

**Institute of Museum and Library Services (IMLS)
Digital Collection and Content Project Phase 2 –
University of Illinois, Urbana-Champaign**

Supporting Statement A

A. Justification

- 1. Explain the circumstances that make the collection of information necessary. Identify any legal or administrative requirements that necessitate the collection. Attach a copy of the appropriate section of each statute and regulation mandating the collection of information.**

The Institute of Museum and Library services (IMLS) is required by the Library Services and Technology Act “to stimulate excellence and promote access to learning and information resources in all types of libraries for individuals of all ages;” “to promote library services that provide all users access to information through State, regional, national and international electronic networks;” and “to provide linkages among and between libraries.” (PL 104-208 as amended, title VII, Subtitle B, Sec. 212).

As efforts to integrate and federate digital resources proceed apace, information scientists are learning more about the problems that emerge at different levels of scale and granularity. Building on prior work of the Digital Collections and Content project (DCC), researchers at the University of Illinois at Urbana Champaign (UIUC) will investigate and implement a systematic approach that confronts these problems and offers robust means for adding value and improving access to existing digital aggregations. Based on over four years of development and research experience, the project has identified new core initiatives that can substantially upgrade the quality of the IMLS digital collection registry and metadata repository for users, as well as, advance the current base of knowledge and practice in the federation of digital collections. Work to date has positioned the project to move into a more highly informed level of development and theoretical work as it continues to test and ground advances in the engineering pragmatics of building, learning, and documenting.

The IMLS Digital Collections and Content (DCC) project has created two products: 1) a publicly available registry of IMLS National Leadership Grant (NLG) and LSTA digital collections; and 2) a repository of item-level metadata available from these collections. The initial round of the project (2002-2005) examined shareable metadata, collection description, and the scalability of OAI-PMH. An extension of the project (2005-2007) included an examination of the necessary requirements for extending these resources beyond 2007 and the possible benefits for long-term maintenance of the portal. In 2007 IMLS announced continued funding through the NLG program to maintain and enhance the resources developed in previous grant cycles, including continued harvesting of metadata and refreshing of the database, while answering research questions encountered during previous IMLS DCC development.

As part of this research, the project will contact all NLG digital projects for the years 2007-2010 and selected LSTA digital projects. Responses from the survey will be used to create brief descriptive records about the collections that will be vetted by the collection managers. Records will initially remain password protected until verified by project managers. If the information in the collection description is sufficient, the collection manager may simply submit the record as it stands. The involvement of the collection managers in the creation and vetting of records for the collection registry will help ensure that the registry is accurate and will promote use of the registry as a discovery tool.

- 2. Indicate how, by whom, and for what purpose the information is to be used. Except for a new collection, indicate the actual use the agency has made of the information received from the current collection.**

The purpose of collection level descriptive records describing digital collections funded through IMLS is two fold. The descriptive records will be part of a searchable database that will be available (for the duration of the project) to IMLS, the UIUC research team, participating IMLS grantees, and the general public. This database will help to promote the discovery of dispersed digital collections. The second purpose is to research the effectiveness of collection level description to facilitate end-user access and to provide additional context for the content within these collections. This is a research goal of the project.

- 3. Describe whether, and to what extent, the collection of information involves the use of automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submission of responses, and the basis for the decision for adopting this means of collection. Also describe any consideration of using information technology to reduce burden.**

Initial survey responses will be collected using an online survey tool. Using this information, a brief record will be created by the project staff, and the collection manager will be asked to verify and add/edit information. These entry/edit forms will be available only through a web interface. This form of data collection allows automated update of the registry database and allows grantees to correct information they have already submitted without going through the project coordinator or other personnel.

- 4. Describe efforts to identify duplication. Show specifically why any similar information already available cannot be used or modified for use for the purposes described in item 2.**

The UIUC team will collect preliminary information from NLG grant applications and collection webpages. Information requested in the survey will augment and update gathered information. Collection managers will verify and augment the collection description information with additional information not otherwise

available or obtainable, except directly through the personnel maintaining these collections.

The second phase of this project seeks to build upon earlier records, and collect new metadata on the grants awarded from 2007-2010. In this way, this collection is not duplicating metadata previously collected.

5. If the collection of information impacts small business or other small entities (Item 5 of OMB Form 83-I), describe any methods used to minimize burden.

While the collection will include small non-profit organizations, no significant impact is expected. The two forms should take approximately 50 minutes combined.

