NTHMP Repository

Version 0.3 (Preliminary)

N:\Loren\Repository Design\Loren’s functional specification outline\_6

Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Version** | **Description** | **Author** |
| 10/13/2010 | 0.1 | First draft template. | Loren Pahlke |
| 10/20/2010 | 0.2 | Formatted and added Scott’s comments. Revised requirement format. | Loren Pahlke |
| 10/21/2010 | 0.3 | Added additional comments from Scott. | Loren Pahlke |
|  |  |  |  |

Table of Contents

1 Introduction 4

1.1 The need for an NTHMP repository 4

1.2 Purpose and scope of this document 4

1.3 Audience for this document 4

1.4 Conventions 4

2 Overall description 5

2.1 Project parameters 5

2.2 Classes of users 7

3 Specifications 8

3.1 Functional requirements 8

3.2 Interface requirements 29

3.3 Data requirements 31

3.4 Metadata requirements 31

3.5 Operational requirements (Open issue) 32

3.6 Maintenance requirements 34

4 Repository policies 34

4.1 Existing policies and guidelines that affect repository design 34

4.2 Policies to be developed 34

5 Concerns 34

6 Appendices 34

6.1 Minimum metadata 34

6.2 Technical metadata 34

6.3 User interface prototypes 34

6.4 Formats 34

6.5 Survey responses 34

6.6 Policy drafts 34

6.7 Glossary 34

7 References cited 35

8 Index 35

# Introduction

This document aims to give tangible form to a vision for a tsunami-related repository that addresses weaknesses in the existing haphazard preservation and distribution system. The repository proposed here preserves the output of organization employees, improves access to tsunami materials, and supports content reuse and re-purposing. This document itself is intended to facilitate discussion of repository policy and management issues, improve stakeholder understanding, help the planning and decision-making process, and to serve the purpose of strengthening support for the repository among participant organizations.

## The need for an NTHMP repository

Part of the mission of the NTHMP is to conduct “ public outreach, local dissemination, planning and education… to communities on all U.S. coastlines.” In support of this mandate, the members of the NTHMP (and other organizations too) have developed a wide array of tsunami-related materials. To protect and leverage the investment in these products, it is becoming increasing important to create an infrastructure that is robust enough, capacious enough, and simple enough to provide a home for these materials. By serving as a central storehouse and archive, the repository protects against data loss, and facilitates the sharing and reuse of existing materials.

## Purpose and scope of this document

The purpose of this document is to identify and prioritize the requirements for a repository designed to hold materials related to tsunami modeling, education, and mitigation. The document focuses especially on the need to create a repository for the items tallied in the NTHMP inventory of tsunami-related materials. The core information presented here consists of functional requirements, but additional information also addresses data and metadata requirements and potential policies to govern the repository.

Web sites, at least as complete functional entities, are not included in the repository and are beyond the scope of this document.

## Audience for this document

This document assumes that readers are generally familiar with the issues surrounding tsunami modeling, education, and mitigation so the document presents no background on those topics. The primary foci of the document are to meet the needs of the NTHMP and of its component organizations, including tsunami warning centers, tsunami modelers, and emergency managers; and to present sufficient detail so that developers can implement the repository and so that testers can verify the functionality.

## Conventions

Requirements are stated in simple present tense, and the priority assigned to the requirement is indicated with one of the letters M, S, or C.

* M(ust) – One of the minimal set of features required for a basic repository implementation.
* S(hould) – A key feature that makes the repository more complete and has significant benefits for users.
* C(ould) – A feature that extends the functionality of the repository but is not essential for a basic implementation.

Each function is marked with one or more of the following letters to indicate which user classes have access to the function. We need a way to associate these directly with each requirement.

* U(ser)
* C(ontributor)
* A(dministrator)

Items colored like this are machine behaviors, not something that a user does.

# Overall description

Conceptually, the NTHMP repository, in common with Open Archival Information System (OAIS) repositories in general, has six functional areas and the interfaces that connect them (CCSDS 4-1). As illustrated below, a contributor provides a submission information package (SIP). After acceptance and any necessary transformations, the resultant archival information package (AIP) becomes available in the repository. Consumers then interact with the repository to identify and request information, which the repository delivers to them in the form of a dissemination information package (DIP). The vertical dashed lines indicate that the repository administration is involved with the various aspects of ingest, data management, storage, preservation, and access.

