

IMLS Award Number LG-24-08-0093-08

May 01, 2009

## Executive Summary

The New Media Consortium (NMC) and the Indianapolis Museum of Art (IMA), with significant support from the Institute for Museum and Library Services (IMLS), have launched the *Steve in Action Project*, a three-year effort to develop, demonstrate, evaluate, and document social tagging tools and methods in a range of cultural heritage settings. Building on the work of a previous research grant funded by the IMLS in 2006, the *Steve in Action* project team will shift the project focus from research to practice. Early results of the research project indicate that social tagging does have the potential to significantly enhance access to online collections of art; our proposed project will support and encourage the mainstream adoption of social tagging tools and methods by cultural heritage organizations of all sizes and types.

At the end of the *Steve in Action* project's sixth month, all of the milestones identified for the first half year of the project have been met, and the project is well ahead of schedule. The project has successfully developed a team of programmers, designers, and technologists drawn from institutions located across the country. The project is being watched with great interest by a number of institutions and groups, and project members have been invited to make eight presentations to national conferences and other forums in the project's first six months alone.

Match spending on the project exceeds IMLS funding for the first six months, with institutional match, at 117% of grant funds expended. Total spending for the first year is on track, with approximately 119% in anticipated match.

## Project Background

*Steve in Action*, referred to informally as Steve, is a three-year effort to advance the mainstream adoption of social tagging tools and methods by cultural heritage organizations of all sizes and types. The NMC's partnership with members of the Steve project, a collaboration of individuals and organizations investigating the value of social tagging to enhance access to collections and enable visitor engagement with collections, thus far has proven to be a powerful partnership. Existing tools developed by the Steve team will be enhanced and extended to support broad-based use and to serve as the core software application for a number of local, institutional implementations.

The NMC, an international consortium of nearly 300 learning-focused organizations dedicated to the exploration and use of new media and new technologies, leads the project; Dr. Larry Johnson, CEO of the consortium, serves as Project Director; Rachel Smith, NMC Vice President, NMC Services, is Project Manager. The NMC brings experience in the leadership of complex, multi-institution projects, including the IMLS-funded Pachyderm 2.0 project, a collaboration with the San Francisco Museum of Modern Art and thirteen university and museum partners. Susan Chun, a founder of the Steve project, and Robert Stein, CIO of the Indianapolis Museum of Art, serve as Project Lead and Technical Lead, respectively. Chun and Stein were the directors of the previous IMLS-funded National Leadership Grant that enabled the research underpinning the toolsets.

## Steve in Action Project Partners

There are currently 10 partner institutions involved in the project:

### Project Leads

The New Media Consortium  
Indianapolis Museum of Art

### Museum Partners

Los Angeles County Museum of Art  
Metropolitan Museum of Art  
MINITEX Library Information Network  
Minneapolis Institute of Arts  
Rubin Museum of Art  
San Francisco Museum of Modern Art  
Walker Art Center

### University Partners

University of California, Los Angeles

## Project Goals

The implementation of social tagging tools on a broad basis could have significant benefits for cultural heritage professionals and the public at large, enabling improved access to collections, new forms of engagement with visitors, and an opportunity for cultural heritage professionals to see their collections through the eyes of visitors. Over the grant period, the *Steve in Action* team will: 1) enhance the software tools originally developed for the research project; 2) develop, implement, evaluate, and document multiple models for adopting social tagging; 3) demonstrate integrations with major information management systems; and 4) continue to encourage discussion within the community about the strengths and weaknesses of tagging as a method for communicating with their audiences.

Specifically, the goals of the grant are: 1) to enhance the existing tagging software tools to make Steve easy to use for museums of all sizes and types; 2) to develop next-generation tagging tools that motivate and engage users, including mobile interfaces that allow tagging in museum spaces; 3) to investigate ways to aggregate tags in order to facilitate cross-collection searching and browsing; and 4) to demonstrate integrations of the Steve tagger with commonly-used museum systems. Each museum partner has pledged to design at least one implementation of the Steve tagger that will demonstrate a way in which tagging can be applied to museum practice.

