

**Centre for Environmental  
Policy and Advocacy**

*An environment and natural resources library for Malawi*

# CEPA Curation Manual



**Ferguson Lynch**

Rosanne Hessmiller, PhD  
Ferguson Lynch

November 9, 2017



## Table of Contents

<b>Defining content curation</b> .....	<b>3</b>
<b>What is Content?</b> .....	<b>4</b>
Content Types .....	4
<b>Copyright and Fair Use applies to content</b> .....	<b>5</b>
Content has a Title and a Description .....	5
<b>Best practices for saving files (file ID naming)</b> .....	<b>6</b>
Follow the file ID naming conventions.....	7
Be specific .....	7
Other tips for file ID naming.....	7
<b>Renaming files</b> .....	<b>8</b>
<b>File ID Naming Convention – CEPA Organization Solution Planning</b> .....	<b>8</b>
<b>Content Has a State</b> .....	<b>9</b>
<b>To change the state of content:</b> .....	<b>9</b>
<b>Sharing Private Content on the CEPA Library Portal</b> .....	<b>9</b>
<b>Sharing with Portal Users</b> .....	<b>9</b>
Sharing with Portal Groups.....	10
<b>Glossary of CEPA Library Portal Terms</b> .....	<b>10</b>
<b>Further reference:</b> .....	<b>12</b>
Standard Rules for Content Curation: The Dublin Core: <a href="http://dublincore.org/documents/profile-guidelines/">http://dublincore.org/documents/profile-guidelines/</a> .....	12
More technology and Ideas and definitions available at: .....	12

## Defining content curation

Ferguson Lynch assisted the CEPA staff with the process involved in successful content curation. For the CEPA organization the content curation process included sorting through the vast amounts of content and presenting it in a meaningful and organized way around a specific theme. Curation of content is ongoing and involves sifting, sorting, arranging, and publishing information. It involves CEPA staff selecting the best quality online resources for their intended audience and organizing and displaying them on their CEPA Library portal so users can access.

Content Curation adds value and insights to the selected content. Strong curation involves making decisions about what is and is not useful to deepening understanding of the subject and helping make content useful for others. Well curated topics and subjects help to inform and allow knowledge transfer to happen at faster rates.

Content must be curated to be successfully uploaded, organized, presented, disseminated, archived, and preserved.

## Digital curation involves maintaining, preserving and adding value to data (content) *throughout its lifecycle.*

Digital curation and content preservation are ongoing processes, requiring considerable thought and the investment of adequate time and resources. You must be aware of, and undertake, actions to promote curation and preservation throughout the data (content) lifecycle. The digital curation lifecycle comprises the following steps:

**Conceptualize:** plan content capture methods and storage options.

**Create:** assign administrative, descriptive, structural and technical archival titles and metadata.

**Access and use (permissions):** ensure that designated users can easily access digital objects on a day-to-day basis. Some digital objects may be publicly available, whilst others may be password protected.

**Appraise and select:** evaluate digital objects and select those requiring long-term curation and preservation. Adhere to documented guidance, policies and legal requirements.