Describes general factors and provides background for requirements. Include: product functions, user characteristics, constraints, assumptions, dependencies.



## Project parameters

This section of the document circumscribes the NTHMP repository by listing relevant objectives, standards, assumptions, and constraints.

### Objectives

The NTHMP repository seeks to provide:

* An attractive, web-based interface (NTHMP Strategic Plan).
* Robust access to content by utilizing well-targeted metadata and efficient search.
* Intuitive presentation of related content.
* The ability to select multiple items (a shopping cart) for access.
* An uncluttered interface that adjusts to the kind of user, whether general public, emergency management, or researcher.
* Easy deposit of content

### Standards (Open issue)

Standards such as “Dublin Core,” “OAI-PMH” etc.

Section 508 user interface accessibility requirements

### Assumptions

Repository policies typically specify what kind of content is appropriate and control how the content is stored and delivered. At this point, there is no agreed upon policy document for the NTHMP repository and so the assumptions in this section are used instead. These assumptions comply with the proposed policies found in Section 6.6.

|  |  |
| --- | --- |
|  | Only digital materials are eligible. |
|  | Web pages are supported only to the extent of allowing the ingress of HTML documents. |
|  | “Look and feel” is secondary to the preservation of content. If it is necessary to migrate content because the existing format is becoming obsolete, the primary migration goal is to preserve the content. |
|  | The repository does not provide any special software required to access repository materials (e.g. for audiovisuals or computer applications). |
|  | The repository can contain objects of different levels of permanence. |
|  | In general, all objects in the repository are available for public use, although specific objects might be restricted or embargoed for specified periods of time. |
|  | The repository does not harvest metadata from other repositories. |
|  | Metadata is included only for digital materials stored within the repository, not for digital materials stored elsewhere. Open issue |
|  | The repository design minimizes duplicate data entry of metadata. |
|  | The repository uses automation whenever possible to extract descriptive and technical metadata. |

### Constraints

The objects in the NTHMP repository embody large investments of time and money, so the repository must implement appropriate security controls and procedures.

Security requirements and interfaces with off-site remote backup facilities.

## Classes of users

Describe end-users, contributors, staff, who require multiple levels of access to the materials. Education, knowledge level, technical expertise. Map user classes to functions they can use.

The primary users of the repository are engaged in creating and using tsunami-related materials for a variety of purposes, from research, to education, to mitigation, to history. As such, they can take on different roles at different times:

* Content creators
* Content contributors
* Content discoverers
* Content users
* Repository administrators

Users can also be envisioned according the roles they play in the tsunami community:

* Public
* Emergency management
* Community planning
* Teaching
* Tsunami research

# Specifications

This section specifies the repository to a level of detail sufficient to design and test the system.

## Functional requirements

### Manage user interface

#### Select preferred localization language (UCA)

|  |  |  |
| --- | --- | --- |
|  | M | **Localize interface**Source: conversations with potential contributorsThe user interface can be localized to English, Spanish, or French |

#### Manage preferences (UCA)

##### Turn tooltips on or off

##### Return to default preferences

##### Provide list of items for which to be notified of new items

##### Sign up for RSS

##### Sign up for emailed information

#### Navigate the website (UCA)

|  |  |  |
| --- | --- | --- |
|  | M | **Navigate website**Source: conversations with potential contributorsThe repository provides the following means of navigating the site: navigation bar, breadcrumbs, search, site map.Comments: Many potential contributors indicated that ease-of-use is a critical overarching requirement for the repository. These diverse navigation tools offer a number of different approaches to finding information in the repository, and are essential to making information easy to locate and reducing a user’s frustration when using the site. |

### Manage users

#### Manage passwords (UCA)

##### Create user account ID and password (newbyA)

###### Provide email address(es)

Verify email address

Select default email address

###### Provide first and last names

###### Specify user’s organization

###### Create user account ID

###### Create password

###### Accept the terms of the license

##### Change password (UCA)

##### Recover forgotten user account ID (UCA)

##### Recover forgotten password (UCA)

#### Authenticate user (UCA)

##### Log in

##### Log out

##### Choose type: default, user, contributor, administrator

#### Manage user types

|  |  |  |
| --- | --- | --- |
|  | M | **Support multiple user roles**Source: review of other requirementsThe repository supports multiple user roles and allows actions based on those roles.Comments: Given other requirements, a method to limit contributions to specific individuals is clearly needed. The means of requesting contributor status should be available within the repository and should not require that users initiate the request via other channels. |

