards

Controlled Vocabulary: Collection
Type: local
Value: Albert Field Collection of Playing Cards
URI: http://id.library.columbia.edu/term/57155a98-0345-489f-bbe7-a7fc
Internal ID: 55
CLIO ID: 13484140
Coming soon: **Hyacinth 3**

**Who is a Hyacinth user?**

- Granular admin permissions
- User specs on a field-by-field basis
- A space for documentation & training standards
Future Developments

- Priority-setting after big projects
- Guard against over-specialization
- Clear and thorough documentation
- User assessment
References

Hyacinth: https://github.com/cul/ldpd-hyacinth
Digital Library Collections: https://dlc.library.columbia.edu
Academic Commons: https://academiccommons.columbia.edu
This presentation: https://doi.org/10.7916/d8-2p4w-0t69
Images from Columbia’s collections, selected by Meredith Self

Copy of a mural in the Kataluva Purvarama Maha Viharaya temple, Sri Lanka. 1971