At the end of the three-year grant period, we expect to have implemented and evaluated a variety of tagging environments with at least thirty partners. Many of these tagging environments will run in several cycles, allowing us to compare the experiences of users from different demographics, or to judge the value of tagging for different types of collections. Our tagging implementations will test the applicability of the Steve tools to historical collections, library collections, archives, and a variety of museum collections; an application to aggregate collected tags will serve as a testbed for looking at cross-collection searching and browsing.

Furthermore, the project seeks to answer questions of user motivation to understand how social tagging engages and rewards the visitor; to gauge the uses and benefits of social tagging for institutions and their visitors; and to measure what kinds of support and resources are required by institutions hoping to implement social tagging practices.

## Summary of Major Activities

**Kick-Off Meeting.** The first two months of the project were devoted to organization and the creation of the essential web and communications tools that would support the project. The *Steve in Action* Kick-Off Meeting was held in San Francisco on December 3-5<sup>th</sup>, 2008, and was attended by 32 representatives of the partner institutions. At this meeting, the partners discussed the project and its goals, collaborated on requirements gathering for the Steve tagger and server, and began building the connections between people that sustain the project between face-to-face meetings.

The meeting, which was structured as a working session, resulted in the development of more than 100 user profiles, 80 scenarios, and over 250 requirements for using the next generation Steve tagging tool. Project partners were encouraged to voice their wishes and desires for the tagging tool, whether for features designed for immediate implementation or for more long-term consideration. Requirements were generated using the profile and scenario method, standard in the field of software design.

The user profiles and scenarios were generated by the group and captured using a tracker tool that provided a collaborative way to develop a variety of potential user personalities and the scenarios in which those users would interact with the Steve tagger. The team divided into small groups of two to five people and were assigned specific user categories in order to cover a range of users and situations.

The resulting user profiles and scenarios may be viewed at:  
<http://steve.nmc.org/2009/03/profiles-and-scenarios/>

Once the scenarios were created, the teams were redistributed and assigned groups of scenarios. Each team examined their assigned scenarios in order to mine them for possible system requirements. The teams tracked their findings in the Requirements Tracker System, which later served as a guide for the Development Team who performed initial scoping of requirements.  
(See <http://steve.nmc.org/tracker/login.php>)

**Requirements Sorting and Prioritizing.** As a result of the team's efforts at the Kick-Off Meeting, over 250 requirements were captured. The development team then sorted the requirements according to the part of the Steve system to which each pertained. This sorting revealed a number of duplicate requirements; the development team made note of these and condensed the remaining requirements. Finally, the team reviewed each requirement with an eye to verifying its classification, estimating the time for development, and determining the requirement's relative priority. Dependencies between requirements were also noted. This work took place from December through March, 2009.

**Steve in Action: Software Architecture Document.** The *Steve in Action: Software Architecture* document was formed from the requirements generated by the project partners. The requirements were grouped into a dozen categories in five broader classifications related to the overall project goals, including:

### **Making Steve Easy**

- Ingest / Import
- Integration / Export
- Administrative Interface
- Installation

### **Technical Components**

- API
- Database
- Search

### **Post-Processing**

- Term Review
- Reporting

### **Extensibility Requirements**

- Mobile Tagging
- Text, Tagging and Trust

### **Useable Implementation**

- Tag Server

Each category was discussed in detail and resulted in the corresponding sections that comprise the document. This document will be used to guide the development process and timeline for the grant. Please see the appendix for the full *Software Architecture Document*.

As the culmination of an extensive consensus-based process that has been the main activity of the first six months, the formal five-page requirements document was released in mid-April to the project partners. The document was posted online in a form that encourages commentary at <http://wp.nmc.org/sia-architecture/>. All project partners were asked to comment, and to pass the document address to others at their institutions who might also have ideas and suggestions. The resulting commentary is permanently attached to the document and will be used to inform a revision, which will then serve to guide the development effort going forward. All documentation, supporting tools, and resources are available on <http://steve.nmc.org>.