##### Request to become a designated contributor (U)

|  |  |  |
| --- | --- | --- |
|  | M | **Request contributor role**Source: review of other requirementsThe repository provides a way for any user to request the contributor role. |

##### Withdraw from being a designated contributor (C)

##### Authorize a would-be contributor (A)

|  |  |  |
| --- | --- | --- |
|  | M | **Notify administrator of request to become contributor**Source: review of other requirementsThe repository provides a way to notify an administrator that a user has requested to become a contributor, and provides a way for the administrator to approve or deny that request. |

##### Authorize a would-be administrator (A)

##### Deauthorize a contributor (A)

##### Request to become a forum moderator (UC)

##### Withdraw from being a forum moderator (UC)

##### Authorize a would-be forum moderator (A)

### Manage data objects

#### Ingest data objects (CA)

##### Contribute data objects offline

###### Contribute on physical media

|  |  |  |
| --- | --- | --- |
|  | M | **Accept data on physical media**Source: similar data management issues at NGDCThe repository administrator can accept data objects on, at a minimum, the following kinds of physical media:* USB portable hard drives
* CD-Rom
* DVD-ROM
* USB flash drives

Comments: At times, especially at the initial ingest of material, any of the following conditions may make it impractical for a contributor to submit objects via the Internet:* A large number of objects to submit
* Objects of large size to submit
* Slow network connection, or connection is unavailable

NOTE: This requires a process that scans the media to ensure no viruses or malware are introduced to the administrator’s workstation or to the repository. |

##### Contribute data objects online

|  |  |  |
| --- | --- | --- |
|  | M | **Web ingest**Source: conversations with potential contributorsThe repository offers a means for contributors to upload material to the repository. Comments: Contributors have stated that the repository must be as easy to use as possible. While it is necessary to have other means of ingesting large volumes of objects or objects whose size precludes uploading via the Internet, in most cases it is far easier for contributors and for the repository administrator if contributors can directly submit materials electronically. |
|  | S | **FTP ingest**Source: The repository offers a means for contributors to FTP material to the repository.Comments: While direct upload via a web form is the preferred method, an alternative (or auxiliary) method is to provide an FTP site. This requires (a) automation to process contributed materials, (b) manual processing by a repository administrator, or (c) some combination of these two processes. |

###### Specify file path to data object

Browse to data object location

Submit file path

##### Provide initial information

###### Input submission information

Grant/accept licenses and policies

Certify data objects comply with copyright requirements

###### Correct submission information as necessary

###### Verify submission information

##### Receive notice of acceptance or rejection (C)

#### Check suitability of contribution (A)

##### Create automatic rejection filter

|  |  |  |
| --- | --- | --- |
|  | M | **Check compliance with format requirements**Source: independent researchThe repository does not allow contributors to submit materials unless the materials are in a recognized and supported format. |

##### Receive notice of new potential contribution (A)

##### Check compliance with policies

|  |  |  |
| --- | --- | --- |
|  | M | **Check compliance with policy against PII**Source: independent research, conversations with potential contributorsThe repository ensures that all submitted materials have any personally identifiable information (PII) correctly redacted. |

##### Check compliance with supported data types

##### Check submission for viruses

##### Check suitability of metadata

##### View or examine proposed contribution

##### Authorize or reject submission

##### Notify contributor

###### Notify of automatic rejection

###### Notify of rejection by human reviewer

#### Create low-res versions of hi-res images

|  |  |  |
| --- | --- | --- |
|  | M | **Create low-res versions of hi-res images**Source: Conversations with potential contributorsWhen returning images in search results, the repository uses thumbnails for efficiency, but disseminates images in the highest available resolutions.  |