6. Describe the consequences to Federal program or policy activities if the collection is not conducted less frequently, as well as any technical or legal obstacles to reducing burden.

The survey will only be administered once.

- 7. Explain any special circumstances that require the collection to be conducted in a manner**
- requiring respondents to report information to the agency more often than quarterly;
 - requiring respondents to prepare a written response to a collection of information in fewer than 30 days after receipt of it;
 - requiring respondents to submit more than an original and two copies of any document;
 - requiring respondents to retain records, other than health, medical, government contract, grant-in-aid, or tax records, for more than three years;
 - in connection with a statistical survey, that is not designed to produce valid and reliable results that can be generalized to the universe of study;
 - requiring the use of a statistical data classification that has not been reviewed and approved by OMB
 - that includes a pledge of confidentiality that is not supported by authority established in statute or regulation that is not supported by disclosure and data security policies that are consistent with the pledge, or which unnecessarily impedes sharing of data with other agencies for compatible confidential use; or
 - requiring respondents to submit proprietary trade secrets or other confidential information unless the agency can demonstrate that it was institute procedures to protect the information's confidentiality to the extent permitted by law.

Not applicable.

8. **If applicable, provide a copy and identify the date and page number of publication in the Federal Register of the agency's notice, required by 5DFR 1320.8(d), soliciting comments on the information collection prior to submission to OMB. Summarize public comments received in response to that notice and describe actions taken by the agency in response to these comments. Specifically address comments received on cost and hour burden.**

Describe efforts to consult with persons outside the agency to obtain their views on the availability of data, frequency of collection, the clarity of instructions and recordkeeping, disclosure, or reporting format (if any), and on the data elements to be recorded, disclosed, or reported.

Consultation with representatives of those from whom information is to be obtained or those who must compile records should occur at least once every 3 years - even if the collection of information activity is the same as in prior periods. There may be circumstances that mitigate against consultation in a specific situation. These circumstances should be explained.

Initial publication in the Federal Register was on March 28, 2008 (Volume 73, Number 61) Page 16725. We received no comments through the publication in the Federal Register.

The schema used for the collection description was posted on the UIUC project web site at http://imlsdcc.grainger.uiuc.edu/CDschema_overview.htm. In addition, the UIUC project team consulted with their steering committee (a group of experts in digital libraries and collections) to gather feedback and solicited feedback about what should be included in a collection description schema from recipients of National Leadership Grants. Comments received generally focused on what elements of the collection description should be required and which should be optional and which controlled vocabularies should be used for specific types of elements (subject, geographic locations). The project team responded to these comments by minimizing the number of required elements and by implementing a controlled vocabulary for both the subject heading (the Department of Education GEM subject heading) and the geographic location (Getty Thesaurus of Geographic Names). We did not receive any comments about the cost and hour burden.

9. **Explain any decision to provide any payment or gift to respondents, other than remuneration of contracts or grantees.**

Not applicable.

10. **Describe any assurance of confidentiality provided to respondents and the basis for the assurance in statute, regulation, or agency policy.**

The initial collection level record created from this information will be password protected and available to only the project team and the collection manager. After the project manager has reviewed the record, it will be made publically available in the collection registry. Within the collection registry, the only personal information we request is the name and email address of the person

responsible for the digital collection. However, this information is for administrative use only and will be suppressed on the public interface of the collection registry.

- 11. Provide additional justification for any questions of a sensitive nature, such as sexual behavior and attitudes, religious beliefs and other matters that are commonly considered private. This justification should include the reasons why the agency considers the questions necessary, the specific uses to be made of the information, the explanation to be given to persons from whom the information is requested, and any steps to be taken to obtain their consent.**

IMLS does not collect sensitive information from grant applicants. Any research project funded by IMLS under this program is subject to federal statutes and regulations governing research on human subjects. The project team has received Institutional Review Board exemption at the University of Illinois at Urbana-Champaign.

