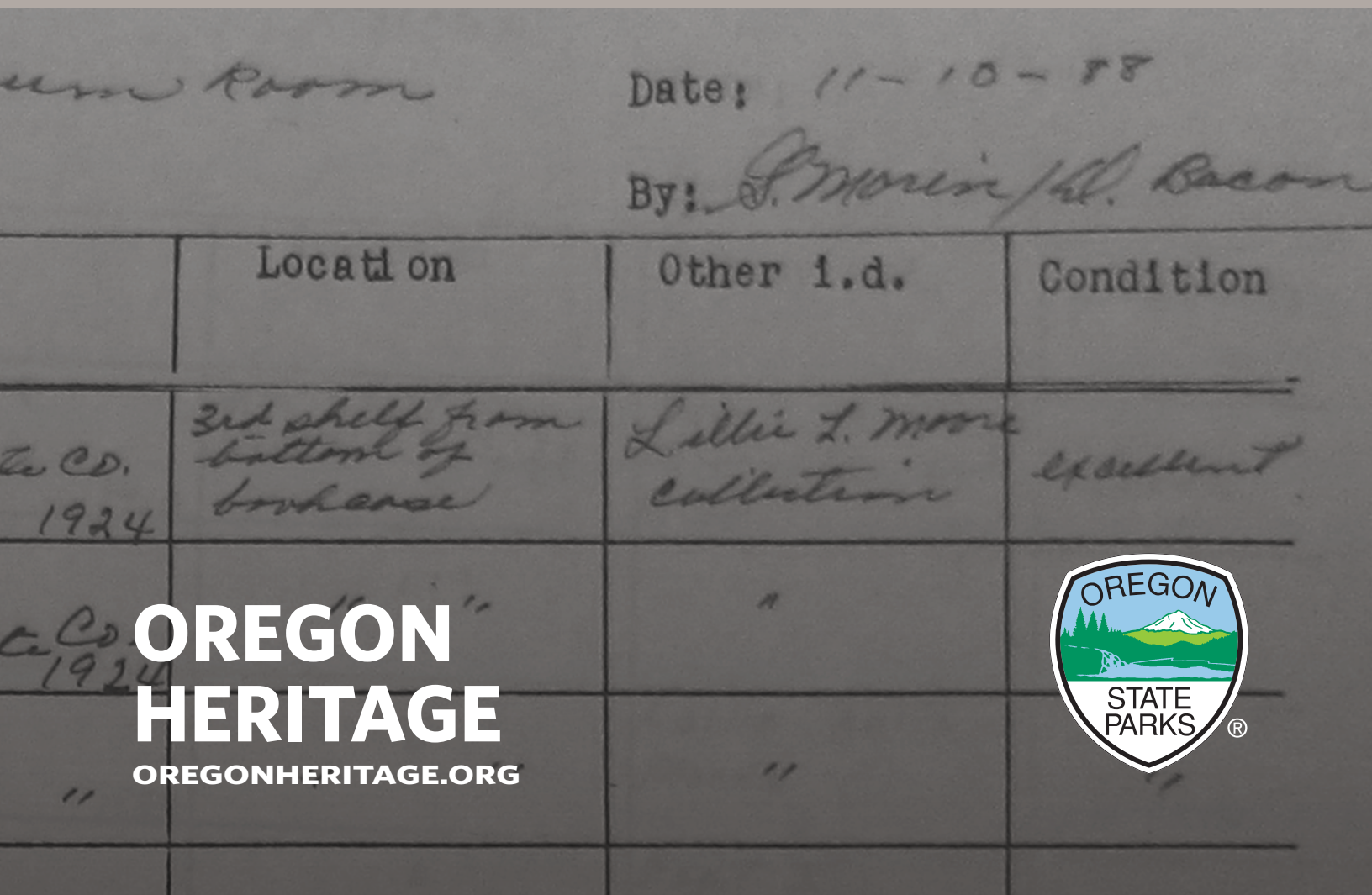
THE HERITAGE ORGANIZATION'S GUIDE TO CREATING AND CLEANING UP COLLECTION INFORMATION

2023 Version 1.0



OREGON HERITAGE

OREGONHERITAGE.ORG

EDITORS NOTE:

This guide is a living document. As feedback is received from users of this guide, information and tools will be changed and added in order to clarify the process even further. Every user will have a different level of knowledge and a different way of receiving information and it is our goal to continually adapt this guide to make it as accessible as possible.

If you have a printed version of this guide, please check the online guide to make sure you have the most recent version. The version is indicated on the front cover. The online version of the guide can be found on the technical resources page at www.oregonheritage.org. Scanning this QR code with your device camera will take you directly to the guidebook on the website.

Oregon Heritage would like to thank the following for their involvement in the development of this guide:

Funding provided by:

CULTURAL TRUST

Content contributor:

Relicura™
Historical Research & Collections Management

Review, feedback, and editing:

Northwest Digital Heritage WA OR

INTRODUCTION

Heritage organizations come in all shapes and sizes with varying degrees of capacity. Some have paid staff, many are volunteer run, but something they all have in common is that there will always be more work than there are resources, such as people and time, to do that work. It is very easy to become overwhelmed, especially when consulting resources on best practices for caring for heritage collections.

If you are reading this guide, you are likely the caretaker or one of the caretakers of your heritage organization's collection. One of the responsibilities of being a caretaker is managing the information or data about the items in your organization's collection.

What do we mean by collection information or collection data? This is information about your collection items that you collect, manage, and use to record, identify, track, locate, interpret, care for, and document its significance. Throughout most of the guide, we will be using the term "data" when talking about information about the collection. Often times, organizations have a system for managing this information. This system can range anywhere from a card catalog, to a written or typed log, to a spreadsheet, or a software/cloud-based database.

BACKGROUND

At Oregon Heritage, one of our primary goals is to listen to heritage organizations, try and identify collective challenges, and provide technical assistance that supports heritage organizations where they are in trying to accomplish their work. Over the past several years during our conversations with heritage organizations, we have heard concerns related to creating a database of collection information and cleaning up a database of collection information that has changed hands over time and requests for guidance. These conversations were made even more relevant as Oregon Heritage partnered with the State Library of Oregon and Washington State Library on the digital access initiative, Northwest Digital Heritage. With more organizations showing interest in having their collections accessible online, the need for creating or cleaning up collection information and figuring out where to start in that process become even more apparent.

The purpose of this guide is to break down best practices and standards for creating collection data and provide an actionable path forward for your organization to improve data quality.

Compounding the issue of figuring out how to move these collection information related projects forward is the obstacle of the complexity of resources related to best practices and standards. It can be extremely difficult, even for staff and volunteers that have worked in this field for several years, to understand and apply best practices and standards to collection information.

This guide is a response to those concerns and as this guide is used and feedback is received, we will adjust it and add tools and resources to help further breakdown the information.

FINAL WORDS

We recognize that everyone has a different way of learning and absorbing information and a different comfort level with technical information, and despite our best effort to try and make this guide as accessible as possible, there will likely still be some moments where you will feel overwhelmed. If that is the case, there are other resources available to help you work through this information:

Oregon Heritage MentorCorps: This is a free technical assistance program. Organizations can apply for a mentor to help them with project planning and training. More information can be found at https://www.oregon.gov/oprd/OH/pages/mentorcorps.aspx.

Northwest Digital Heritage: An initiative that provides digitization and collection data advice and training. More information can be found at https://www.northwestdigitalheritage.org.

How to use this guide

If your organization is interested in creating a system to manage information (data) about your collection, you can use this guide to:

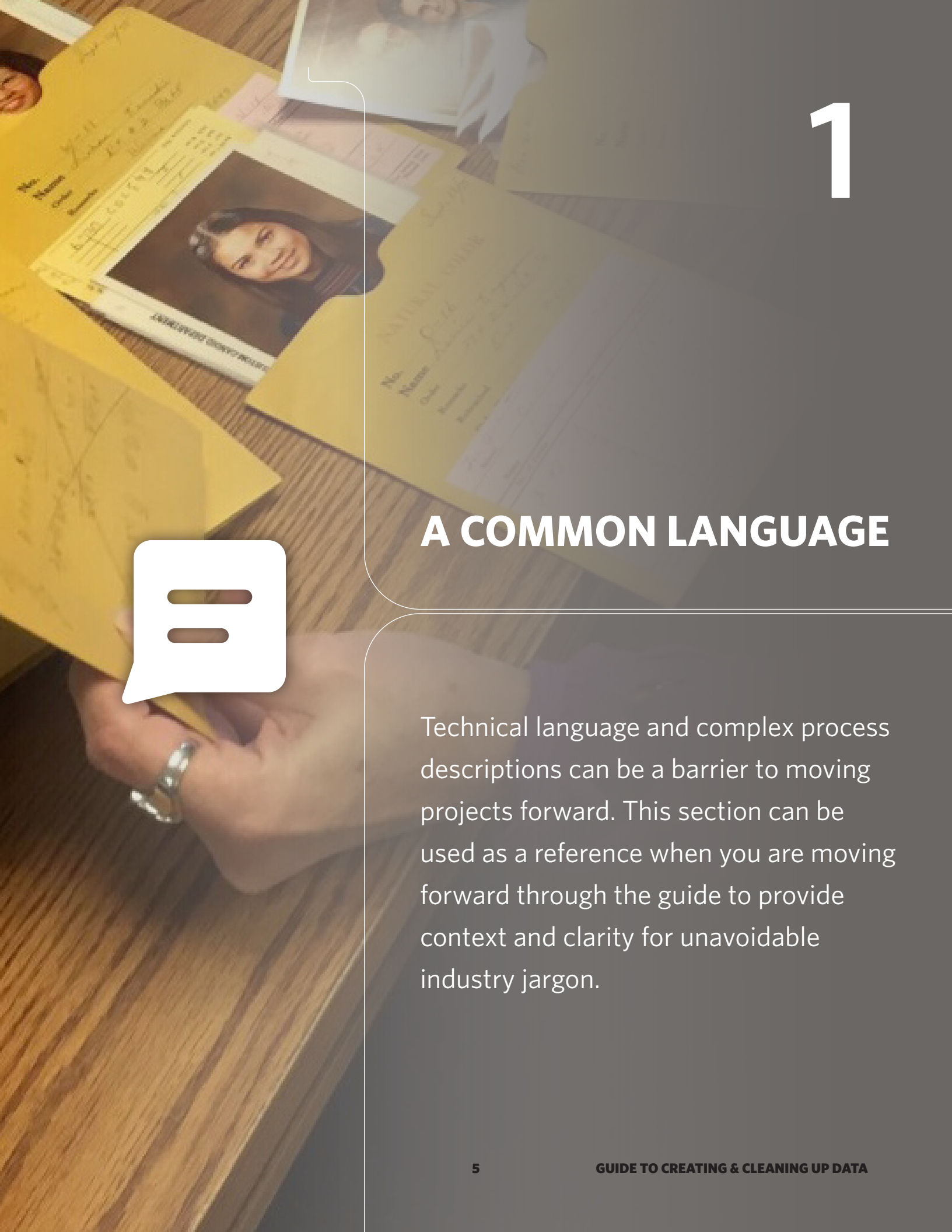
- Understand the importance of documenting consistent data about your collection.
- Make decisions on the structure of your system to create and manage collection data.
- Create a guide for your organization on how to put information into that structure.
- And if you are interested in choosing a collection management system, this guide can provide guidance on what to consider.

If your organization is already using a system, you can use this guide to:

- Understand the importance of documenting consistent information about your collection.
- Make decisions on the structure of your system to create and manage collection data.
- Create a guide for your organization on how to put information into that structure.
- Review, assess and prioritize your current collection information compared to your guide.
- Choose and implement a strategy for cleaning up your collection information
- And if you decide you want to change your data management system, this guide can provide guidance on what to consider.

CONTENTS

KEY TO BOXES	
	The box contains a relevant tool such as an application, software, or program.
	The box contains a tip that may help the process being discussed.
	The box contains a relevant resource related to the topic.
	The box directs readers to a worksheet in the appendix that may help organizations work through a part of the process.



1

A COMMON LANGUAGE

Technical language and complex process descriptions can be a barrier to moving projects forward. This section can be used as a reference when you are moving forward through the guide to provide context and clarity for unavoidable industry jargon.

ESTABLISHING A COMMON LANGUAGE

Every professional field comes with a set of industry standard language to facilitate understanding and discussion.

For the purposes of this guide, professional language will be used to help establish each topic and will shift to common language as much as possible.

Because Collection Data is the main focus, we will start with "data" and then move to alphabetical system for the rest of the industry terminology.

DATA:

Information about anything that is systematically gathered for the purposes of analysis and decision-making. For heritage organizations, information types can include data on donors, acquisitions and accessions, conservation treatments, exhibition planning and loans, reference requests, etc. This data is typically captured in the Collections Management System (see: Collections Management System) and placed in the corresponding acquisition, accession, condition, exhibition, donor, or descriptive catalog records. In this guide, "data" is used interchangeably with "collection information" and "collection data".

CATALOG RECORD:

A catalog record is a record created in a database (see: Collections Management System) that captures information (see: data) about an item, a group of items, or the collection as a whole.

COLLABORATIVE PORTAL:

Multiple heritage organizations that either use the same database (see: Database) or a database aggregator (see: Database Aggregator). Typically, collaborative portals pull together data hosted online from different heritage organizations using a metadata harvester and then present the aggregated data together. Collaborative portals are becoming increasingly popular as they serve as a powerful tool to search and view data across multiple collections. For example, Northwest Digital Heritage[1] is a collaborative portal.

COLLECTION:

A group of materials (in any format) that were created or collected together. A heritage organization can refer to its entire contents as a collection or reference a specific collection among its many collections.

COLLECTION ACCESS:

Broadly, this is how the general public or audience your organizations serves can search, view, or interact with collections or collection information.

[1] Northwest Digital Heritage, accessed March 7, 2023, via https://www.northwestdigitalheritage.org/.

This can be through exhibits, interpretation strategies, publications, direct requests, or online. Related specifically to collection data access, this is the ability to search, find, and view data about items in a collection, typically through a database.

COLLECTIONS CARE:

Collections care refers to the physical and digital care of collections. The care is focused on maintaining the overall health of items by tracking them and preserving them with appropriate storage, monitoring, and any related maintenance. When an organization accepts materials into their collection, they assume the legal and ethical responsibility to care for it in perpetuity as best they can. Part of this responsibility can be supported through functionality available in many Collections Management Systems (CMS) (see: Collections Management Systems).

COLLECTIONS CARETAKER:

A collections caretaker is an individual tasked with the care of all or a portion of the collection. Caretakers ideally possess knowledge and experience (formal and informal) on the specific standards, best practices, and techniques to best care for the collection based on capacity and resources. Caretakers can be staff, contractors, volunteers, and interns. They may hold the title of archivist, collections manager, curator, director, librarian, registrar, and more.

COLLECTIONS MANAGEMENT:

Management of a collection includes the collection's physical and intellectual control. The management of a collection typically includes creating and adhering to processes associated with collections such as accessioning (the process of accepting items into the permanent collection), documenting information about the collection items, maintaining care (see: Collections Care), and deaccessioning (the process of removing items from the collection). Capturing data about the collection is included in managing a collection. Populating collection data into a Collections Management System (see: Collections Management System) can aid in the caretaker's effectiveness in managing the collection and help keep track of the collection.

COLLECTIONS MANAGEMENT SYSTEM (CMS):

The CMS can be a software or web-based application used to enter, interact with, search, and display collection information, also known as data. The data in the CMS makes the CMS a database and helps collections caretakers make informed decisions related to the care and management of the collection. A CMS is critical to effective collections care and management since any heritage organization can have hundreds, thousands, or even millions of items. Commonly used collections management systems: Access to Memory (AtoM), CatalogIt, CONTENTdm, Omeka, and PastPerfect.

DATABASE:

A database is a computer system that captures and organizes data in a structured format. In addition to structured data, the database can support the relationships that exist among different data points. This allows users to search the database in a variety of ways for the purposes of tracking, searching, and interacting with the collection. For example: A search for "How many items were accessioned in 2019?" will pull all accessioned items with a 2019 date. A further search of "How many items were accessioned in 2019 and were included in our 2022 exhibit?" leverages the relationships that exist among the data and can deliver results for complex queries. Users can add and update data, transform data through activities the database supports (e.g., collection processing), and provide access to areas of data to a larger public (if desired).

DATABASE AGGREGATOR:

A tool that collects and compiles data or content from various sources (see: Collaborative Portal).

DATA FIELDS:

A database consists of records that contain fields. Every piece of data on a collection or object has a corresponding field to record in the database. The type of organization (archives, library, heritage organization, etc.) and the type of collection will dictate which fields are of the most use. For those using mechanisms such as a card catalog, written or typed inventory, or a spreadsheet, data fields would be categories of information you are collecting.

DATA EXPORT:

To export data from a database is to extract a copy of it and deliver it in a human and/or computer-readable format. Export of data from a database can be used to help evaluate the data or share it with a different database.

DATA SCHEMA:

The schema is the structure of a database, the order of information, and the relationships present. Heritage organizations tend to follow a specific data schema (such as Dublin Core) which provides a list of designated data fields to use with specific instructions on what types of data should live where in the database. Knowing the structure of your database may also be helpful when you are accepting items from donors and collecting information from those donating the items.

DISCOVERABILITY:

The ability to find information when searching for something. This relates to being able to track collection items through accurage collection information and when staff/volunteers/the public are searching for items related by type, subject, date, etc. Accurate collection information increases discoverability.

HERITAGE ORGANIZATION:

A term used to encompass various organizations that help preserve and share Oregon's culture, heritage, and history. A heritage organization is typically a nonprofit or government entity and can take the shape of an archive, community organization, museum, historical society, historical site, or library. Organizations can also be hybrids where more than one type is present. For example, an archive that is embedded with a library or museum; or a historical society with a library and special collections.

METADATA:

Is commonly described as "data about the data." Essentially, the fields used, their identification, data type, format, and description are all helpful information about the data that goes into the field. For the purposes of this guide, metadata is the catalog record (fields and data included).

For example, Title is a field name, and we enter an object title in the field. The field "Title," its identification, data type, format, and description—otherwise known as the field's structure—helps us understand the data we find within the field. In this case, the object title.

FIELD:	**TITLE**
The "Title" field is used to capture the title of an object. It's a free-text entry and it is a unique, non-numerical identification of the item.	

TITLE:	**VICTORIAN TEA SET**
The metadata provided for the field helps us to know that the data entered for the title "Victorian Tea Set," is a unique and non-numerical identification of the object being described.	

METADATA HARVESTING:

The act of extracting copies of metadata from virtually any database type. Typically, this requires a computer application (i.e., a tool) to collect the specified data and metadata.

OPEN ARCHIVES INITIATIVE PROTOCOL FOR METADATA HARVESTING:

An established protocol to extract a copy of data from one database to place it in another, typically to incorporate multiple systems' worth of data in one place.

SPECIFICATIONS:

A set of requirements regarding database functionality.

2

THE POWER OF GOOD DATA

The information we have about the items in our collections is just as important as the items themselves. Doing the best you can to have complete and consistent information about each collection item can help the organization meet its mission.

THE POWER OF GOOD DATA

Information (data) about an object is what provides context for its significance.

Gathering information and creating data to pair with the object or item helps support and communicate its significance. An object displayed without information runs the risk of being difficult to interpret or loses its meaning and any associated significance and is a missed opportunity for your organization and those accessing your collection.

Data intake helps us understand purpose, meaning, and significance. Knowing who created or used the object, when, how, and why, offers a more comprehensive understanding of the item. While any object can be believed to be important, a clear catalog record is necessary to illustrate that significance. The more information that's recorded and assigned to the object, the more accurately the object can be interpreted and understood.

Which object conveys more meaning?

With, or without data?

Look at this photo above. We can make intuitive guesses about the meaning of their positions, the approximate years the photograph may have been taken, etc. However, it isn't until we view the information (data) about the item that we know the setting is a tea ceremony taught at the Hood River Japanese Community Hall. The data tells us who was there, where they were, and what they were doing. These details would be hard to know if they weren't presented with the item.

Densho Digital Repository

Home Browse Names About densho.org

Home / Collections / ddr-densho-259 / 135

Search...

Tea ceremony demonstration

Download large (160.8 KB) Download full-size (3.8 MB)

Free re-use available This object is protected by copyright, but the rights holder has allowed us to make it available to you for non-commercial, educational projects.
Learn more...

PARTNER
Densho
Visit partner

OBJECT ID
ddr-densho-259-135

PARENT COLLECTION
Yasui Family Collection

DESCRIPTION
Caption by Homer Yasui: "[Shidzuyo Yasui] teaching a tea ceremony class at the [Hood River] Japanese Community Hall. Most of the women are dressed in kimono, but all of them probably had been taking lessons."

DATE

LOCATION
Hood River, Oregon

OBJECT GENRE
Photographs

OBJECT FORMAT
Still Image

TOPICS
Community activities -- Festivals, celebrations, and holidays
Arts and literature

FACILITY

CREATORS

PERSONS/ORGANIZATIONS
Yasui, Shidzuyo
Kajita, Mary
Tamura, Alice
Takasumi, Chieko
Iwasa, Haruye
Fujimoto, Matsuyo
Yamaki, Hama
Nakamoto, Kane
Tanimoto, Teru
Yasui, Shizue
Morita, Ruth
Hirasawa, Mary
Inouye, Matsue
Munemoto, Kosumi
Iwatsuki, Shizue

CONTRIBUTOR
Densho

PREFERRED CITATION
Courtesy of the Yasui Family Collection

RIGHTS
© Copyright restricted

Twelve ways good data can powerfully assist heritage organizations.

Good data can be leveraged in many ways, making any effort related to creating and cleaning up data worth it. But it can take time and resources so knowing the various ways good data can support your organization's work can help justify the effort. In addition to supporting staff in collections care and management, here's how good data can assist heritage organizations:

#1 OBJECT INTAKE:

Information (data) captured when the organization first takes ownership of an object helps heritage organizations to fulfill their responsibilities as caretakers of that object.

#2 ACCOUNTABILITY:

Accurate data capture helps with accountability regarding the object's ethical acquisition, care, location, and management.

#3 GAPS IN THE HISTORICAL NARRATIVE AND PAST HARM:

Data can help fill gaps in the historical narrative, acknowledge harmful past practices, and correct past misrepresentations and effects from the dominant culture perspective.

#4 LEGAL & ETHICAL QUESTIONS:

Data specific to an object's provenance will support future inquiries and transparency regarding legal or ethical questions specific to object acquisition, treatment, and access.

#5 CONSERVATION:

Data can help highlight items in need of treatment before the next exhibit by checking conservation reports or when prioritizing conservation needs.

#6 EXHIBITION USE:

Data can demonstrate how often collection objects are on display or not by chronicling exhibition history; this can inform audience interest and exhibit creation.

#7 COPYRIGHT AND INTELLECTUAL PROPERTY PROTECTION:

Data helps proactively protect any copyright or intellectual property claims. Or designate which items are available for use under Creative Commons Licenses[2]. Tracking this protects the organization related to restrictions and also provides opportunity for access and use to less restricted items.

#8 ENGAGEMENT:

Data supports the public's enjoyment of, education from, and engagement with objects.

#9 DISCOVERABILITY:

Consistent data improves searches and creates a greater possibility for the discovery of additional information.

#10 CONTENT CREATION:

Data encourages the incorporation of objects in research, educational initiatives, general content creation, and programmatic content. It also allows relationships to be identified to enhance interpretation.

#11 MITIGATES RISK OF LOSS:

Data decreases the risk of lost or misplaced objects by recording locations, moves, inventory, and auditing.

#12 PROOF OF VALUE (USE):

Data can prove collection use by tracking statistics of in-person use and digital viewership and determining traffic and demand. This information can be used as proof of value when advocating for additional resources internally and when engaging community partners, potential funders, and local decision makers.

#13 DISASTER RESILIENCE:

Nothing can replace the original items but should there be a disaster and those original items are no longer viable, the data of those items still persist and their significance can still be shared and continue to add context to the historical record.

Sharing the Value of Heritage Toolkit

Talking about the value of your organization's work is essential in order to engage the community, funders, and decision makers.

The Oregon Heritage Commission has committed to gathering data and creating messaging tools to help organizations talk about their value.

You can find this toolkit by visiting:

oregon.gov/oprd/OH/pages/tools.aspx

Assuming responsibility for accepting items into your organization's collection should be a careful, structured decision given the legal and ethical duties your organization assumes as caretakers. Consider reviewing the following resources related to ethics and duties of heritage organizations preserving objects and archives:

American Alliance of Museums (AAM)

- **Core Standards for Museums**: https://www.aam-us.org/programs/ethics-standards-and-professional-practices/core-standards-for-museums/
- **Code of Ethics for Museums**: https://www.aam-us.org/programs/ethics-standards-and-professional-practices/code-of-ethics-for-museums/

Society for American Archivists (SAA) Core Value Statement and Code of Ethics: https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics

American Association for State and Local History (AASLH) Statement of Standards & Ethic: https://learn.aaslh.org/products/aaslh-statement-of-standards-and-ethics

[2] Creative Commons, About The Licenses, accessed March 7, 2023, via https://creativecommons.org/licenses/.

Collection Management Systems versus spreadsheets.

