

# Metadata and digital objects: To embed or not to embed?

Rachel Jaffe  
Metadata Librarian  
University of California, Santa Cruz

Edward M. Corrado  
Associate Dean for Library Technology  
Planning & Policy  
University of Alabama

## Abstract

Libraries, Archives, and Museums (LAMs) are curating an increasing amount of digital objects. Embedded metadata places valuable identifying and descriptive information within the object itself. However, embedding metadata after acquisition changes the object, potentially raising questions about its authenticity. This poster reports on a survey conducted in order to explore the benefits and costs of embedding additional (i.e. LAM) metadata into digital objects, to the end of identifying and evaluating current practice, and defining best practices.

## Preliminary Observations (aka the chicken-&-egg scenario):

- No clear use cases
- Systems aren't doing anything with embedded metadata
- People don't have resources or knowledge, but would welcome best practices
- Consensus around the idea of embedding tombstone metadata

*"Core metadata should be embedded in files because it facilitates data exchange: that is, it extends the usability of images, texts, and other collection items in a variety of environments."*

*"There is no reason not to embed it other than if/when the proper staff or tools aren't in place and it becomes less of a priority due to higher priority activities taking up all the resources."*

## Selected Results

### Is your LAM embedding metadata into digital objects?

#### Response (134/134)

Yes: 58 (43.3%)  
No: 58 (43.3%)  
Not Sure: 20 (15.2%)

## Initial Assumptions

- LAMs are now, more than ever, creating, managing, and providing access to born-digital and digitized content
- There is a lack of (general, widely-adopted) best practices
- Conversations and practices vary among different communities and practitioners
- There is interest in embedded metadata and there are communities of interest

## Methodology

Developed and analyzed a 25-question, web-based survey, consisting of a mix of closed and open ended questions. The survey was disseminated via professional listservs.

## General Findings

- 134 respondents
- Demographic information
- 97.0% of respondents are creating/managing digital objects
- 85.9% of respondents are actively digitizing content.
- Respondents evenly working with both of born digital and digitized objects (89.6%/94.8%). Predominately text (82.8%) and images (91.0%), but also audio (70.9%), moving images (70.1%), and datasets (35.8%). Most are creating both item and collection level metadata (95.5%/64.9%; 62.7% are doing both)
- Most are managing objects created by the institution (91.8%)

## Why does your LAM choose not to embed metadata into digital objects?

#### Response (64/134)

Time/Resources/Lack of support: 17 (26.5%)  
Lack of training/knowledge/policy: 13 (20.0%)  
System limitations/requirements: 9 (14.1%)  
No convincing argument for doing so: 8 (12.5%)  
Unsure/Don't know: 7 (10.9%)  
Other: 7 (10.9%)  
Actively considering embedding metadata: 5 (7.8%)  
Not a priority: 4 (6.3%)  
Workflow issues: 4 (6.3%)  
Metadata versioning/editing concerns: 4 (6.3%)

## Why does your LAM embed metadata into digital objects?

#### Response (55/134)

Metadata and files travel together: 17 (30.9%)  
Preservation: 11 (20.0%)  
Other: 11 (20.0%)  
Discovery/Access: 9 (16.4%)  
Context/Self-documenting objects: 8 (14.5%)  
Rights management: 7 (12.7%)  
Part of established workflow/  
local best practice: 4 (7.3%)

*"For technical and administrative metadata, I think that embedded metadata will be absolutely essential to the long-term preservation of the digital object. However, embedding descriptive metadata, at this point, seems to be superfluous."*

*"Not doing so is like not including wall tags in a museum exhibition or publishing an image without a caption. [...] Information, even if out-of date, is still better than no information."*

*"I think that in principle it is a good idea. However as our systems and our workflows don't support it we haven't been able to do it. This questionnaire is making me think about looking to change that."*

## What arguments would you make for or against the practice of embedding metadata?

### For

- Self-describing
- Portability: facilitates exchange and sharing of metadata and objects
- Rights management
- More discoverable by search engines
- Minimize data loss and reduce future work
- Future proofing? (May assist future users and archivists)

### Against

- Time/Resources/Lack of knowledge
- Return on investment (ROI)
- File corruption/file integrity
- Metadata editing/versioning
- Users can't/won't use it; systems can't/won't use it
- Visibility/retention of embedded metadata as it moves across systems