- 12. Provide estimates of the hour burden of the collection of information. The statement should:**
 - **Indicate the number of respondents, frequency of response, annual hour burden, and an explanation of how the burden was estimated. Unless directed to do so, agencies should not conduct special surveys to obtain information on which to base hour burden estimates. Consultation with a sample (fewer than 10) of potential respondents is desirable. If the hour burden on respondents is expected to vary widely because of differences in activity, size, or complexity, show the range of estimated hour burden and explain the reasons for the variance. Generally, estimates should not include burden hours for customary and usual business practices.**
 - **If the request for approval covers more than one form, provide separate hour burden estimates for each form and aggregate the hour burdens in Item 13 of OMB Form 83-i.**
 - **Provide estimates of annualized cost to respondents for the hour burdens for collections of information, identifying and using appropriate wage rate categories. The cost of contracting out or paying outside parties for information collection activities should not be included here. Instead, this cost should be included in Item 13.**

IMLS estimates the following burden for collection of this information:

Number of Grantees Asked to Complete Survey: 360
Estimated response time – .3 hours
Estimated cost per respondent: \$7.50 (.3 hours x \$25 per hour)
Estimated total burden: .3 hours x 310 = 108 hours, \$2700
Estimate based on: UIUC research team estimate

Collection Registry Entry Edit Forms:

Number of Grantees Asked to Review/Edit Forms: 360

Estimated response time – .5 hours

Estimated cost per respondent: \$12.50 (.5 hours x \$25 per hour)

Estimated total burden of individual interviews: .5 hours x 310 = 180 hours, \$4500

Estimate based on: UIUC project team estimate

- 13. Provide an estimate of the total annual cost burden to respondents or recordkeepers resulting from the collection of information. (Do not include the cost of any hour burden shown in items 12 and 14).**
- **The cost estimate should be split into two components: (a) a total capital and start-up cost component (annualized over its expected useful life) and (b) a total operation and maintenance component. The estimates should take into account costs associated with generating, maintaining, and disclosing or providing the information. Include descriptions of methods used to estimate major cost factors including system and technology acquisition, expected useful life of capital equipment, the discount rate(s), and the time period over which costs will be incurred. Capital and start-up costs include, among other items, preparations for collecting information such as purchasing computers and software, monitoring, sampling, drilling and testing equipment; and record storage facilities.**
 - **If cost estimates are expected to vary widely, agencies should present ranges of cost burdens and explain the reasons for the variance. The cost of purchasing or contracting out information collection services should be a part of this cost burden estimate. In developing cost burden estimates, agencies may consult with a sample of respondents (fewer than 10), utilize the 60-day pre-OMB submission public comment process and use existing economic or regulatory impact analysis associated with the rulemaking containing the information collection, as appropriate.**
 - **Generally, estimates should not include purchases of equipment or services, or portions thereof, made: (1) prior to October 1, 1995, (2) to achieve regulatory compliance with requirements not associated with the information collection, (3) for reasons other than to provide information or keep records for the government, or (4) as part of customary and usual business or private practices.**

Recordkeeping burden for respondents caused by these data collections: None. Data will be provided by participants' based on their knowledge of their NLG or LSTA project. Collection-level descriptive information generally does not fluctuate greatly and the collections described are stable entities. We expect that data collection will occur once and/or as a collection changes substantially

- 14. Provide estimates of annualized cost to the Federal Government. Also, provide a description of the method used to estimate cost, which should include quantification of hours, operational expenses (such as equipment, overhead, printing, and support staff), and any other expenses that would not have been incurred with the collection of information. Agencies may also aggregate cost estimates from paragraphs 12, 13, and 14 in a single table.**

These data collections are part of IMLS Cooperative Agreement LG-06-07-0020, \$975,903 to UIUC for three years.

15. Explain the reasons for any program changes or adjustments reported in Items 13 and 14 of the OMB Form 83-i.

Not applicable

16. For collections of information whose results will be published, outline plans for tabulation and publication. Address any complex analytical techniques that will be used. Provide the time schedule for the entire project, including beginning and ending dates of the collection of information, completion of report, publication dates, and other actions.

Grant period is October 1, 2007 through September 30, 2010. Survey will be distributed once OMB approval is acquired and will end at the completion of the grant project. The collection registry (which displays the descriptive information verified and augmented by the grantees) is publically available. Publication of research into the effectiveness of the collection registry and collection description will follow through presentations at the professional conferences such as American Society for Information Science and Technology (ASIST) and Joint Conference on Digital Libraries (JCDL) and publications in the scholarly journals such as Journal of American Society for Information Science and Technology (JASIST) and D-lib., but dates are not yet established.. **IMLS PUBS**

17. If seeking approval to not display the expiration date for OMB approval of the information, explain the reasons that display would be inappropriate.

Not applicable

18. Explain each exception to the certification statement identified in Item 19, "Certification for Paperwork Reduction Act Submissions," of OMB Form 83-I).

Not applicable

B. Collections of Information Employing Statistical Methods

Not applicable