A NOTE ON DATA IN COLLECTIONS MANAGEMENT SYSTEMS

Collection data is typically captured in a Collections Management System (CMS). The CMS is a database where collection caretakers can add, update, interact with, and review data on the collections. It was initially created to replace the card catalog, where limited object information was constrained to a 3x5 index card and could only be retrieved in person. A CMS isn't the size of an index card which means more information such as provenance, description, conservation reports, and exhibition histories are also documented in a CMS. However, the CMS is much more than just the collection information it maintains, it's a powerful tool that helps collection caretakers with their responsibility to care for and manage the collection to effectively meet the heritage organization's mission. Entering data into a CMS helps to maximize the power of good data.

Generally speaking, heritage organizations benefit from a centralized place where staff and volunteers can instantly access current collection information. A CMS fulfills this need as it can centralize a great amount of object information captured in a standardized format. Depending on the CMS and how well it's been implemented, staff can leverage the automated information a CMS provides to improve the care and management of the heritage organization collection. Information derived from the CMS can assist staff in making critical operational decisions and influence resource allocation.

BUT POSSESSING A CMS ISN'T NECESSARY

A CMS is not a requirement in order to create and maintain good data. Many heritage organizations do not possess a CMS for a variety of reasons. If resources are limited and the complex collections management needs aren't there (e.g., museum loans, conservation work, lack of capacity, volunteer turnover, etc.), then a good process paired with the use of a spreadsheet is an excellent option for your organization. If the option is to capture collection data in a spreadsheet or no data capture at all, then the answer unequivocally is: use a spreadsheet! The Spreadsheet format[4] is excellent for beginning and maintaining data capture and is in an ideal format to transform and transfer into a CMS tool should your organization make that decision down the road. Using a spreadsheet does not discount participating in a collaborative portal so when we talk about types of access in the next section, keep this in mind.

[4] Specifically, a .CSV file format.

Here's an example of a spreadsheet template:

identifier	publisher	title	creator	date	type
This is a unique identifier field. This field is intended to capture a unique identification number that's tied to the object being described. If there's an accession numbering practice in place then use the assigned accession number. Otherwise, devise a numbering sequence that's incremental and long enough to service numbering the collection for the foreseeable future. For example: "0000101".	The name of the archives, heritage organization, museum, or historical society where the materials reside on a permanent basis.	This is the title or name of the object.	The name(s) of the people and/or organization responsible for creating the item being described.	The date span represented by the earliest and latest dates of creation for the item being described. If the date is unknown, review the item for any context clues that could indicate an era or date approximation.	The type of item or object classification.
19221220_ FireAdjustersofAstoria.jpg	Astoria Public Library	Fire Adjusters of the Astoria Fire	Frank Woodfield	December 20, 1922	image

format	description	language	subject	rights	relation
The physical medium or dimensions of the resource.	A description of the subjects represented within the item, the historical or cultural significance, the predominate formats of the materials used, and an overall sense of size. The following "Sparta, Oregon" description does an excellent job orienting the person to what's depicted in the image as well as provides information on the item itself.	Indicate (as appropriate) the language(s) found within the item.	Select subjects that help to describe the item and assist external audiences to find the item via browsing by subject. The subject area is intended to capture the subject matter of the content depicted by the image or in the document, as well as the format of the material. Please see the "Portland, Oregon" record on the next page for an example of how multiple subjects are used.	The rights statement specific to the item or collection and how external audiences may or may not use it. For example: copyright or Creative Commons licenses. Please	A link or citation of other collection items that are related to the item being described.
4 x 6 inches; black and white photograph	This is the back side of the photo of the fire adjusters. There is a list of 55 numbers and next to most of them is a name and/or their job title or insurance company. These numbers correspond with the numbers above each man's head in the previous photo.	en	Astoria, Astoria Public Library, Fire Adjuster, Insurance Adjuster	This image is provided under the Creative Commons Attribution-NonCommercial CC BY-NC license.	

If you are currently using or considering using a spreadsheet to manage your collection information, there are tools within many of the spreadsheet applications to help control what data is being entered. As you progress in this guide and make decisions on what data you will be entering for each field, investigate your spreadsheet tool (whether it's Excel, Google Sheets, Smartsheets, etc.) to see if there are opportunities for creating drop down menus to select from a controlled list for some fields, date formatting, and other ways to put some controls in place to achieve the goals of consistent, accurate collection information.

Let's have a conversation about access generally, and digital access specifically

Broadly speaking, an organization provides access to their collection in many ways: exhibits, interpretation, tours, special events, publications, communications, research, online, etc. Providing opportunities for the people you serve to view, interact with, and learn from your collection is considered access. For heritage organizations that are nonprofits or publicly owned and operated, allowing access to the collection is considered part of it's duty of public service.

Having a conversation about what kind of access your organization would like to provide is a best practice as it can help your organization focus time and resources on effective ways to provide whatever level of access your organization chooses. If or when you decide to have that conversation, here are some industry standards and ethics to consider in that conversation:

AMERICAN ALLIANCE OF MUSEUMS (AAM)

- Core Standards for Museums: https://www.aam-us.org/programs/ethics-standards-and-professional-practices/core-standards-for-museums/
- Code of Ethics for Museums: https://www.aam-us.org/programs/ethics-standards-and-professional-practices/code-of-ethics-for-museums/

SOCIETY FOR AMERICAN ARCHIVISTS (SAA)

- Core Value Statement and Code of Ethics: https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics

AMERICAN ASSOCIATION FOR STATE AND LOCAL HISTORY (AASLH)

- Statement of Standards & Ethic: https://learn.aaslh.org/products/aaslh-statement-of-standards-and-ethics

Good collection data contributes to all forms of access to collections as it is used to create content for exhibits, interpretation, communication, etc. Direct access to the collection information (data) itself is the focus in the next section.

Collection Data Access

If your organization would like to consider widening opportunities for the public to access and interact with your collection information, digital access is a common consideration. There are many ways and mechanisms available to do this based on your organization's capacity and resources and below are just a few of them:

- Social Media
- Blogs
- Online photo galleries

- An on-site computer station for visitors to have a view-only opportunity of your database
- Online hosting of your collection database

If digital access is something your organization desires, there are resources available to help you decide on a tool based on your organization's capacity and resources:

- Talk to other organizations of a similar size and ask them what they do and if they have any lessons learned to pass on.
- Consider applying for a mentor through the Oregon Heritage MentorCorps to help you talk through your options and evaluate your capacity and resources. They can also connect you with similar organizations that may already use some of the tools mentioned.
- Do some research online to narrow it down or attend workshops and conferences on this topic.
- Consider hiring a consultant to help you make a digital access plan based on your current capacity and resources with recommendations for the future should capacity and resources increase.

If during your organization's conversation about access you decide you want the broadest digital access possible, then your organization should consider joining a collaborative portal such as Northwest Digital Heritage (https://www.northwestdigitalheritage.org/).

WHAT IS A COLLABORATIVE PORTAL?

A collaborative portal is a space that exists online with a unified front-facing web page that encourages public audience members to browse or search across collections from dozens (sometimes hundreds) of heritage institutions. This expansive searchability is made possible by harvesting collection information from the preexisting online hosted databases at participating heritage organizations. There are many benefits of contributing to a collaborative portal.

BENEFITS OF CONTRIBUTING TO A COLLABORATIVE PORTAL INCLUDE:

- It provides a seamless experience to public audience users.
- It increases the discoverability of the collections owned by the heritage organization.
- It facilitates the discovery of collection items with related objects in another organization's collection, allowing for new and deeper observations.
- Participation in a collaborative portal is typically very attractive to grant funders as it increases access to and exposure to the collections.
- Additional resources such as access to portal manager expertise, technical resources, training and guidelines, and increasing staff capacity are all possible.

Heritage organizations that contribute to collaborative portals typically already have a CMS in place that is hosted online, and it's the CMS that's harvested for collection data. However, depending on the portal partnerships, it may be possible to contribute collection data

via spreadsheet format for direct inclusion into the portal. This will depend upon how the portal is structured—specifically, whether it can accept data directly uploaded and stored in its system. Additionally, resource-protective policies may be in place for the portal manager that would preclude the direct uploading of collection data.

DATA PREPARATION REQUIRED FOR CONTRIBUTING TO A COLLABORATIVE PORTAL

If contributing to a collaborative portal is your organization's goal, then an evaluation of content eligibility and collection data will be necessary. All collaborative portals require data adheres to the adopted standards.[7] These standards encompass data schema (fields) and data content (what and how data appears in the fields) and will be further explored in the next chapter.

Northwest Digital Heritage and Digital Public Libraries of America are Collaborative Portals

The Northwest Digital Heritage (NWDH)[5] program is a collaborative portal supported by the Oregon Heritage Commission, the State Library of Oregon, and the Washington State Library. The focus for this collaborative portal is the collections from heritage organizations in Oregon and Washington. Heritage organizations can join one of two ways:

PLATFORM PARTNERS:

Organizations with digital collections hosted by NWDH instead of hosting their own CMS.

CONTENT PARTNERS:

Organizations with their own CMS and digital collections are harvested from the CMS by NWDH to contribute to the NWDH collaborative portal.

MEMBERSHIP BENEFITS:

- NWDH offers heritage organizations support in the form of guidelines and advice;
- The harvesting of collection data from the heritage organization's CMS for inclusion in the portal; and
- NWDH provides the collection data to an even larger collaborative portal: the Digital Public Library of America (DPLA).[6]

For more information on NWDH, membership benefits, and eligibility, please visit: northwestdigitalheritage.org

[5] Northwest Digital Heritage (NWDH), accessed February 1, 2023, via https://www.northwestdigitalheritage.org/.

[6] Digital Public Libraries of American (DPLA), accessed February 21, 2023, via https://dp.la/.

[7] If you're interested in learning more about NWDH's data standards please see the Washington Rural Heritage Metadata Guidelines, Version 3.1, August 2018, accessed April 6, 2023, via https://washingtonruralheritage.org/digital/collection/wrh/id/266. Guidelines prepared 2018 by Nikki Chiampa, Digital Projects Librarian, Washington State Library and Evan Robb, Digital Repository Librarian, Washington State Library.

Is participating in a Collaborative Portal right for your organization?

For most heritage organizations, the option to have the collections be more discoverable is a good thing. The more available a collection is, the more it is accessed and used. Consider the following prompts:

Who are the current collection users, and how do they currently access the collection?

Who are the potential collection users, and how might they want to access the collection?

For some heritage organizations, a hyper-local community base is already well served with digital collections hosted by the organization. Other heritage organizations may find that the collection serves local, statewide, and international audiences or want to promote their history to larger audiences and greater access or discoverability via a collaborative portal makes sense.

RISKS

If publishing collection content online (to be publicly available) and contributing to a collaborative portal (to be more broadly discoverable) is of interest, consider this next set of prompts to evaluate the potential risks and benefits:

Does the collection contain harmful depictions or descriptions of historically mistreated communities?

Does the collection contain content that could cause or trigger trauma?

Does the collection contain personally identifiable information, medical information, or financial information?

Does the collection contain materials it doesn't own or possesses items it doesn't have the copyright to?

Are the materials at risk of theft or vandalism if their existence and location are made more easily discoverable?[8]

If the answer is "Yes" to any of these questions, then contributing collections content to a collaborative portal has potential or perceived risk. Further evaluation and conversation within your organization and with others in the field can help you evaluate the risks and potentially find solutions. Organizations can elect which content they contribute to the portal, and one solution may be to craft criteria to help the organization determine which collection content is eligible for contributing to a collaborative portal.

Worksheet: Is a Collaborative Portal right for your organization? found in Appendix 3 will help you consider and evaluate your organization's goals of access versus potential or perceived risk of access.

[8] Please note: When organizations elect to contribute content to a collaborative portal it is usually the case that the collection data is already available publicly. However, when collection content joins a collaborative portal, it is more easily discoverable.

3

DATA STANDARDS OVERVIEW

Talk of standards can be overwhelming, but understanding and adopting standards and incorporating them into processes and procedures will benefit the collection and the organization currently and into the future. It ensures consistency through time and continued value toward the preservation of the collection.

Let's break down data standards a bit to better understand them

Standards are created because they help heritage organizations achieve best practices by creating consistency across the field. Establishing data best practices for the heritage organization collection has multiple benefits:

- It ensures consistent data (in quality and quantity) across all heritage organization object records.
- It empowers collection caretakers and allows them to feel confident in their data creation.
- It mitigates data errors that could spill over into research, exhibit, and educational programming endeavors.
- It saves collection caretakers' time spent on unnecessary data cleanup and protects against burnout.
- It increases how often the collection data is used because the data provided is substantive.

However, standards can be complicated and hard to implement especially when more than one set of standards are involved. Additionally, staff and volunteers working in heritage organizations have a wide spectrum of experience and training related to standards.

With that said, here are three levels of break down for standards:

- Baseline Level
- Medium Level
- High Level.

Start with the Baseline level and if that is sufficient enough for you, stop there and move on to the next section. If you decide you need a bit more context, proceed to the Medium Level and if you need even more, there is the High Level. Don't let this section derail your project and if you feel that happening, there are resources available through Oregon Heritage and Northwest Digital Heritage for someone to talk through this information with you and your organization.

Baseline Level

Consistent, complete, and accurate information is the goal. In order to achieve that, you need the following:

- A generally agreed upon set of fields that determine the minimum required information to record for every collection item.
- A generally agreed upon format for how that information is entered into those fields.

A template of minimum required fields, recommended format, and common issues to keep in mind during your review process starts on pg. 29. This table uses the Dublin Core data schema (the generally agreed upon set of fields) and DACS as the content descriptive standard for archives and CCO for museum artifacts (the generally agreed upon format for how that

information is entered into the fields).

If you are using a CMS, you may find some of these data fields go by different names. Check the instruction manual for the CMS or portal host to understand how the data field names may translate to match with the Dublin Core field names (the generally agreed upon set of fields i.e. schema).

If your are satisfied here using the table on pg. 29 as your guide for your data creation or cleanup, then proceed to the next chapter. If you'd like additional context, move on to the Medium Level.

Medium Level

Let's start by using two pieces of information from the Baseline Level and adding to them:

- A generally agreed upon set of fields that determines the minimum required information to record for every collection item is the **DATA SCHEMA**
- A generally agreed upon format for how that information is entered into those fields are **CONTENT STANDARDS**

DCMI: DUBLIN CORE™ METADATA ELEMENT SET IS THE MOST WIDELY ADOPTED SCHEMA

Cultural heritage organizations often contain collections that span the traditional library, archives, and museum collections. This is partly why the Dublin Core™ Metadata Element Set[9] is the most universally adopted, with its straightforward approach to capturing core data across all collection types. It's also the de-facto standard schema to support Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)—the ability to have collection data "harvested" and placed in a collaborative portal.

CONTENT STANDARDS

Content standards inform what content goes into the fields and how Two widely used content standards include:

- Describing Archives: A Content Standard (DACS)[10] is one of the most widely used for archival collections
- Cataloging Cultural Objects (CCO)[11] for museum artifacts. To inform what content goes into the catalog and how, please refer to either of these common content standards.

CROSSWALK: HOW THE STANDARDS RELATE TO EACH OTHER

The phrase "crosswalk data" comes up often in discussions of schema and standards. To crosswalk data is to relate individual standards to one another. The following table illustrates

[9] DCMI: Dublin Core™ Metadata Element Set, Version 1.1, accessed February 15, 2023, via https://www.dublincore.org/specifications/dublin-core/dces/.

[10] The Society of American Archivists, Describing Archives: A Content Standard (DACS), Version 2019.0.3. accessed February 28, 2023, via Link: https://files.archivists.org/pubs/DACS_2019.0.3_Version.pdf.

[11] Cataloging Cultural Objects (CCO), A Guide to Describing Cultural Works and Their Images, accessed August 4, 2023, via https://www.vraweb.org/cco

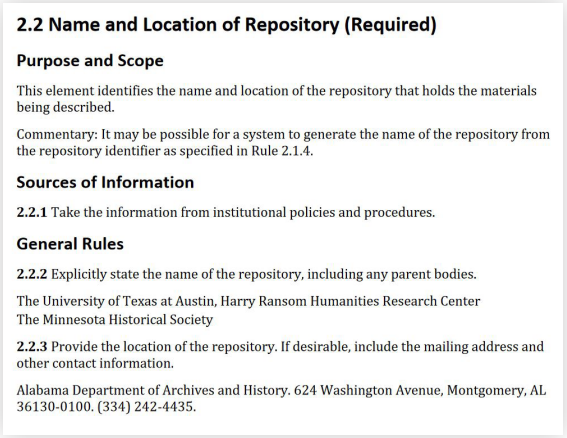
how the Dublin Core schema utilizes the data content standards (ingredients) from both DACS and CCO. This table also highlights how multiple content standards relate to one another—making it possible to catalog mixed collection types within one system. For example, a historical society that wants to catalog its museum artifacts in the same system as the archival collections. It also means that data from both archives and museums can be harvested and contributed to a collaborative portal filled with mixed collection types. As with any standard, both CCO and DACS suggest required and recommended (or "added value") content fields. This table captures the requirements for each with a few highly prioritized recommended areas.

CROSS WALK CHART: The numbers that follow each DACS or CCO standard are to indicate the chapter identification of that standard in the over standards guide. An in-depth guide using DACS with examples can be found in Appendix 2 – Describing Archives: A Content Standard (DACS) Reference.

DUBLIN CORE FIELDS	DACS CONTENT STANDARDS	CCO CONTENT STANDARDS
Identifier		
Publisher	Name and Location of Repository Element (2.2); Reference Code Element (2.1)	Repository Numbers (21.2.3)
Title	Title Element (2.3)	Title (3.1)
Creator	Name of Creator(s) Element (2.6); Administrative/Biographical History Element (2.7 added value)	Creator (4.1); Creator Role (4.1.1)
Date	Date Element (2.4)	Creation Date (4.2)
Type	Extent Element (2.5)	Work Type (1.2)
Description	Scope and Content Element (3.1)	Subject Matter (16)
	Conditions Governing Access Element (4.1)	
Language	Languages and Scripts of the Material Element (4.5)	
Subject	Access points (See Overview of Archival Description)	Subject (16.2); Classification (2.1)
Rights	Conditions Governing Reproductions and Use (4.4 added value)	
Relation	Related Archival Materials (6.3 added value)	
Format		Measurement (3.2.1); Materials and Techniques (3.2.2)

Here is a practical application example using that crosswalk chart:

You are the collection manager for a small museum and you are inputing data into your database . You use Dublin Core fields and are in the PUBLISHER field and you are using DACS for the Content Standards. Looking at the crosswalk chart, the PUBLISHER field corresponds to the "Name and Location Repository Element" in section 2.2 of the DACS standards. You would access those standards online, https://files.archivists.org/pubs/DACS_2019.0.3_Version.pdf, or have them printed out and look up section 2.2 and this image to the right is what you would find to help decide what and how to put information (data) in the PUBLISHER field.



2.2 Name and Location of Repository (Required)

Purpose and Scope

This element identifies the name and location of the repository that holds the materials being described.

Commentary: It may be possible for a system to generate the name of the repository from the repository identifier as specified in Rule 2.1.4.

Sources of Information

2.2.1 Take the information from institutional policies and procedures.

General Rules

2.2.2 Explicitly state the name of the repository, including any parent bodies.

The University of Texas at Austin, Harry Ransom Humanities Research Center
The Minnesota Historical Society

2.2.3 Provide the location of the repository. If desirable, include the mailing address and other contact information.

Alabama Department of Archives and History. 624 Washington Avenue, Montgomery, AL 36130-0100. (334) 242-4435.

Pg. 12 of the The Society of American Archivists, Describing Archives: A Content Standard (DACS), Version 2019.0.3, https://files.archivists.org/pubs/DACS_2019.0.3_Version.pdf.

If that is enough for now as it relates to the relationship of the standards, move on to the next section to learn more about fields and formats. For those that want further knowledge about Data Schema and Content Standards, move on to the High Level.

Are you using PastPerfect for your collection management system? Check out Appendix 7 for a chart that matches up PastPerfect fields with the crosswalk chart referenced in this section. This was part of a Case Study involving an organization cleaning up data in their PastPerfect database.

High Level

For a more thorough description of standards and their relationship with each other, visit **Appendix 1 – Data Schema and Content Standards.**

A guide to creating consistent data collection content.

Different Types of Data Field Formats

When thinking about data content, it's important to understand what format is expected for each field you intend to use and that the heritage organization data in each field match what best practices dictate. Knowing this information greatly informs data creation and cleanup. The following are the main data field formats:

FREE TEXT: Any type of data in the form of text or numbers.

- Alphabet-only (letters only)
- Numeral-only (numbers only)
- Alpha-numeral (both letters and numbers)

CALENDAR: Usually in a uniform version of MM/DD/YYYY; DD/MM/YYYY, etc.

FLAT VOCABULARY: One level of information that is not categorized or layered like hierarchical vocabulary (definition to follow). Flat vocabulary is linked to a controlled vocabulary in the system or a list developed by the organization and documented.

- **What is a controlled vocabulary?** A list of words and phrases used for the purpose of consistency, accuracy, and searchability.

HIERARCHICAL VOCABULARY: a structure of broader and narrower terms from an authority vocabulary source.

Nomenclature for Museum Cataloging found at https://page.nomenclature.info/ is a hierarchical vocabulary list and is useful for the TYPE field for physical objects to have consistent terminology for searchability. Below is an example from this resource illustrating the hierarchy of terms related to 'desk'.

All categories
▼ Category 02: Furnishings
▼ Furniture
▼ Storage & Display Furniture
▼ desk
campaign desk
computer desk
▶ drop-front desk
kneehole desk
partner's desk
quarter-cylinder desk
rolltop desk
school desk
secretary
standing desk
tambour desk
writing desk
▶ writing table

Accessed at https://page.nomenclature.info/parcourir-browse.app?id=1280&lang=en&ws=INT&wo=N

DROP-DOWN {SELECT TYPE}: Fields with predefined and selectable values, typically dictated by the CMS platform and can be changed based on the heritage organization's use.

REFERENCE FIELDS: Pulls and displays linked data from another area in the CMS.

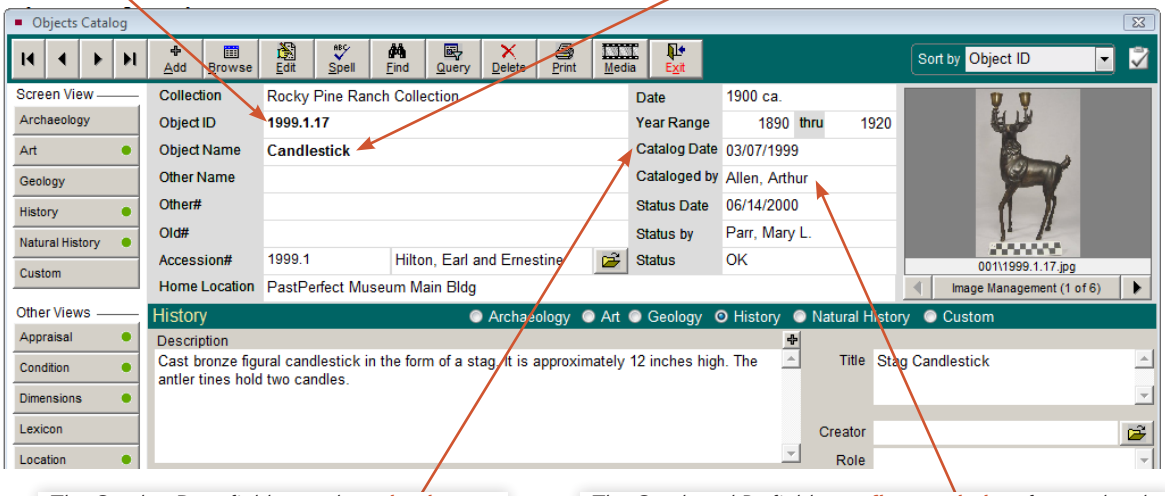
3

If your heritage organization uses a CMS, you can access a field guide that lists all available fields in the system and what data values they accept. This can be very helpful in instances where you're considering the uses of previously untapped fields.

Below is an example of a record from the PastPerfect CMS Version 5.0 with examples of the field format types indicated in this guide and using their Object Field Description Guide found here: https://museumsoftware.com/download/pp5fielddescrips/Objects.PDF

This Object ID field is the **free text alpha numeric** format type. It accepts letters and numbers depending on the numbering system used by the organization.

This Object Name field uses **hierarchical vocabulary**. PastPerfect uses Nomenclature 3.0 for Museum Cataloging as the authority list.



The Catalog Date field uses the **calendar** format type. Users of PastPerfect are able to select the format of this when setting up their database.