**Communication and Project Management.** To support internal communication, the *Steve in Action* team has created several online groups through Google Groups to facilitate easy communication and document sharing. To communicate with the public about the activities occurring as a part of *Steve in Action*, an open blog has been established at <http://steve.nmc.org/>. This blog is updated regularly by the project team, and links to all public documents and resources are posted there as they emerge. The project's core team attends a bi-weekly update call where decisions are made collaboratively and action items are assigned.

## Development Team Report

The current members of the Steve development team are:

Robert Stein, Indianapolis Museum of Art, Technical Lead  
Ed Bacht, Indianapolis Museum of Art, Lead Developer, T3  
Charlie Moad, Indianapolis Museum of Art, Lead Developer, Steve in Action  
Alan Levine, The New Media Consortium, Senior Technology Advisor

**Web Applications Developer.** With the Software Architecture Document in place, the team is ready to hire a Web Applications Developer to perform activities in support of the core operations of the *Steve in Action* project. The Chief Information Officer at the IMA will provide day-to-day supervision, working in partnership with the NMC and the *Steve in Action* Project Director.

The Web Application Developer's priority will be the development of social tagging tools for *Steve in Action*. Specific development tasks will be outlined in architecture documents pertaining to the project. Specific job duties are listed below:

- Design and implement web interfaces for tools belonging to the Steve toolset
- Design and construct installer packages for Steve tools
- Design and develop reports for objects and tags in the Steve database
- Modify the Steve database as required to interact with web interfaces
- Document development tasks, including specific changes, for use in project reports and technical documentation
- Provide support for the Steve tag server
- Provide additional support and development as directed by the IMA and NMC

This position will be located at the Indianapolis Museum of Art in Indianapolis, Indiana.

## Web Sites and Other Resources

*Steve in Action* makes extensive use of technology to support the work of its distributed teams, and to ensure that interested parties have access to complete information about the project and its activities. The project makes use of technological resources that are publicly accessible as well as private resources limited to team members only. All relevant reports, documents, and other materials are routinely published to the web so that they may be accessed by the team and interested members of the general public. The following list describes the most important of these tools and provides links to each of them.

- **Steve in Action Project Website.** This is the official web site for the project. In addition to a project description and listing of partners, it contains links to the IMLS, to the *steve.museum* site, to the NMC website, and to the IMA website. Project documents, photographs, and the public Google Group are also linked from this site. Additionally, a blog updates the public about the progress of the project. <http://steve.nmc.org>
- **Steve.museum.** This multi-author blog is the official communication center for all Steve projects, past and present. It is the best source of project news and up-to-the moment detailed information. <http://steve.museum>
- **Software Architecture Document Comment Site.** In order to solicit feedback from the partner institutions, the Software Architecture Document was posted to a comment press site that allows for collaborative analysis and debate. This document remains available for review. <http://wp.nmc.org/sia-architecture/>
- **Requirements Gathering Tracking Tool.** This tool was created for use at the Kick-Off Meeting in December for the participants to scope, track, and prioritize requirements for the Steve tagger. <http://steve.nmc.org/tracker/login.php>

Additionally, the Steve in Action Project team has set up several Google groups to facilitate communication between the core team members, with the museum partners, and with the developers. If desired, and upon request, any IMLS personnel will be granted access to these private groups.

## Conferences and Presentations

The work of the *Steve in Action* project has been presented at the following conferences and online events during the first six months of the project:

- *Social Tagging: Implications for Cultural Heritage Professionals*  
Susan Chun. A presentation in the "Discovery This Side of the Virtual Wall" panel at the Art Libraries Society of North America (ARLIS) Annual Conference, Indianapolis, April, 2009.
- *Steve in Action: Social Tagging Tools and Methods Applied*  
Susan Chun, Tiffany Leason, Rob Stein, and Bruce Wyman. A Workshop at the Museums and the Web Conference, Indianapolis, April, 2009.
- *Improving Search and Discovery of Digital Resources Using Topic Modeling*  
Susan Chun. Presentations for Advisory Board of the IMLS-funded collaboration between Yale, the University of Michigan, and the University of California. A post-presentation meeting concerned ways in which the Steve project and the Yale project might share resources or information, including the Steve tagging software, data, or evaluation methods. March, 2009.