#### Store data objects

##### Ingest packages

##### Normalize by converting into a supported format

##### Assign a unique identifier to every object

###### Manage “complex” or bundled objects

|  |  |  |
| --- | --- | --- |
|  | S | **Collections or packages**Source: Conversations with potential contributorsThe repository allows multiple related items to be identified as a "collection" or "package.”Comments: Some contributors indicated that they have certain types of objects (such as evacuation maps) that are part of a package of materials. Some users might be interested in acquiring the entire package, but other users might only be interested in acquiring the map itself.  |
|  | S | **Visibility of items within packages**The individual items included within packages are searchable and appear in search results. |
|  | S | **Ability to download individual items from packages**The repository allows users to download individual items from packages (e.g., a single evacuation map). |
|  | S | **Allow enforced bundling**Source: conversations with potential contributorsThe repository can prevent users from downloading certain types of materials (such as inundation maps) without accompanying materials that ensure appropriate use (such as a technical report). This requirement enhances requirement 3.1.4.2.3.8.1. |

###### Guarantee uniqueness of identifier

##### Provide a stable reference (persistent URL)

#### Manage personally submitted data objects (CA)

##### View list of personally submitted data objects

#### Archive AIPs

|  |  |  |
| --- | --- | --- |
|  | M | **Archive all repository data objects**Source: NTHMP Strategic Plan, conversation with Jenifer Rhoades, Sue McLean, et al.All objects deposited in the repository are archived. |

#### Manage obsolescent data objects (A)

#### Perform file migrations (A)

|  |  |  |
| --- | --- | --- |
|  | M | **Migrate obsolete data formats**Source: conversations with ESRL repository managerThe repository allows for upgrading object formats when they become obsolete. |

#### Withdraw data objects leaving tombstone behind (A)

|  |  |  |
| --- | --- | --- |
|  | M | **Make data objects inaccessible**Source: conversations with potential contributors, independent researchThe repository provides a way for contributors or administrators to mark a data object as inaccessible, without affecting the visibility of the associated metadata.Comments: At times, it is necessary to prevent any users from acquiring a particular object from the repository, but not to lose track of the metadata about that object. Rather than deleting the object from the repository, there should be a way for contributors or administrators to mark it as inaccessible.. |

#### Expunge data objects, not leaving tombstone behind (A)

### Manage metadata

#### Implement metadata schema or schemas (A)

#### Contribute metadata (CA)

##### Import existing metadata

###### Import from XML source

|  |  |  |
| --- | --- | --- |
|  | M | **Accept metadata from XML sources**Source: conversations with potential contributors, independent researchThe repository supports the import of metadata from XML. |

##### Use crosswalk to translate metadata

##### Create new metadata (CA)

|  |  |  |
| --- | --- | --- |
|  | M | **Create metadata via provided forms**Source: conversations with potential contributorsThe repository’s upload process presents contributors with web forms that capture all the required metadata for the submission. |
|  | M | **Allow administrators to manage metadata**Source: The repository provides a means for an administrator to manage metadata. |

###### Specify names

###### Specify dates

###### Specify descriptions

###### Specify keywords

###### Specify access restrictions

###### Specify disclaimers

|  |  |  |
| --- | --- | --- |
|  | M | **Accept disclaimers**Source: conversations with potential contributorsThe repository is capable of including appropriate disclaimers for data objects that might be used incorrectly. The repository provides a way for a contributor to include his own disclaimer. (See also: Requirement 3.1.3.4.3.1.4.) |

###### Specify origin, source of data objects

###### Specify expected user type

###### Specify additional or approved sources for the data object

|  |  |  |
| --- | --- | --- |
|  | S | **Specify additional sources for data objects**Source: conversations with potential contributorsThe repository allows contributors to indicate that items are available from other sources for a fee. Contributors can provide links to the items so users can order the items if desired.Comments: Some contributors indicated that they make some materials available for a fee, generally to recoup the cost of providing the materials to the public. It is important that the repository does not create a back-door means of acquiring those materials for free, unless authorized by the contributor. Instead the contributor should be able to indicate that the material is available for a fee and provide links to purchase or request the materials. |

#### Edit metadata (CA)

|  |  |  |
| --- | --- | --- |
|  | M | **Edit metadata in situ**Source: conversations with potential contributors, independent researchThe repository allows in situ editing of metadata via web forms. |