The Cataloged By field uses **flat vocabulary** for one level of information and PastPerfect allows organizations to create/edit User Authority Files, which are essentially **controlled vocabulary** lists

(Image from the PastPerfect 5.0 Object Catalogue Tutorial; https://museumsoftware.com/WebHelp/Chapters/PP5-6d.htm)

Fields, Format, Common Issues, & Tips to Avoid Common Issues

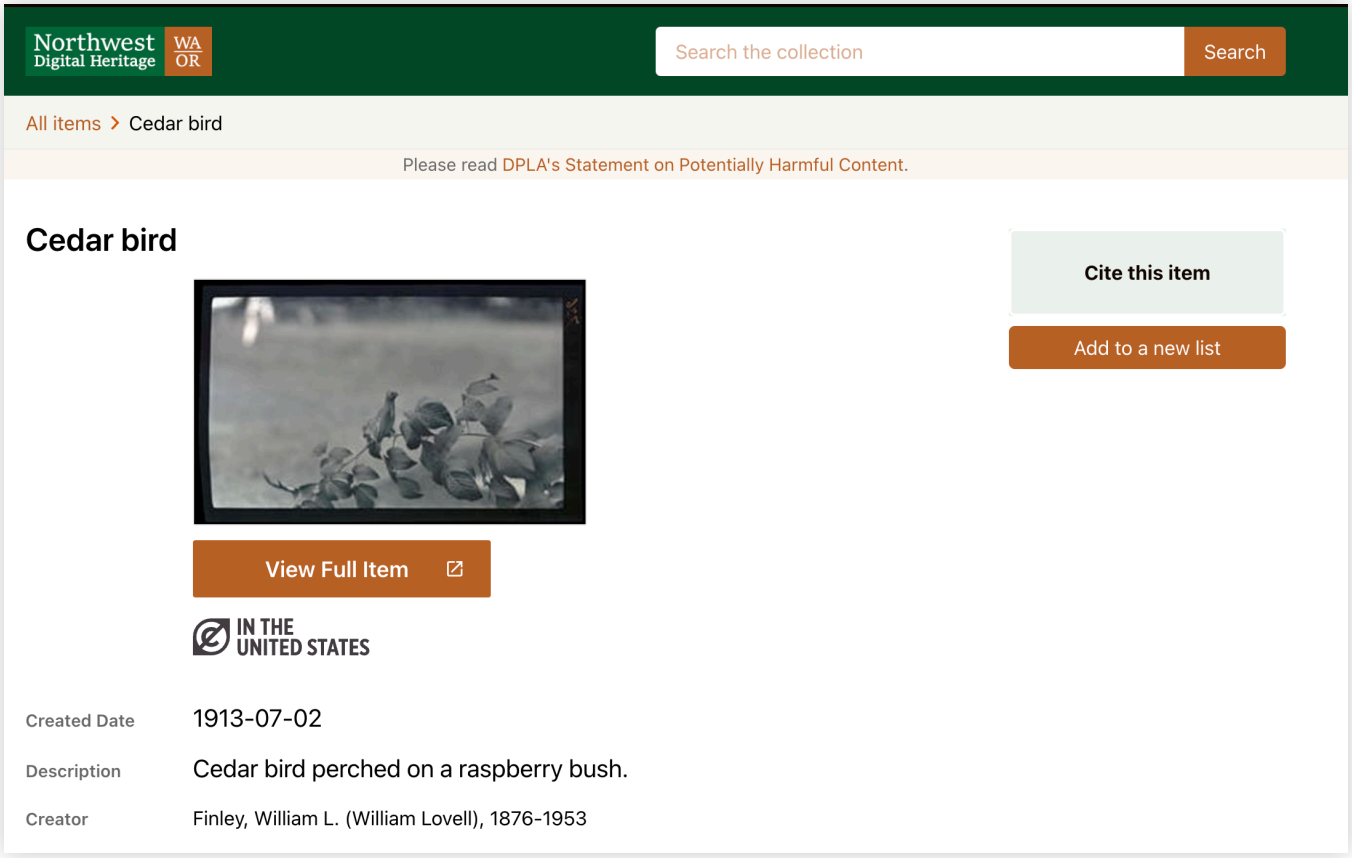
The following chart pairs Dublin Core data schema (the minimal field set as defined by Dublin Core) with DACS for the content standard. Depending on your CMS and/or the collaborative portal you are participating in, you may find that some of these data fields go by different names. Check the instruction manual of the CMS or portal host to understand how to crosswalk the fields. This Chart also identifies common issues that come up related to the field and format and some considerations to help avoid issues. The next chapter will help guide you in a process of creating a data standards guide for your organization. To prepare, use the following chart to determine the answer to this question: **What fields are you using and what are you putting in those fields?**

If your organization already has a guide, consider reviewing it alongside this chart to help you consider areas that you may need to add to or adjust for the sake of clarify and consistency.

DATA TYPE	FORMAT	COMMON ISSUES	TIPS FOR YOUR GUIDE
IDENTIFIER This is a unique identifier field tied to the described object/ material. If an accession numbering practice is in place, use the assigned accession number. Otherwise, devise a numbering sequence that's incremental and long enough to service numbering the collection for the foreseeable future. For example, "0000101."	Alpha-numeric or numeric-only.	Missing data, inaccurate or obsolete accession numbers.	Put a plan in place for when the accession number is missing. Will you first search for the physical object and see if the accession number exists but was not recorded in the database? Will you devise a numbering system for those records without an identifying number?
PUBLISHER The name of the archives, heritage organization, museum, or historical society where the materials reside on a permanent basis. When contributing to a collaborative portal, this element is sometimes called the "provider" or "data provider."	Alphabet-only; sometimes a controlled vocabulary.	Informal or shortened names for the organization.	Outline in the instructions what name should be used in this field to ensure consistency.

3

DATA TYPE	FORMAT	COMMON ISSUES	TIPS FOR YOUR GUIDE
TITLE This is the title or name of the object.	Alphabet-only; sometimes a vocabulary (in the instance of museum object names). Some repositories choose to use journalism-style titling, while others prefer simple titles, as seen in the "Cedar bird" example[12] below.	Inconsistent title style creation, inaccurate titles, titles that are descriptive sentences, vague titles.	Consider guidelines in the instructions to help avoid issues. Prompts like the ones below by medium might be beneficial: Who or what is the focus of the photograph? What event is being captured in the photograph? What is the object? Is it associated with a person or family?



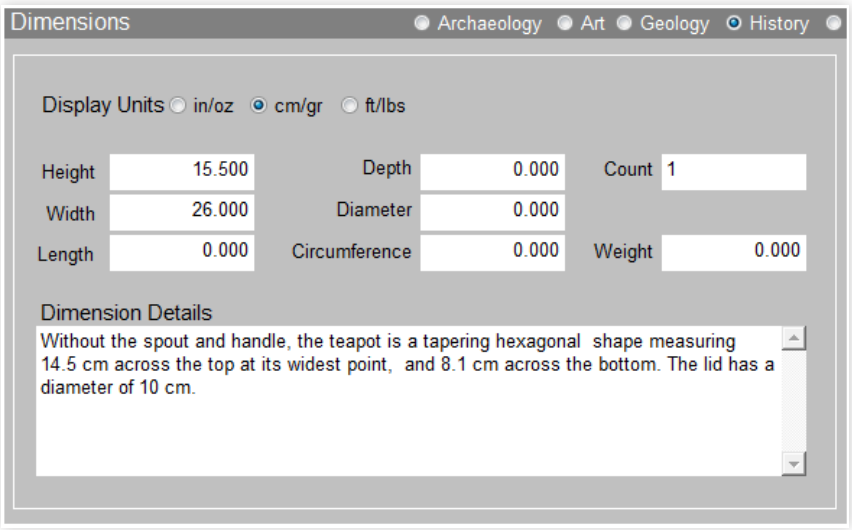
[12] Finley, William L. (William Lovell), 1876-1953. Cedar bird. 1913-07-02. Repository: Oregon Historical Society Research Library. Retrieved from the Digital Public Library of America, https://digitalcollections.ohs.org/cedar-bird-8. (Accessed March 16, 2023.)

AN OVERVIEW OF DATA STANDARDS continued

DATA TYPE	FORMAT	COMMON ISSUES	TIPS FOR YOUR GUIDE
CREATOR The name(s) of the people and/or organization responsible for creating the item being described.	Flat vocabulary. Depending on chosen descriptive standards, the creator's name can be Last Name, First Name, First Name, Last Name, or both, followed by the birth and death dates of the creator (if a person and if known). In the previous "Cedar bird" record example, the creator is entered as Last Name, First Name, Middle Initial. (First Name Last Name), birth date-death date.	Inconsistent or missing data. If the creator is unknown, the organization should determine if the field should be left blank or enter "Unknown".[13] inconsistent creator name can lead to multiple creator name variations (same creator, multiple name authority records. For example, "Jane Smith", "Smith, Jane", "Jane Smith", "Jayne Smyth" could all be created at one point in time though there is only one "Jane Smith).	This field can be tricky for objects that are notable for who owned them or time period rather than who created them. Make decisions about the use of unknown and in what circumstances it is used for each type of object/item.
DATE The date or date span represented by the earliest and latest dates of creation for the item being described. If the date is unknown, review the item for context clues indicating an era or date approximation.	Calendar. In the "Cedar bird" example, the date is entered in the following format: YYYY-MM-DD. Formats will vary depending upon both the content standards and what the CMS platform can handle. YYYY-MM-DD and MM-DD-YYYY are the most common. If using an approximate date, circa should precede the estimated date, for example, circa 1990s.	Missing data, especially when a "circa" date can be determined.	Provide in the instructions tips for dating unknown items to at least try and identify an era or time period. Be clear about how and when circa should be used.

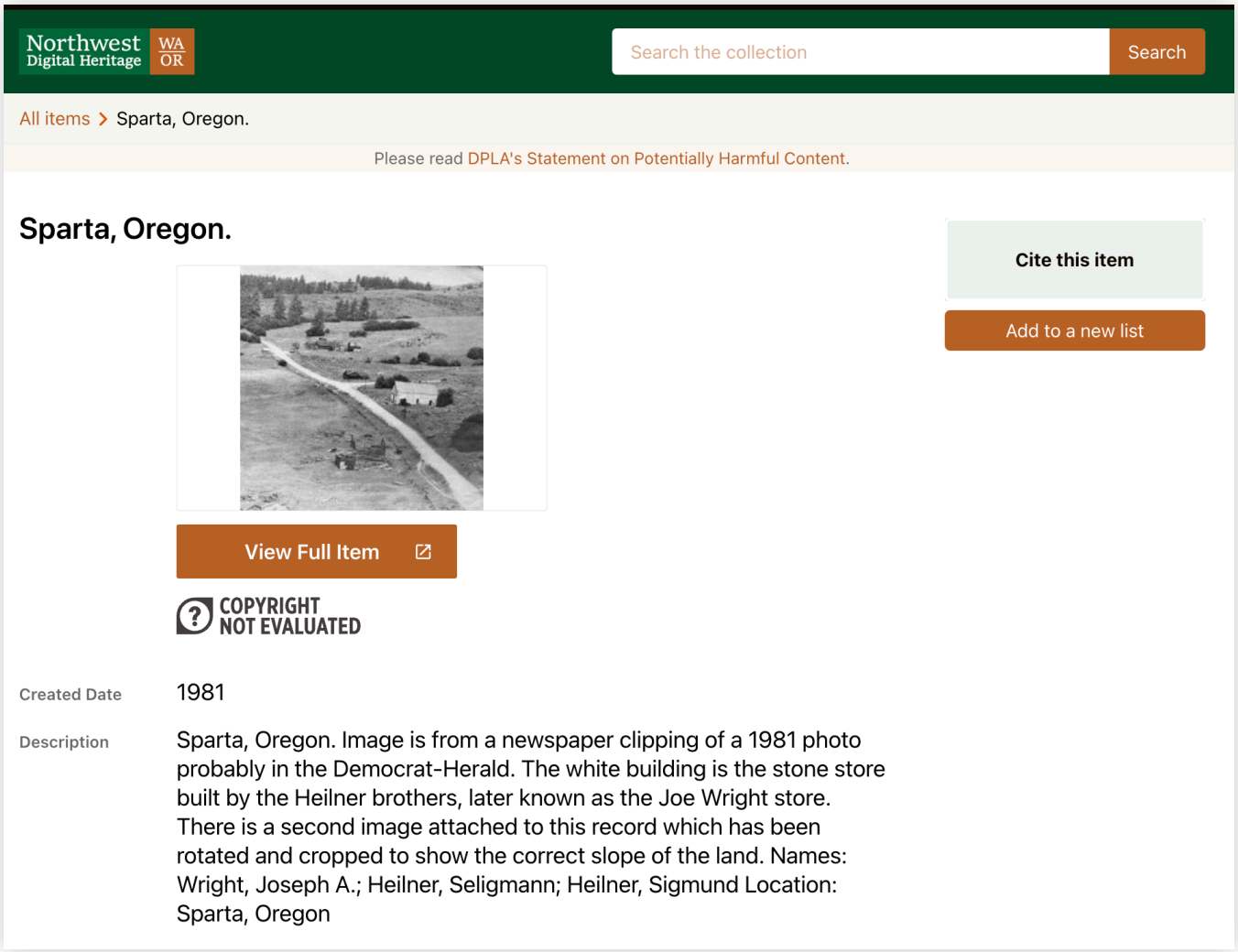
[13] Refer to the descriptive standards (DACS, CCO, etc.) your cultural heritage organization has elected to follow, or consult with your CMS vendor to decide what your local best practice will be.

DATA TYPE	FORMAT	COMMON ISSUES	TIPS FOR YOUR GUIDE
TYPE The type of item or object classification.	Hierarchical vocabulary.	No classification selected. A classification should be selected if it is obvious from viewing the object. If you don't know what the object it, you may need to ask other institutions or consider crowdsourcing.	If you use a CMS, it may come with a lexicon for this. Make sure your CMS system has controls in place to minimize additions to the type list without approval. If you are not using a CMS, Consider this free resource: Nomenclature for Museum Cataloging found at https://page.nomenclature.info/ If you don't know what the object it, you may need to ask other institutions or consider crowdsourcing.
FORMAT The format of the item being described, along with any dimensions. Below is an example of dimentions from PastPerfect 5.0.[14]	Vocabulary, drop-down value selection with numeral-only (for dimensions).	Missing data. At the very least, dimensions should be captured.	Depending on the database there may be multiple fields to fulfill this standard such as dimensions, material, technique.



[14] Dimensions subscreen for PastPerfect 5.0 accessed in the PastPerfect 5.0 Users' Guide (revision 3) on Aug. 30, 2023: https://museumsoftware.com/WebHelp/index.htm#t=Chapters%2FPP5-6f.htm

DATA TYPE	FORMAT	COMMON ISSUES	TIPS FOR YOUR GUIDE
DESCRIPTION A description of the subjects represented within the item, the historical or cultural significance, the predominant formats of the materials used, and an overall sense of size. The following "Sparta, Oregon"[15] orients the person to what's depicted in the image as well as providing information on the item itself.	Free text.	Missing, inaccurate, or incomplete data.	Consider providing a guide for what you want to capture in this field and how much is objective or interpretive. If you have objects and archives, have examples and a guide for the various types of items.



[15] Sparta, Oregon.. 1981. Retrieved from the Digital Public Library of America, https://bakerlib.pastperfectonline.com/photo/D195DF84-8430-4A7C-836C-576913673723. (Accessed March 18, 2023.)

DATA TYPE	FORMAT	COMMON ISSUES	TIPS FOR YOUR GUIDE
LANGUAGE OF MATERIALS Indicate (as appropriate) the language(s) found within the item.	Flat vocabulary.	Missing data. Even if it seems irrelevant, any viewable language should be captured.	
SUBJECT Select subjects that help to describe the item and assist external audiences in finding the item via browsing by subject. The subject area captures the subject matter of the content depicted by the image or in the document, as well as the format of the material. Please see the "Portland, Oregon"[16] records on the next page for an example of how multiple subjects are used.	Hierarchical vocabulary.	No subjects selected, or so few and broad subjects that they are minimally helpful in discovery searches.	Consider spending some time identifying subjects for this category. A resource to help you with a controlled vocabulary is the Library of Congress guide http://www.loc.gov/rr/print/tgm1
RIGHTS The rights statement specific to the item or collection and how external audiences may or may not use it. For example, copyright or Creative Commons licenses.[17]	Free text.	No rights statement provided. There is language for every imaginable rights situation, and a rights statement should always be provided.	Consider visiting RightStatements.org to find a standardized list of 12 rights statements to choose from for each record.
RELATION A link or citation of other collection items that are related to the item being described.	Reference field.	Missing links to data that exists inside the home museum and related peer museums. This is the most often skipped required field.	Brainstorm how relationships may span the collection and which items should indicate the other

[16] Christian (Photographer). Portland, Oregon. 1947?. Retrieved from the Digital Public Library of America, https://digitalcollections.ohs.org/portland-oregon. (Accessed March 18, 2023.)

[17] About Creative Commons Licenses, Creative Commons, accessed March 18, 2023, via https://creativecommons.org/about/cclicenses/.

Portland, Oregon

View Full Item

UNKNOWN RIGHTSHOLDER

Cite this item

Add to a new list

Created Date	1947?
Description	Negative of a photographic postcard showing a street scene in downtown Portland. Several businesses are visible, including the Congress Hotel, a Lincoln Mercury car dealership, and a battery shop. An automobile is parked in the left foreground and its license plate indicates the year 1947. There is a large tree on the sidewalk. Handwriting on negative reads, "Portland, Oregon. Christian 11-833."
Creator	Christian (Photographer)
Partner	Northwest Digital Heritage
Contributing Institution	Oregon Historical Society Research Library
Collection	Gilliam Portrait Studio and Camera Shop negatives
Publisher	Gilliam Portrait Studio and Camera Shop
Subjects	Black-and-white negatives Photographic postcards Buildings--Oregon--Portland Streets--Oregon--Portland Automobiles
Location	Portland (Or.) Oregon
Type	image
Format	Black-and-white negatives Photographic postcards
Language	English
URL	https://digitalcollections.ohs.org/portland-oregon
Standardized Rights Statement	**In Copyright - Rights-Holder(s) Unlocatable or Unidentifiable:** This Item is protected by copyright and/or related rights. However, for this Item, either (a) no rights-holder(s) have been identified or (b) one or more rights-holder(s) have been identified but none have been located. You are free to use this Item in any way that is permitted by the copyright and related rights legislation that applies to your use.

4

CODIFY, REVIEW, & ASSESS YOUR COLLECTION DATA

Let's create your organization's data standards guide and walk through how to review, assess, and prioritize what needs to be cleaned up.

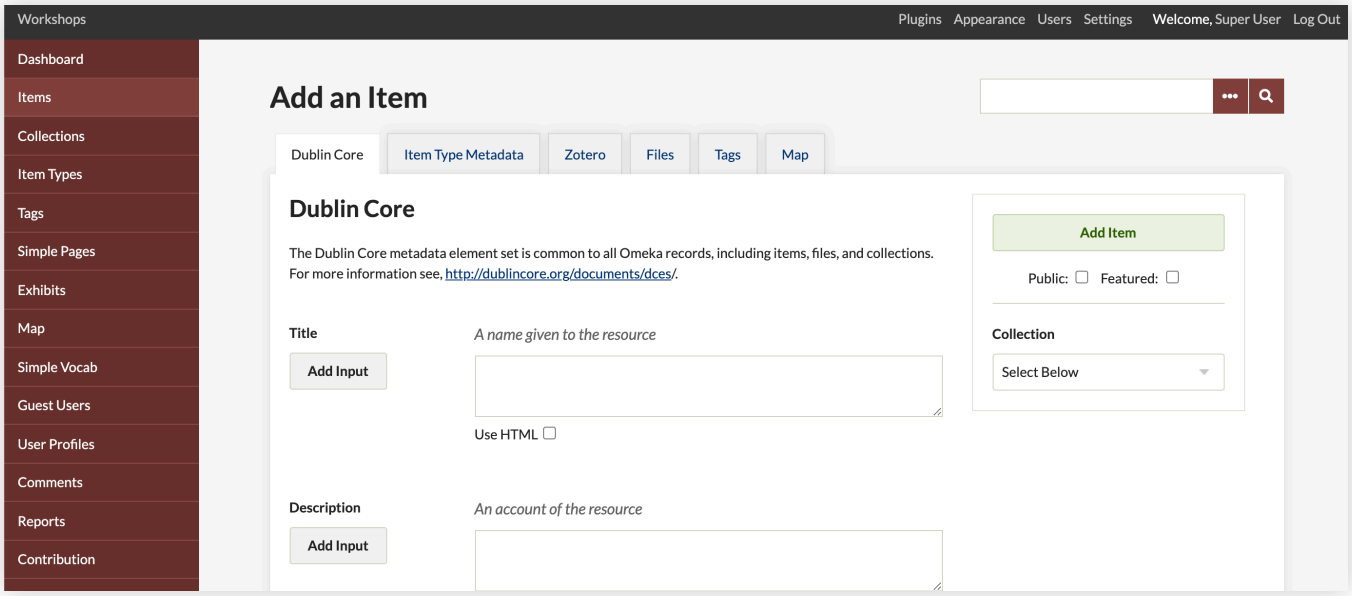
CODIFY, REVIEW, & ASSESS YOUR COLLECTION DATA

In order to review and assess, your organization needs a guide.

Creating a data standards guide

The first step to data review and assessment is to create a set of data standards specific to your heritage organization. Codifying a set of standards will enable you and your team to compare and contrast the current CMS data with standards the heritage organization wishes to follow. This will help identify gaps and inefficiencies in the current data and provide examples of how these issues can be addressed.

These steps are the foundation of data best practices for any heritage organization and will serve to empower your organization toward continued data excellence. If you find yourself stuck with what fields to choose or standards to follow, take a look at a peer heritage organization's CMS for inspiration. Additionally, don't forget to check out what catalog templates (informed by standards) come pre-loaded in your CMS.



[20] "Add an Item," Omeka Classic User Manual, accessed April 6, 2023, via https://omeka.org/classic/docs/Content/Items/.

Worksheet: Creating your organization's data standards guide in Appendix 4 expands on these instructions and provides different ways of documenting your data standards guide.

CREATE AND CODIFY A SET OF DATA STANDARDS

1. Select the data schema (fields) and content standard(s) (what goes into the fields) that make the most sense for your collection type and your organization.

Review Chapter 3 for guidance on this. You can either list the fields mentioned in the previous chart and how you want information to be entered into those fields, or you can make the decision to follow specific standards. For example, your organization may decide to use Dublin Core for its field standards and DACS as the content descriptive standard for archives and CCO for museum objects.

2. If you are using a CMS, determine what standards your CMS is using so you can crosswalk the fields and document that in your instructions.

Below is an exert of a chart from an organization using PastPerfect. The full chart can be found in Appendix 7 from the PastPerfect case study. It cross walks the Dublin Core field standards with the content standards (DACS and CCO) and then the corresponding PastPerfect CMS field.

Dublin Core Fields	DACS Content Standard	CCO Content Standard	PastPerfect
Identifier			Object ID; Other Number
Publisher	Reference Code Element (2.1); Name and Location of Repository Element (2.2)	Current Location (5.2.1)	Collection; Home Location
Title	Title Element (2.3)	Title (3.1)	Object Name; Title

3. Now that you have decided and documented which standards you are using (fields and what is going in the fields), you will now create instructions that capture the following for each field:

- Field Name, definition, and format
- Whether this field is required or recommended
- Whether there a controlled vocabulary or list for this field
- Instructions, standards, things to avoid for this field. Use the table in Chapter 3 to help with the instruction on this focusing on the common issues (things to avoid) and the column "Tips for Your Guide".
- Good examples for what to put in this field

4. Once you have your data standards guide in place, you can use it to create, cleanup, and continue care for your organization's collection data.

Reviewing your data

We've gone over data standards, you have an existing guide or have created a guide for your organization's data standards, now you will use that guide to review and assess. In order to do that, you need to get the data into a format that makes review easier.

It is easiest to view data collectively and a spreadsheet view is a good way to do this. This makes it easier to spot missing data, incorrect data formats, and any data inconsistencies.

- If your heritage organization is currently using a spreadsheet for data capture (and doesn't have a CMS), then you're ready to evaluate your data in its current form and can move on to Review and Assess.
- If you have a CMS, most CMS tools will allow exporting your data as a .CSV, which can be viewed as a spreadsheet. Most heritage organization collections have hundreds if not thousands (or tens of thousands) of objects, so it's best to export and analyze the data in chunks.

Some CMS tools can provide pinpointed queries and reporting that you can use to target areas of the data that you know are weak. For example, if you suspect that a field is missing or usually missing from the object records, you can query the CMS to report back with all missing records.

If you're not sure what your CMS is capable of then try the following:

- Reach out to the CMS vendor or refer to available CMS instructional documents.
- Reach out to a colleague who works with the same system to ask for advice.[21]
- Search for and view helpful tutorials, presentations, and trainings, that may be available online via YouTube or other platforms.

For those with an existing CMS platform (especially those that are online) there are a few formats data could be exported to, but the most universal form is the Common Separated Value (CSV) file. Here are the main three export formats:

- Comma Separated Value (CSV)
- JavaScript Object Notation (JSON)
- Extensible Markup Language (XML)

If you're not sure which export format your data was exported to, look to the file extension at the end of the file name. Here are examples: dataexport.csv; dataexport.json; dataexport.xml.

Assessment

The goal of data cleanup is to improve the data that currently describes collection items to

[21] Local heritage networks as well as online professional communities are a great place to crowd-source information. For example, reaching out to Oregon Heritage Commission's MentorCorps, Northwest Archivists, social media channels, etc.

be fuller, more accurate, consistent, and meet standards.

WHAT DO WE MEAN BY "QUALITY" DATA?

There are a few markers of quality data to consider:

- The data for each object is accurate and complete.
- The data is located in the appropriate fields of the CMS and in the correct format.
- There's data in each of the minimum requirement fields.
- The data is consistently used and formatted where appropriate (e.g., dates, locations, vocabularies).
- The data follows best practices set forth by the heritage organization field and any established internal practices specific to your heritage organization.