- Presentation for and discussion with the Princeton University Digital Assets cross-functional group.  
Susan Chun. The assembled group also included curatorial and other museum staff. The Princeton team, a cross-campus team of librarians, archivists, museum technologists, and IT staff, is now thinking about ways to implement tagging in a number of ways on campus and online. Princeton, NJ, March, 2009.
- *Social Tagging and Museums: Hearing the Visitor's Voice*  
Susan Chun. Speaker at the Museums in Conversation Conference (Museums Association of New York/Upstate History Alliance Annual Meeting), Tarrytown, NY, March, 2009.
- *Exploring Social Tagging for Enhanced Access to Art Historical Materials*  
Susan Chun. Panelist in the Committee on Intellectual Property session at the College Art Association Annual Conference, Los Angeles, February, 2009.
- *Should You Care about Social Tagging? Findings and Recommendations from Steve.museum*  
Robert Stein, Jennifer Trant, and Susan Chun. A presentation at the Museum Computer Network Annual Conference, Washington, DC, November 2008.
- *Steve.Museum: Social Tagging of Art Images*  
Susan Chun and Daniel Starr. A presentation for the Columbia University Digital Library Seminar Series, New York, November, 2008.

## Evaluation Activities

Evaluation in the first six months has focused on two critical goals. The first is that Steve be well-designed and appropriate to its planned purposes. It has been developed as a user-centered tool from its very inception given that its main purpose is to bring the visitor's perspective to into a discussion about art. As a result the first tools developed in the project — the requirement tracker and the persona and scenario building tools — were designed to document the needs of Steve users in a wide variety of contexts. These tools were essential to developing the specifications for the planned enhancements to the Steve toolset, especially at the initial meeting of partners, where the bulk of this work was done. A wide variety of perspectives were represented via this process, including museum personnel of all types, librarians, educators, students, technology professionals, and media-makers.

The second component of the evaluation in this period was that communications, resources, and web sites are effective and meeting the needs of the core team, the partners, and the public. A subset of this area is the efficacy of face-to-face meetings. Meeting evaluations, based on interviews with participants, have been uniformly positive, as has feedback on the quality of the project web resources, reports, and communication tools and processes.

## New Collaborations

The project team continues to be approached about the possibility of collaborating with other institutions and projects. Among the key collaborations we have pursued in the first six months of the grant period are a project to integrate the Steve tagging software with Omeka, the open source content publishing system for collections content funded by the IMLS, the Sloan Foundation, and the Kress Foundation. Omeka shares a target user group (museums, libraries, archives, and scholars) and an open-source, community-focused development philosophy with Steve.

In addition, the project team has devoted considerable time to studying areas of convergence between the *Steve in Action* project and *T3: Text, Tags, Trust* — the Steve project's IMLS-funded research component.

The two projects share a number of team members. We are examining the development roadmap and the research project design closely to determine whether collaboration around tool development, evaluation strategies, and data collection are likely to be practical or useful, and expect to have developed a concrete plan for areas of shared interest by the end of the first year of work (September 2009).

Lastly, the project has been approached by the Museum of Modern Art (MoMA) in New York to add them as a project partner. An invitation is planned to bring them formally aboard.

## **Conclusion**

At the end of the sixth month of activity, the *Steve in Action* project is fully on track with its stated goals and objectives. The members of the project have designed a highly productive approach to working in a distributed fashion, and have built a range of collaboration tools and resources to support the work and the sense of being part of a team. These efforts have resulted in a very high level of participant satisfaction and productivity.

All the milestones for this period have been achieved, and significant progress on several major milestones for the next period has been made as well, and it is anticipated that several of these will be reached ahead of schedule. Financially, the project is also on track, with expectations for match significantly exceeded.

