
Sharon E. Dennis, M.S.
Spencer S. Eccles Health Sciences Library
University of Utah
sdennis@lib.med.utah.edu

Chris S. Candler, M.D.
University of Oklahoma College of Medicine
chris-candler@ouhsc.edu

Sebastian Uijtdehaage, Ph.D.
David Geffen School of Medicine at UCLA
bas@ucla.edu

Shona R. Dippie, M.L.S.
Spencer S. Eccles Health Sciences Library
University of Utah
sdippie@lib.med.utah.edu

Sandra A. McIntyre, M.Ed.
David Geffen School of Medicine at UCLA
smcintyre@mednet.ucla.edu

Abstract

Health sciences educators are increasingly incorporating multimedia (including images, animations, and videos) into educational materials such as PowerPoint lectures, Web sites, and interactive quizzes and cases. Educators continue to "reinvent the wheel" and develop costly duplicates of multimedia resources, despite new opportunities offered by the Internet to share resources. The Health Education Assets Library (HEAL) is designed to provide health sciences educators with freely available, high quality multimedia materials to augment health sciences education. We describe an XML schema that we created to index the health sciences multimedia resources in HEAL. The metadata schema provides a common mechanism by which remotely located distributed systems may share metadata records, allowing the end user to search many collections through one interface.

1. Introduction

Rapid advances in technology have created new opportunities for health sciences educators to integrate multimedia resources into the curriculum. A high quality image with annotations, an animation depicting a difficult concept, or a video of a lecture that can be reviewed "on demand" can enrich the educational experience for both faculty and students. Web-based course management systems now make it possible for faculty to offer students a wide variety of course materials in a customized environment. Educators can offer students interactive experiences through online cases and quizzes that are not possible with traditional educational methods. The use of educational technology by health sciences educators has increased steadily for years; as the Internet has become ubiquitous, use has increased dramatically [1 2 3 4 5]. Despite the increased use of technology in health sciences education, many faculty members continue to "re-invent the wheel" by creating "yet another" brain atlas or video archive of neurologic signs and symptoms. This occurs either because the faculty member is unaware of similar resources or does not have access to materials that may be re-purposed and adapted to their particular curricular needs and teaching style. The need for establishment of consortia and standards to facilitate collaborative development and sharing of educational resources is apparent [6 7 8]. Educators would obviously benefit from an international multimedia repository that would offer "one-stop shopping" for freely available, high quality multimedia resources in a variety of subject areas.

2. Methods

In 2000 the authors received an award from the National Science Digital Library Initiative (NSDL) [9] of the National Science Foundation (NSF) to plan, design, and develop a web-based application to facilitate sharing of healthcare education materials. [10] The Health
Education Assets Library (HEAL) [11] was subsequently established as a core component of the NSDL whose purpose is to provide educators with free, quality multimedia materials to augment health sciences education. The goals of the HEAL project are fourfold: i) to improve access to high quality multimedia resources, ii) to promote sharing of educational resources between faculty and institutions, iii) to foster interoperability through the creation of standards, and iv) to establish a sustainable infrastructure. Through collaboration with numerous faculty, medical schools, the Association of American Medical Colleges, and the National Library of Medicine, the HEAL project strives to meet the educational resource needs of healthcare educators across the educational continuum.

The HEAL application is available to the public at our Web site (http://www.healcentral.org). An initial prototype collection of 2500 items was made available beginning June 2003. In 2002, the authors received a second NSF grant to substantially increase the breadth and depth of the HEAL collection. We anticipate offering upwards of 50,000 items by the fall of 2004.

The HEAL application allows users to browse the collection using a hierarchical controlled vocabulary tree (Figure 1) or to conduct a simple keyword search. In addition, users may conduct an advanced search that includes options for narrowing a search by author, subject heading, or health science specific characteristics (specimen type, radiograph type, etc.) (Figure 2). The search results page includes a thumbnail depiction of each item in addition to a description, a reference to the source collection, the file format and size; users can add an item to a download folder for later batch downloading.

While the HEAL digital library accepts contributions from individuals on our Web site, for the collection to grow in subject breadth and depth as envisioned a proactive collection development strategy was required. To that end, we formed partnerships with other institutions and individuals that had already developed high quality collections of multimedia resources. To date over 100 partners have expressed an interest in contributing a collection to HEAL. Since uploading multimedia materials from these external collections onto the central HEAL catalog independent of their desire to contribute some or all of their multimedia resources (Figure 3).