#### Validate metadata against schema (CA)

#### Associate metadata with data objects (CA)

### Locate data objects

#### Specify external reference, such as a data object’s handle (UCA)

#### Search the whole web site, not just data objects (UCA)

#### Search for data objects (UCA)

##### Simple search without options

##### Advanced search with options

|  |  |  |
| --- | --- | --- |
|  | M | **Flexible search**Source: conversations with potential contributorsThe repository supports searches of the following kinds:* Map-based search for georeferenced information
* Keywords
* Object type (e.g. map, emergency response plan, educational game, etc.)
* Object format (e.g. PDF, Word document, PowerPoint presentation, etc.)
* Language
* Author(s)
* Contributor
* Dates
 |

###### Use Boolean search

##### Ascertain persistent URL and unique ID for each data object

##### Sort and display results

|  |  |  |
| --- | --- | --- |
|  | S | **Search items per page**Source: SchiffUsers can choose to display 10, 25, or 50 items per page of returned search items. |
|  | S | **Browsing through pages of results**Source: SchiffUsers can browse back and forth through the pages of results. Users are given information as to the number of pages of results that there are for a given search. |
|  | S | **Result item information** (Open issue)Source: SchiffEach item in a result set includes the following information:* Author
* Title
* Keyword in context in results
* KWIC Pics, where an image snippet appears when hovering over keyword in context results, for PDFs and ???
* Linked abstract
* Similar items, that display when the Similar items link is clicked.
* User reviews or ratings (Need requirement for these.)
* Download this item, which when clicked allows user to print or save the content.
 |

###### Sort results by relevance

|  |  |  |
| --- | --- | --- |
|  | S | **Default sort is by relevance**Source: SchiffThe default sorting of search results is by relevance. Users can choose among the following sorting criteria: publication date, reverse date, author, title, and back to relevance. Open issue |

###### Sort results by date

|  |  |  |
| --- | --- | --- |
|  | S | **Present most recent version first**Source: conversations with potential contributorsWhen multiple versions of a data object are returned in a result set, the repository, by default, presents the most current version of the data object first.Comments: This requirement is not intended to address overall presentation so much as presentation of objects with multiple versions (i.e. “current” vs. “archived”). Since the repository should provide a means to retrieve older versions from the archive, it should nevertheless present the most current version as the first choice.One design choice would be to only return the current version in basic search results, and implement an advanced search capability to also search archived versions. |

###### Sort results by location

###### Sort results by expected users

##### Find related items

##### Support search by software agents

#### Browse for data objects (UCA)

##### Browse by subject

##### Browse by location

##### Browse by title

##### Browse by author or contributing organization

##### Browse by date

##### Browse by image

#### Access lists of featured data objects (UCA)

##### Receive “what’s new” alerts of recently added data objects (UCA)

|  |  |  |
| --- | --- | --- |
|  | S | **Provide “What’s New” information**Source:The repository provides the following means for users to discover what is new in the repository:* Automatically-updated "What's New" content
* RSS feed
* E-mail subscription (each vs. digest)

Comments: In a typical portal implementation the automatically-updated “What’s New” content would probably be delivered in a channel (portlet, gadget) of its own. The RSS feeds should be compatible with RSS 2.0 and Atom 1.0 specifications. |

##### Receive notice of interesting data objects selected by the administrator (UC)

##### Receive info about the most popular downloads (UCA)

#### Select items to operate on from set of located data objects (UCA)

|  |  |  |
| --- | --- | --- |
|  | M | **Shopping cart capability**Source: independent researchThe repository allows users to select multiple items to acquire before determining the acquisition method.Comments: Other portals have implemented a “download cart” methodology where users can select multiple objects to download before initiating the process. This allows users to choose the method of receiving the objects:* Individually
* ZIP archive
* On physical media (presumably CD-ROM or DVD-ROM)

The repository might also implement some form of download manager software (typically an applet or plugin that users must download and install before downloading the desired content. There should be compelling reasons to implement this type of capability, since the process of installing the required software can be frustrating to a user even when it goes smoothly. Even if this capability is implemented, the repository should offer alternatives that do not require users to install additional software. |

### View or download data objects

#### Control access to data objects

#### View selected data objects on the web (UCA)--???

|  |  |  |
| --- | --- | --- |
|  | S | **Display metadata along with data**Source: conversations with potential contributorsThe repository displays proper attributions as part of the metadata for data objects that are returned in search results.Comments: Proper attributions must be part of the corresponding metadata. The repository should display the attributions when returning objects in search results. |