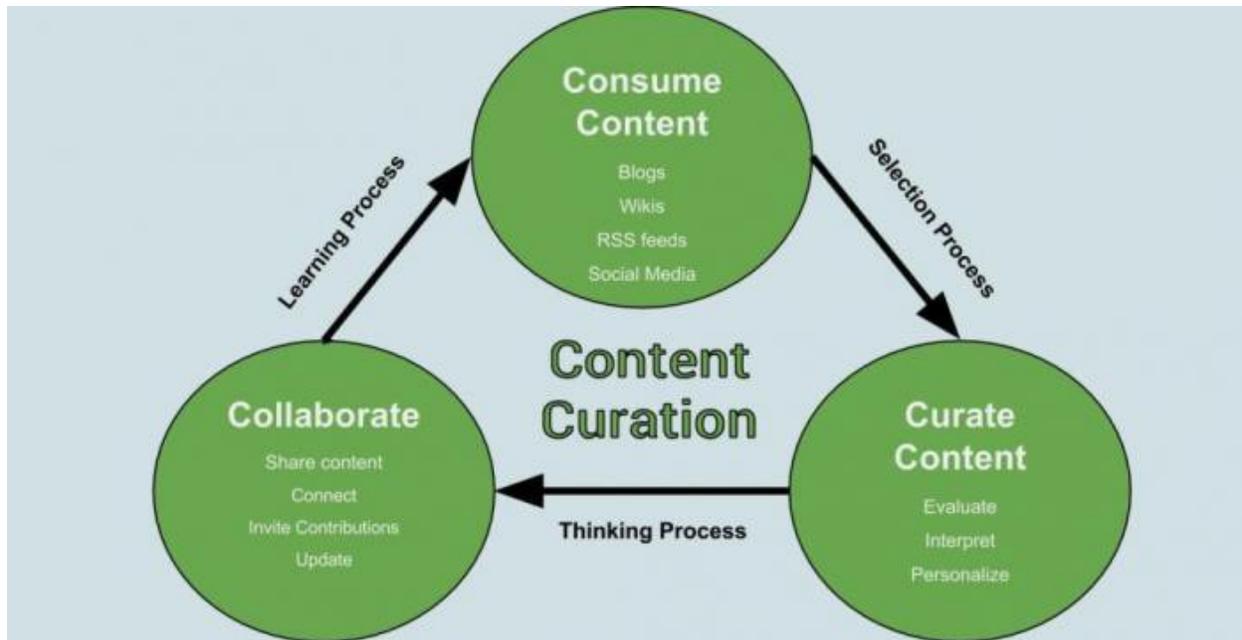
**Preservation action:** undertake actions to ensure the long-term preservation and retention of the authoritative nature of digital objects.

**Store:** keep the data (content) in a secure manner as outlined by relevant standards. Include **Backup and performance logs.**

**Access and reuse:** ensure that content is accessible to designated users for first time use and reuse. Some material may be publicly available, while other data (content) may be password protected.

**Transform:** create new digital objects from the original, for example, by migration into a different form.

**Track:** Document versioning of the content in your system



## What is Content?

Content are web pages, documents, audio files, videos, and links that can be posted, edited, stored, and deleted on the portal, and indexed for easy search. Content can be Private, Published, or in an intermediate state called Public Draft.

## Content Types

Several types of content can be added to the CEPA Library portal are:

- **Page** - web pages made up of text and images, stored on the portal.
- **Link** - URL links to external web sites and pages and internal portal content.
- **File** - file content like Microsoft Word, Excel, Access and PowerPoint files, Adobe Acrobat files, and audio and video files.
- **Image** - digital images that are in a web friendly file format like .jpg, .gif, or .png.
- **Video** - digital videos typically streamed from YouTube or Vimeo

- **Audio** -Radio programming, recordings, Audio files
- **Event** - happenings to be listed on a portal public or private group calendar.
- **News Item** - text news items of interest to the portal community or to a private group of users.

## Copyright and Fair Use applies to content

Before posting content, be clear who owns the copyright, and what terms and conditions might apply to reproducing that content on the portal. Fair usage of copyrighted materials does not necessarily mean you are free to repost that material to the portal, though you may be allowed to repost limited excerpts.

If in doubt:

- Ask the owner of the copyrighted material for permission to reproduce (copy).
- If permission cannot be obtained, it is generally acceptable to link to that copyrighted content, unless the owner specifically prohibits linking or places other restrictions on access.
- Review these portal guidelines on copyright and fair use.

## Content has a Title and a Description

Description of the Title and Description fields for content on the portal.

Content that you **add** to the portal **must** have a **Title** and a **Description**.

1. A **Title** is used to identify your **content** to portal visitors, and in searches.

Some rules of thumb for a **title** are:

- Short but descriptive.
- Uses proper grammar.
- Does not use abbreviations
- Uses characters of the alphabet, numbers, and certain special characters ( , . : . )
- Does not use the characters ! @ # \$ % ^ & \* ( ) = < > ? / \ | + " " " in the title. These characters can have special meanings and cause problems with searching for them.

Examples of good **titles** are:

- Wood Quality Implications of Forest Management
- UA BARA Meeting Minutes, March 12, 2006
- HOWTO: Copyright and Fair Use

2. A **description** is a one-paragraph text abstract that accurately and concisely describes content though inclusion of key words.

Accurate keywords in the description ensure that the document will be found during a search.