# Appendices



sparkling innovative  
learning & creativity

# Steve Kickoff Meeting Agenda

December 1-3, 2008 | Grand Hyatt, San Francisco

Monday, December 1, 2008	
8:30 – 9:00	Gathering, continental breakfast
9:00 – 9:45	Welcome and Introductions (Larry Johnson) <i>Larry sets the context for the project and introduces the core project team; other participants introduce themselves with the two-breath rule. Larry describes the roles of those present (and of any others who are not able to be there).</i>
9:45 – 10:00	Project Overview, Demo, and Research Results (Susan Chun, Rob Stein) <i>Susan and Rob describe the Steve Project to date and outline the overarching goals for the next three years. Susan discusses the research results from the first Steve grant, and Rob demonstrates the engine.</i>
10:00 – 10:15	Break
10:15– 11:45	Project Overview, Demo, and Research Results (Susan Chun, Rob Stein) <i>(continued)</i>
11:45 – 12:00	Review of Agenda and Workplan (Rachel Smith) <i>Rachel goes over the extended meeting agenda, describes the short-term meeting goals, and explains briefly the processes to be used.</i>
12:00 – 1:00	Lunch
1:00 – 2:00	User Profiles (All) <i>The group reviews existing user profiles and develops new profiles for the requirements effort. Rachel facilitates.</i>
2:00 – 2:15	Break
2:15 – 5:00	Scenarios (All) <i>The group reviews existing scenarios and develops new ones for the requirements effort. Rachel facilitates.</i>
5:00 – 6:30	Personal time
6:30 until...	Group Dinner <i>Osha Thai Restaurant, 149 2nd St. San Francisco, CA</i>

Tuesday, December 2, 2008	
8:30 – 9:00	Gathering, continental breakfast
9:00 – 9:15	Review of Yesterday (Larry, Susan, Rachel) <i>Briefly review yesterday's work and set the context for today's.</i>
9:15 – 10:30	Requirements: Mining the Scenarios (All) <i>The group begins to mine the scenarios for requirements, entering them into the Tracker. Rachel facilitates.</i>
10:30 – 10:45	<i>Break</i>
10:45 – 12:00	Requirements: Mining the Scenarios (All) <i>(continued)</i>
12:00 – 1:00	<i>Lunch</i>
1:00 – 2:30	Requirements: Mining the Scenarios (All) <i>(continued)</i>
2:30 – 2:45	<i>Break</i>
2:30 – 5:00	Requirements: Review & Classification (All) <i>The group reviews and classifies the requirements and begins to prioritize and scope them. Rachel and Rob facilitate.</i>
5:00	<i>Dinner on your own</i>

Wednesday, December 3, 2008	
8:30 – 9:00	Gathering, continental breakfast
9:00 – 9:15	Review of Yesterday (Larry, Susan, Rachel) <i>Briefly review yesterday's work and set the context for today's.</i>
9:15 – 10:15	Requirements: Prioritizing & Scoping (All) <i>The group continues to prioritize and scope the requirements. Rachel and Rob facilitate.</i>
10:15 – 10:30	<i>Break</i>
10:30 – 11:00	Requirements: Wrap-Up (All) <i>The group deals with any remaining issues with the requirements classification. Rachel and Rob facilitate.</i>
11:00 – 11:30	Meeting Wrap-Up & Next Steps (Larry and Susan) <i>Larry and Susan wrap up the meeting and clarify next steps for the workplan, communications, future meetings, and assignments.</i>
11:30	Meeting Adjourns
11:30 – 12:30	<i>Optional Lunch</i>

# Steve in Action: Software Architecture

Last revised: Charles Moad, Indianapolis Museum of Art, Mar. 25<sup>th</sup>, 2009

1	Introduction.....	2
2	Architecture Overview .....	2
2.1	Making Steve Easy.....	2
2.1.1	Ingest / Import.....	2
2.1.2	Integration / Export .....	2
2.1.3	Administrative Interface.....	3
2.1.4	Installation.....	3
2.2	Technical Components .....	4
2.2.1	API.....	4
2.2.2	Database.....	4
2.2.3	Search .....	4
2.3	Post-Processing.....	4
2.3.1	Term Review .....	4
2.3.2	Reporting .....	5
2.4	Extensibility Requirements.....	5
2.4.1	Mobile Tagging .....	5
2.4.2	Text, Tagging and Trust (T3) .....	5
2.5	Useable Implementation.....	5
2.5.1	Tag Server .....	5

## Introduction

This document provides an overview of the software components and requirements for the IMLS grant, Steve in Action. It is designed to serve as a reference describing – at a high level – the software development work to take place over the course of the grant. In the development of this software, efforts will be made to ensure that the steve architecture uses primarily open source software tools that have large and active user communities. This will result in the most maintainable and extensible software toolkit possible.