#### Present data objects to software agents

|  |  |  |
| --- | --- | --- |
|  | M | **Visibility to internet search engines**Source: conversations with potential contributorsThe repository allows internet search engines to crawl its contents.Comments: In order to make the repository as useful as possible, it should allow Internet search engines to crawl its contents. This helps to advertise the repository to a wider audience and helps it become a primary resource for tsunami information. |

#### Disseminate selected data objects (UCA)

There are several factors that govern how a user might wish to acquire materials from the repository: the speed of their network connection, the number and aggregate size of the materials they wish to acquire, and the tools they prefer to use to manage the information. No one method satisfies all users’ needs equally, so the repository should offer alternatives to downloading one object at a time via a web browser.

|  |  |  |
| --- | --- | --- |
|  | M | **Disseminate objects via web browser**Source: independent research, conversations with potential contributorsThe repository allows users to download one object at a time via a web browser. |

#### FTP selected data objects (UCA)

|  |  |  |
| --- | --- | --- |
|  | S | **Disseminate objects using FTP**Source: independent research, conversations with potential contributorsThe repository allows users to download objects using an FTP site. |

#### “Zip and ship” selected data objects

|  |  |  |
| --- | --- | --- |
|  | M | **Disseminate objects using “zip and ship”**Source: independent research, conversations with potential contributorsThe repository is capable of using a “zip and ship” approach to disseminate objects. Comments: This could occur as a download (via either of the previous methods) or could involve writing the selected objects to physical media and mailing the media to the user. |

### Access metadata (UCA)

#### Control access to metadata

#### View user-comprehensible metadata (UCA)

#### Export metadata in standard formats (UCA)

##### Export to XML

|  |  |  |
| --- | --- | --- |
|  | M | **Allow export of metadata to XML**Source: conversations with potential contributors, independent researchThe repository is capable of exporting metadata in XML format. |

#### Provide software-harvestable metadata

### Manage contributor profiles

#### Create contributor profiles (CA)

|  |  |  |
| --- | --- | --- |
|  | S | **Ability to create personal profile**Source: conversations with potential contributors, independent researchThe repository allows each user to construct a personal profile.Comments: Personal profiles allow users to provide some basic information about themselves and to advertise their organizations. Contributions to the repository can be linked to the contributors’ personal profiles (see requirement 3.4.1.) |
|  | S | **Self manage personal profile**Source: conversations with potential contributors, independent researchThe repository allows users to manage their own profiles. |
|  | S | **Administrative management of personal profiles**Source: The repository allows administrators to manage other users’ profiles. |

##### Provide links from profile to contributor-chosen data objects

#### Revise contributor profiles (CA)

#### View contributor profiles (UCA)

### Manage repository documentation

#### Add documentation (A)

#### Delete documentation (A)

#### View documentation (UCA)

##### Access guidance information for using the repository

###### View summaries of available documentation

###### View complete documentation texts

###### View video instructions

###### Access FAQs

###### View tooltips

##### Learn “about repository”

###### Learn about hosting organization

###### Read announcements

###### Read about staff

###### Learn about disclaimers

###### Learn about policies

Learn how cookies are used

###### Learn about metadata conventions

Formats used for exported metadata

###### Learn how to cite items in the repository

###### Learn how to link to items in the repository

##### Find out how to contact repository personnel

##### Learn what to do if there is a problem with the repository

#### View use statistics (CA)

|  |  |  |
| --- | --- | --- |
|  | S | **Contributors have access to usage statistics** (open issue)Source: Schiff, independent researchThe repository provides content contributors with, at a minimum, the following usage statistics:* Number of downloads
* Number of object views
 |

#### Audit security (A)

#### Audit performance (A)

#### View ingestion and data serving log (A)

### Provide feedback about the repository or about data objects

#### Send email to repository administrator (UC)

#### Comment via social media (UC)

|  |  |  |
| --- | --- | --- |
|  | S | **Provide means for users to comment or discuss**Source: conversations with potential contributorsThe repository offers a means (e.g., forums) for users to comment on or to discuss materials. |
|  | S | **Utilize social media**Source:The repository ties in with, at a minimum, the following social media websites: Facebook, Twitter, Digg.Comments: These methods allows users to advertise links or materials of interest to their respective communities: Facebook, Twitter, Digg, Others?Question: If the repository implements a means of sharing items on Facebook, should there be a corresponding “fan page” for the repository? |

#### Operational scenarios

What are the inputs to a function, the system interactions, and the outputs for all major user classes and functions.