HERE'S WHAT TO LOOK FOR

Evaluating the heritage organization data can be overwhelming, especially for organizations with more than a few hundred objects. However, there's a short and reliable list of things to look for as you evaluate your data. With your organization's data standards guide that you just created, consider the following prompts as you review:

- Do all the required fields contain data?
- Is that data in the correct field?
- Is the data accurate and complete?

Capturing Assessment Findings

Create a spreadsheet dedicated to capturing your assessment findings. Consider the following column headers below:

ISSUE NUMBER	FIELD FOR CLEANUP	QUANTITY	PRIORITY	CLEANUP INSTRUCTIONS	ATTACHMENTS
As you assess your data, document each field type where you're noticing data gaps or deficiencies when comparing actual data to your codified data standards document. If a you discover more than one issue in a field mark each issue as a separate entry as it may be that you have to take different approaches with each issue.		Attempt to get an estimate of how many records are impacted by each issue. Quantifying the issue will help you determine what's achievable for the available capacity of your staff.	Determine the level of priority for each item to the best of your ability. Things to consider when prioritizing	Cleanup will depend on the issue, staff knowledge and capacity, and the CMS. See section 5: Data Cleanup Strategies: Application for more information.	If there were any queries, reports, or other documentation you used for data assessment, make a note of them here so that any staff member can see if there's related documentation needed for effective cleanup.

A Note on Cleanup Priorities

Before you embark on cleanup activities, and as part of the assessment process, it's important to contemplate how prioritization will be based. Priorities for what's tackled first in data cleanup can vary dependent upon the heritage organization's goals and may also take into account the scale of cleanup needed. The following are criteria to consider when assigning a priority level. It is possible for the organization to hold multiple priorities, but they should be ranked in order to be of use in dictating what cleanup efforts are performed first.

"DATA CLEANUP IN THIS AREA WOULD..."

- Improve search result accuracy
- Improve the searcher's ability to correctly determine relevancy of the result to their query.
- Improve audience understanding of the collection, subject, and/or community represented.
- Improve historical accuracy and representation of community partners.
- Improve the ability to participate in a collaborative portal.
- Improve the proper attribution of the items to the organization and adherence to any rights or use restrictions.
- Help to connect audiences to related materials through documented relationships.
- Improve data accuracy and the ability for the system to function well.
- Offer an "easy win" or fill-in short periods of downtime.

Worksheet: Collection Data Review, Assessment, & Prioritization in Appendix 5 provides different ways to document the issues you discover during your review and assessment process.

PastPerfect Case Study of a cleanup project can be found in Appendix 7.

This case study is a great resource to review for a play by play of assessment and prioritizing a data cleanup project.

Electronic versions of the results are available at www.oregonheritage.org. This may be especially beneficial if you'd like a template for the assessment review spreadsheet.

5

DATA CLEANUP STRATEGIES: APPLICATION

Cleaning up data can be very overwhelming so the goal is to break it down into small steps and accept that it will take time. Having a plan in place and giving yourself grace during the process is essential.

DATA CLEANUP STRATEGIES: APPLICATION

Before you begin, consider the resources available for the data cleanup project.

Some strategies will be "do-it-yourself" without system assistance, and others will leverage common core functionality within many Collections Management Systems (CMS). The emphasis will be on practical steps that don't require specialist knowledge.

It's important to consider the resources available to help scope the cleanup work appropriately. A data cleanup project that is too large to undertake can easily burn out staff and volunteers and lead to even more messy data which will just compound the problem in the long run.

PROJECT RESOURCE CONSIDERATIONS

- Number of Staff and Volunteers available for the project.
- How much time per week/month can be dedicated to the project.
- Amount of funding (if any) to access for outsourcing considerations

SUPPORT CLEANUP EFFICACY WITH BITE-SIZED WORK PLANS

Data cleanup is the most effective when it's performed with laser focus. For example, focusing on a subset of the collection, and even just one field[22] commonly used in every record, will help to parcel out the cleanup work into easily completable sections. This approach helps to keep the data cleanup team focused and allows for the inevitable breaks in cleanup to perform other activities.

TECHNICAL SKILLS AND TRAINING

For staff and volunteers involved in the project, evaluation of techinical skills and training are essential in order to make sure they are comfortable with the project and to increase success of the project:

- Are they knowledgeable in data content standards for your organization?
- Do they know how to use the CMS?
- Consider the knowledge and skills needed in order to perform data cleanup accurately. What are reasonable expectations?
- What documentation or instructions are needed? Or does assistance need to be brought in from an external source?

[22] To offer further specifics, you may choose to update all empty Creator fields with "Unknown".

5

There are four common strategies for data cleanup.

This section will delve deeper into the execution of each strategy listed and offer advice, resources, and examples. While each CMS is different, the overall approach and logistics to data cleanup are the same. We'll review the four most common strategies and depending on the CMS and the data issues present, you may find that only one strategy works best for your collection. Or that several of these strategies can be used based on the type of data cleaning you intend to undertake.

Spreadsheet Strategy

Record-by-Record Strategy

In-System, Find and Replace Strategy

CMS Vendor Supported Strategy

[23] Learn more or download OpenRefine via https://openrefine.org/.

[24] Screencasts, GitHub, accessed May 07, 2023, via https://github.com/OpenRefine/OpenRefine/wiki/Screencasts.

Spreadsheet Strategy

The spreadsheet strategy can provide a comprehensive view of all of your data. It makes it easy to spot records missing data in required fields and data that was entered incorrectly. In addition to reviewing data, the spreadsheet can facilitate any data edits needed. It's as simple as editing the spreadsheet directly. If you are using a spreadsheet to manage your data than you are already set up with this strategy. For those using a CMS, The majority of CMS platforms allow users to export their data into an editable format, typically .CSV. For example, this strategy can be used with Omeka and Access to Memory (via command line). Check your CMS documentation to see if data export is a supported function, and review the steps you need to take for that system.

POSSIBLE PITFALLS

This strategy isn't without its risk. When editing a spreadsheet, there are no controls that would prevent further mistakes from being made, such as misspellings, data in the wrong field, deletion of data accidentally, etc. This strategy relies on a strong review of the data in the spreadsheet BEFORE ingesting the data back into the system.

SPREADSHEET STRATEGY DATA CLEANUP STEPS:

1. Export data from your CMS into a .CSV (spreadsheet) file.
2. Evaluate the data for each record. In a spreadsheet, each row is a record entry.
3. Identify required fields that appear to have data missing, have incorrect data, such as data in the wrong field or entered incorrectly—for example, misspellings, incorrect date format, improper use of a controlled vocabulary term, etc.
5. Make a game plan for how to tackle the identified data issues. The most effective way is focusing on one field or a subsection of records.
6. Fill in data that is missing but fairly easy to create. For example, the Creator and Date fields shouldn't be blank if the standard for the heritage organization is to use "Unknown."
7. Correct the data that was identified as incorrect. Make a note of any records that will require future research. These records can be worked at a later stage in cleanup.
8. After the editing is complete, perform a review of all the changes made to ensure accuracy. You don't want to import new mistakes or errors into the system's data.
9. Import the record back into the system. Because you used a spreadsheet generated from the CMS, it should be able to recognize its data and what has changed and then apply those changes to the corresponding records.

OpenRefine[23] is an open-source tool that supports the cleaning and refining of data. This tool can be powerful once learned. Before diving in, we recommend viewing the GitHub provided screencasts[24] for instructions and to evaluate if this tool is a good fit. Please note that OpenRefine is only an option if your CMS can both export and import data via a spreadsheet file format.

Record-by-Record Strategy

The record-by-record strategy is a strategy to use if your CMS doesn't support data export into a spreadsheet and it doesn't support data cleanup tools. It can be a slow and painful strategy, so it's incredibly important to develop a data cleanup plan and break the cleanup into achievable sections. In order for this work to be sustainable, a doable pace will need to be established. This strategy can be used with collections management systems that don't allow the export of data or data cleanup tools. This is becoming rarer and rarer, but there are a few older systems out there, or if the tool is open source, then the ability to perform more robust cleanup work requires more technical knowledge than is available at your heritage organization.

POSSIBLE PITFALLS

The pitfalls for this record cleanup stem from the labor-intensive act of reviewing and performing cleanup on a record-by-record basis. The monotony paired with reviewing extensive information can be mentally exhausting and should only be performed in short sprints for the best results.

RECORD-BY-RECORD STRATEGY DATA CLEANUP STEPS:

1. Make a game plan for how to evaluate each record for data imperfections. The most effective way is typically focusing on a discrete subsection of records.
2. Evaluate the data for each record.
3. Identify required fields that appear to have data missing. Because this work has to happen record-by-record, you can make these corrections now. If a record requires more research, mark the record and the data needed so that you can return to it later.
4. Identify fields that appear to have incorrect data, such as data in the wrong field or entered incorrectly. For example, misspellings, incorrect date format, improper use of a controlled vocabulary term, etc. Because this work has to happen record-by-record, you can make these corrections while you're in the record.
5. Perform a review of the data corrected and added before you move on to the next record.

In-System, Find and Replace Strategy

The find and replace strategy is one to use if your CMS or locally grown database (e.g., FileMaker Pro) provides data cleanup tools such as find and replace (like global replace or bulk edit). Using in-system data cleanup tools can be a huge timesaver. For example, if multiple records are missing a "Creator" entry because the creator is unknown, then those records can be selected and updated on mass with "Unknown" to fill in that field. This strategy is best applied in universal update scenarios. For more intensive individual cleanup—such as providing fuller descriptions—the cleanup will have to be executed record-by-record or (if available) the spreadsheet strategy. This strategy can be used with any system that offers find and replace (or similar) functionality. Check your system documentation to see if this functionality is supported, and review the steps you need to take for that system.

POSSIBLE PITFALLS

While this approach is powerful and effective for bulk cleanup, it loses usefulness the moment the cleanup becomes more specific and individual. Additionally, the bulk change feature can create other problems if the query isn't specific enough and if the review processes are skipped before executing the change.

IN-SYSTEM, FIND AND REPLACE STRATEGY DATA CLEANUP STEPS:

1. Execute a review of your data and identify where universal data cleanup can be performed. This review can be done record by record, from a query or report, or via an export of data from the CMS or local database into a .CSV (spreadsheet) file.
2. Identify required fields that appear to have data missing. If the missing data can be filled with an "Unknown" (e.g., Creator or Date fields) or a set of universal data (e.g., Rights Statement field), then mark these as prime candidates for a bulk edit via find and replace.
3. Identify fields that appear to have incorrect data, such as data in the wrong field or entered incorrectly—for example, misspellings, incorrect date format, improper use of a controlled vocabulary term, etc. Misspellings and incorrect controlled vocabulary can be candidates for find and replace.
4. Make a game plan for how to tackle the identified data imperfections. For this strategy, each universal find and replace action will need to be predefined by crafting a specific query for the exact set of records you intend to update.
5. Fill in missing or incorrect data using the find and replace (i.e., bulk edit) tool. Double-check your query and your "find" and "replace" entries to ensure everything is correct BEFORE executing the change. Some CMS tools have an "undo" action, but not all of them do, nor do they necessarily undo all the harm inadvertently afflicted.

CMS Vendor Supported Strategy

This strategy is, in some ways, easier because you have a strong technical partner to lean on for data cleanup support. This option is only available to proprietary CMS products that a functional company actively services. Part of what you pay for with proprietary products is access to technical experts and tools to help you get the job done faster.

POSSIBLE PITFALLS

CMS vendors differ in their support and how "freely" they give it. Depending on the vendor, additional costs may be associated with their help.

CMS VENDOR SUPPORTED STRATEGY DATA CLEANUP STEPS:

1. Contact your CMS vendor and learn about the tools available to review and clean up your data. Also, learn what type of data cleanup they can help execute that may not be available to you. For example, some systems don't allow the export or import of data on your end, but the vendor can do it. This is the case for PastPerfect.
2. Make a game plan for how to evaluate records for data imperfections and classify which cleanup actions can be supported by the vendor.
3. Create a data cleanup plan that documents all the data that needs to be cleaned or updated.
4. Work with the vendor to identify the right records and review the cleanup needed.
5. Perform spot checks of the work to make sure no mistakes were made in the preparation or execution phase.

Continued care of your collection in important. If you are reading this guidebook, that means you are considering spending a lot of time cleaning up your data and you don't want to have to do this again. Cementing your data standards guide into your policies and procedures for when new data is entered into your system can help prevent future large scale collection data cleanup projects. Be aware of any approval processes within your organization related to policies and procedures. They may need to be reviewed and voted on by a committee or board.

6

NEW CMS CONSIDERATIONS

If this process of creating and cleaning up data has you considering new a new CMS, it's okay and even encouraged to constantly evaluate the tools you use and do a scan of what's out there. Knowing what collection information you are managing, how it is being managed, and what you want to do with it will help you determine the system best for your organization.

Kam Wah Chung State Heritage Site

NEW CMS CONSIDERATIONS

Thinking of a new CMS? Here is what you need to know...

Sometimes the process of data cleanup can cause heritage organizations to wonder if they are using the right system for their collection data. This section will provide an overview of Collections Management System (CMS) options, how to craft specifications, and how to use a decision-making matrix to support selecting the right CMS. The good news is that many CMS platforms are getting both easier to use and more affordable. While financial resources are always tight, many CMS tools can now be obtained for a few hundred to a few thousand dollars a year. With the right type of system in place paired with strong data documentation, heritage organizations can leverage their modest staff and volunteer resources and provide world-class management of their collections.

A NOTE ON FREE COLLECTIONS MANAGEMENT SYSTEMS

Collections managers or heritage organization leadership may insist that the heritage organization consider free CMS platforms. A free CMS can be a suitable choice in the right circumstances. However, if a free CMS is to be considered, the heritage organization must have staff with the expertise needed to implement, run, and support it. When considering a free CMS, heritage organization staff need to be aware of—and prepared to solve—the following challenges:

CHALLENGES INHERENT IN FREE SOFTWARE:

- It often takes experts to implement the free software correctly.
- It requires experts to modify or improve the software.
- Only the user community or hired experts can offer trouble-shooting support, updates, and tool development.
- The software expiration is unknown, and the death of a free CMS can happen suddenly and without exit support.

The majority of CMS platforms are now offered as Software as a Service (SaaS). This means that the software is offered as a subscription that's paid for annually, and the annual cost covers technical support and system maintenance. The benefits of a SaaS system mean that the heritage organization no longer must maintain a software program on its own computer or server or fear the program will become obsolete in just a few short years. This also usually means that the software can be accessed and used anywhere there's an online connection—which became a critical piece of functionality to have during the global pandemic.

How to Identify CMS Specifications

CMS specifications will depend upon heritage organization, collection, staff, and future vision. The decision makers need to review every current and anticipated CMS use in order to tease out what tools are required for heritage organization staff to do their jobs effectively. Asking questions is a great way to explore what CMS specifications your heritage organization should include. For the following example questions, the word "tool" represents any CMS feature or function that a person would use as a tool to do their job.

Questions you should consider are:

- What tools do I currently use that are required to do my job?
- What tools do I not have currently that would make my job easier?
- What tools do I consider as mandatory versus recommended?
- What tools are so important that it's a deal breaker if the CMS doesn't offer that tool?

Next are questions that may apply to how the heritage organization operates and intends to use the CMS.

Example questions:

- Does the digital file backup meet preservation standards?
- Are there digital file size storage requirements?
- What administrative control levels are available?
- What customization is needed (on both the front and back end), and can the CMS accommodate those needs?
- What reports or other automation tools are desirable to support heritage organization staff?

In addition to staff/volunteer tool requirements which are all internal, it's equally important to review what tools may be required for external use.

Questions staff/volunteers should consider together as they think about the external heritage organization stakeholders:

- Is streaming video a requirement?
- Is the CMS ADA-compliant?
- Can the CMS navigation and object information be translated into another language on the fly?
- Can the CMS gather user contributions to object records?
- Can the CMS external search include objects from a different CMS at a peer heritage organization?
- Is there a zoom feature or watermark feature for images?
- Can users create and save their object searches?
- Are users able to download object images and data or share them via social media?

Make sure to capture all responses that result in a requirement of the new CMS tool. These will be important during the evaluation of CMS options.

How to Create CMS Specifications

Once the heritage organization has identified all possible CMS requirements, it's time to put them together and determine which tools are required and which are recommended but OK to do without. This part of the process can be the most difficult as staff often have conflicting opinions on what is a required tool. To assist in deliberation, each person should indicate to what degree they believe the tool is required. One easy way to do this is to assign an importance scale. For example, here is a sample list of required tools:

- Cloud-based (can access online form anywhere); Score: important
- Customization of catalog template; Score: very important
- Exhibit reports configuration; Score: somewhat important
- Curated lists (or pre-selected searches); Score: not very important

IDENTIFY AND RANK CMS SPECIFICATIONS

Create a form for each person to identify their CMS tool requirements and what they perceive the broader heritage organization requirements to be. Next to each requirement, have the staff assign a ranking of how important it is with the following scale: not very important, somewhat important, important, very important, extremely important.

Once everyone has identified and ranked their required tools, the organization can see which tools received the highest ratings. This exercise will help everyone understand how important a tool is to their colleagues and can see which ones have the most obvious appeal. Once the tools have been identified and ranked, the organization can construct a specifications sheet to present to CMS companies and to evaluate how each CMS ranks.

EXAMPLE OF CMS SPECIFICATIONS

CMS	1	2	3
Required: Cloud-based			
Required: Customizable Catalog			
Required: Report Creation			
Recommended: Streaming Media			
Recommended: Curation of Content (Digital Exhibit)			
Within Budget			

In addition to features, there are three important criteria to keep in mind when shopping for a new CMS: the ability to contribute to a collaborative portal, affordability, and the ability to exit the platform.

CONTRIBUTING TO A COLLABORATIVE PORTAL

When multiple institutions agree to participate in a portal together, they don't necessarily need to have the same CMS, though that is an option. They only need a CMS allowing external access to "harvest" its data and digital files. Publishing collection materials online

will (in most cases) require a CMS. Publishing them to a collaborative portal could require a CMS that possesses Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) functionality.

AFFORDABILITY

When selecting CMS options for initial consideration, knowing if the CMS can fit within the heritage organization budget is important. For the majority of organizations, the CMS takes up a large percentage of the operating budget, which adds pressure to the CMS decision-making process. To avoid costly missteps, the heritage organization should have a predetermined budget in mind when shopping for a CMS. The budget for the CMS will need to be workable within the larger operating budget, but it must also be realistic in terms of what a CMS can cost. If the heritage organization can't afford a CMS well suited to the stated heritage organization needs, then the budget isn't realistic. The heritage organization will need to re-evaluate the budget, the identified CMS needs, or both. Once the heritage organization has settled on a realistic budget, staff can begin to search for CMS options within that budget. If the cost of a specific CMS is nowhere close to the heritage organization budget, then the CMS will need to be removed from consideration. Additionally, though the heritage organization may be able to afford CMS "X" now, it's important to determine (to the best extent possible) if the heritage organization can maintain the annual and sometimes increasing cost of the CMS in the future.

THE ABILITY TO EXIT

This may seem like an odd place to start when shopping for a new CMS, but that's exactly the point. CMS migration and evolution are inevitable, and it's critically important to have a CMS allowing an organization to extract its data and maintain its structural integrity. No heritage organization should have to start from scratch nor spend countless hours massaging data to fit into a new CMS because the export of data from the previous CMS is unusable. Doing so would be extremely costly to the heritage organization regarding staff time and labor.

How to Identify CMS Options

It's hard to know where to begin when first considering a new heritage organization CMS. While the prospect can seem daunting, there are some easy steps to identify relevant CMS options quickly. Here are four easy-to-follow recommendations:

- Ask a friend or respected colleague at a different heritage organization what CMS they use and if they would recommend it.
- Visit peer heritage organization websites to view what CMS they use. The CMS name can be branded on the page or listed as part of the website address. If it's not immediately apparent, then reach out to the collections manager and ask.
- Review professional association pages to see if they provide a recommended vendor list or if they post CMS company advertisements.
- When attending conferences, note which CMS companies are in the exhibit hall, sponsored a meal or published an advertisement in the conference program.
- CMS Exploration at the Conference Exhibit Hall. Take advantage of the exhibit hall time to visit CMS companies when attending conferences. This is a great way to meet company staff informally, have a tour of the CMS platform, and ask questions.

WHERE TO FIND OFF-THE-SHELF CMS OPTIONS

There are a few good sources that track currently available collections management systems. The resources below each provide a list of the options, the collections they cater to, and the functionality offered.

- Canadian Heritage Information Network CMS Vendor List
- Car Library and Heritage organization
- Collections Trust (UK)
- Software Comparison Sites such as Capterra and G2

COMMON CMS VENDORS WITH AFFORDABLE SOFTWARE VERSIONS

- CatalogIt: CatalogIt is emerging as a major competitor to PastPerfect clients and even offers free data migration for PP clients.
- Collector Systems: Collector Systems offers a robust CMS at an affordable price. It's easy to use on the back and front ends.
- Lucidea: Lucidea offers a suite of programs under its "Essentia" brand and is a robust enough product to facilitate multiple accounts and a shared portal.

Consider an Off-the-Shelf CMS

Staff/volunteer capacity, technical knowledge, and desire to use heritage organization tools all factor into what CMS platforms will work for an organization. Sometimes a large, completely customized, robust CMS makes sense. Other times, the modestly sized and pragmatically built CMS tool is the better choice. There are several pieces of information to consider when discussing options:

- What tools are staff/volunteers currently using to manage their collections?
- What's the staff/volunteer capacity for learning and using new software?
- Who will use the CMS, and how?
- What are the desired outcomes of using an organization's CMS?

Technical knowledge, available budget, and desired outcomes all play a role in determining the right heritage organization CMS option for an organization. An off-the-shelf heritage organization CMS option is often the best option. This post will review the top four situations where an off-the-shelf CMS makes sense and where you can find them.

WHEN BUYING AN OFF-THE-SHELF CMS MAKES SENSE:

1. You need a CMS that's affordable. This is the #1 thing to consider as the heritage organization needs to afford any initial implementation cost as well as annual licensing fees. Off-the-shelf CMS products are affordable because they are an off-the-shelf, plug-and-play option. Minimal set-up or maintenance is required due to how the off-the-shelf product is built, and, as a result, it's more affordable. If your heritage organization budget is tight, you need an affordable CMS product, and an off-the-shelf tool may be right for you.
2. You need a CMS that's convenient and easy to set up. This comes down to staff technical knowledge and capacity. If heritage organization staff have limited technical knowledge and/or they don't have the time for an intricate heritage organization CMS implementation, then an off-the-shelf product is the better choice.
3. You need a CMS that's intuitive and easy to use for different skill levels. For heritage organization staff with limited technical prowess or organizations that rely heavily on volunteers and interns, an off-the-shelf, easy-to-use product is a necessity. Buying a large, complicated, and highly technical CMS doesn't do any good if no one at the heritage organization can use it.
4. You have limited in-house IT and no money to outsource technical support. Heritage organization CMS implementation, maintenance, troubleshooting, and updates all require staff with technical knowledge beyond the typical catalog entry work. If your heritage organization lacks IT staff and can't afford to outsource technical support, then choosing an off-the-shelf option with easy implementation and minimal maintenance is the best way to go.

CONCLUSION

In the end, if you are ready for a CMS, any CMS is better than no CMS. If you have the time, take it to thoughtfully identify the collection needs and assess how each CMS option may help meet your needs.

7

FUNDING RESOURCES

After scoping out your project and creating a plan, you might determine that hiring someone to help you execute the plan is the best course of action. Knowing some options for grant funding and tips for crafting a competitive application will increase your chances of success.

FUNDING RESOURCES

First and foremost, having project goals and a plan are essential for fundings asks.

If you follow this guidebook, you will end up with a plan and that in of itself increases your chances when applying for or asking for funding.

Funding resources such as grants can provide support for moving your heritage organization's collection data work forward. Grant resources can be offered by local, state, federal/national entities, and private foundations.

Keep in mind that funding opportunities can vary by the types of collections that an organization has, subject matter of the collection, and community partnerships and initiatives.

Below is not an exhaustive list, but a place to start looking for grant opportunities that may apply to a collection data cleanup project.

LOCAL GOVERNMENT GRANTS

- There are some cities that may have a grant funds for community organizations. It's worth asking your city staff or officials if one exists for your city. At the very least it gets you talking to your city government and letting them know what your organization is doing.
- Every County and Tribe has a **cultural coalition** that receives a certain amount of money from the Oregon Cultural Trust every year to be responsive to local needs related to heritage, art, and culture. Find out how to contact your local County or Tribal Cultural Coalitions visit https://www.culturaltrust.org/about-us/coalitions/

STATE GOVERNMENT GRANTS

Oregon Heritage

Oregon Heritage is a division of Oregon Parks and Recreation Department that serves anyone doing heritage preservation work in Oregon. There are two grant programs that could apply to your project:

- The Oregon Museum Grant supports Oregon, public and nonprofit, museums in projects for the collection and management of heritage collections, for heritage-related tourism, and heritage education and interpretations. This grant is open every spring.
- The Oregon Heritage Grant provides matching grants to non-profit organizations, federal recognized tribal governments, universities and local governments for projects that conserve, develop or interpret Oregon's heritage. This grant is open the fall of every odd year.

All Oregon Heritage grant information can be found at https://www.oregon.gov/oprd/OH/Pages/Grants.aspx.