3. Metadata Significance

Consistent metadata and controlled vocabulary standards for describing each multimedia resource in the HEAL catalog facilitate retrieval of relevant items from the database and the management of items within the system. The descriptive information about each item is referred to as "metadata" (data that describes the content, form, and other characteristics of information using a standardized structure and vocabulary). A metadata standard consists of a precise set of descriptive, technical, and administrative elements; for instance, "title" to describe the title of the resource, "author" to describe the author, etc. A common standard is necessary to ensure that there is no confusion regarding the meaning of each element; for instance, an element called "title" could be interpreted in multiple ways (i.e., a faculty “title” vs. the “title” of the resource). A controlled vocabulary standard further enhances the accuracy and consistency of metadata element values within and between systems. Standards establish common rules and principles to improve the quality of the metadata, provide a mechanism for keeping the metadata consistent across systems, help to ensure that the resources can be accessed in the future, and facilitate the exchange of metadata between disparate systems [13].

The idea of describing resources with metadata, while not a new one, takes on added importance in today's distributed digital environment. As Weibel and Lagoze note, "The association of standardized descriptive metadata with networked objects has the potential for substantially improving resource discovery capabilities by enabling field-based (e.g., author, title) searches, permitting indexing of non-textual objects, and allowing access to the surrogate content that is distinct from access to the content of the resource itself." [14]

A metadata record consists of a set of elements, or attributes, necessary to describe the resource in question. For example, a metadata system common in libraries -- the library catalog -- contains a set of metadata records with elements that describe a book or other library item: author, title, date of creation or publication, subject coverage, and the call number specifying location of the item on the shelf.

The linkage between a metadata record and the resource it describes may take one of two forms: elements may be contained in a record separate from the item, as in the case of the library's catalog record; or the metadata may be embedded in the resource itself. Many metadata standards in use today, including the Dublin Core standard [15], do not prescribe either type of linkage,
4. The HEAL Metadata Schema

Before developing the HEAL metadata schema, we researched a number of metadata standards developed by national and international standards organizations, including the Dublin Core, Institute of Electrical and Electronics Engineering Learning Objects Metadata (IEEE LOM) [17], and the Educause Instructional Management Systems schema (IMS) [18]. The Dublin Core, originally designed to describe Web resources in a simple manner, consists of fifteen elements; we concluded that the Dublin Core was not robust enough to describe health sciences multimedia resources. Educause IMS, which consists of seventy nine elements and builds on both the Dublin Core and IEEE LOM, proved to be a better fit for the needs of the HEAL project.

Although the IMS elements capture most of the important information needed for describing a multimedia resource (e.g., author, title, keywords, copyright information, etc.), we determined that HEAL would need additional elements to reflect the unique needs of health science educators. We extended IMS to include additional health sciences related elements, including specimen type, radiograph type, orientation, magnification, annotations, disease process, and clinical history. Figure 4 illustrates the hierarchical relationship between Dublin Core, IEEE LOM, IMS, and the HEAL metadata standard. Because of the extended number of health sciences specific elements the HEAL metadata standard allows users to precisely specify their search criteria.

In general, the breadth and depth of any metadata standard is subject to continual debate due to the variety of user perspectives and needs. The goal of the HEAL metadata specification is to find a balance between manageability and adequacy that both serves the diverse needs of health sciences educators and the interoperability requirements of the HEAL system. Although the HEAL metadata schema contains over seventy elements, the vast majority of elements are entirely optional and the required set of elements for any given item is quite minimal.

Other organizations may further extend the HEAL metadata standard if they desire further refinements. To ensure that the HEAL collection is suitable for all levels of health sciences education the metadata schema includes information regarding educational level for which the resource is designed and flags materials that may be inappropriate for minors. Although the current standard is designed to describe multimedia items, we anticipate adding elements to describe other learning object formats (e.g., lectures, cases, and quizzes) in the future.

When a faculty member contributes a resource to HEAL, he or she is required to provide basic metadata (e.g., author and copyright information, title of the resource, media type, suggested keywords and/or description of the resource, and usage rights). The suggested keywords and description are used to further catalog the resource with a controlled vocabulary standard; this is necessary in order to ensure that different keywords are not used to describe the same concept (e.g., CT scan, CAT scan, X-Ray Computed Axial Tomography). HEAL resources are currently indexed with the Medical Subject Headings (MeSH) controlled vocabulary. MeSH consists of a hierarchical tree of health sciences related terms; this tree is available in the HEAL application for use by searchers and contributors.

While MeSH provides a good framework to begin indexing resources, it is often not specific enough to describe resources in many sub-specialties of medicine. To address this problem, we are investigating the possibilities offered by the Unified Medical Language System (UMLS) [19]. The UMLS includes a metathesaurus that maps terminology from many different vocabularies to a central concept, including domain-specific vocabularies such as the Systematized Nomenclature of Medicine (SNOMED).