## Interface requirements

### User interfaces

|  |  |  |
| --- | --- | --- |
|  | M | **Comply with ADA requirements (Open issue)**Source: independent research, government requirementsThe repository user interface complies with:* The W3C Checklist of Checkpoints for Web Content Accessibility Guidelines 1.0: http://www.w3.org/TR/WAI-WEBCONTENT/full-checklist.html
* WebAIM Section 508 Checklist: http://www.webaim.org/standards/508/checklist.php
 |
|  | S | **Provide “What’s New” information**Source: The repository provides the following means for individuals to discover what is new in the repository:* Automatically-updated "What's New" content
* RSS feed
* E-mail subscription (each vs. digest)

Comments: In a typical portal implementation, automatically-updated “What’s New” content would probably be delivered in a channel (portlet, gadget) of its own. RSS feeds should be compatible with RSS 2.0 and Atom 1.0 specifications. |
|  | S | **Allow direct URL link to specific objects in the repository**Source: conversations with potential contributors**The repository allows users to link directly to specific objects within the repository.**Comments: **Implication: do not use HTML frames as an implementation****Once users discover certain materials in the repository, they may wish to refer to them directly in the future. The repository should allow a user to discover the item’s URL in order to bookmark it. Some implementations mask URL’s making it difficult or impossible to save the item’s location** for future reference. |
|  | S | **Links to NTHMP documents**Source: conversations with potential contributorsThe repository user interface provides links related to working with the NTHMP (such as grant application forms, etc) |
|  | C | **Branding**Source: SchiffAll pages in the user interface have a branding area at the top that includes at least the NTHMP brand. |

#### Required screens

#### Required reports

#### Online help

|  |  |  |
| --- | --- | --- |
|  | S | **Tooltips or “hover help”**Source: independent researchThe repository uses tooltips (or "hover help") where appropriate for help and additional information. |

#### Documentation

|  |  |  |
| --- | --- | --- |
|  | S | **Document guidelines and instructions for contributors**Source: conversations with potential contributorsThe repository documents clear instructions for contributing materials. The instructions are viewable online, and are available for download as PDF documents. |
|  | S | **Internal documentation**Source: SchiffThe repository back end is documented to address internal management and service disruption issues. General areas of documentation include:* Ingest process for all content sources
* Metadata generation and management
* Indexing process
* Production environment, start-up and shut-down routines
* Usage statistics generation and delivery
* Workflows
 |

### Software interfaces

Name of application with which the repository must interface. State details of interface if determined by the other application.

|  |  |  |
| --- | --- | --- |
|  | M | **Browser support for common browsers**Source: independent researchThe repository works correctly with the Firefox, Google Chrome, and Internet Explorer browsers. |
|  | S | **Browser support for other browsers**Source: independent researchThe repository works correctly with the Opera and Safari browsers. |

### Communication interfaces

To other systems or devices, such as LANs.

## Data requirements

Data entities, their decomposition, their definitions. Data structures.

|  |  |  |
| --- | --- | --- |
|  | S | **Links to training resources**Source: conversations with potential contributorsThe repository includes information to guide users to training resources.Comments: One of the primary reasons to implement the repository is to provide resources for modelers, emergency managers and community planners in areas where tsunami preparation is minimal. Some of the contributors we interviewed mentioned that the repository should provide links to any available training materials, especially those that might not be appropriate for inclusion in the repository itself. |

## Metadata requirements

About the metadata…

|  |  |  |
| --- | --- | --- |
|  | S | **Proper credit**Source: conversations with potential contributorsProper attributions are included in the metadata for repository data objects. The repository displays the attributions when returning objects in search results.Comments: Several potential contributors indicated that being given credit for their contributions would be an incentive to continue contributing materials to the repository. (See requirement .) |
|  |  | Formats accepted |
|  |  | Check that the required minimal set is provided |
|  |  | Metadata should be stored in a single, common format |
|  |  | Allow for metadata updates |
|  |  | Ability to search, display metadata for humans and machines |

