

# Glossary

## Glossary - DRAFT

### Collection

The term "collection" is used for the user-facing interfaces. It may be modified with an adjective to differentiate between Administrative Set and Discovery set in the UI, but that is still TBD.

The term "Collection" has two definitions (usages) in Hypatia.

- Collection as Administrative Set: A group of items bound together by shared descriptive metadata and policies. A simple descriptive metadata record – "title" and "description" – applies to each member item of the set. (An item also has more extensive item-level metadata). Policies, set by the Collection Administrator, that bind together a set include: access restrictions (e.g., embargo, discoverability, downloadability) ; collection management (contact information, who can deposit and administer). A collection (administrative set) can have one or more Administrators.
- Collection as Discovery Set: A group of items arbitrarily grouped together for purposes of discoverability. Items of a discovery set may originate from one or more administrative sets. Some descriptive metadata (title ... what else?) may be applied to this kind of collection in order to facilitate discovery and use; however, no administrative or usage policies may be defined and applied to a discovery set as a whole. The policies of an item's administrative set govern the item in all contexts, regardless of which discovery set(s) it may be a part of.

**MM: How does this differ from the archival concept of the "collection"? It seems more like an administrative set just because I don't see how a discovery set would apply.**

- **MO: Administrative set in my mind most closely matches our understanding of an 'archival' collection but is probably oversimplified. We'll need administrative sets that can be clustered in parent child relationships with one administrative set (the collection) having additional administrative sets that live under it (series). I'm still trying to get a better understanding of Discovery set and wonder if this definition is off in one important respect; policies for discovery sets will need to be varied and can't be universally inherited from a single administrative set policy as these will vary. Some series will need to be discoverable, metadata only discoverable, or not discoverable at all. Thoughts? MO**

### Item/Object

A cohesive unit, containing one or more files, that constitutes a single information resource. An item is required to be a member of a collection (administrative set) (even if that collection contains only one item). An item may optionally also be part of a collection (discovery set).

Note: In some cases an item has been termed an "object" by developers but will always be termed "item" in the user-facing UI

Examples of an item include:

- a forensic or logical disk image. Disk images can contain many items.
- a single file (text, audio, video, image)

## Dashboard / Administrative Console

### 1. intellectual arrangement

Each archival collection will have its intellectual arrangement, that is the arrangement of the material in a hierarchical nature that intends to reflect its original creation or arrangement within a recordkeeping system. Over time additional material may be received and these accessions will be integrated into the collection and the intellectual arrangement will be updated. The arrangement is used to portray and distinguish critical elements of context.

### 2. descriptive metadata [in-progress]

To facilitate discovery and access to the archival collection archivists like to provide descriptive metadata about the archives. This typically includes an overview of the entire collection with a biographical / historical introduction to place the collection into its wider context (for example about the person, family or organisation), the scope, span dates and format of the material and details of related material held in other collections (within the same archival repository or held elsewhere).

At a lower level within the intellectual arrangement the descriptive metadata will often include title, description, dates, format and extent information. In some cases each distinct physical item will have its own description but this is not always the case for example we might assign a single reference number to a bundle of correspondence and clarify the extent as 1 bundle (52 letters). With born digital material the sheer volume means that there is a high likelihood that a single descriptive entry will be used for a series of related digital assets

Although there are different national cataloguing standards (DACS in North America and ISAD(G) in UK) there is wider commonality in sharing this data using EAD [see below] which clearly identifies possible fields at collection and component levels.

### 3. technical metadata [todo]

### 4. EAD [in-progress]

EAD2002 - Library of Congress website <http://www.loc.gov/ead/index.html>

Schema - <http://www.loc.gov/ead/eadschema.html>

Tag Library - <http://www.loc.gov/ead/tglib/index.html> or as [PDF](#)