Rules of thumb for a **description** are:

- An abstract that quickly describes the content to a casual viewer.
- Uses keywords that concisely describe the contents of the content.
- Uses standard references like the document source, authors(s), date of publication, publishers, copyright information, and web reference, etc.
- Use a single paragraph. Line and paragraph breaks are not supported in a description.

A good example **description** is:

*"Wood Quality Implications of Forest Management" is an extension publication from Oregon State University, which looks at the basics of conifer tree growth, characteristics that define wood quality, and the effects of silvicultural management practices on wood quality. Of relevance to agroforestry practices that aim to grow valuable timber are the various ways that managers can influence wood quality, e.g. planting density, thinning, pruning, and fertilization. The publication is relevant not only to western conifers, but also to softwood species in other regions as well. Original source of article obtained from {<http://eesc.oregonstate.edu/agcomwebfile/edmat/pnw576.pdf>} Also available from OSU Extension & Station Publications, Oregon State University, 422 Kerr Administration, Corvallis, OR 97331, "Tree Growth, Forest Management and Their Implications for Wood Quality" (PNW 576).*

## Best practices for saving files (file ID naming)

How you organize and name your files will have a big impact on your ability to find those files later and to understand what they contain. You should be **consistent and descriptive** in naming and organizing files so that it is obvious where to find specific data and what the files contain.

It's a good idea to set up a clear directory structure that includes information like the project title, a date, and some type of unique identifier. Individual directories may be set up by date, researcher, experimental run, or whatever makes sense for you and your research.

File names should allow you to identify a precise experiment from the name. Choose a format for naming your files and use it consistently.

You might consider including some of the following information in your file names, but you can include any information that will allow you to distinguish your files from one another.

- Project or experiment name or acronym
- Location/spatial coordinates
- Researcher name/initials
- Date or date range of experiment
- Type of data
- Conditions
- Version number of file

- Three-letter file extension for application-specific files

Another good idea is to include in the directory a readme.txt file that explains your naming format along with any abbreviations or codes you have used.

## Follow the file ID naming conventions

Use descriptive file names for easy identification and retrieval but don't go overboard - file/path names have length limits which vary between operating systems.

Under Windows the maximum full path length for a file (e.g. the drive letter + folder names + file name) is 260 characters. Use common abbreviations wherever possible.

## Be specific

Give files logical, specific names and include dates in file names if possible. The goal when naming files is to be able to tell what the file is about without having to open it and look. So, if the document is a letter to a customer reminding him that payment is overdue, call it something like "overdue\_20120115"; rather than something like "letter". How will you know who the letter is to without opening it?

If you are sharing files via email or portable devices you may want to have the file name include more specific information, since the folder information will not be included with the shared file. For example, if your document resides in My Documents\Invoices\2014\Customers\Doe\_John\_20140416.doc and the file is shared or emailed all the recipient will see is the Doe\_John\_20140416.doc and may not be able to tell that the file is a customer invoice without opening it.

Avoid extra-long folder names and complex hierarchical structures but use information-rich filenames instead.

Abbreviate the content of elements whenever possible.

## Other tips for file ID naming

- A good format for date designations is YYYYMMDD or YYMMDD. This format makes sure all of your files stay in chronological order, even over the span of many years.
- Try not to make file names too long, since long file names do not work well with all types of software.
- Special characters such as ~ ! @ # \$ % ^ & \* ( ) ` ; < > ? , [ ] { } ' " and | should be avoided.
- When using a sequential numbering system, using leading zeros for clarity and to make sure files sort in sequential order. For example, use "001, 002, ...010, 011 ... 100, 101, etc." instead of "1, 2, ...10, 11 ... 100, 101, etc."