## Operational requirements (Open issue)

How the repository runs and communicates with operations personnel.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Audit trail | Activities (and the data) to be recorded in the audit. |
|  |  | Availability | Hours available to use. When is peak usage?  |
|  |  | Capacity | Expected volumes in user terms. (The system must be able to ingest so much per month, handle so many users simultaneously) |
|  |  | Fault tolerance | Probably not applicable. Assume when a portion fails, all becomes unusable. |
|  |  | General performance | Response time for queries and activities. Expected rate of user activity (transactions per hour) |
|  |  | Recoverability | Ability to restore function and data after failure. How soon? To what level of currency? If site is destroyed, how soon can it be restored? |
|  |  | Reliability | What damage can occur if the repository fails? What is the measurement of reliability to be used? |
|  |  | Safety |  |
|  |  | Security | Who can view and alter data. Consequences of erasure or changes to data, consequences of disclosure of information about individuals. State type of security required: access to facility, classes of users, access by class, access to system functions by user class, accreditation? |
|  |  | Software quality |  |

## Maintenance requirements

### Testability

### Maintainability

# Repository policies

## Existing policies and guidelines that affect repository design

## Policies to be developed

# Concerns

# Appendices

## Minimum metadata

### Title

### Author

### Organization making the resource available

### Identifier/name of the resource contributor

### Creation date

### Date object became available from the repository

### Type of data (format)

### Physical description (extent, size, duration)

### Rights, copyright

### Language of the resource

## Technical metadata

## User interface prototypes

## Formats

## Survey responses

## Policy drafts

## Glossary

**Administrative Metadata**—“Provides information to help manage a resource, such as when and how it was created, file type and other technical information, and who can access it” (NISO 1).

**AIP**—See “Archival Information Package.”

**Archival Information Package (AIP)**— “Information packets of content converted to the archive standards….Generating the AIPs deals with transforming or repackaging of data to the current…archival standards” (NLM 20).

**Descriptive Metadata**—“Describes a resource for purposes such as discovery and identification. It can include elements such as title, abstract, author, and keywords” (NISO 1).

**DIP**—See “Dissemination Information Package.”

**Dissemination Information Package (DIP)**—“The version of the information package that is delivered from the repository to the user in response to an access request” (NLM 49).

**Ingest**—“The process through which objects are added into the Digital Repository” (NLM 50).

**Metadata**—“Structured information that describes, explains, locates, and otherwise makes it easier to retrieve and use an information resource” (NISO 16).

**OAIS**—See “Open Archival Information System.”

**Open Archival Information System (OAIS)** —“An archive, consisting of an organization of people or systems, that has accepted the responsibility to preserve information and make it available for a Designated Community. The term OAIS also refers, by extension, to the ISO OAIS Reference Model for an OAIS. This reference model is defined by recommendation CCSDS 650.0-B-1 of the Consultative Committee for Space Data Systems; this text is identical to ISO 14721:2003” (Open Archival Information System)

**SIP**—See “Submission Information Package.”

**Structural Metadata**—“Indicates how compound objects are put together, for example, how pages are ordered to form chapters” (NISO 1).

**Submission Information Package (SIP)** —“The version of the information package that is transferred from the content producer …prior to its ingest into the repository” (NLM 52).

**Technical Metadata**—“A form of administrative metadata dealing with the creation or storage encoding processes or formats of the resource” (NISO 16).

# References cited

**CCSDS** (Consultative Committee for Space Data Systems). “Reference Model for an Open Archival Information System (OAIS).” CCSDS Secretariat, Program Integration Division (Cods M-3), National Aeronautics and Space Administration, Washington, DC. Blue Book, Issue 1, January 2002. Web. <http://public.ccsds.org/publications/archive/650x0b1.pdf>

**Schiff**, Lisa. “eRep Feature Requirements Document.” eScholarship: University of California. Personal communication, April, 2009.

**NISO** (National Information Standards Organization). “Understanding Metadata.” NISO Press, 2004. Web. <<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>>

**NLM** (NLM Digital Repository Working Group). “National Library of Medicine Digital Repository: Policies and Functional Requirements Specification, Version 1” 2007. Web. <<http://www.nlm.nih.gov/digitalrepository/NLM-DigRep-Requirements-rev032007.pdf>>

“**Open Archival Information System**.” Wikipedia. 2010. Wikimedia Foundation, Inc. 14 Oct 2010 <http://en.wikipedia.org/wiki/Open\_Archival\_Information\_System>

# Index