# Towards Contextually Descriptive Embedded Metadata

Matthew Miller
Pratt Institute, US
mmille18@pratt.edu

Christopher J. Mullin
Pratt Institute, US
cmullin@pratt.edu

**Keywords:** metadata; embedded metadata; Adobe XMP; open source systems; Dublin Core

## Abstract

Contextually descriptive embedded metadata has been utilized in a number of applications. Many HTML webpages contain meta tags with information reflecting the content of the page. Likewise, microformats are continually striving to make web documents more accessible by embedding descriptive data. One widely known application of contextually descriptive embedded metadata is found in digital music files such as the MP3 format; this format contains embedded data such as artist and song title information. However, digital images, ubiquitous throughout digital archives, are a medium that has been slow to experience this technology. While recent advancements in using embedded metadata in administrative and structural work flow situations have been made (Christensen & Dunlop, 2010) the focus of our work is on contextually descriptive metadata. This type of information would represent metadata such as an artist name or the title of a painting. Advocating and promoting the creation and usage of contextually descriptive embedded metadata supports the digital push towards semantic-based, readable archival programming and the subsequent effective ease of use it provides file creators and end users alike.

To explore embedded metadata in images and other digital assets, a system that understands and leverages embedded metadata to demonstrate its utility and serves as a foundation for further applicative research was conceived and actualized. The first step taken was the construction of the Embedded Metadata Explorer (EME), an online embedded metadata editor. This tool (located at http://embedmydata.com) uses Adobe's XMP technology to allow a user to embed and edit a Dublin Core record into a digital image. Furthermore, the tool enables the creation of embedded images without the need to have commercial software on the user's computer and provides an overview of all the embedded meatadata contained in an image. Currently the EME is the only XMP editor available online. Though it exists as a static, non-integrated tool, EME readily demonstrates how simply and effectively descriptive metadata can be created and utilized. Creation of this tool allowed us to solve the technical requirements raised by such a project by implementing Adobe's XMP functionality in a fully functional, open source application. Though XMP is an emerging standard for embedding information and allows for multiple namespaces and/or metadata schemes to co-exist in a single digital asset, we favored usage of Dublin Core's namespace for its ability to simply and effectively describe a resource. We also selected it for its long-standing, established acceptance and usage within cataloguing discourse, software development and project initiatives.

Building from EME's foundational position, we proceeded to create a digital archive system that understands and makes use of contextually descriptive embedded metadata. To accomplish this, we developed a plug-in for the Omeka web publishing system (http://omeka.org/), allowing it to utilize embedded metadata in digital assets. Though in initial stages of development, the system is able to ingest Dublin Core embedded content and use that information to populate the record's metadata. Likewise, any changes or additions to the record metadata is updated in the embedded content of the asset as well as all of its derivatives, such as thumbnails (Fig 1). This ability demonstrates the interoperable functionality inherent when using embedded metadata. Digital assets then have the ability to act as self-contained records. The asset can now exist as its own complete record with migratory capabilities to any system or application equipped to utilize embedded data.
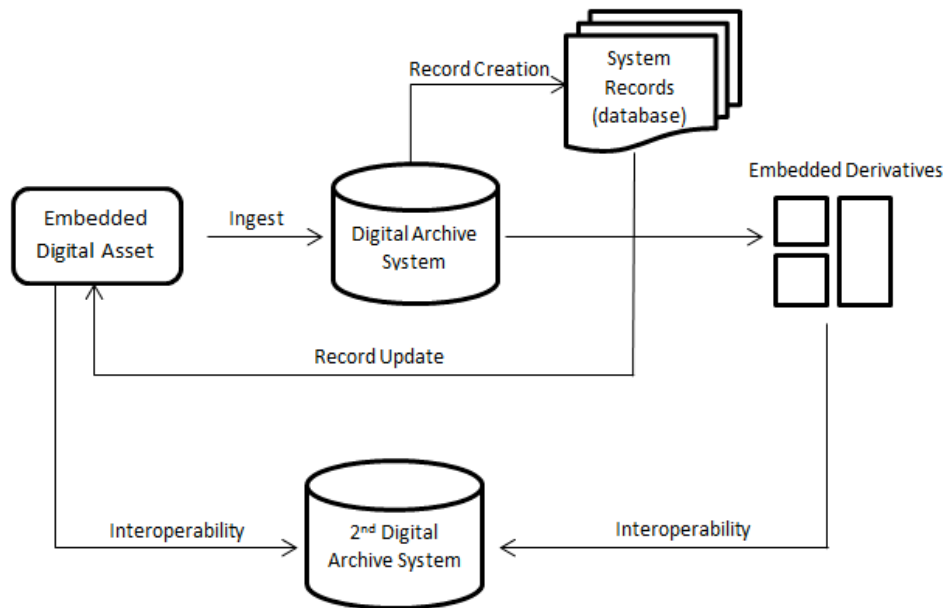
FIG. 1. Interoperability

As the project develops we hope to further explore the benefits and opportunities created by a system that is able to utilize contextually descriptive embedded metadata. One defined future goal includes examining the possibility of a "database-less" digital archive that operates solely on and from the embedded metadata housed within the archive's holdings. Deliverables as well as demonstrations produced from the project will be made available on the Embedded Metadata Explorer's website as research develops.

## References

Adobe. (2010). XMP SPECIFICATION PART 1. Retrieved April 20, 2011, from http://www.adobe.com/content/dam/Adobe/en/devnet/xmp/pdfs/XMPSpecificationPart1.pdf

Christensen, S., & Dunlop, D. (n.d.). The Case for Implementing Core Descriptive Embedded Metadata at the Smithsonian. Article, . Retrieved May 1, 2011, from http://si-pddr.si.edu/jspui/handle/10088/11123.

DCMI. (2003). Guidelines for implementing Dublin Core in XML. Retrieved April 20, 2011, from http://dublincore.org/documents/dc-xml-guidelines/.

Federal Agencies Digitization Guidelines Initiative. (2010). Still Image Working Group, Embedded Metadata. Retrieved from http://www.digitizationguidelines.gov/stillimages/subcommittees.html.

VRA Embedded Metadata Working Group 2011. Retrieved from http://metadatadeluxe.pbworks.com.