

Final summary report on Harvard Scholarship Indexing Project
November 16, 2012

PROJECT SUMMARY

The long-range goal of the Harvard Scholarship Indexing Project (HSIP) is to build a lightweight metadata-only repository to catalogue and link to scholarly artifacts created at Harvard, to provide:

- A panoramic institutional record of Harvard scholarship
- A one-stop process for metadata entry and claiming of auto-populated records
- Standard data formats to ensure reusability and system interoperability
- Bilateral data exchange among the HSIP repository, School-based systems, and external data sources
- A foundation for data-driven decision processes in Library strategy and other University-wide academic planning endeavors

As an initial planning phase, the Library Lab funded a study to gather requirements and assess options for creating within the Library a University-wide bibliographic metadata repository for data management, discovery, and distributed access to works produced by Harvard scholars.

The Library Lab award was used to hire two consultants with expertise in library technologies, standards, and metadata to gather and analyze requirements and produce a report that addresses:

- Assessment of overall need and feasibility
- School, Library, and IT requirements
- Data formats and standards
- Implementation options and recommended next steps

The resulting report reiterates the need for the HSIP service and provides a roadmap forward. Harvard Faculty Finder, based on Catalyst Profiles, is highlighted as the best Harvard-internal technology option, given: (1) its use of the now-standard VIVO ontology for representation of bibliographic and other metadata elements, and (2) University support for the Harvard Faculty Finder project, which is already indexing and importing a wide range of faculty related information including structured publication data. That said, Profiles/Finder is not optimized for end-user input and may need to be enhanced or integrated with other proposed technologies.

As the next phase in the development of the service, the report recommends one or more pilot projects to evaluate feasibility and demonstrate:

- The advantages of harvesting bibliographic entries from external systems to save faculty time and effort and to provide quality citations
- A manual data-entry interface and HSIP data format
- Integration with the DASH repository
- An alert mechanism for auto-updates

The pilot opportunity encouraged in the report is to partner with an interested School to implement an integration of Profiles/Finder and OpenScholar or Symplectic Elements. In the proposed prototype, the former serves as the data aggregation platform and is enhanced to provide more robust data export capabilities, based on expanded sources of data, and the latter functions as the user interface and data entry component.

Among the report appendices is a detailed outline of the metadata elements required for the proposed HSIP service to ensure re-usability and interoperability, in order to fulfill its potential to exchange data among the HSIP repository, School-based systems, and outside data sources.

ACCOMPLISHMENTS

Report completed; next steps mapped out.

CHALLENGES

Less detailed metadata format recommendations than hoped for.

NEXT STEPS

One or two pilot projects with willing Schools or departments.