Faculty may contribute items to HEAL by completing a simple form that captures metadata about the contributed resource. In order to ensure that the item is indexed accurately and that the correct MeSH terms are
employed, the HEAL team has hired a professional librarian to catalog resources. We believe that the use of professional catalogers is imperative to ensure the level of quality control and consistency desired for HEAL.

Broad acceptance from the health sciences education community of the HEAL metadata standard is crucial for facilitating the growth of the HEAL collection. We issued a Request For Comments (RFC) in February 2002 soliciting feedback about our proposed metadata standard. The RFC was sent to the health science education community at large and to the organizations that had expressed interest in partnering with HEAL. The subsequent comments and revisions led to Version 1 of the HEAL metadata schema which was formally released in April 2002.

5. Components of the HEAL Metadata Schema

The HEAL core metadata elements are based on the content and structure of the widely-accepted Instructional Management Systems (IMS) metadata specification and were implemented using the eXtensible Markup Language (XML). They include i) selected elements from the IMS Metadata v.1.2 specification [20], and ii) several non-IMS elements that have utility within the health sciences (e.g. Specimen Type, Radiograph Type, Disease Process, etc.). The XML-based HEAL metadata schema follows the Best Practices recommended by the IMS Specification [21].

The elements in the HEAL metadata schema (including the core elements based on IMS and the HEAL extensions) are organized into eight sections:

- Metametadata
- General
- Lifecycle
- Technical
- Educational
- Rights
- Relation
- Classification

These sections are described in detail below. The HEAL metadata specification documentation and examples can be downloaded from the HEAL Web site [22].

Metametadata. The metametadata section includes information about the metadata itself, including the catalog entry (a unique metadata record and collection ID), the contributor of the metadata, the language of the metadata, and the version of the HEAL metadata schema. The contributor of the metadata does not have to be the same as the contributor or creator of the resource described in the Lifecycle section.

General. The General section includes elements that can be used to describe the content of the resource. Elements adapted from IMS include catalog entry (unique identifiers for the collection and the resource), title and description. This section also includes extensions specific for the health sciences, which are described in more detail below.

Lifecycle. Lifecycle contains contact information for the contributor or resource creator, which may be described using the vcard standard for contact information, and the contribution date.

Technical. The technical section describes the digital characteristics of the multimedia item, including the file extension, file height and width, running time (for videos and audios), file size, file location (URL), and software requirements. Since the resources in HEAL are Web resources by definition, the software requirements are used to describe particular browser versions and/or operating systems required to view the item. The technical section also describes the format of the item. The format is a restricted vocabulary based on a partial list of the Ilumina formats [23]. Formats specified in the HEAL metadata schema are: Audio, Animation, Chemical Structure, Executable, Image, Java Applet, Portable Document, Presentation, Video, Web Page, and Word Processing Document.

Educational. The educational section includes elements such as interactivity type, interactivity level, learning resource time, and typical learning time; these elements do not directly apply to the granular multimedia items available in HEAL, but are included for IMS compliance. The context element includes a description of the target audience(s) for which the multimedia item is designed. The restricted vocabulary for target audience within the HEAL metadata schema is: K-12:primary; K-12:secondary; higher education, patient education, and health profession education.

Rights. The rights section includes information about the copyright holder as well as a description of usage rights for the item. To be included in HEAL, a rights statement must include, at a minimum, the right to freely use the materials for educational purposes. The copyright holder retains the copyright and may offer additional rights if desired. In the future, the HEAL application will offer the contributor an option to use the Creative Commons tools [24] to develop a detailed rights statement.
Classification. The classification section includes a taxon (controlled vocabulary or thesaurus) source and keywords. Examples of taxon sources include any standard controlled vocabulary such as MeSH or the Systemized Nomenclature of Medicine (SNOMED), or may reference a custom thesaurus developed by the contributor. More than one taxon source may be used in the HEAL metadata schema; the choice of a vocabulary source is determined by the particular needs of the subject areas described within the resource. Records may use SNOMED rather than MeSH, or may specify a UMLS Concept Identifier to indicate the appropriate concept that can be mapped from many different vocabularies. Although keywords are not a required field, all items in HEAL will be cataloged by a professional librarian using a controlled vocabulary to supplement contributed metadata.

HEAL Extensions. The IMS schema provided the basis for most of the HEAL metadata schema. While the IMS elements were adequate for the general and technical descriptions of the resource, we extended the schema to include elements specific to health sciences education, as well as elements that we felt were critical to meet the functional requirements of the HEAL system.

The health science related extensions were based on our experience with a previous health sciences multimedia database used to describe videodisc images [25]. The health sciences specific elements in the HEAL metadata schema are described in Table 1.