- Do not use spaces. Some software will not recognize file names with spaces, and file names with spaces must be enclosed in quotes when using the command line. Other options include:
  - Underscores, e.g. file\_name.xxx
  - Dashes, e.g. file-name.xxx
  - No separation, e.g. filename.xxx
  - Camel case, where the first letter of each section of text is capitalized, e.g. FileName.xxx
  - **NOTE: Adding the FILE ID to the footer or bottom of the body is a helpful tip for finding exact document later.**

## Renaming files

You may already have a lot of data collected for your project and wish to organize and rename these files for easier data management. If you have too many files to rename them all by hand, try one of the following applications for renaming your files:

- [Bulk Rename Utility](#) (Windows, free)
- [Renamer 4](#) (Mac)
- [PSRenamer](#) (Linux, Mac, or Windows, free)

## File ID Naming Convention – CEPA Organization Solution Planning

Curation depends on the effort inside CEPA to institute file / folder naming and storage policies across the organization and to ensure success it is extremely important that such policies be dictated by the stakeholder with the highest authority. This individual must appreciate and instill to all participants that electronic files:

- are information assets that are the property of the organization.
- are shareable resources although sometimes for restricted users only.
- are critical components to operations that affect productivity and profitability.
- The goal of electronic file management is to ensure that you can find what you're looking for, even if you're looking for it years after its creation. Proper organization of electronic documents is especially critical in a shared environment - if one of your employees is absent (temporarily or permanently!) you should be able to easily locate any documents created or managed by that person.

### Execution

Once the file naming and storage policies are finalized, they should be documented and made available to all CEPA staff to follow. Everyone should be sensitized to the importance of diligently applying the naming policies and that deviating or compromising in policy adherence could have costly consequences to the organization. Therefore, it is incumbent on management that policy adherence be monitored as much as possible because there is always a natural tendency for participants to stray away mostly because of human factors.

## Content Has a State

Description of the private, public draft, and published 'state' of content on the CEPA Library portal.

Content, by default is created in a **private** state. Only the person who created it (**owner**), the users and groups that the owner chooses to share with, and portal managers can see it.

### To change the state of content:

1. To submit content for review, locate the **state dropdown menu**, and select **make visible**. This sets the content **state** to **public draft**. *Note that users who are not logged into the portal cannot see public draft content. Users who are logged in to the portal can see public draft content.*

2. To **publish** content, i.e. to make publicly available for viewing, locate the **state dropdown menu**, and select **publish**.

*Note that you WILL have permission to publish content in you're my folder, that you MAY have permission to publish content in a group that you are associated with, and that most likely you WILL NOT HAVE permission to publish content in the **shared library, training, links, and help folders** of the portal.*

3. *Note that content cannot be edited in the **published state**, except by portal managers. You must **retract** the content's state to public draft or private to edit the contents.*

To retract content, **navigate** to the content, located the **state dropdown menu**, and set the document **state** to **public draft** or **private**. If you have permission to edit the content, that is, you are the owner, an edit tab will become available, allowing you to modify the content.

## Sharing Private Content on the CEPA Library Portal

In the CEPA Library portal you may share private content with other portal members and groups.

### Sharing with Portal Users

To share private content with a portal member(s), locate the content, and click on its **sharing** tab.

1. Scroll down to **Add sharing permissions to users**.

2. Click the **Search by** dropdown menu to search by **User Name**. Type a user's name, and click the **perform search** button. If the user is found on the portal, they will be listed, with a checkmark placed next to their name.

3. Select a role to apply to the user or group to share with. Selecting the **owner** role will let

that user or group edit the content, selecting reviewer will let that user or group review and publish the content. Both may be selected by ctrl-clicking as many roles as you want.

4. Click the **assign local role to selected user(s)** button. The content is shared. At the top of the page you can see a listing of the share name, type, inherited role and local.

5. Repeat the procedure to add users.

## Sharing with Portal Groups

1. To share content with portal group(s), locate the content, and click on its **sharing** tab.

2. Scroll down to **Add sharing permissions to groups**.

3. Note that all portal groups are listed. Place a checkmark beside the group(s) that you want to share the content with.

4. Scroll down to **Role to assign** and click or ctrl+click to select one or more **roles**, and click the **assign local role to selected user(s)** button. The content is shared at the role level(s) you have selected. At the top of the page you can see a listing of the share name, type, inherited role and local.

## Glossary of CEPA Library Portal Terms

Description of terms used in the portal

**Portal** - A website [link to website] that combines a broad range of information for the visitors of that website. The home page of a website is often described as such because it serves as a launching pad to the variety of information contained within the website.

