Open Metadata Provisioning Investigation Charter

I. Problem/Value Statement

Problem Statement:
The Harvard Library creates and maintains descriptive metadata about its vast holdings according to several different standards and formats in a variety of systems, and repurposes and shares that metadata through a range of mechanisms. Over the past decade the systems in which the metadata is created, the approaches to sharing metadata, and the needs of those using Harvard Library metadata have all changed, and the technical environment, workflows, and data flows among them have become ever more complex. It is time to reassess the ways Harvard shares its descriptive metadata to align with the requirements of today's users, and to do so in a way that is scalable and technically sustainable.

Background:
Since the mid-1980s, Harvard libraries have created and acquired metadata describing materials in their collections with the goal of supporting cooperative reuse and chose not to acquire metadata with access restrictions. This enabled the libraries to participate in domain-based union catalogs with other institutions, for example, such as for theology and for law. Since 2012, Harvard Library has had an explicit policy to provide open access to library metadata, subject to legal and privacy factors. Various portions of the Library's descriptive metadata have been made openly available without registration or cost through bulk download, via APIs including LibraryCloud and DASH, and as sets from an OAI data provider. The Library contributes metadata to collaborative discovery platforms such as the Digital Public Library of America (DPLA) and resource sharing initiatives such as ReCAP. Library systems also support internal reuse, some of which includes the small proportion of records that may not be shared due to legal constraints.

Business Value:
The investigation will inform future development to meet several goals:

- to facilitate both programmatic and creative reuse of Harvard Library metadata by faculty, students, researchers, and staff, within and beyond Harvard, by providing metadata with the scope, in the formats, and via the mechanisms that will support their work, and
- to rationalize, clarify, and, ideally, simplify the provision of metadata to improve maintainability, responsiveness to change, and the accuracy, completeness, and timeliness of the metadata.

II. Vision and Approach

Describe the solution:
Investigate current needs and desiderata in support of reuse of library metadata to inform and help shape a future technical project to modernize Harvard Library's open metadata infrastructure.

Deliverables/Work Products:
A report encompassing

- a review of Harvard Library's current open metadata offerings
- capabilities of Harvard's metadata source systems
- a landscape review of metadata sharing infrastructure at peer institutions
- an inventory of representative use cases, and
- functional requirements for an updated metadata provision service
- prioritized recommendations to improve open metadata offerings for Harvard library and archives metadata

Defining Done:
Review and acceptance of the report by the executive business owner.

In Scope:

- Harvard Library's current open metadata offerings
- Available outputs of descriptive and holding data from the major systems of record for Harvard Library metadata (Alma, ArchivesSpace, JSTOR Forum, HGL, plus DASH, Dataverse, and Mongo)
- Use cases for open metadata and/or metadata sharing, including functional requirements (e.g., user-defined bulk downloads, partner-specific customizations)
- Metadata sharing practices by peer institutions
- Metadata standards, formats, and serializations
- Mechanisms for sharing metadata (e.g., OAI, APIs, bulk downloads, custom downloads)

Out of Scope:

- Metadata storage
- Technical specification and development
- Linked data metadata projects
- Administrative, technical, structural, and presentation metadata about, or embedded in, DRS content
- Inventory data (e.g., HD)
- OCR
- Playlists
III. Stakeholders and Project Team

Glossary

Stakeholders

Who is sponsoring the work? Who is funding the work? Who will accept the work? What organizations, departments, or people will benefit from this work (for medium and large projects)? Link to related governance structure wiki page(s) where relevant.

<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Title</th>
<th>Participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claire DeMarco</td>
<td>Associate University Librarian for Discovery and Access</td>
<td>Executive Business Owner</td>
</tr>
<tr>
<td>Stu Snydman</td>
<td>Managing Director of Library Technology Services</td>
<td></td>
</tr>
</tbody>
</table>

Project Team (tentative proposed)

Roles: Project Manager, Business Analyst, Quality Assurance Analyst, Architect, Software Engineer, Systems Engineer, UI Designer, Metadata Analyst, Subject Matter Expert

<table>
<thead>
<tr>
<th>Team Member</th>
<th>Role(s)</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robin Wendler</td>
<td>Project Manager</td>
<td>HUIT / Library Technology Services</td>
</tr>
<tr>
<td>Corinna Baksik</td>
<td>Subject Matter Expert</td>
<td>HUIT / Library Technology Services</td>
</tr>
<tr>
<td>Ceilyn Boyd</td>
<td>Subject Matter Expert</td>
<td>HL / Research Data Management</td>
</tr>
<tr>
<td>Michelle Durocher</td>
<td>Subject Matter Expert</td>
<td>HL / ITS - Metadata Management</td>
</tr>
<tr>
<td>Marc McGee</td>
<td>Subject Matter Expert</td>
<td>HL / ITS - Metadata Management</td>
</tr>
<tr>
<td>Vanessa Venti</td>
<td>Subject Matter Expert</td>
<td>HL / User Experience and Discovery</td>
</tr>
</tbody>
</table>

IV. Cost and Schedule

There are no out-of-pocket costs associated with the investigation project.

Report to be submitted no later than December 17, 2021

IV. Key tasks and outcomes

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Outcomes</th>
<th>Responsible Parties</th>
</tr>
</thead>
</table>

VI. Assumptions, Risks, and Constraints

Assumptions:

- Stakeholders have identified the appropriate subject matter experts to participate in the Working Group and who can accurately and completely define the business requirements for the project
- Stakeholders will have made available the time required to participate in project activities and to complete tasks as requested
- Project sponsor and other stakeholders are empowered to make the decision required for the project to be a success
- Project sponsor will provide written approval to move forward with system development when requested as part of incremental/iterative system demonstrations

Dependencies:

Risks (description, plan, impact, owner):

<table>
<thead>
<tr>
<th>Description</th>
<th>Plan</th>
<th>Impact</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Update during course of project as needed.