Oregon Cultural Trust

The Oregon Cultural Trust's mission is to lead Oregon in cultivating, growing and valuing culture as an integral part of communities. We do this by inspiring Oregonians to invest in a permanent fund that provides annual grants to cultural organizations.

- The Cultural Development Grant recognizes and supports significant cultural projects that preserve and enhance Oregon's diverse arts, history, heritage, preservation, and humanities efforts.

Information for Oregon Cultural Trust Grants can be found at https://culturaltrust.org/grants/what-we-fund/.

State Library of Oregon

If you are a library that serves as the caretaker of a heritage related collection or a museum that partners with a library in preserving heritage related collections, you may be eligible to apply for a State Library of Oregon LSTA grant. More information for this can be found at https://libguides.osl.state.or.us/lstagrants/programs.

FEDERAL GRANT

Institute of Museum and Library Services (IMLS)

The mission of IMLS is to advance, support, and empower America's museums, libraries, and related organizations. They offer a variety of grant programs to help fund museum and library projects that benefit the public. For more information visit https://www.imls.gov/grants/grant-programs

PRIVATE FOUNDATIONS IN OREGON

Private Foundation grant opportunities greatly vary by region and foundation priorities. Be sure to search for local foundations and inquire about their funding priorities.

General grant tips

Every grant application varies but here are a few general grant related tips to consider:

Call or email grant program contacts.
Most grant program administrators are happy to talk to you about your project, whether it qualifies, and how to make your application more competitive. Some even offer the opportunity to submit a draft for feedback in advance of the deadline.

Review all of the guidelines and resources.
Many funders offer a variety of resources, checklists, webinars, FAQ's, and more to help grantees with their applications. Be sure to review these resources.

For example, **Appendix 6** is a checklist of information to include by project type that Oregon Heritage offers.

Network with other organizations.
Asking other organizations how they funded similar projects can help you identify funders or even help you identify some creative fundraising opportunities.

Have someone else unrelated to the organization read your grant application.
Grant review panelists come from a wide variety of disciplines related to the priorities of the grant program. Be sure that you are telling your project story in a way that most people can understand what you are trying to accomplish.

Visuals and attachments go a long way.
For collections related projects, pictures of the collection or screenshots of the collection data issues can really help your case making in the grant application. Attachments of any plans related to the outcome of this guidebook would be compelling to reviewers.

Clear and concise projects goals and measurables.
If you follow the guidance in this guidebook, you have the ability to create very clear goals and measurables. You will know how many records you need to clean up, you will know the biggest issues, and you will know the importance of having good data. If you plan to make your collection accessible through Northwest Digital Heritage, that too would be a compelling goal in terms of access.

Don't be afraid to phase projects.
It's been stated many times in this guidebook that a cleanup project will take time. Funders would much rather fund a reasonable phase of a project that can be accomplished within the grant period rather than fund an impossibly large project in the same period when it's obviously something that will take time. They want grantees to be successful and finding clear phases for the project can be very compelling. If you receive funds for one phase, that could even increase your chances of being successful in another grant round because you have proven success in initial phases of the project.

Do research on fair compensation.
If you are looking to hire an intern or a temporary position to complete this project, make sure you are looking at the going rate for similar positions and duties and planning to pay a fair and equitable wage.

Do a workflow test run.
Spend a couple hours cleaning up some data according to your plan but document what you are able to accomplish during that period of time. That will help you establish a reasonable amount of time spent and work accomplished when projecting project time and costs for a grant application. Be sure to include this workflow test run in the application.

This guide provides a lot of information that would help make your grant application more competitive.

For grant application questions related to the value and impact of the project, review Section 2: The Power of Good Data.

For grant application questions related to project need, goals and details, be sure to review Section 3-5 which will result in your data standards guide and clean up plan.

For grant questions related to ability to complete the project, if you did a workflow test this would help make the case for realistic accomplishment of measurable goals.

For attachments, be sure to include any of the outcomes of the exercises on creating a guide, review of the data, and cleanup plans and priorities.

Check out Appendix 6 to see an example of a checklist resource of what to include in your application based on project type for the Oregon Heritage Museum Grant and the Oregon Heritage Grant.

A

APPENDICES

APPENDIX 1 - DATA SCHEMA & CONTENT STANDARDS

This section discusses data schema and data content standards, how data content applies to an established schema, and how standards can cross-walk or translate with each other. To begin, we start with an analogy:

A recipe tells us what dish we're making. It contains a total list of required ingredients and may even recommend some optional "add-on" ingredients for the dish. The recipe dictates what ingredients are used and when they should be added.

If data schema is our recipe, then the data content for each field is the ingredients.

In order to make the best recipe, the ingredients need to be of the best quality and prepared according to the instructions, i.e. what format the ingredients have to be in—chopped, sliced, cubed, etc. The ingredients can be used for any recipe that calls for them, meaning data content standards can be used with various data schema.

Data Schema: Our Recipe

To begin, it's important to understand which data schema is currently used (or will be used) in your Collections Management System (CMS). This is the framework of your data.

DCMI: Dublin Core™ Metadata Element Set is the Most Widely Adopted Schema

Cultural heritage organizations often contain collections that span the traditional library, archives, and museum collections. This reality is partly why the Dublin Core Metadata Element Set[27] is the most universally adopted, with its straightforward approach to capturing core data across all collection types. It's also the de-facto standard schema to support Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)—the ability to have collection data "harvested" and placed in a collaborative portal.

Data Content: Our Ingredients

Given the typical collection types for Oregon heritage organizations, this guidebook references the two most popular content standards: Describing Archives: A Content Standard (DACS)[28] for archival materials and Cataloging Cultural Objects (CCO)[29] for museum artifacts. To inform what content goes into the catalog and how, please refer to either of these common content standards.

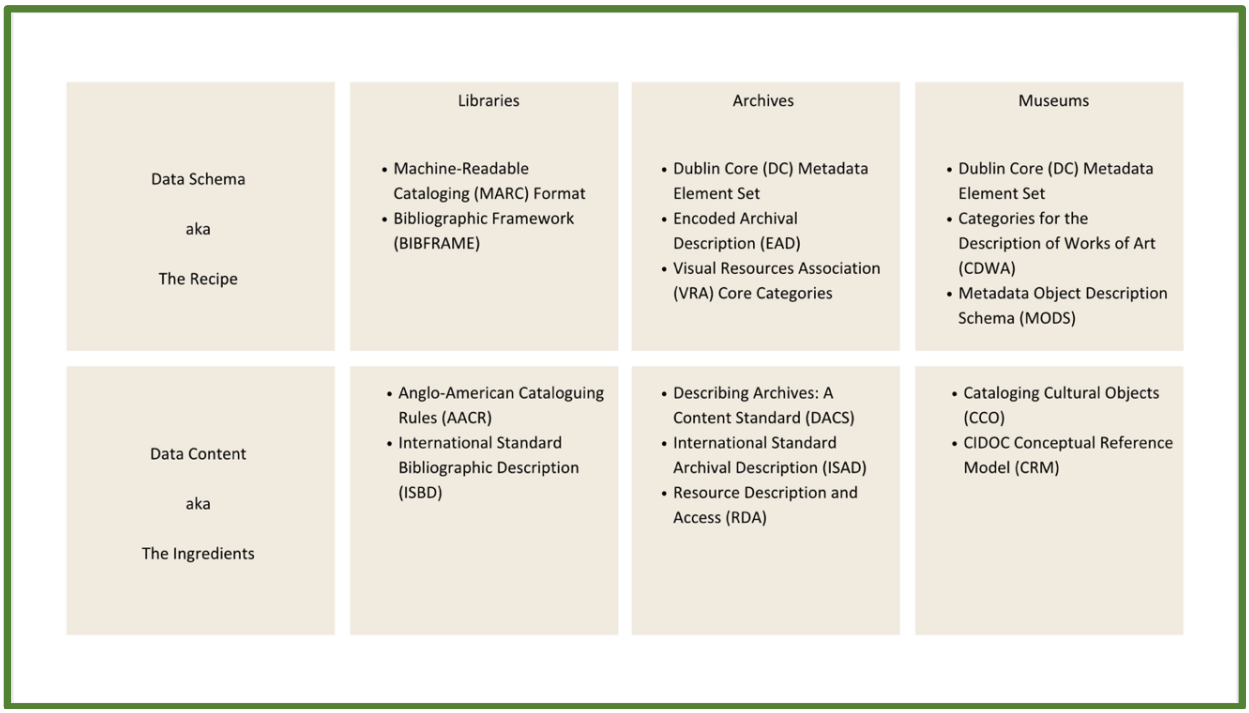
---

[27] DCMI: Dublin Core™ Metadata Element Set, Version 1.1, accessed February 15, 2023, via https://www.dublincore.org/specifications/dublin-core/dces/.

[28] The Society of American Archivists, *Describing Archives: A Content Standard (DACS),* Version 2019.0.3. accessed February 28, 2023, via Link: https://files.archivists.org/pubs/DACS_2019.0.3_Version.pdf.

[29] Cataloging Cultural Objects (CCO), *A Guide to Describing Cultural Works and Their Images*, accessed February 28, 2023, via https://vraweb.org/wp-content/uploads/2020/04/CatalogingCulturalObjectsFullv2.pdf.

For reference, here are the most common data schema and data content standards for libraries, archives, and museums:



	Libraries	Archives	Museums
Data Schema aka The Recipe	• Machine-Readable Cataloging (MARC) Format • Bibliographic Framework (BIBFRAME)	• Dublin Core (DC) Metadata Element Set • Encoded Archival Description (EAD) • Visual Resources Association (VRA) Core Categories	• Dublin Core (DC) Metadata Element Set • Categories for the Description of Works of Art (CDWA) • Metadata Object Description Schema (MODS)
Data Content aka The Ingredients	• Anglo-American Cataloguing Rules (AACR) • International Standard Bibliographic Description (ISBD)	• Describing Archives: A Content Standard (DACS) • International Standard Archival Description (ISAD) • Resource Description and Access (RDA)	• Cataloging Cultural Objects (CCO) • CIDOC Conceptual Reference Model (CRM)

The Data Crosswalk

To crosswalk data is to relate individual standards to one another. The following table illustrates how the Dublin Core schema (recipe) utilizes the data content standards (ingredients) from both DACS and CCO. This table also highlights how multiple content standards relate, making it possible to catalog mixed collection types within one system. For example, a historical society that wants to catalog its museum artifacts in the same system as the archival collections. It also means that data from both archives and museums can be harvested and contributed to a collaborative portal filled with mixed collection types. As with any standard, both CCO and DACS suggest required and recommended (or "added value") content fields. This table captures the requirements for each with a few highly prioritized recommended areas.

Note: The numbers that follow each DACS or CCO standard are to indicate the chapter identification of that standard in the over-standards guide. An in-depth guide using DACS with examples can be found in ***Appendix 2 – Describing Archives: A Content Standard (DACS) Reference****.*

Dublin Core Fields	DACS Content Standard	CCO Content Standard
Identifier		
Publisher	Name and Location of Repository Element (2.2); Reference Code Element (2.1)	Repository Numbers (21.2.3)
Title	Title Element (2.3)	Title (3.1)
Creator	Name of Creator(s) Element (2.6); Administrative/Biographical History Element (2.7 added value)	Creator (4.1); Creator Role (4.1.1)
Date	Date Element (2.4)	Creation Date (4.2)
Type	Extent Element (2.5)	Work Type (1.2)
Description	Scope and Content Element (3.1)	Subject Matter (16)
	Conditions Governing Access Element (4.1)	
Language	Languages and Scripts of the Material Element (4.5)	
Subject	Access points (See Overview of Archival Description)	Subject (16.2); Classification (2.1)
Rights	Conditions Governing Reproductions and Use (4.4 added value)	
Relation	Related Archival Materials (6.3 added value)	
Format		Measurement (6); Materials and Techniques (7.1)

RESOURCE: *The following resource comprehensively accounts for data schema and content standards. Jenn Riley, Glossary of Metadata Standards Content, published 2009, http://creativecommons.org/licenses/by-nc-sa/3.0/us/. This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License and was funded by the Indiana University Libraries White Professional Development Award.*

APPENDIX 2 - DESCRIBING ARCHIVES: A CONTENT STANDARD (DACS) REFERENCE

Describing Archives: A Content Standard, Version 2019.0.3.
Link: https://files.archivists.org/pubs/DACS_2019.0.3_Version.pdf

DACS Standard	Required or Recommended?	Notes
2.1.3 Repository Identifier	Recommended, if available	Not required, but helpful if submitting to larger, consortia databases. Must be requested from Library of Congress.
2.2 Name and Location of Repository	Required	Determine the repository's name and keep it consistent across the records. If your records ever join consortia records from other repositories this name will help identify where the collection or item "lives".
2.1.1 Local Identifier (Local Call # or Accession #)	Required	Determine a naming convention that is sustainable per collection and digitized item. Document this standard and be consistent. Most catalog systems, and consortia systems will not support duplicate identification numbers.
2.6 Name of Creator (LOC authority or local authority)	Required, if available	Name of the creator of the collection/item, if known. Import the name from the Library of Congress Name Authority file. If an adequate name cannot be found then a local authority may be created - make sure to document this decision. If the creator is unknown, then use Unknown.
2.3 Title	Required	Full name of the collection creator/donor if known, with the first letter of each word capitalized, i.e. First, MI, Last, Suffix. Follow the proper name with (in lower case) “collection” if mixed materials, or “papers” if primarily written documents.
2.4 Date	Required	The time period spanned for material in the collection. For open collections which are still receiving material, give the earliest date of the material followed by a hyphen and the word “ongoing” in square brackets. EX: 1875-[ongoing].
2.5 Extent (Physical Description)	Recommended, if appropriate	The quantity or number of units described.
Physical Description - Type of Unit	Required	The unit of measurement: linear feet, cubic feet, boxes, items.

2.7 Administrative History or Biographic Note	Recommended	Example: Burton E. Ashley, geologist, worked for the British South Africa Company (1930-1933). He received a M.A. degree from the University of Minnesota in 1936 and worked for the Texas Company (1936-1943), Phillips Petroleum Company (1943-1950), U.S. Geological Survey (1950-1957) and the U.S. Bureau of Mines (1957-1960). In 1961, he became minerals officer for the U.S. Department of State, serving in Australia. After retiring, Ashley served as a volunteer in the Mineral Sciences Department of the National Heritage organization of Natural History.
3.1 Scope & Content (Summary/Description)	Required	Collection description, primarily in terms of content (vs. format). Provide a narrative description of the scope and content of the collection, for Minimum Level Processing the description should be in the form of an abstract. For Optimum Description the summary should be a full length. Summaries should include the creator and/or donor of the collection items, the major subjects captured by the collection items, countries or regions represented in the collection materials, and the span of years represented in the collection.
4.1 Conditions Governing Access	Required, if restrictions exist	Used for physical access restrictions. This includes the terms governing access and the physical access provisions Use a standard text line in all collection and subsequent level records.
4.4 Conditions Governing Use	Recommended	Used for intellectual restrictions on use including reproduction. Use the following standard text lines when appropriate in all collection level records. Leave blank when no restrictions exist. Use this field for when copyright or duplicates exist.
4.5 Language of the Materials	Required, if applicable	What languages appear in the collection?

5.1 Custodial History/5.2 Immediate Source of Acquisition (Provenance)	Recommended	Recommended if known. Mostly helpful for backend accounting and transparency. Donated, Purchased, Loaned or Transferred from whom or from where and year. When purchased, do not mention the name of the seller. Examples: Donated by Volkmar Wentzel, 1999.; Purchased, 2000.; Collected by the Archives Staff, Willamette Falls Heritage Foundation Archives.
Access Points (Personal Name)	Recommended	Recommended best practice is to use controlled vocabulary, such as Library of Congress Authorities, Getty Thesaurus, Art & Architecture Thesaurus
Access Points (Topical Term)	Recommended	Recommended best practice is to used controlled vocabulary, such as Library of Congress Authorities, Getty Thesaurus, Art & Architecture Thesaurus
Access Points (Geographical Name)	Recommended	Recommended best practice is to used controlled vocabulary, such as Library of Congress Authorities, Getty Thesaurus, Art & Architecture Thesaurus
Access Points (Culture)	Recommended	Recommended best practice is to used controlled vocabulary, such as Library of Congress Authorities, Getty Thesaurus, Art & Architecture Thesaurus

APPENDIX 3 -

Worksheet: Is a Collaborative Portal right for your organization?

The intent of this activity is to consider what kind of access you want the public to have to your organization's collection information and evaluate the risk involved with more access and how participation in a collaborative portal like Northwest Digital Heritage might fit in with your organization's priorities.

Who in the organization should do this activity?

This would depend on the type of governance structure of your organization. For larger organizations or departments in a larger organization, talk to your manager about the decision making process.

For smaller organizations with minimal or no staff, it is advised that ultimately the decision to join a collaborative portal should be a board level decision. If staff/volunteers primarily manage the collections OR your organization has a collections committee, either one of these groups could do this activity and provide recommendations to the board. Alternatively, if the organization is all-volunteer and board members also manages the collection, it would be a good idea to consider having the entire board do this exercise.

What organization information should you review before doing this activity?

- Any organization core documents such as Mission, Vision, Values, Ethic statements, Collections Policy
- If your organization does not have them, consider industry ethics standards such as:
    - American Alliance of Museums (https://www.aam-us.org/programs/ethics-standards-and-professional-practices/code-of-ethics-for-museums/)
    - Society of American Archivists -(https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics)
- Who can currently access your collection information and in what ways (online, on site, general public, members only, researchers only, etc.)
- Instructions on your collection management system your organization uses to manage your collection, this may either be a database or a spreadsheet depending on the system you use.

Additional considerations

This activity is meant to promote internal conversation by generally assessing your organization's ideal level of access to your collection information and what risk factors to consider based on that ideal level of access. It does not provide legal advice or absolution of risk associated with access.

Ultimately your organization assesses and decides what risk you are willing to take. But, just because risk exists, does not mean steps can't be taken to mitigate risk. If during this exercise you find that there is risk, please consider conversation and resources to clearly define and address real or perceived risks factors.

Additional resources for any recommendations or next steps

- Oregon Heritage MentorCorps - a technical assistance program that matches organizations with someone trained in heritage preservation and organization management skills. Applications to request assistance open in July.
- Northwest Digital Heritage - a collaborative portal that also provides technical assistance related to digitization and access for online collections.

Step 1: Follow the directions for each section below related to 'Access' and 'Risk'.

Access

Mark the box next to each statement your organization agrees with regarding the type of desired access you ideally want the public to have to your collection.

- ☐ Anyone can access the collection online.
- ☐ Only the local community should access the collection.
- ☐ The collection can be accessed on-site at a publicly accessible research station.
- ☐ The collection should NOT be accessed online.
- ☐ Only Staff/Volunteers should be allowed to access the information and control what others can view.
- ☐ We want the public to engage with collections we are not able to display for reasons such as space and preservation risk.
- ☐ We don't have the capacity to answer research requests and would like the public to be able to search the collection on their own.
- ☐ If teachers and students are unable to travel to our site, we would like them to engage with our collection online to support their curriculum.
- ☐ We want to limit online access because we depend on fees for research requests and image requests.
- ☐ By having metrics of how many people access our collection online, we would be able to leverage that information with local decision makers, grant funders, donors, and partners.
- ☐ Our organization is private and only exists to serve a particular audience therefore access is restricted to a membership.

Scoring instructions: Count up the number of statements in black you marked. Subtract one for each statement in red you marked. For example, If you marked 2 statements in black and 3 red statements, your total would be -1. Put the total in this box ⟶	**ACCESS TOTAL**

Risk

Mark the box next to any of the statements that are true for your organization regarding risk of access to your collection information.

- ☐ The collection contains harmful depictions or descriptions of historically mistreated communities.
- ☐ We would consider working with the community on addressing harmful language in our collection.
- ☐ The collection contains content that could trigger trauma for those engaging with the information.
- ☐ We would consider a process to determine triggers for trauma that exist in our collection information.
- ☐ The collection contains personally identifiable, medical, or financial information.
- ☐ Our collections system allows for us to flag records that have harmful language or personal information.
- ☐ The collection contains materials that it doesn't own, or has possession of items it doesn't have the copyright to.
- ☐ Our organization is willing to have internal conversation on the level of risk associated with items that have unknown copyright.
- ☐ Materials are at risk of theft or vandalism if their existence and location (via repository identification) are made more easily discoverable.
- ☐ We know we can improve the security of our collection but need guidance and resources.

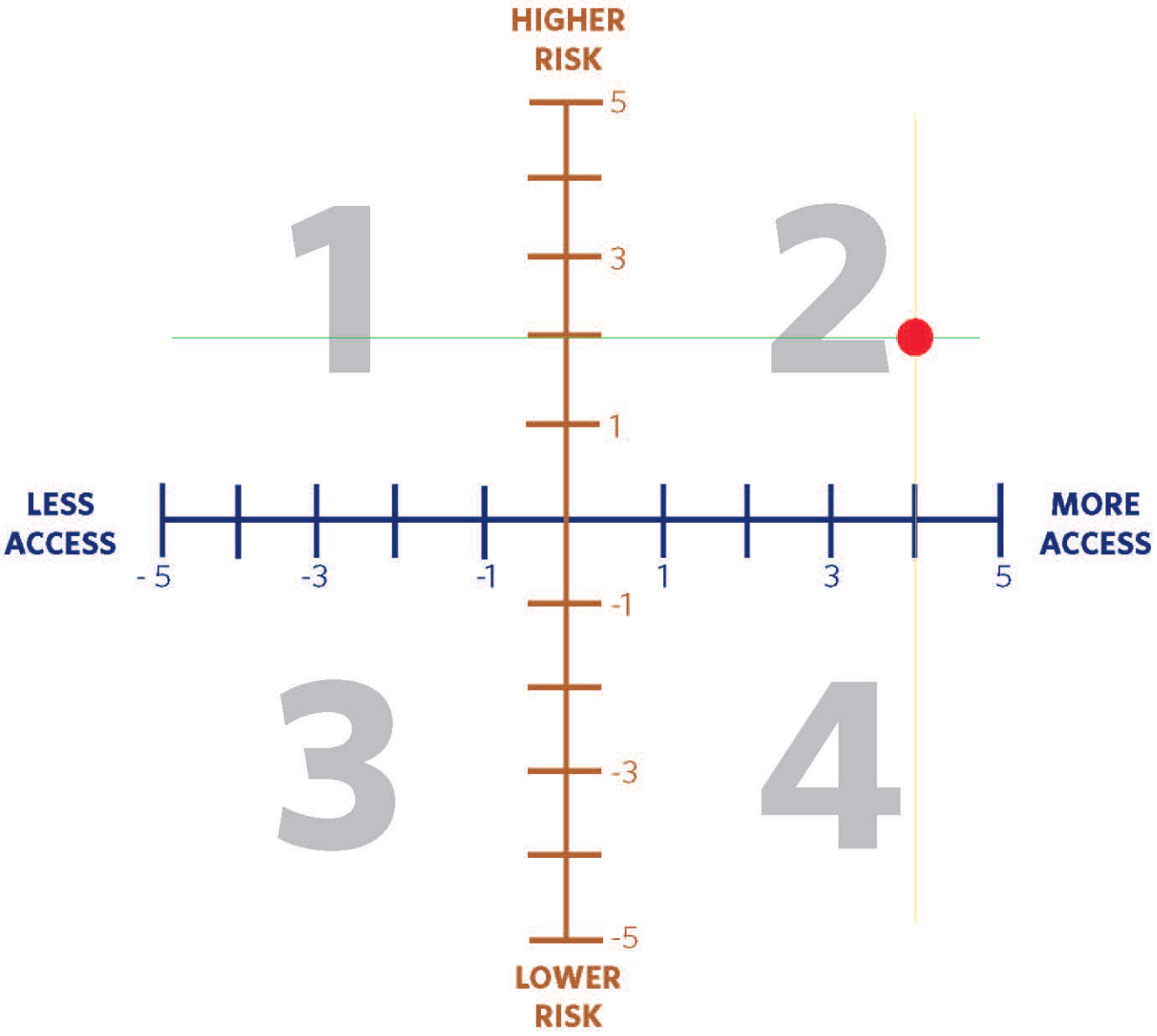
Scoring instructions: Count up the number of statements in black you marked. Subtract one for each statement in red you marked. For example, If you marked 4 statements in black and 1 red statement, your total would be 3. Put the total in this box ⟶	**RISK TOTAL**

Proceed to the next page for the next step.

Step 2: Take the totals you recorded in Step 1 and place them on the appropriate axis. Plot the point at which they would meet and record the Quadrant number this point falls within. If you are having trouble, use the following example as a guide.

EXAMPLE: Let's say an organization had the following totals: Access = 4 & Risk = 2

1. First, find 4 on the blue axis and draw a vertical line going through that axis (shown in yellow below).
2. Next, find 2 on the orange axis and draw a horizontal line through that axis (show in green). The red dot is where they intersect and it falls in quadrant 2.



What quadrant did your point fall within? Record that number here and continue to the next page to find out recommendations based on these results. (if your score is 0 on either of the axis, record the two quadrants you fall between and review both of them on the next page)	QUADRANT

Step 3: Find the quadrant that matches your results from the previous page and review recommendations and considerations.

1 Less Access/Higher Risk

A collaborative portal may not be the right choice for your org. at this time.