**Portlet** - A sub-portion of an entire web page. Derived from the word "Portal" [link to definition of Portal]. As in, many "Portlets" make up the entirety of the "Portal." Common "Portlets" in Plone [link to definition of Plone] could be dynamic content [link to dynamic content] (i.e. an automatically updated "Calendar", a "Recent Updates" listing, or a "Random photo") or could be static content [link to static content] (i.e. a very seldom changing "Mission" statement "Portlet"). "Portlets" can appear either in the main body of the page or in a right or left column of any or all sections of your website

**Dynamic content** - Content that is auto-generated, rather than maintained by hand, based on some predefined behaviors for the website. An example of dynamic content might be an auto-updating site map that displays all content available on a website, so that content editors need not update this manually each time a new page or article is added to the website.

**Static content** - Content that is maintained by hand and will be display exactly as the user has entered it, rather than being enhanced by some auto-generated text. A page listing an organization's mission is most likely to be static content.

**RSS** - (Rich Site Summary or RDF Site Summary or Real Simple Syndication) A commonly used protocol for syndication and sharing of content, originally developed to facilitate the syndication of news articles, now widely used to share the contents of blogs. RSS is an XML-based summary of a web site, usually used for syndication and other kinds of content-sharing. There are RSS "feeds" which are sources of RSS information about web sites, and RSS "readers" which read RSS feeds and display their content to users.

**Anchor** – Allows quick jumps down a page from Index or Table of Contents

**Mailto Link** - A link that opens a new mail message in an email program

**Landing Page** The Page you arrive on when clicking on Tabs in the site navigation

**Breadcrumbs** - A series of links that provide a "trail" to the section of the site that you are currently viewing.

**CSS (Cascading Style Sheet)** - A standard for specifying the appearance of text and other elements. CSS is typically used to provide a single "library" of styles that are used over and over throughout a large number of pages in a web site. A CSS file might specify that all numbered lists are to appear in italics. By changing that single specification, the look of a large number of documents can be changed.

**Text Editor** - A simple application that provides minimal support for text formatting. Like NotePad for Windows, and TextEditor for Macintosh.

**HTTP** Hypertext Transfer Protocol

**FTP** A method used to upload and download files form your site.

**Browser** A browser is an application used to view web pages.

**Cookie** A cookie is a temporary file used to store user preferences for a web site.

**Domain Name** - The unique name that identifies an Internet site. Domain Names always have 2 or more parts, separated by dots. The part on the left is the most specific, and the part on the right is the most general. A given machine may have more than one Domain Name but a given Domain Name points to only one machine. For example, the domain names for several well-known sites are:

- nytimes.com
- google.com
- npr.org

**Hyper-Text Markup Language (HTML)** HTML is markup language used to create web pages.

**FAQ** Frequently Asked Questions. A section of your site that will assist your visitors with the

common questions they may have about your site.

**JPEG/GIF/PNG** images **MB/KB** size of files

**PDF** Portal Document Format. A cross platform file type that allows you to distribute documents and ensure they display and print properly.

**Hyperlink** A link in a web page that points to another website or web page.

**Uniform Resource Locator (URL)** The address used to access a website or particular web page.

**Content Management System (CMS)** (from Wikipedia) A content management system (CMS) is a computer software system for organizing and facilitating collaborative creation of documents and other content. A content management system is frequently a web application used for managing websites and web content, though in many cases, content management systems require special client software for editing and constructing articles.

**Plone** Plone is a content management system. **TinyMC** The webeditor used by Plone.

**Site Map** A site map is a directory layout of your site. A well-constructed site map allows you to find information in the same way you would use the index of a book.

**Relative Links** A link that omits the domain name in the link address. You should only create this type of link for content within your site. This ensures that if you change domain names, you will not have to recode your links. Example: `<a href="/about/staff">staff</a>`

**Absolute Links** A link that included the domain in the link address. This type of link is not as flexible as a relative link and is best used for creating external links.

**Search Engine Optimization (SEO)** The practice of designing web pages so that they rank as high as possible in search results from *search engines*.

## Further reference:

Standard Rules for Content Curation: The Dublin Core:

<http://dublincore.org/documents/profile-guidelines/>

Using Pinterest:

<https://www.forbes.com/sites/georgeanders/2016/04/11/pinterest-wins-fans-in-academia-helping-users-analyze-tough-novels/#68da94b97568>

More technology and Ideas and definitions available at:

<http://www.matisse.net/files/glossary.html>