## 1 Architecture Overview

This section describes the high level components of the steve software architecture and breaks down each section into more detailed requirements and specifications of the system.

### 1.1 Making Steve Easy

In order to achieve one of the primary goals of this grant, steve must be able to easily interact with other systems. Import and export of museum metadata and existing tags should be well documented, tested, and easy for a non-technical staff member to perform.

**1.1.1 Ingest / Import.** Steve must be able to import comma separated value (CSV) and XML files representing object metadata and tags as long as they adhere to a to-be-defined well specified format. Steve should also be able to import tags from external sources, including but not limited to TagCow and the legacy steve-tagger.

To ensure that metadata remains up to date, steve will support a polling model of importing and updating object metadata. There will be several methods of adding new objects as valid data sources for steve. The first will be the use of documented HTML markup formats that will allow steve to identify information to scrape from a webpage in order to ingest or update a work, its metadata, and imagery. An alternative will be to expose a collection by using tools like Dapper (<http://www.dapper.net/>), which allows users to create API's to access to their collection simply from existing web pages. Steve will periodically poll these data sources in order to sync objects and their metadata.

Images will be specified as a metadata field on objects. A URL to the image will be required and optionally a thumbnail URL can be provided. Steve will maintain its own copy of all thumbnails. If a thumbnail URL is not specified initially, steve will create a thumbnail immediately. Multiple images can be provided per object and an order can be specified to control display preference.

Steve will provide an interface for specifying object metadata fields. The steve metadata model will be flat and based on existing standards where possible. The Dublin Core standard will be explored as a basis for the schema. It will support extensibility to allow end-users to create their own metadata fields of varying type. Standard mappings or crosswalks will be defined in order to support existing formats such as CDWA, MARC, etc... The use of semantic web technologies like RDF will be explored as a way to express relationships between metadata fields.

**1.1.2 Integration / Export.** Steve will be able to export CSV and XML files representing object metadata and tags. Steve will support its standard XML schema and attempt to provide

transformations to other XML standards to help support integration with existing collection management systems (CMS) and digital asset management (DAM) systems.

1.1.3 **Administrative Interface.** Steve will provide a web administration tool. The admin tool will provide a metadata builder in order to allow custom metadata on objects. The tool will allow for creating, updating, deleting, and searching of objects, users, and term reviews. Users will also be able to define and manipulate sets of objects and users.

Steve will support the creation and management of static and dynamic (smart) sets. This support will allow users to create traditional sets of objects as well as sets that will auto-update based on the object metadata, such as saved searches or other rules-based selections.

Administrators will be able to specify per-user or per-group access permissions to objects using access-control lists. Term reviews in addition to other tag properties such as frequency or source can be used to control a default display policy on tags.

Application environments will be controlled in the admin tool and will describe the different contexts in which tags are added to the steve tag server. Users will be able to describe an application for collecting tags and will be issued application keys that will be used to identify and authenticate tags coming from that application.

The admin tool will also provide users code snippets (widgets) for accessing object metadata and tags and embedding this information in their own web pages. The snippets will be customizable in a fashion similar to the embedding of YouTube videos or Flickr badges. Example customizations include the controlling of metadata presentation and filtering tags by display frequency.

1.1.4 **Installation.** There are two primary types of installation of the steve software. Users will be able to host their own tag server along with all the components described in this document. In addition, a shared tag server will be provided in order to prevent all users from needing to host the software. In this case, users will be able to install client tools to make use of the steve software without hosting their own instance of the software. We envision that the primary use of the steve-tagging tools will be via this hosted tag server deployment.