It could be that the level of risk is impacting your organization's desire for the public to have more access to your collection. There are tools and resources to help address each of the risks listed when your organization is ready.

If capacity is the main issue for increasing access to your collection, consider contacting Oregon Heritage MentorCorps to help with prioritization or project planning. However, if the organization does not value access to the collection, more internal conversation may need to happen about your organization's role as a collecting institution and how it fits into ethics, standards, and duties related to public service and collections. Consider reviewing the industry ethics, standards and duties on the first page of this activity.

2 More Access/Higher Risk

You want to increase access but need guidance on lowering the risk.

It appears your organization would like more access to your collection but are concerned about the level of risk.

This does not discount participation in a collaborative portal, but there are resources to help address the risk factors.

It is recommended you consider the following actions:

- Contact Oregon Heritage MentorCorps to get guidance on potential steps to address concerns of risks related to collection management and security
- Contact Northwest Digital Heritage to request guidance on harmful language, copyright, and personal information concerns.

3 Less Access/Lower Risk

Consider more internal conversation about access to collections.

Despite having a lower risk level for your organization's collection, access does not seem to be a high priority. This may be because your organization exists to serve a very particular audience or interest group.

Generally, for museums/libraries/archives that are nonprofits or publicly owned and operated, allowing access to collection information is considered part of it's duty of public service.

It is recommended that your organization's governing body review industry standards and ethics related to collection access such as those listed on the first page of this activity.

If capacity is the primary reason for less access, it may benefit your organization to take advantage of strategic planning processes or free project planning services such as Oregon Heritage MentorCorps.

4 More Access/Lower Risk

Your organization should consider a collaborative portal!

Access to your collection information seems to be a major priority for your organization and risk for your collection is on the lower side.

Cleaning up your collection data is the right step if you want to join a collaborative portal. It is recommended you pursue a collection data clean up project. Also contact Northwest Digital Heritage to request information on how to participate so it can be incorporated into your planning efforts to increase access to your collections.

If in the risk section you selected any of the red statements, Northwest Digital Heritage can help provide guidance on addressing those risks.

If your collection is not currently available online, Northwest Digital Heritage can provide a hosting option.

Step 4: Regardless of which quadrant you landed in, consider continuing the conversation in your organization related to access and risk. Resources and contacts for any recommendations or next steps can be found on the first page of this activity.

APPENDIX 4 -

Worksheet: Creating your organization's data standards guide

The intent of this worksheet is to help you complete a Data Standard Guide before you review and assess your data. Be sure to review section 3-4 of this guidebook ending on the steps to create your organization's data standards guide before attempting to work through this worksheet.

DOCUMENTING YOUR CHOSEN STANDARDS AND CREATING A CLEAR QUICK REFERENCE GUIDE FOR THEM.

Step 1 & 2 on the instructions asked you to choose your fields and field formats (standards), crosswalk them with your CMS fields if you are using a CMS, and document them at the beginning of your data standards guide. Below is an example from an organization using the PastPerfect CMS.

Dublin Core Fields	DACS Content Standard	CCO Content Standard	PastPerfect
Identifier			Object ID; Other Number
Publisher	Reference Code Element (2.1); Name and Location of Repository Element (2.2)	Current Location (5.2.1)	Collection; Home Location
Title	Title Element (2.3)	Title (3.1)	Object Name; Title
Creator	Name of Creator(s) Element (2.6); Administrative/Biographical History Element (2.7 added value)	Creator (2.2.1); Creator Role (2.2.2)	Artist, Author, Creator, Photographer
Date	Date Element (2.4)	Creation Date (4.2.3)	Date
Type	Extent Element (2.5)	Work Type (1.2)	Type
Description	Scope and Content Element (3.1)	Description (8.2.1)	Description
	Conditions Governing Access Element (4.1)		
Language	Languages and Scripts of the Material Element (4.5)		n/a
Subject	Access points (See Overview of Archival Description)	Subject (4.1); Classification (7.2.1)	People, Subjects, Classification Search terms
Rights Management	Conditions Governing Reproductions and Use (4.4 added value)		Notes & Legal > Legal Status or Web Rights
Relation	Related Archival Materials (6.3 added value)		Relations
Format		Measurement (3.2.1); Materials and Techniques (3.2.2)	Dimensions; Material; Technique
Current Location		Current Location (5.2.1)	Location > Temporary location

If the standards discussion is still scary, the most important thing is to document what fields of information you are collecting. This next step will help you really nail down what consistent information and format you will have for each field.

CREATE INSTRUCTIONS FOR EACH FIELD

Be sure to review the table is section 3. The goal is to document instructions for each field so that you have that nailed down and can reference this guide in your review. This can be done in several formats and will vary based on the individual decisions you may make for each field your organization chooses. Below are some templates to consider.

Document format

Write down or type a document with the following information:

Field Name:
Field Definition:
Field Format:
Recommended or Required?
Is there a controlled vocabulary for this field? If yes, include reference to that list or attach it.
Are there instructions, standards, things to avoid for this field?
Provide good examples of data for this field

Spreadsheet format

Create a spreadsheet in excel or a similar program that has instructions for each field similar to the example below:

Field Name	Field Definition	Field Format	Recommended or Required	Is there a controlled vocabulary	Are there instructions, standards, things to avoid for this field?	Provide good examples of data for this field

Table Format

Create a table like the one below for each field.

Field Name:	
Field Definition	
Field Format	
Recommended or Required?	
Is there a controlled vocabulary for this field? If yes, include reference to that list or attach it.	
Are there instructions, standards, things to avoid for this field?	
Provide good examples of data for this field	

FINAL STEPS

Once you have done all of this work, be sure to keep it all together for you review process. You may also consider attaching it to any policies or procedures associated with your collection as you have also just completed a guide to entering new information to your database to avoid future problematic collection data issues.

APPENDIX 5 -

Worksheet: Collection Data Review, Assessment, & Prioritization

The intent of this activity is to offer a structure on how to review your current data to identify the major issues and how to prioritize a clean-up project of your collection data. Before you view the information below, be sure to review section 3-4 of this guidebook.

SETTING THE SCENE

After working through sections 3-4 of this guidebook, you will have already accomplished the following

- Made decisions about what standards your organization is using (fields and content formats)
- Created your organizations data standards guide outlining specific information for each field and what and how information will entered into each field.
- Have your collection information ready to view in a comprehensive format such as a spreadsheet

Once all of that is in place, you ready to review your data.

REVIEWING & ASSESSING YOUR DATA

Using your organizations data standard guide and the review recommendations provided in section 4 of this guidebook, start working your way through your data and recording the issues you are finding. One logical way to do this is working field by field and document your findings based on the following criteria:

Does the field contain data?
Is the information accurate and complete?
Is the format of the data consistent?

There are different ways you can document your review and assessment:

- **Document narrative** by field note the issues related to the above questions, what actions to take, and priority level. You can see an example of a narrative assessment in the PastPerfect Case Study in Appendix 7
- **Spreadsheet format** with the following column heading suggestions. There is also a really great example of a spreadsheet like this in the PastPerfect Case Study, Appendix 8.

ISSUE NUMBER	FIELD FOR CLEANUP	QUANTITY	PRIORITY	CLEANUP INSTRUCTIONS	ATTACHMENTS
As you assess your data, document each field type where you're noticing data gaps or deficiencies when comparing actual data to your codified data standards document. If a you discover more than one issue in a field, mark each issue as a separate entry as you may have to take different approaches with each issue.		Attempt to get an estimate of how many records are impacted by each issue. This will help you determine what's achievable for the available capacity of your staff.	Determine the level of priority for each item to the best of your ability. Things to consider when prioritizing	Cleanup will depend on the issue, staff knowledge and capacity, and the CMS. See section 5: Data Cleanup Strategies: Application for more information.	If there were any queries, reports, or other documentation you used for data assessment, make a note of them here so that any staff member can see if there's related documentation needed for effective cleanup.

- **Table format** like the example on the next page can be helpful if you need to record it by hand while viewing the spreadsheet.

Field Name	Circle the number that fits each statement based on how many records fit the criteria for the field.					Assign priority level (high, med., low) & justification
	Field contains data:	1 All	2	3 Some	4	5 None	
	Data is accurate & complete :	1 All	2	3 Some	4	5 None	
	Format is consistent:	1 All	2	3 Some	4	5 None	

What are the issues for this field and corresponding actions for cleanup?

Issue	Cleanup Instructions

PRIORITIZATION

After your review and assessment, take a minute to assign prioritization to your cleanup tasks you have now created. Review "A note on Prioritization" at the end of Section 4 for how you might develop your own prioritization rubric. You may also want to review the PastPerfect Case Study in Appendix 7 to see how that organization assigned priority level to the issues.

APPENDIX 6 - OREGON MUSEUM & HERITAGE GRANT CONTENT CHECKLIST

The following is an excerpt from Oregon Museum Grant 2023 Guidelines, produced by the Oregon Heritage Commission.

To be as competitive as possible, be sure to include the following information in your grant application. Check out our Heritage Bulletins and MentorCorps for additional technical assistance.

Collections Cataloging

- ✓ Current condition
- ✓ How the objects to be cataloged in the project were selected and prioritized
- ✓ If it is a previously undocumented or researched history, describe how it was excluded and why it should be included now.
- ✓ Historical and organizational significance of the collection
- ✓ Cataloging process details
- ✓ Expertise that will be used (experts, resumes, training, etc.)
- ✓ Estimated number of objects, boxes, and linear feet to be cataloged.
- ✓ Photo of collections
- ✓ If a collection represents the history people frequently excluded from historic interpretation. Examples include community populations based on race, ethnicity, LGBTQ+, gender, etc., or themes like labor, immigration, etc. Include collaboration and letters of participation with organizations that represent those people now—for example, the tribes, Oregon Black Pioneers, local youth organizations, etc.

Collections Housing and Storage

- ✓ Current condition, including photos
- ✓ How the objects to be rehoused in the project were selected and prioritized
- ✓ Historical and organizational significance of the collection
- ✓ If it is a collection related to a population or theme with limited documentation or research history, describe how it was excluded and why it should be included now
- ✓ The process and materials that will be used
- ✓ Expertise that will be used (experts, resumes, training, etc.)
- ✓ Estimated number of objects, boxes, and linear feet to be cataloged.
- ✓ If a collection represents the history people frequently excluded from historic interpretation. Examples include community populations based on race, ethnicity, LGBTQ+, gender, etc. or themes like labor, immigration, etc. Include collaboration and letters of participation with organizations that represent those people now—for example, the tribes, Oregon Black Pioneers, local youth organizations, etc.

Collections Digitization

- ✓ How the objects to be digitized in the project were selected and prioritized
- ✓ Historical and organizational significance of the collection
- ✓ If it is a collection related to a population or theme with limited documentation or research history, describe how it was excluded and why it should be included now

- ✓ The process that will be used (standards of digitization, metadata, size of a file, etc.)
- ✓ The storage method and level of public access (online, at the organization, etc.)
- ✓ Expertise that will be used (experts, resumes, training, etc.)
- ✓ Estimated number of objects, boxes, and linear feet to be cataloged.
- ✓ Materials to be used (equipment specifications)
- ✓ Sample image of collections to be digitized
- ✓ If a collection represents the history people frequently excluded from historic interpretation. Examples include community populations based on race, ethnicity, LGBTQ+, gender, etc., or themes like labor, immigration, etc. Include collaboration and letters of participation with organizations that represent those people now—for example, the tribes, Oregon Black Pioneers, local youth organizations, etc.

University of Oregon Digital Newspaper Program – adding local newspaper content to this online program.

- ✓ How the objects to be digitized in the project were selected and prioritized
- ✓ Historical and organizational significance of the collection
- ✓ If it is a collection related to a population or theme with limited documentation or research history, describe how it was excluded and why it should be included now
- ✓ The process that will be used (standards of digitization, metadata, size of a file, etc.)
- ✓ A letter from the University of Oregon program affirming the project is on the schedule for the grant period.
- ✓ The storage method and level of public access (online, at the organization, etc.)
- ✓ Expertise that will be used (experts, resumes, training, etc.)
- ✓ Estimated number of objects, boxes, and linear feet to be cataloged.
- ✓ Materials to be used (equipment specifications)
- ✓ Sample image of collections to be digitized
- ✓ If a collection represents the history people frequently excluded from historic interpretation. Examples include community populations based on race, ethnicity, LGBTQ+, gender, etc., or themes like labor, immigration, etc. Include collaboration and letters of participation with organizations that represent those people now—for example, the tribes, Oregon Black Pioneers, local youth organizations, etc.

Education Projects – working with schools and school programs

- ✓ Information that will be included, learning goals, how and why these were selected
- ✓ If it is interpreting a previously untold or limited history of a population or theme, describe how it was excluded and why it should be included now
- ✓ Teaching methods to be used
- ✓ Planning with schools, districts, teachers
- ✓ Expertise that will be used (experts, resumes, training, etc.)
- ✓ How the program meets the targeted state core requirements
- ✓ Evaluation of learning beyond teacher survey
- ✓ If the program represents the history people frequently excluded from historic interpretation. Examples include community populations based on race, ethnicity, LGBTQ+, gender, etc., or themes like labor, immigration, etc. Include collaboration and letters of

APPENDIX 7 - PASTPERFECT CASE STUDY

OREGON HERITAGE
OREGONHERITAGE.ORG

PastPerfect 5 & Online Case Study

An Assessment Summary and Cleanup Plan for Deschutes County Historical Society

Deschutes
Historical Museum

v. July 29, 2023

CULTURAL TRUST

This project was made possible with funding from the Oregon Cultural Trust.

Relicura™
Historical Research & Collections Management

Created by Rachael Cristine Woody, Founder + Director of Relicura™ LLC

Table of Contents

Introduction

The data in your Collections Management System (CMS) is important. The quality of it impacts how well you can manage your collection and provide access to it. Any cultural heritage organization that's existed longer than a decade knows the challenges that can exist with changing technologies, shifting best practices, reliance on volunteers and interns, and staff transitions. Any one of these items can contribute to inconsistent data creation in a Collections Management System (CMS) and it's safe to say 99% of heritage organizations each have a part of their collection catalog that needs data refinement (aka data cleanup) in order to meet current needs and best practices. It's important to known and get comfortable in the knowledge that messy data exists and we all have messes we feel embarrassed about regardless of whether we're the ones that created the mess. Another truth to get comfortable with is that attempting data cleanup can absolutely feel overwhelming. Where do you start? How much is there? How long will it take? And, How do you even do it? These are big questions and this isn't a topic typically covered in graduate school or post-grad workshops. So, if you're reading this and you don't know the answer to any of the posed questions in this section: great. You're in the right place.

Any Data is Better than No Data: *This section focuses on data cleanup and enhancement, but is in no way implying that data should be 100% correct and complete before being published in the catalog. Any data (any point of access) is better than no data. No data means the collection is completely inaccessible and that's worse than messy data.*

The Case Study

The following is a case study-style report of how Relicura™ LLC worked with Deschutes County Historical Society (DCHS) to perform an assessment of their PP 5 and PP Online data, and create a data cleanup plan. This case study is intended to be paired with the Oregon Heritage Commission produced *A Guide to Collection Data Cleanup*. As we make our way through the case study sections of the Guide will be referenced where you can find more information on a particular topic. While this study is intended to provide you with specific examples, some of the non-essential detail may be omitted for clarity or out of deference to DCHS.

Desired Outcome

The most obvious answer is: Clean and complete data! But, it's more nuanced than that. The goal is data cleanup performed in an effective way and at a sustainable pace. Yes, we'd love to have 100% of our data clean at the end of this, but that's the ultimate "future" goal. Not our "now" goal. We need to make sure we make best use of limited staff time and that the pace is sustainable so as not to overburden staff or have the work done at the expense of other priority items.

Good Data: *For more information on the importance and helpfulness of "good data" please see the "The Power of Good Data" section in A Guide to Collection Data Cleanup.*

Cleanup Limitations

Every CMS will have data cleanup tools and limitations. As both software and user expectations evolve, data cleanup *should* become easier and easier for us, however, no tool is the same. Nor, is any collection or data mess the same. It's important to understand both the possibilities and limitations of your CMS in order to determine the cleanup strategies available to you. The PP product is limited in its support of data cleanup tools or functions. As a result, the majority of data cleanup has to take place at the item-level record and progress item by item. There is an option for bulk data conversion, but it must go through PP and is (currently) billed at $100/hour. Unfortunately, the impact of this limitation disempowers the PP user and can prevent meaningful cleanup from taking place.

Capacity and Realistic Expectations

The museum, archives, and cultural heritage organization sector suffers from chronic under-funding and low staffing levels. As such, the staff in place are doing their best to perform all required duties for collection care, management, and access—and they are already at capacity. With that in mind, the only way data cleanup can happen is if: 1. Capacity is added via additional staff, or grant-funded or otherwise temporary staff; and/or 2. The current slate of priorities are changed and something is taken off the to-do list—even temporarily—in order to accommodate the time needed for data cleanup.

In addition, know that data cleanup is likely a long-term, multi-year effort. In order to proceed it's important to break the cleanup into achievable, short-term chunks. This approach requires performing data assessment up front and the intentional partitioning of cleanup activities for future work. There are additional benefits to this approach: 1. It supports easy "grab and go" cleanup for a window of staff availability or a trusted intern; and 2. It can serve as a framework for a potential grant or donor-funded project.

Finally, sometimes we just need an easy win. With the pre-identification of cleanup areas in your data, you'll have a sense of how pervasive a particular data mess may be. In many instances there are small and confined data messes that require only 5-15 minutes to correct. For example: Updating (or consolidating duplicates of) a creator authority record. Or, perhaps block out 1-hour a week to tackle any data cleanup sets that can be performed in that amount of time. Employing one or more of these tactics will greatly support your overall data cleanup effort and help keep the work moving at a sustainable pace.

Limited Capacity and Realistic Expectations Requires Us to:

- Perform an assessment of your data and identify areas of cleanup.[1]

[1] Information and guidance on how to perform data assessment please see *A Guide to Collection Data Cleanup*, section titled "A Review of Data Cleanup Strategies."

- Carve up the areas of cleanup by what needs to be done, by field, and (if there's a lot) by record set. A record set can be an arbitrarily decided number or another narrowing criteria factor.[2]
- Pair each cleanup batch with guidance on what's needed and how the cleanup can be performed.[3]
- Identify the quantity of records per cleanup batch, estimate the time it would take to perform the data cleanup (factor in research if the data is missing), and assign a priority level.[4]
- Finally, indicate if the work can be performed by a volunteer, intern, or staff member. This will depend on the knowledge and capability of the person as well as who is trained to use your CMS.

It is entirely possible to have non-CMS users assist with data cleanup by performing necessary research, imaging the collection, and providing data for review in a spreadsheet format for a staff member to copy/paste or import into the CMS at a later date. This case study and many of the cleanup strategies recommended in *A Guide to Collection Cleanup* rely on spreadsheets for data analysis, data preparation, and (if possible) data import. The *Data Cleanup* section in this case study provides further examples on how to incorporate other team members in the cleanup process.

The Assessment Summary

The Deschutes County Historical Society (DCHS) uses PastPerfect 5 (PP 5) as the primary CMS. Currently, there are approximately 75,000 records in the database with 559 of those records provided through PastPerfect Online (PP Online). This study will offer a summary assessment for both (PP 5 and Online) and the cleanup specified for each.

The best way to evaluate data is as an aggregate and when using PP, the only way to effectively do so is with generating a query and exporting the results as a spreadsheet. A spreadsheet view will display each entry and all of the data in each field. This makes it easier to spot missing data, incorrect data formats, and any data inconsistencies.

Data Review via Spreadsheet: *For more information on the importance and helpfulness of "good data" please see "A Review of Data Cleanup Strategies" section in A Guide to Collection Data Cleanup.*

In PP it's currently impossible to perform an export of more than a few hundred records at a time. In order to review data in the aggregate for DCHS, queries were ran in order to generate spreadsheet reports that could provide an aggregate view of the selected records. Given that DCHS

[2] See the examples in this case study for how we divided the data cleanup work.
[3] What's needed will depend on the field and the perceived data gap or discrepancy. Best practices can help inform what's needed and other items might require research into the collection or item.
[4] How you determine the priority level is up to you. Use this case study for examples in what was prioritized and why.

has 75,000 records, this review could be performed on the 500 PP Online records and using a sample size of PP 5 records from each of the main collecting areas in PP: Archive, Library, Object, and Photograph. The gaps in data outlined below were present in the data reviewed and are likely indicative of missing data across the collections.

Data Standards

The following data standards were used to help conduct the assessment[5]: Dublin Core™ (DC) Metadata Element Set[6] and Cataloging Cultural Objects[7] (CCO). DC is the most universally adopted with its straightforward approach to capturing core data across all collection types. It's also the de-facto standard schema to support Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)—the ability to have collection data "harvested" and placed in a collaborative portal. Each schema has designated fields that are considered the minimum required fields. In addition to what fields to use as a schema, there are also descriptive standards (and subsequent fields) to consider. For many museums and cultural heritage organizations the standard typically used is CCO or (and especially if archives are present) DACS.

Data Standards: *For more information on data standards, please see the "An Overview of Data Standards" section in A Guide to Collection Data Cleanup. For an introduction to the language used while discussing collection data, please see the "Establishing a Common Language" section.*

CCO minimum fields to use:

- Work Type
- Title
- Creator; Role (controlled vocabulary- can be local, an authority, or both)
- Creation Date
- Subject (use of Nomenclature for a controlled vocabulary)
- Current Location
- Measurements
- Materials and Techniques (use of Nomenclature for a controlled vocabulary)

The Cross-Walk from Data Schema, to Descriptive Standards, to PP Fields

The following grid is a cross-walk[8] among the data schema fields (per Dublin Core™), the two descriptive standards (DACS and CCO), and the corresponding field in PP. The numbers found in the DACS and CCO columns refer to the chapter and section of where you can find more

[5] For more information on data standards, please see *A Guide to Collection Data Cleanup,* "An Overview of Data Standards."
[6] DCMI: Dublin Core™ Metadata Element Set, Version 1.1, accessed February 15, 2023, via https://www.dublincore.org/specifications/dublin-core/dces/.
[7] Cataloging Cultural Objects (CCO), *A Guide to Describing Cultural Works and Their Images*, accessed February 28, 2023, via https://vraweb.org/wp-content/uploads/2020/04/CatalogingCulturalObjectsFullv2.pdf.
[8] Cross-walk in this case is used in reference to the equating of standard when there are multiple standards or best practices in play. For example, one field title in Dublin Core may be called something else in PastPerfect. Another example is that some standards may require specific fields not included in other standards. By viewing the standards side-by-side in this grid, you will be able to see which fields are considered required and what that field name is across the different standards, and in PastPerfect's case, the name of the field in the database.

information on the field via the standards reference guide. Some boxes are absent of content when there's not a corresponding field to use or standard to reference.

Dublin Core Fields	DACS Content Standard	CCO Content Standard	PastPerfect
Identifier			Object ID; Other Number
Publisher	Reference Code Element (2.1); Name and Location of Repository Element (2.2)	Current Location (5.2.1)	Collection; Home Location
Title	Title Element (2.3)	Title (3.1)	Object Name; Title
Creator	Name of Creator(s) Element (2.6); Administrative/Biographical History Element (2.7 added value)	Creator (2.2.1); Creator Role (2.2.2)	Artist, Author, Creator, Photographer
Date	Date Element (2.4)	Creation Date (4.2.3)	Date
Type	Extent Element (2.5)	Work Type (1.2)	Type
Description	Scope and Content Element (3.1)	Description (8.2.1)	Description
	Conditions Governing Access Element (4.1)		
Language	Languages and Scripts of the Material Element (4.5)		n/a
Subject	Access points (See Overview of Archival Description)	Subject (4.1); Classification (7.2.1)	People, Subjects, Classification Search terms
Rights Management	Conditions Governing Reproductions and Use (4.4 added value)		Notes & Legal > Legal Status or Web Rights
Relation	Related Archival Materials (6.3 added value)		Relations
Format		Measurement (3.2.1); Materials and Techniques (3.2.2)	Dimensions; Material; Technique
Current Location		Current Location (5.2.1)	Location > Temporary location

Overview of Data Gaps in PP Online Data

All 559 PastPerfect Online records are from the Photograph collection module in PP. These records are fairly modest in data capture with the consistent use of the following four (4) PastPerfect fields:

- Object ID (OBJECTID)
- Collection
- Type
- Description (DESCRIP)

When comparing the current PP Online fields used to current data schema standards (Dublin Core™) we find the following fields absent:

- Identifier (this is the individual item number)
- Title (of the item being cataloged)
- Creator (controlled vocabulary- can be local, an authority, or both)
- Date (creation date)
- Format (use of Nomenclature for a controlled vocabulary)
- Language (when applicable)
- Subject (use of Nomenclature for a controlled vocabulary)
- Rights (any copyright or other retained rights)
- Relation (when applicable, indicate a relation to other known items in the DCHS or peer museum collection)

Overview of Data Gaps in PP 5 Data

Across all four collections the following data gaps were found (see below). Many of them are part of the data elements (fields) that should be present in a complete record.

