Documenting a Metadata Standard for the Performing Arts:

An Application Profile for the Global Performing Arts Database (GloPAD)

The early development of the Global Performing Arts Database (GloPAD), a database of digital objects and data related to the performing arts from across the globe and over time, highlighted the need for a metadata standard for the performing arts. Unlike most other databases, a collection of data on the performing arts must interact with numerous types of media as well as deal with the ephemera nature of the art, and the complex issue of multiple creators. A scholarly global database must also deal with the various languages involved in the many performances it records. To do this metadata standards are crucial. With support from an IMLS grant (2002-2006), we have addressed the lack of an existing standard by creating a sophisticated metadata schema that accommodates the particular descriptive requirements of performance. This schema goes far beyond anything currently available; however, it will continue to be refined and expanded.

GloPAD provides access to intellectual and artistic works through digital objects, which are digital representations of various types of recordings (e.g. photos, sound tracks, videos) of the performance itself and of the artifacts created for or about that performance (costumes, posters, photos of creators, etc.) Because the performing arts exist only in the moment of their performance, an image, audio, video, or complex media object, whether born digital or digitized, must stand in for the performance moment itself. And the artifacts surrounding the performance, which may still be extant, are also entered in the database by digital representations. While there are artifacts and activities, common to many types of performance (such as stage designs, scripts, and costumes), they are often named and categorized in disparate ways; therefore a standardization scheme of labels is needed to contextualize these items within a database so that they can be appropriately retrieved and compared.

To ensure that the hundreds of individuals involved in the project maintain a high degree of consistency, we need to control the vocabulary input into most database fields. This requires researching existing controlled vocabularies in order to make use of standardized terms whenever possible. The Library of Congress, Getty, and International Standards Organization (ISO) name authorities, subject headings, language and country listings, and specialized thesauri have been particularly useful.

The metadata schema we have developed includes eleven major tables covering people, places, objects, and events related to the performing arts. It has been implemented in a PostgreSQL relational database management system with a web-based, password-protected, multilingual input interface for editors—the scholars, curators, librarians and others who enter the digital objects and accompanying metadata--and a public interface currently providing access to over 6,000 digital objects from the collections of GloPAC participants. The metadata schema is

realized most strictly in the editors' interface, where templates for data entry in several languages (currently Russian, Japanese, Chinese, and German) as well as in the base language of English allow editors to upload digital objects and describe them and their complex relationships. The public interface is one expression, among many possible ones, of the connections between the Digital Objects and their related data. In creating our metadata schema, we incorporated a number of existing schemas. The most universal is the Dublin Core Metadata Initiative (http://dublincore.org/documents/2005/06/13/dcmi-terms/), but we also looked to more specialized schemas developed for the visual arts, such as the Visual Resources Association Core 3.0 (http://www.vraweb.org/vracore3.htm), the Getty Categories for Describing Works of Art http://www.getty.edu/research/conducting_research/standards/cdwa/2_overview/index.html), and the Art Museum Image Consortium Data Specification (http://www.amico.org/AMICOlibrary/dataDictionary.html). We also used existing schemas created for a single format among the many represented in GloPAD—e.g. NISO Metadata for Images in XML Schema, at http://www.loc.gov/standards/mix/--and even those originally intended for physical rather than digital objects (U.S. Machine-Readable Cataloging). The following application profile reveals how GloPAD maps to these other schemas. The application profile also documents the GloPAD metadata schema. According to the *Dublin Core Application Profile Guidelines*, "An Application Profile is a form for documenting which terms a given application uses in its metadata, with what extensions or adaptations, and specifying how those terms related both to formal standards such as Dublin Core as well as to less formally defined element sets and vocabularies."[1]

The GloPAD application profile serves as an interchange format between various systems using different metadata standards. The attached version includes the most granular metadata mappings we have, arranged in alphabetical order by the name of the element. In order to allow the metadata records to be harvested using OAI/PMH (Open Archives Initiative Protocol for Metadata Harvesting), we also created a simplified version, using unqualified Dublin Core (the most basic metadata language), of just the descriptive metadata relating to database records. The simplified records lose the richness of the full records, but gain the ability to be discovered alongside other digital collections in the University of Illinois at Urbana-Champaign's IMLS Digital Collections Registry (http://imlsdcc.grainger.uiuc.edu/collections).

The first page of the Application Profile for GloPAD Performing Arts Metadata Schema contains versioning information and a list of all of the namespaces for other schemas referenced in the document. This document has gone through many modifications, because of changes in the Global Performing Arts Database and changes in the standards themselves. We will continue to update it as appropriate.

In this document each metadata term (element) is described in tables with the format shown below, which includes additional annotations:

[1] Comité Européen de Normalisation. *Dublin Core Application Profile Guidelines*. CWA 14855, November 2003. Retrieved from ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/MMI-DC/cwa14855-00-2003-Nov.pdf on February 12, 2006.

Term Name	Proper name given to the metadata element in a particular metadata schema (also known as a namespace).
Term URI	A Uniform Resource Identifier used to identify the term. URIs are recommended as a best practice but thus far Dublin Core is the only namespace using them, so this field is often blank.
Namespace	The namespace (metadata schema) from which the metadata element derives.
Data Values	How the data should be formatted, and from which controlled vocabulary or vocabularies it has been taken.
Definition	The element's definition from its namespace.
Mapping	Elements from other standards this term maps / relates to
GloPAD Interface Name	The field on the GloPAD Editor Interface form represented by this element.
GloPAD definition	An expansion of the namespace's definition highlighting the implementation and usage of the element in GloPAD.
Obligation	Indicates whether the element is required to always or sometimes be present (Mandatory, Conditional, or Optional).
Occurrence	Indicates any limit to the repeatability of the element—e.g. 1 for once, and unbounded for unlimited occurrences.
Example Data	GloPAD example of how the element should be populated.

This application profile was initiated by Kari Smith and developed by Rachel Howard.