Deciding on Fields to be used in ICO Collections

Collection Configuration

The cataloging tool used by Image Collections Online is called Photocat. Photocat is infinitely configurable to allow the collection of metadata that's appropriate for any collection. As part of the process of setting up (or revising) a collection, a set of metadata fields must be determined to meet the needs of the materials in the collection. If collections contain wildly disparate materials it may be more fitting to treat materials of different types as different collections rather than present irrelevant or confusing fields to catalogers.

When metadata already exists (even just the filenames assigned images as scanned) it should be preserved when moving those items into the Photocat system. For metadata that is poorly structured, unnormalized, or of ambiguous meaning it is often best to import these values into the most fitting field and then clean up the values once in Photocat. Alternative values may be imported for reference only. While those values would be associated with the record (and searchable) they may be read-only and wouldn't be presented to users.

Field Selection

All fields used by all existing collections are defined in an XML Field Definitions file (see the current version here). A field may consist of multiple parts, and may be multivalued.

Some of the more popular fields that are applicable to most media types are:

- Title
- Date taken/created
- Topical Subject
- Geographic Subject
  - Country
  - State
  - County
  - City

For photograph collections the following fields are often exposed:

- Photographer
- Print size
- Film Type
- Person pictured
- Event

If there is no current field that is appropriate for a piece of information that you've collected or would like to collect for items in a collection, a new field can be defined. To define a field we need to determine its structure, attributes and the semantic meaning of data entered into it.

Vocabulary Sources

To facilitate easier and more accurate metadata entry, most field types can be loosely or tightly coupled with a controlled vocabulary source.

The following thesauri are currently exposed in a way that allows for easy integration with Photocat:

- Locally hosted/mirrored Thesauri
  - TGMI
  - TGMII
  - Digital Library of the Commons Thesaurus
  - AFS Ethnographic Thesaurus
- OCLC Terminology Services(subject to some restrictions)
  - LCGFT
  - BISAC®
  - FAST subject headings
  - Form and genre headings for fiction and drama

Providing a list of possible values for a field is another option.

Alternately, though not strictly a "controlled" vocabulary, fields may be configured to remember the values entered into them and suggest those values when users begin typing terms.

Field Usage Guidelines

Resulting from discussions with stakeholders and metadata experts, the Photocat Core Fields are now required as part of the setup of new collections in Photocat, the IUB Libraries' online image collection cataloging application. This field selection represents a unified endorsement of those elements widely shared by image collections. All new collections added to Photocat will automatically contain these fields and collections are encouraged to make use of them, adhering to the definitions for each field.

Special Fields

Submitted Filename:
Collections imported from sets of images for which the image filenames contain useful descriptive metadata may be configured to display the original name of the file in this field. This value may not be edited, but is useful in searches or for uncataloged item identification.

Last Modifying users:
The username of the individual who last saved the record in Photocat will be presented in this read-only field.

STATUS:
Besides being used internally to track the status of a metadata record for an item, the status field is also tied to the public-facing web application to determine whether a particular field is public. Status fields are defined as:

- auto-generated: initial submission (you can't save a modified record with this status)
- in progress: indicates a record is being worked on and should not be published
- minimal: published, but has the minimal amount of metadata necessary
- pending completion: some checks may be necessary, not published
- cataloged: metadata is complete - published

We encourage you to use the status field as a guide for determining which records are complete and which need more work.