Arch Records

- CatBy field infrequently used
- Date field infrequently used
- EarlyDate and LateDate columns usually have a “0”
- HomeLoc has at least one row without an entry
- PubPlace and Publisher columns are inconsistently used – either both, one or the other, or neither are filled out
- Row 6 doesn’t have an entry in the Title column
- Sterms [Search Terms] field barely used
- Subject field is barely used (probably less than 10 entries
- Rows 72-76 are missing info in the ImageFile column (possibly more)

Lib Records

- Creator field infrequently used
- CallNo field infrequently used
- CatDate field infrequently used
- CatBy field infrequently used
- Date field infrequently used
- ImageFile field infrequently used
- Language field infrequently used
- PhysDesc field infrequently used
- PubPlace field infrequently used
- Publisher field infrequently used

Object Records

- AccessNo field infrequently used
- Creater field infrequently used
- Row 1 and 2 of CatDate field are missing info
- CatBy field infrequently used
- Date field infrequently used
- Row 4 HomeLoc missing info (possibly more)
- ImageFile field infrequently used
- People field infrequently used

Photo Records

- CatBy field infrequently used
- Date field infrequently used
- Descrip field infrequently used
- HomeLoc field infrequently used
- ImageFile field infrequently used

The Data Cleanup Summary

The PP product is limited in its support of data cleanup tools or functions. As a result, the majority of data cleanup has to take place at the item-level record and progress record by record. With this in mind, it's critically important to spend time analyzing the data, identifying areas of cleanup, prioritizing cleanup tasks, and creating a strategy before any major cleanup effort is made in PP.

This data cleanup plan will outline how DCHS staff can effectively conduct data remediation of the 559 online records and will include coverage of the cleanup process, strategies for approach, suggested prioritization, possible tools or strategies to use, and offer any additional advice to help ensure a smooth and supportive cleanup process.

The Cleanup Strategy

The record-by-record cleanup strategy is the strategy we selected to inform cleanup actions. This strategy was chosen because PP doesn't support data changes via import of a spreadsheet, nor support any data cleanup tools. Record-by-record cleanup can be a slow and sometimes frustrating process, so it's incredibly important to develop a data cleanup plan that breaks the data cleaning tasks into achievable sections. In order for this work to be sustainable, a doable pace will need to be established.

Cleanup Strategies: *For a refresher on data cleanup strategies please see the "A Review of Data Cleanup Strategies," and the "Practical Application of Data Cleanup Strategies" sections of A Guide to Collection Data Cleanup.*

Leverage a Spreadsheet for Data Work Pre-PastPerfect

While data must be directly entered into PP, it's often easier to gather the information, review it, and create the data ahead of time. This allows time for any research or file pulling that might need to happen, and it can help with data quality control. By entering or editing the data into the spreadsheet first you'll be able to see patterns more easily and benefit from data already created.

For example, many collection items share affinities such as the same creator (Creator field), descriptive details (Description field), and subjects (Subject field). Using a spreadsheet as your template will allow you to easily copy and paste relevant data and help to ensure you're providing a consistent level of detail for each related entry. Additionally, spellcheck can be ran across the entirety of entries, and a secondary review can be conducted efficiently with the spreadsheet before any (potentially erroneous) data enters PP.

Essentially, using a spreadsheet as your draft template helps to save you time and supports quality control of the data before it ever touches the database. Finally, it speeds up the eventual data entry time as the data can be transposed (copy/pasted) into each record with full confidence.

Spreadsheet Template Summary of Benefits:

- Allows you time to gather the information you need.
- Supports the easy comparison of multiple records and makes patterns easier to spot.
- Facilitates the copy/paste function of shared data elements across records, lessening the need to recreate repeat data and helps to ensure consistency.
- Helps to maintain correct field use with each field present as a column header in the spreadsheet, and reenforced with each entry correctly distributing data into each field.
- Provides further quality control mechanisms such as spell check and review of entries before data ever enters the system.
- Speeds up the eventual data entry process into PP.

Approaches to Cleanup

When approaching the spreadsheet for pre-PP work or once ready to enter data into PP, the following two approaches can be selected and used:

Approach A: Fill out each record and strive for record completeness. This looks like working in each record entry and filling out or editing data in each of the required fields. This approach is for brains that like to fully complete one item (record) before they move onto another.

Approach B: Pick a field and work on that field exclusively across all of the records. Then pick the next field and repeat. This approach is for brains that like to see quick progress and the ability to focus on one piece of data at a time.

Either approach is effective and with limited downsides so it really is up to you and how your brain likes to work!

Prioritization and Workflow

Approach A Prioritization

Prioritization can look like selecting a batch of records that are the most interesting and/or most important records. By selecting a prioritizing criterion, you can focus your efforts on records that are immediately beneficial to the heritage organization. This also helps to keep the scope small and achievable, which is an important practice when working with hundreds (and sometimes thousands) of records. Given the content of the PP Online records, we recommend that the priority batches are created based on the shared location or event depicted in the photographs. This will allow the cataloger an opportunity to build consistent, quality descriptions.

Approach A Workflow Steps

1. Select your criteria. Potential areas of focus could be a portion of the collection that is highly request, could aid in an upcoming project, or is “hidden” and could benefit from improved data.
2. Narrow the scope of the records you intend to look at and break it into batches of no more than 100 records (aka items) at a time.
3. Run a query in the system that will pull up the records that match your criteria and select the first 100 item records to export into a spreadsheet file format.
4. Review the spreadsheet and remove any entries that meet best practices and don’t need data cleanup. Then highlight each field (column) where every record entry remaining could use further review and data cleanup or enhancement.
5. Make your way through each record (represented as a row in the spreadsheet) and cleanup or fill-in missing data. As you complete an entry, fill the row in a light gray color so that you can easily track which entry you’re working on and what’s been completed thus far.

6. Once you've completed your work on this set, transfer the cleaned or enhanced data into PP by pulling up each record and copy/pasting the corrected data into the system. (Unfortunately, PP doesn't support data import for cleanup purposes).

Approach B Prioritization

Prioritization is inherent in this approach as you can determine the fields you work on first as priority fields. These can be fields that are the most impactful to good search results such as Date, Description, or Subject. Selecting fields by this prioritization will help ensure your work is immediately focused on the most impactful areas of the data. The following sections titled *An Assessment of the Data* and *An Assessment on the Completeness of Record*, offers a recommendation and priority level for each of the required fields. They are listed here for reference:

Existing Data Priority Areas:

- Object Name (High)
- Description (High)
- Collection (Low)

Absent Data Priority Areas:

- Identifier (Low)
- Publisher (Low)
- Title (High)
- Creator (Medium)
- Date (High)
- Format (and Dimensions) (Low)
- Language (Low)
- Subject (High)
- Rights (Medium)
- Relation (Medium)
- File Name (Low)

Approach B Workflow Steps

1. Select your priority area of focus aka the priority field.
2. Narrow the scope of the records you intend to look at and break it into batches of no more than 100 records (aka items) at a time. For this approach we can start at the beginning of your accession numbers.
3. Run a query in the system that will pull up the records that match your criteria and select the first 100 item records to export into a spreadsheet file format.
4. Review the spreadsheet and remove any entries that meet best practices for your field of focus and don't need data cleanup. Then highlight the priority field (column) to focus on for review and data cleanup or enhancement.
5. Make your way through each record (represented as a row in the spreadsheet) and cleanup or fill-in missing data for the priority field. As you complete an entry, fill the row

in a light gray color so that you can easily track which entry you're working on and what's been completed thus far.

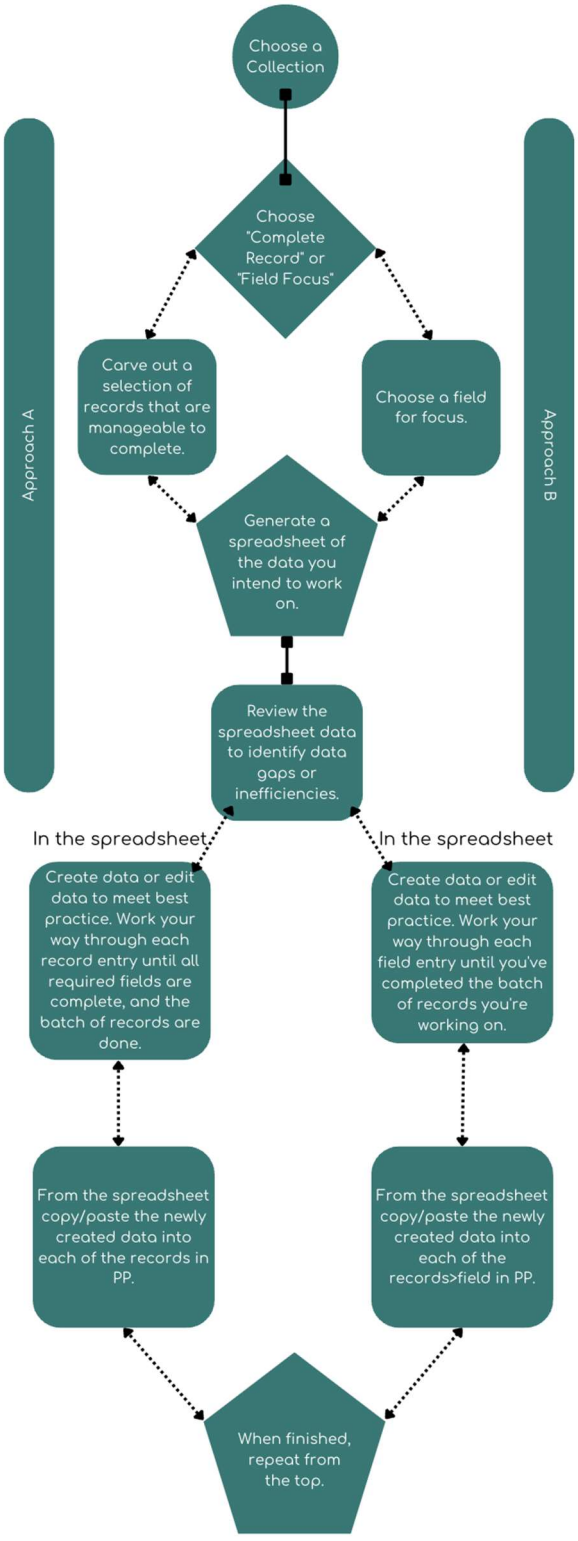
6. Once you've completed your work on this set, transfer the cleaned or enhanced data into PP by pulling up each record and copy/pasting the corrected data into the appropriate field in the system. (Unfortunately, PP doesn't support data import for cleanup purposes).

Working with a team? Regardless of which approach you use, follow the same steps but instead make sure the spreadsheet being worked on is a shared spreadsheet using OneDrive, GDrive, Dropbox or similar document sharing and collaboration tool. Make sure everyone follows step #5 and grays out the rows they've completed so that work isn't duplicated.

Or, you can provide each member of the team (volunteers included) their own spreadsheet with set of 100 to work on.

If the other team members have access to and know how to correctly enter their data in to PP, then they can complete step #6 on their own. If not, then have the team submit their completed spreadsheets to you for PP input.

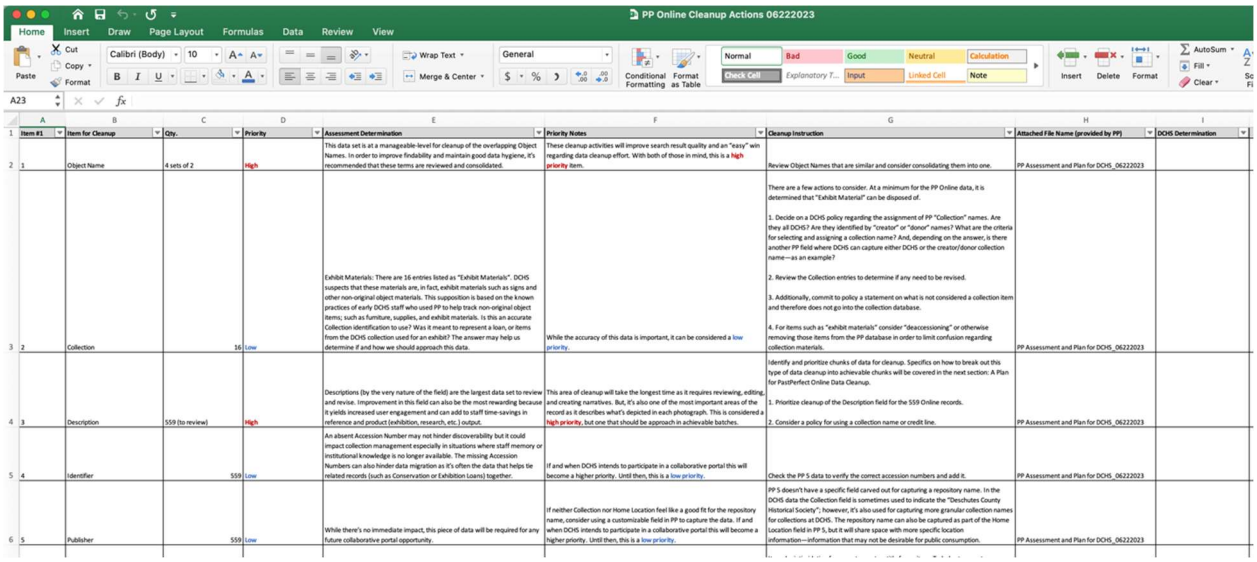
Suggested Workflows – A Visual



An Assessment of the Data

The following sections are an assessment of the data contained in the four fields used in DCHS's 559 PP Online records: Object ID, Object Name, Collection, and Description. Each of these fields has a corresponding descriptive standard to refer to and has helped to inform this assessment.

For Brains That Like Spreadsheets: *If you prefer to see this information in spreadsheet format, please see PP Online Cleanup Actions 06222023.xlsx document on the Oregon Heritage Commission website https://www.oregon.gov/oprd/OH/pages/technicalresources.aspx#datacleanup*



PP Online Cleanup Actions 06222023

	Item #1	Item for Cleanup	Qty.	Priority	Assessment Determination	Priority Notes	Cleanup Instruction	Attached File Name (provided by PP)	DCHS Determination
2	1	Object Name	4 sets of 2	High	This data set is at a manageable-level for cleanup of the overlapping Object Names. In order to improve findability and maintain good data hygiene, it's recommended that these terms are reviewed and consolidated.	These cleanup activities will improve search result quality and an "easy" win regarding data cleanup effort. With both of those in mind, this is a high priority item.	Review Object Names that are similar and consider consolidating them into one.	PP Assessment and Plan for DCHS_06222023	
3	2	Collection	16	Low	Exhibit Materials: There are 16 entries listed as "Exhibit Materials". DCHS suspects that these materials are, in fact, exhibit materials such as signs and other non-original object materials. This supposition is based on the known practices of early DCHS staff who used PP to help track non-original object items; such as furniture, supplies, and exhibit materials. Is this an accurate Collection identification to use? Was it meant to represent a loan, or items from the DCHS collection used for an exhibit? The answer may help us determine if and how we should approach this data.	While the accuracy of this data is important, it can be considered a low priority.	There are a few actions to consider. At a minimum for the PP Online data, it is determined that "Exhibit Material" can be disposed of. 1. Decide on a DCHS policy regarding the assignment of PP "Collection" names. Are they all DCHS? Are they identified by "creator" or "donor" names? What are the criteria for selecting and assigning a collection name? And, depending on the answer, is there another PP field where DCHS can capture either DCHS or the creator/donor collection name—as an example? 2. Review the Collection entries to determine if any need to be revised. 3. Additionally, commit to policy a statement on what is not considered a collection item and therefore does not go into the collection database. 4. For items such as "exhibit materials" consider "deaccessioning" or otherwise removing those items from the PP database in order to limit confusion regarding collection materials.	PP Assessment and Plan for DCHS_06222023	
4	3	Description	559 (to review)	High	Descriptions (by the very nature of the field) are the largest data set to review and revise. Improvement in this field can also be the most rewarding because it yields increased user engagement and can add to staff time-savings in reference and product (exhibition, research, etc.) output.	This area of cleanup will take the longest time as it requires reviewing, editing, and creating narratives. But, it's also one of the most important areas of the record as it describes what's depicted in each photograph. This is considered a high priority, but one that should be approach in achievable batches.	Identify and prioritize chunks of data for cleanup. Specifics on how to break out this type of data cleanup into achievable chunks will be covered in the next section: A Plan for PastPerfect Online Data Cleanup. 1. Prioritize cleanup of the Description field for the 559 Online records. 2. Consider a policy for using a collection name or credit line.	PP Assessment and Plan for DCHS_06222023	
5	4	Identifier	559	Low	An absent Accession Number may not hinder discoverability but it could impact collection management especially in situations where staff memory or institutional knowledge is no longer available. The missing Accession Numbers can also hinder data migration as it's often the data that helps tie related records (such as Conservation or Exhibition Loans) together.	If and when DCHS intends to participate in a collaborative portal this will become a higher priority. Until then, this is a low priority.	Check the PP 5 data to verify the correct accession numbers and add it.	PP Assessment and Plan for DCHS_06222023	
6	5	Publisher	559	Low	While there's no immediate impact, this piece of data will be required for any future collaborative portal opportunity.	If neither Collection nor Home Location feel like a good fit for the repository name, consider using a customizable field in PP to capture the data. If and when DCHS intends to participate in a collaborative portal this will become a higher priority. Until then, this is a low priority.	PP 5 doesn't have a specific field carved out for capturing a repository name. In the DCHS data the Collection field is sometimes used to indicate the "Deschutes County Historical Society"; however, it's also used for capturing more granular collection names for collections at DCHS. The repository name can also be captured as part of the Home Location field in PP 5, but it will share space with more specific location information—information that may not be desirable for public consumption.	PP Assessment and Plan for DCHS_06222023	

Object ID

For the bulk of the PP Online data the numbering appears in this format: YYYY.NNN.NNNN

Y = Year and N = Number

However, there are variations. Those variations are:

YYYY.NNNX.NNNN (in this case with a literal "X" inserted); example: 1999.029X.0043

YYYY.NNN.NNNNL and YYYY.NNNX.NNNNL (L = letter, and it's added at the end); examples: 2003.000.0226A and 1992.018X.0001A

Letters

Sometimes "A" and "B" is used to refer to the front and back of image. Other times A, B, and other letters of the alphabet are used. The current practice at DCHS is to use letters (in alphabetical order) to indicate multiple parts of one unit.

Example: A letter with an envelope. The letter receives an “A” at the end of its Object ID, and the envelope receives a “B”.

A few Object IDs from the 1980s are: YYYY.NNN.NNNN.NNN. These numbers are a combination of “Object ID” number and the “Other Number”.

Finally, there’s: YYYY.DHS.NNNN; example: 2011.DHS.0141. The “DHS” is used in reference to the Deschutes County Historical Society.

Determination: While the varied numbering system has some quirks, it ultimately doesn’t prohibit (to a detrimental extent) the usability of the data and the findability of the item records. With this in mind, it’s recommended that this data is left as-is. While it’s important to document the different numbering schemas used throughout the years, the data doesn’t need to be updated. The only exception to this determination is if there are duplicates of Object ID numbers being used as that can impact the integrity of the data with risk of conflating two separate items.

Action: No action.

Priority: N/A

Object Name

Across the 559 records there are the following terms found in the Object Name (OBJNAME) field:

OBJNAME List

- File, digital (2)
- Image, digital (13)
- Lithograph (1)
- Negative, Film (24)
- Negative, film (22)
- Photocopy (1)
- Photograph, Cabinet (1)
- Postcard (1)
- Postcard, photo (73)
- Print, photographic (119)
- Print, Photographic (301)
- Tintype (1)

There’s some consolidation that can occur in the OBJNAME field. For example, the sets listed below are most likely duplicates either due to a capitalization difference or a slight variation on identification.

File, digital (2)
Image, digital (13)

Negative, Film (24)
Negative, film (22)

Postcard (1)
Postcard, photo (73)

Print, photographic (119)
Print, Photographic (301)

Determination: This data set is at a manageable-level for cleanup of the overlapping Object Names. In order to improve findability and maintain good data hygiene, it's recommended that these terms are reviewed and consolidated.

Action: Review Object Names that are similar and consider consolidating them into one.

Priority: These cleanup activities will improve search result quality and are an "easy" win regarding data cleanup effort. With both of those in mind, this is a high priority item.

Collection

There are six records assigned to the Deschutes Pioneers' Association[9]:

- 1980.050.0059.001
- 1980.050.0059.003
- 1980.050.0106.040
- 1980.050.0106.067
- 1980.050.0106.089
- 1980.050.0108.011

DCHS would like to retain what came over as Deschutes Pioneers' Association.

Exhibit Materials

Additionally, there are 16 entries listed as "Exhibit Materials". DCHS suspects that these materials are exhibit materials such as signs and other non-original object materials. This supposition is based on the known practices of early DCHS staff who used PP to help track non-original object items; such as furniture, supplies, and exhibit materials. Is this an accurate Collection

[9] In the early 1940s or 1950s the Deschutes Pioneers Association was established and operated as a small history museum for the area. When DCHS was created in 1975 there was a coalescing of efforts and the Association transferred their collection holdings to DCHS. The Association continued to exist as an organization until January 2022, when the Association dissolved and transferred their remaining assets to DCHS.

identification to use? Was it meant to represent a loan, or items from the DCHS collection used for an exhibit? The answer may help us determine if and how we should approach this data.

Collection Names

In the larger PP 5 data sets there is a variation in Collection name assignment. Some are a variation of DCHS. Others are names from the creator or donor of the collection. Either is appropriate, but should be definitively decided and formalized into DCHS policy.

Determination: The PP Online data set has minimal cleanup needed, but does open the door to policy questions for all DCHS data.

Action: There are a few actions to consider. At a minimum for the PP Online data, it is determined that "Exhibit Material" can be disposed of.

1. Decide on a DCHS policy regarding the assignment of PP "Collection" names. Are they all DCHS? Are they identified by "creator" or "donor" names? What are the criteria for selecting and assigning a collection name? And, depending on the answer, is there another PP field where DCHS can capture either DCHS or the creator/donor collection name—as an example?

2. Review the Collection entries to determine if any need to be revised.

3. Additionally, commit to policy a statement on what is *not* considered a collection item and therefore does *not* go into the collection database.

4. For items such as "exhibit materials" consider "deaccessioning" or otherwise removing those items from the PP database in order to limit confusion regarding collection materials.

Priority: While the accuracy of this data is important, it can be considered a low priority.

Description

The descriptions available across the 559 records vary in length, style, and quality. The description field is typically the largest field and takes the longest time to construct and review. While having any data can meet the minimum data schema requirements, many of the entries conflict with descriptive standards. The following data quality issues were found in these records:

- Abbreviations
- Fragmented sentences
- A list of items in the photograph instead of a narrative description
- Inconsistency in length or quality of content provided across similar image sets
- Informal notes on where the item was found
- Concatenated data

- Referencing item numbers that don't appear to be captured in the PP Online records
- Informal commentary on the quality of the item
- Donor notes that may not be appropriate for inclusion
- Provision of an entire history of the photograph's point of interest versus a focus on what the photograph has captured
- Grammar and spelling errors (throughout)

Abbreviations

Abbreviations should be fully spelled out for both consistent quality and to support effective keyword searching across multiple fields (including the Description field).

Examples:

- "photo" for photograph
- "B&W" for "black and white"
- "Hwy" for highway
- "Lbr Co" presumably lumber company?

Fragmented Sentences

Fragmented sentences make descriptions difficult for general users to read and limit the multi-purpose use on behalf of the museum. Descriptions are used to help inform exhibits, research, and curriculum (to name a few examples) and the quality of description can impact the usability of the data.

Example: *George Bradetich place: color photo of horse drawn equipment, potato harvest. Potatoes on ground, in mesh buckets, in bags.*

Corrected Example: This is the George Bradetich residence during potato harvest. This color photograph depicts horse drawn equipment and potatoes on the ground, in mesh buckets, and in bags.

Lists

Other descriptions appear to be more list-like rather than descriptive. Some entries don't even contain a "." at the end of the sentence. This may not impact searchability, but it does impact user interpretation and experience.

Example: *Pilot Butte Inn: lawn, patio, tables, umbrellas, director's chairs*

Corrected Example: This scene was taken at Pilot Butte Inn. Pictured in the foreground is the patio and surrounding lawn, complete with tables, umbrellas, and director's chairs.

Inconsistent Length or Quality Among Related Image Sets

The majority of the descriptions are fairly modest—which is both fine from a best practices perspective, as well as understandable from a staff capacity perspective. However, keep in mind the balance and consistency of longer descriptions. For example, this longer description is also connected to Pilot Butte and contains a good chunk of information that should be considered for inclusion in all of the Pilot Butte item records. As many of these images are related, it may be helpful to view the descriptions all together and then standardize the information that is provided in each image.

SW Corner of Wall and Newport with Pilot Butte Inn seen on right edge of photo. Wall Street in Bend in 1904, showing post office. In 1885, a filing was proposed for a post office to be called Farewell Bend. The post office was established January 18, 1886 with only the name of Bend allowed as there was already a post office of the former name in a community along the Snake River. In 1904 the Bend Post Office was located at the southwest corner of Wall and Newport. The corner of the Pilot Butte Inn porch appears on the right.

Informal Notes on Where the Item Was Found

This description includes an informal note on how the item was found. In previous cataloging practices, catalogers opted to add the majority of collection information in the Description field instead of using more specific and appropriate fields. This type of note is more appropriate as an internal note as it serves no findability or usability purpose to those external to DCHS.