1.1.4.1 **Server Level Installation.** In addition to traditional software downloads, a virtual machine image or live CD will be available to those who wish to host their own tagging server. It will contain all the steve software and prerequisites required. The tagging interface will most likely be implemented using the CMS, Drupal (<http://www.drupal.org>). A Drupal installation profile will be provided. There is no planned solution for providing communication between disparate tag servers.

1.1.4.2 **Client Level Installation.** The preferred method using the steve tools will be to make use of the steve client libraries. This should be as simple as including the API by cutting and pasting some small amount of JavaScript code into your webpage. Additionally, users may choose to use a PHP include of the steve libraries in their own software.

Code snippets and tutorials will be provided in the admin interface to demonstrate how to include the steve software in web pages. Simple customizations can be performed on these

code snippets using the admin GUI to simplify the embedding process. These snippets can also be used as a starting point for more advanced applications.

An API sandbox or explorer will be provided to allow administrators and developers to try out code before it is ever installed on their web pages. This tool will also allow them to become familiar with the underlying steve software libraries.

## 1.2 Technical Components

This section describes components of the systems that should not necessarily be used by tagger or object administrators of steve. It contains components that developers would use in order to develop and deploy custom tagging applications.

1.2.1 **API.** All interfaces to steve, administrative or end-user, will be based directly off the API and never directly interact with the database. The API will allow create, read, update, delete, and search operations for metadata and tags. The API will also support the extensible metadata architecture described in the ingest section of this document.

The entire API will be accessible through a REST interface supporting JSON, XML, and possibly other serializable data protocols. Upgrades to the REST interface will not require updates to client interfaces or software.

The API documentation will be accessible at all times via the admin interface on the hosted tag server.

1.2.2 **Database.** The data schema will support the storage of several core steve data types: objects, images, institutions, environments, etc. The database will support tags with different data types which may include text, integers, and decimals. Tags are n-ary vectors (having one or more components), hence a single tag could contain a longitude and latitude. In addition, the tag vector might be heterogeneously typed. Examples might include recording a text tag at an x, y coordinate of an image.

The database should support clustering of terms and/or term relations. RDF storage solutions will be investigated to support this.

1.2.3 **Search.** A full-text search engine will be available for steve metadata and tags. The most likely candidates for this support include Lucene (<http://lucene.apache.org/>) and SOLR (<http://lucene.apache.org/solr/>). This component will provide a number of text processing utilities such as stemming and support for diacritics. The T3 grant will explore more advanced tools used for search and retrieval.

## 1.3 Post-Processing.

This section describes tools that will be developed to support the post processing and understanding of tags collected.

1.3.1 **Term Review.** Term review access will be controlled through access-control lists in a manner similar to that of objects mentioned above. End-users can be term reviewers. A standard set of review categories will be defined and administrators will have the ability to

add their own custom review categories. Simple workflow support will be included to allow users to send of one or more tags to another user for review. The term review tool should allow for detailed annotation of misperceptions.

1.3.2 **Reporting.** Basic reporting will be provided for objects and tags. Reports can be exported in a CSV and possibly XML format. Reports will be accessible at a global, institutional, or application level of granularity.

## **1.4 Extensibility Requirements.**

This section describes extensibility features that steve must contain in order to support third-party integration and applications.

1.4.1 **Mobile Tagging.** Collection of tags using mobile devices should be performed using either the API directly, the REST interface, or a mobile web browser.

1.4.2 **Text, Tagging and Trust (T3).** The T3 IMLS grant should be supported to allow for additions to the core steve architecture. These include, but are not limited to, faceting, taxonomies, ontologies, user trust, advanced stemming, disambiguation, and tag extraction.

## **1.5 Useable Implementation.**

The steve architecture will include a fully functional interface for tag collection and browsing by an end-user.

1.5.1 **Tag Server.** A hosted tag server will be supported as part of the grant. It will include the ability to tag works and explore works by metadata or tags. User accessibility requirements will be taken into account during the development of the user interfaces for tagging. Custom themes will be supported through logos and color choice. A cross-collection tagging / browsing experience will be provided to bring together all the publicly available works in any particular steve tag server. End users will be able to edit and delete their own tags. This shared implementation will have simple user profile support.