Table 1. Health Sciences Extensions

<table>
<thead>
<tr>
<th>Element Name</th>
<th>Format</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specimen Type</td>
<td>Standard Pick List</td>
<td>Restricted vocabulary: cell, tissue, organ, organ system, organelle</td>
<td>“organ system”</td>
</tr>
<tr>
<td>Radiograph Type</td>
<td>Standard Pick List</td>
<td>Radiology technology used to generate the multimedia item. Restricted vocabulary: angiogram, CT, MRI, nuclear, PET, Plain, ultrasound</td>
<td>&quot;MRI&quot;</td>
</tr>
</tbody>
</table>

Magnification   Free text Magnification of microscopic image “100x”

Disease Process   Standard Pick List; Multiplicity allowed Indicates disease process displayed, discussed, or implied in the item. Restricted vocabulary: vascular, inflammatory, neoplastic, degenerative, intoxication, congenital, allergic/autoimmune, traumatic, endocrine, social, legal, environmental, economical, psychological "endocrine"

Clinical History   Free text; Multiplicity allowed Clinical history of patient “Fifty-year old male with family history of congestive heart failure presented with acute chest pain and shortness of breath.”

In addition to the health sciences specific extensions, we identified elements that were needed to satisfy the functional requirements of the HEAL system. These elements are described in Table 2.

Table 2. HEAL Elements to Satisfy Functional Requirements

<table>
<thead>
<tr>
<th>Element Name</th>
<th>Format</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annotated</td>
<td>Boolean; no is the default value</td>
<td>“Yes” indicates that the multimedia is labeled or</td>
<td>&quot;Yes&quot;</td>
</tr>
</tbody>
</table>
6. Extending the HEAL Metadata Schema

The XML-based HEAL metadata schema was designed to describe a variety of common multimedia items (e.g., images, animations, videos, etc.). This core metadata may also, to a limited extent, be used to describe more complex resources such as problem-based learning cases, practice quiz questions, etc. Future versions of HEAL metadata will provide additional metadata elements aimed at these more complex resources.

The element definitions are not necessarily bound to a particular implementation technology; organizations are free to implement the HEAL metadata specification through a variety of means including database and XML applications. Technical implementers should note that the HEAL specific (i.e., non-IMS) elements have been added according to the IMS mechanism for extension outlined in the IMS Learning Resource Meta-data Best Practice and Implementation Guide. Locally developed applications may further extend this HEAL specification through the same mechanism.

7. Future Work

According to the HEAL Web site logs, the HEAL metadata schema has been downloaded by over 300 users since its release. Over a half a dozen institutions that are developing internal multimedia resource databases have adapted the schema for local use. We are currently collecting feedback from our users and plan to add a metadata registry to our Web site so that we may track downloads and follow up with users of the metadata schema. We anticipate releasing an updated version based on this feedback by the fall of 2004. In addition, we plan to extend the HEAL metadata schema so that is suitable for indexing other types of resources, including interactive cases and quizzes.

8. Conclusion

HEAL has been designed to facilitate the sharing and accessibility of a wide variety of high quality, freely available multimedia resources currently located on many remote servers. This application provides educators with a single but powerful search index and interface through which they can simultaneously search multiple collections for high quality materials that will enhance teaching and learning.

The HEAL metadata schema is the foundation on which the HEAL distributed system is built; the schema is based on international standards and includes extensions specific to the health sciences. Users may implement the HEAL metadata schema on their local systems or may further extend the schema to meet local needs. The metadata schema provides an important mechanism for the HEAL central server to share data with partner collections located on remote servers. As feedback is received from the health sciences community, the HEAL metadata schema will be improved and updated.
HEAL will have a positive impact on the health education community as the team is motivated by an open, collaborative spirit and is determined to share its resources widely; as such, all content, metadata, standards, and application code are made freely available to all audiences and users.

This project has been made possible by grants from the National Science Foundation, DUE-0085660 and DUE-0225360, and from the National Library of Medicine, 1 G08 LM007877-01.


Figure 1. The Browse interface allows users to peruse the collection via the hierarchical MeSH tree. The number of items contained within each heading is indicated in parentheses.
Figure 2. The Advanced Search interface allows users to conduct precise searches using a combination of specific search criteria.
Figure 3: The HEAL metadata harvesting system. Using the Open Archives Initiative (OAI), institutional collections can elect to share all, none, or a portion of their catalog metadata with the central HEAL server, all while obtaining the HEAL catalog metadata for their own uses.

Figure 4: The hierarchical relationship of the HEAL metadata schema compared to other common metadata schemas. Each layer extends the subset of its parent schema through additional elements which provide more specificity.