Group photo of children,adults, and dog in woods. Possibly logging camp residents. Backing says, "Rowena M. Hogan 121 Oregon St. Bend, Ore." From The Hogans Photo Studio (JJ and Rowena M Hogan).. portrait and framing. Framed photo was found in back cupboard in Historical Center 05/2003

Concatenated Data

This and a few other entries appear to have concatenated the data in a way that has duplicated the content.

Referencing Item Numbers that Don't Exist

This example also references a related item with No. 74. As this number doesn't appear in any of the PP Online data, how do we know what this number is referring to? If a connection can be made then it may be worth keeping this data. If not, then this data should be removed to limit future confusion.

Wagons; Grain; Harvesting; Horses; Farm Equipment. Farm wagons with horses, grain harvesting, 1919; farm equipment. Harvesting, 1919. Farm equipment in the field and horses with wagons. Grain harvesting, 1919. Farm equipment in the field and horses with wagons. Horses with wagons, farm equipment, involved with grain harvesting, 1919. Harvesting grain, 1919. Horses with wagons shown. Copied by Don Ellis from Roy Newell Collection, 1980. Print 3.5" x 5"

Same as No. 74: Railroad Day, Downtown Redmond Sept 30, 1911.

Informal Commentary on Quality

This example includes references to a related item and number, and it offers informal (subjective) commentary not consistent with descriptive best practices.

Big snow, 1919, downtown Redmond, Oregon; same as No. 104, but better.

Donor Notes

Unless specifically stated in the donor's Deed of Gift that public acknowledgement of the gift should be made in the catalog, donation information is otherwise internal-only information. This is an area DCHS can develop policy around. For example, there can be inclusion of a field to capture a collection name or a credit line.

Brooks Scanlon Lumber Co. Logging Camp #1, Wet Weather Springs. Camp families pose in Sunday best, with camp homes, logging office and 1920's automobiles in background. Donated by Delmar Davis. Related to photo#1980.050.0059.001.

Provision of Entire History

This description captures an entire history which is unusual for an item-level description.

St. Charles Hospital, downtown Bend, Shrine donated by Klondike Kate (Kate Rockwell), 1937. Article published in Bend Bulletin December 9, 1937 states that the grotto was constructed by Ray Williams and "the impressive grotto, some 15 feet high and holding a statue of St. Joseph and the Christ child, has been built against the new wing of the hospital". According to the article, Klondike Kate (referred to here as Mrs. John Matson) donated the grotto after making a promise to Sister Louise years before the installation. All stones used were collected by Kate, outside of a few purchased from local rock collectors. In another article from July 1, 1957, it was reported that the St. Joseph statue in the grotto was smashed. "A statue in a niche of gemstones at St. Charles Memorial Hospital was smashed Friday afternoon and police were called on to investigate. The state was that of St. Joseph and was given to the hospital a number of years ago by the late "Klondike Kate", Mrs. W.L. Van Duren. The cemented stones forming thje acover were not damaged, but the plaster of which the statue was made was broken into fragments. The vandilism occurred Friday afternoon, possibly around 3 pm. The statue was in front of the old hospital unit. Two men were seen on the hospital grounds at the time and attendants indicated they may have been intoxicated. Possibility that the vandilism might be associated with the recent entry of St. Francis Catholic church and telephoned threats to the pastor were considered by officers. It was learned that Trinity Episcopal Church was entered at the same time that the Catholic church was burglarized. All candles on the Episcopal Church alter were lighted and burned out. Entry of the Episcopal church was not discovered for several days, the pastor being out of town."

Determination: Descriptions (by the very nature of the field) are the largest data set to review and revise. Improvement in this field can also be the most rewarding because it yields increased user engagement and can add to staff time-savings in reference and product (exhibition, research, etc.) output.

Action: Identify and prioritize chunks of data for cleanup. Specifics on how to break out this type of data cleanup into achievable chunks will be covered in the next section: *A Plan for PastPerfect Online Data Cleanup*.

1. Prioritize cleanup of the Description field for the 559 Online records.

2. Consider a policy for using a collection name or credit line.

Priority: This area of cleanup will take the longest time as it requires reviewing, editing, and creating narratives. But, it's also one of the most important areas of the record as it describes what's depicted in each photograph. This is considered a **high priority**, but one that should be approached in achievable batches.

An Assessment and Plan on the Completeness of Record

Data Schema Standards

Data schema standards outline which data elements (fields) should be present in a complete record. When comparing the current PP Online fields used to current data schema standards (Dublin Core™) we find the following fields absent:

- Identifier
- Publisher
- Title
- Creator
- Date
- Format (and Dimensions)
- Language
- Subject
- Rights
- Relation

The following section will outline an approach to data enhancement for missing or anemic data. It will describe the potential impact each field has on item management and use, share any relevant PP 5 data notes, and offer a recommendation with priority level assignment.

Identifier

Identifier: The identifier is an individual item number. For object-based collections it is typical for the Accession Number to also serve as the Object ID. However, anytime there's a multi-item accession—more typical of Photograph and Archives collections—there's the need to apply individual item numbers as well. These can be based on the Accession Number, but should always be unique.[10]

Example A: a tea set

The tea set is composed of dozens of pieces, but the set as acquired as a whole and accessioned as a whole, receiving just one official accession number: 20120706.11

If item-level or component-level cataloging needs to occur—as is typical for many cultural heritage institutions—then the accession number remains the same, but the component can receive its own object ID: 20120706.11.a.

Example B: a photograph collection

A photograph collection is comprised of 100s of photographs. The collection received an accession number: 20230827.01

If each photograph is cataloged at an item-level, the accession number (20230827.01) remains the same, but each photograph receives its own object ID, starting at 001, 002, and so on.

DHCS: The PP Online data captures the Object ID, but is lacking the Accession Number.

Potential Impact: An absent Accession Number may not hinder discoverability but it could impact collection management especially in situations where staff memory or institutional knowledge is no longer available. The missing Accession Numbers can also hinder data migration as it's often the data that helps tie related records (such as Conservation or Exhibition Loans) together especially if the CMS-assigned identifier (e.g. record ID) is impossible to migrate with the related data.

PP 5 Data: In reviewing sample Photograph collection data in PP 5, it appears both Accession Number and Object ID was used in nearly every entry.

Recommendation: If and when DCHS intends to participate in a collaborative portal this will become a higher priority. Until then, this is a low priority.

[10] Unique numbers are required in order to avoid accidental merging or conflating of two separate items.

Publisher

Publisher: The publisher field is intended to capture the person or organization who is publishing these materials to the catalog. In other words, the repository name. While this may seem obvious when cataloging material that all belong to the heritage organization, it's a great practice to get into as it will be necessary when contributing collection content to a larger collaborative portal—such as Northwest Digital Heritage or Digital Public Libraries of America.

Potential Impact: While there's no immediate impact, this piece of data will be required for any future collaborative portal opportunity.

PP 5 Data: PP 5 doesn't have a specific field carved out for capturing a repository name. In the DCHS data the Collection field is sometimes used to indicate the "Deschutes County Historical Society"; however, it's also used for capturing more granular collection names for collections at DCHS. The repository name can also be captured as part of the Home Location field in PP 5, but it will share space with more specific location information—information that may not be desirable for public consumption.

Recommendation: If neither Collection nor Home Location feel like a good fit for the repository name, consider using a customizable field in PP to capture the data. If and when DCHS intends to participate in a collaborative portal this will become a higher priority. Until then, this is a low priority.

Title

Title: The title of the item being cataloged. This field is particularly helpful (and a required field) in Archives and Photograph collections because it can convey key information on the item.

It can be intimidating for some to create a title for an item. To help steer past any roadblocks consider the following prompts:

- What's the most important thing about this photograph?
- Who or what is the focus of the photograph?
- What event is being captured in the photograph?

Examples:

- The Exterior of Pilot Butte Inn Featuring the Patio
- Mrs. Jane Smith with Juniper Elementary School 1st Graders
- A Harvesting Scene with Farm Equipment, Grain, and Horses

Potential Impact: Titles serve to increase discoverability of materials and their use. Titles are one of the first (and sometimes only) fields a user reads before they make an adjudication on the item's usefulness to their search and move on.

PP 5 Data: In reviewing the PP 5 records there appears to be Titles for approximately 40% of the sample size. A comparison of the PP 5 records with the 559 PP Online records may help with completing any data gaps or offer examples of past titles used.

Recommendation: Given the immediate attention a title can command when reviewing items in search results, this is recommended as a high priority area to focus on.

Creator

Creator: The creator is most often a controlled vocabulary- can be local, an authority, or both.[11] In the case where a creator is not known, DCHS has elected to confirm this with the use of "Unknown" in the Creator field.[12]

Examples:

- Unknown
- Smith, Jane
- The Bend Bulletin

Remember: Use the "People" classification in PP to establish the correct names of Creators so that they can be searched for and applied consistently.

Potential Impact: Capturing creator names help provide contextual information to both collection workers and collection users. Using the Creator field provides a tie across collection items that share a creator and can offer insight into the inter-relationship of the collection.

PP 5 Data: There is Creator field data populated in the PP 5 database for the Photograph collection. A comparison of the PP 5 records with the 559 PP Online records may help with completing any data gaps.

Recommendation: This is a helpful field but not quite as critical as Date or Subject. It's a medium priority field.

[11] In PartPerfect 5 this is the "Artist/Author/Creator/Photographer" Authority File though when viewing the data via spreadsheet, the column (field) name is "Creator".

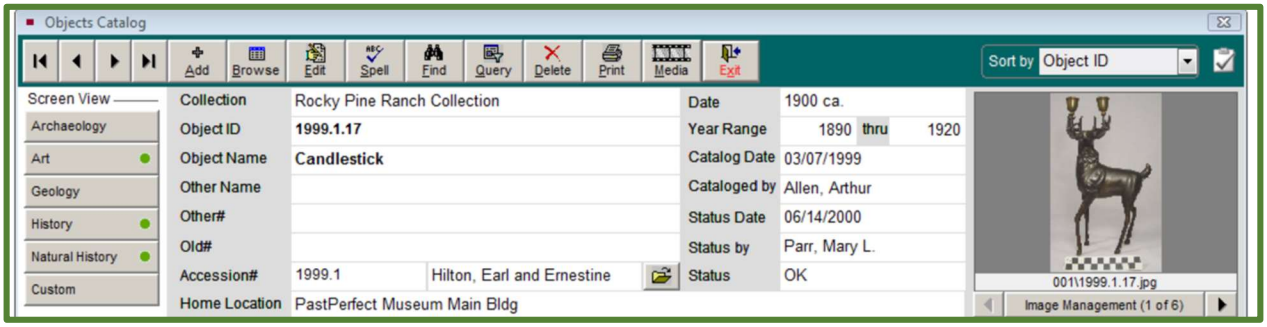
[12] The use of Unknown in the creator field of any database can depend on the descriptive standards followed and individual heritage organization policy. Some prefer to leave the field blank because they consider the use of "Unknown" as junk data. Those who elect to use Unknown typically choose to do so in order to definitively capture that the creator is truly unknown versus an omission of detail as it can lead to future questions and (ultimately) dead end research.

Date

Date: The creation date of the object. Many of the descriptions capture a date or era and can be used to populate the Date field.

In PP there are two date[13] capture fields:

- Date: free text can capture circa, month names, etc.
- Year Range: YYYY-YYYY



Examples to Use in Date:

- August 15, 2010
- August 2010
- 2010
- Summer 1964
- circa 1990s

Examples to Use in Year Range:

- 1890-1910 (if spanning multiple items; e.g. an archival collection)

Examples Not to Use in Date:

- 1919?
  - Correct: circa 1919
- c. 1941
  - Correct: circa 1941
- ca. 1941
  - Correct: circa 1941

[13] PastPerfect Date Style: PastPerfect supports a variety of date formats, including ANSI (2009.04.28), American (04/28/2009), British (28/04/2009), German (28.04.2009), Italian (28-04-2009), Japan (2009/04/28), or USA (04-28-2009). Since dates are stored as Julian dates, changing the date style will reformat the display and printing of all new and previously entered dates. Read more at https://museumsoftware.com/WebHelp/Chapters/PP5-3f.htm, accessed May 26, 2023.

- 1947-48
    - Correct: 1947-1948
- 09/00/1915
    - Correct: September 1915
- 1940s
    - Correct: 1940-1949 (if multiple items, then put in Year Range); or circa 1940s
- Circa 1925
    - Correct: circa 1925
- Aug-42
    - Correct: August 1942
- possibly 1937
    - Correct: circa 1937

Potential Impact: Date is one of the top fields we (and our audiences) use to help limit materials we're searching for and intend to work with. With that in mind, any date provided, even an estimation based on contextual information, can be incredibly helpful.

PP 5 Data: There is Date field data populated in the PP 5 database for the Photograph collection. A comparison of the PP 5 records with the 559 PP Online records may help with completing any data gaps.

Recommendation: It's recommended that when using an approximation, pair it with a year range. Not all systems are able to recognize "circa 1910" as an acceptable record for search results when someone wants an item created between 1905 and 1920. Given the potential helpfulness this data lends to item discoverability, this is a high priority area to focus on.

Format

Format (use of Nomenclature for a controlled vocabulary): For Archives and Photograph collections, this is where we would typically expect to see what is currently in Object Name; such as "Print, Photographic". DCHS may also elect to use the Materials vocabulary. Dimensions can also be included as part of the Format for data entry. However, in PP there is a separate Dimensions field to use.

Potential Impact: While format is an important field to understanding the item being viewed, it's doesn't immediately impact management or use of the item.

Recommendation: This is a lower priority item and one to reconsider if and when DCHS wishes to participate in a collaborative portal—as facet searching based on format is a commonly used tool by our audiences.

Language

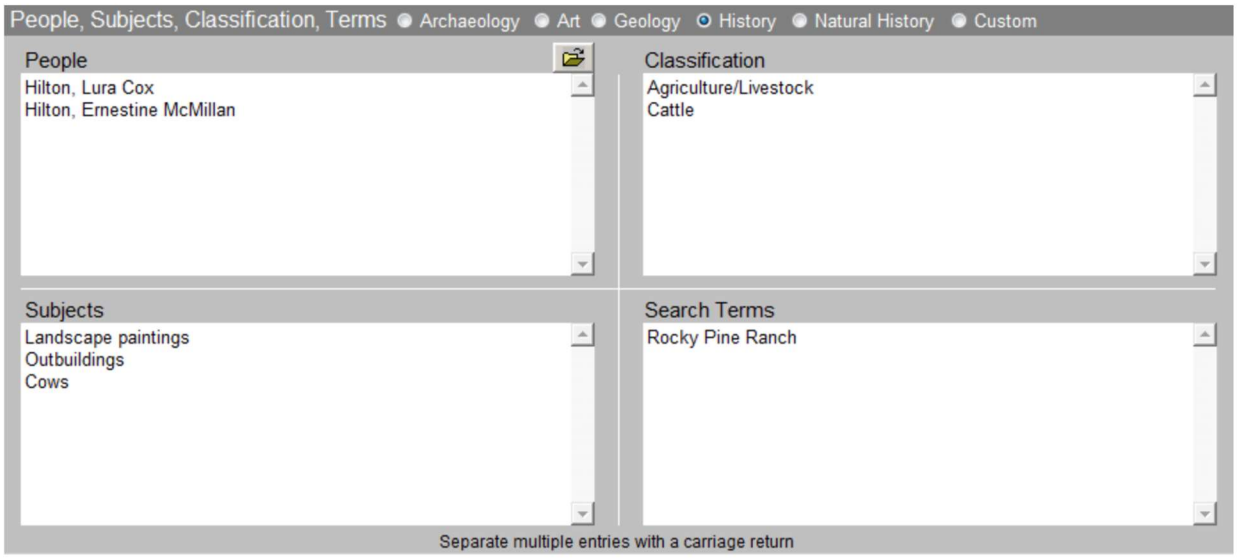
Language (when applicable): This field is most used in the Archives collection module; however, it can increase accessibility when used to indicate if there is written language present within the image.

Potential Impact: For heavily visual collections this is a "would be nice" requirement. It doesn't immediately impact discovery or use in a significant way and can wait until other high priority data areas have been addressed.

Recommendation: This is a low priority area.

Subject

Subject (use of Nomenclature for a controlled vocabulary): Subjects are used as search tools to indicate thematic relationships and meaning among items. There doesn't appear to be any Subject field use in the 559 PP Online records. In this case, it's suggested that DCHS refer to Nomenclature and utilize the People, Subjects, and Classification search terms to improve findability. Below is the corresponding PP subscreen:



Potential Impact: Subject and Search Terms are incredibly powerful tools to aid in item discovery and use. If there are no subjects or inconsistent subjects in a CMS then it makes the items harder to be found—relying heavily on Title (if it exists) and Description fields. No or inconsistent Subject or Search Term use also hinder the ability to browse the collection as there's no thru-line in the material tied together by subject matter.

PP 5 Data: When viewing a sampling of PP 5 material we can see the Subject and Search Term fields were used at one point, but not with controlled vocabulary.

Recommendation: This is a **high priority** field that can yield immediate benefits to the discovery and use of the collection.

Rights

Rights (any copyright or other retained rights): This field indicates to the external audience what rights are tied to this item and whether or not the item is eligible for use and what type of use. It also indicates to the user that there should be an appropriate citation if the item is used. This field can be found in the Notes, Legal, and Provenance sub-screen. The Web Rights field will most likely be the most common field to use; however, Legal can also be used if there are other legal or rights considerations beyond a DCHS copyright.

Potential Impact: This statement is a proactive risk mitigation tool to gently educate audiences that there are different rights to using collection content. The inclusion of this statement is of particular importance if the collections are to be published, made discoverable, and shared online. It makes it clear who the photograph belongs, if it's covered under copyright or CreativeCommons[14] license, or can be used with the stated credit line.[15] While the impact of this statement isn't immediate nor does it impact discoverability, it is an important piece of data to include for all collection items online.

Recommendation: Select a copyright, Creative Commons, or other statement approved by DCHS and apply it to all applicable items in the collection. This is a medium priority for collection items that are available online.

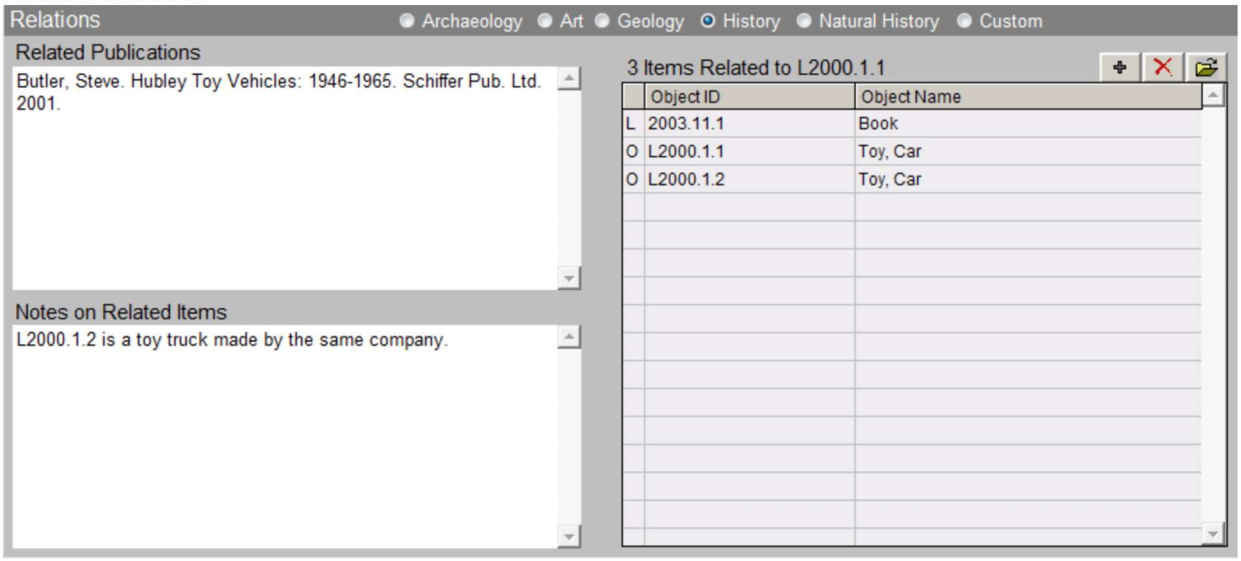
Relation

Relation (when applicable, indicate a relation to other known items in the DCHS or peer museum collection). Given that several of these records refer to other related items, it makes sense to use this field to indicate relationships. The Relation field can also be used to capture a relationship to items at a different heritage organization. *Note: There's not necessarily going to be a relationship to record for every item, so this field is indicated as required when applicable.*

[14] About Creative Commons Licenses, Creative Commons, accessed March 18, 2023, via https://creativecommons.org/about/cclicenses/.

[15] For example, using a credit line can attribute creatorship to a photographer or *The Bend Bulletin* while also including the repository (DCHS) name.

Relations subscreen



Potential Impact: Taking the time to indicate relationships among items is an investment that can have tremendous benefit. While Creator or Subjects can assist by inferring a relationship, the use of the Relation field can help indicate relationships among items that aren't as obvious as Creator or Subject. Using this field can save future time searching for future exhibits and research.

Recommendation: Brainstorm how relationships may span the collection and which items should indicate the other. This field is a medium priority.

Data Descriptive Standards

The remaining field not already included in the Dublin Core™ minimum data element set and is included in the CCO minimum fields to use set is: Current Location.

Current Location: Unless the photographs move frequently, this field is not helpful to add to the record. This field is most appropriate for the object collection.

Value-Added Fields to Consider

The following field isn't required for minimum best practice, but it does offer improvement in collection management and user engagement: File Name.

File Name (to reference for uploaded digital file): The image file name can serve as a helpful reference to quickly and easily access a digital file for exhibit or researcher use. It can also be immensely helpful in any future database migration when both data and digital files are migrated from one system to another and need to be matched back up.

Potential Impact: This field may not have immediate impact, but it will if the data and related digital files ever need to move to a different CMS platform, or be contributed to a collaborative portal. Knowing which record each digital file belongs to—and the actual file name as it exists outside of PP—can be critical to maintaining the pair.

PP 5 Data: In reviewing the data from PP 5 there does appear to be use of the IMAGEFILE field. A comparison of the PP 5 records with the 559 PP Online records may help with completing any data gaps for this field.

Recommendation: It's not an immediate priority, but it is incredibly important as cataloging continues. Start capturing the file name for attached digital files moving forward. As time allows, retroactively populate the file names in a custom filed that you can call upon later when you need it. This is a low priority for cleanup.

Potential Use of Existing Data

In the DCHS PP 5 data for the Photograph collection, it looks like many of the required fields are already present in the data. What's unclear is if the PP Online records have a carbon copy with a fuller dataset in PP 5. If that's the case it's always easier to use existing data, even if it needs some enhancement or editing.

If there is a fuller set of data in PP 5, then the following is recommended:

1. Run a query in PP5 from the Photograph collection for all photographs that transitioned to PP Online. (The PP Online data is a selection of photographs from 1979-2014).
2. Once the query is executed generate a spreadsheet.
3. Review the data in the spreadsheet to see if each required field is filled in. Highlight any required field that's empty.
4. Review data for quality issues such as incomplete sentences, inconsistencies in details, or factual errors.
5. Select which approach (A or B) you intend to take for data cleanup. Then follow the suggested cleanup prioritization for that approach.
6. When data augmenting is complete: run spellcheck.
7. Using the Object ID, work through each record entry and copy the refined data from your spreadsheet (informed in part from the PP 5) into the PP Online record.

A Guide to Creating Data Content

Please see the "An overview of Data Standards," section in *A Guide to Collection Data Cleanup.*

Item #1	Item for Cleanup	Qty.	Priority	Assessment Determination	Priority Notes	Cleanup Instruction
1	Object Name	4 sets of 2	High	This data set is at a manageable-level for cleanup of the overlapping Object Names. In order to improve findability and maintain good data hygiene, it's recommended that these terms are reviewed and consolidated.	These cleanup activities will improve search result quality and an "easy" win regarding data cleanup effort. With both of those in mind, this is a high priority item.	Review Object Names that are similar and consider consolidating them into one.
2	Collection	16	Low	Exhibit Materials: There are 16 entries listed as "Exhibit Materials". DCHS suspects that these materials are, in fact, exhibit materials such as signs and other non-original object materials. This supposition is based on the known practices of early DCHS staff who used PP to help track non-original object items; such as furniture, supplies, and exhibit materials. Is this an accurate Collection identification to use? Was it meant to represent a loan, or items from the DCHS collection used for an exhibit? The answer may help us determine if and how we should approach this data.	While the accuracy of this data is important, it can be considered a low priority.	There are a few actions to consider. At a minimum for the PP Online data, it is determined that "Exhibit Material" can be disposed of. 1. Decide on a DCHS policy regarding the assignment of PP "Collection" names. Are they all DCHS? Are they identified by "creator" or "donor" names? What are the criteria for selecting and assigning a collection name? And, depending on the answer, is there another PP field where DCHS can capture either DCHS or the creator/donor collection name—as an example? 2. Review the Collection entries to determine if any need to be revised. 3. Additionally, commit to policy a statement on what is not considered a collection item and therefore does not go into the collection database. 4. For items such as "exhibit materials" consider "deaccessioning" or otherwise removing those items from the PP database in order to limit confusion regarding collection materials.
3	Description	559 (to review)	High	Descriptions (by the very nature of the field) are the largest data set to review and revise. Improvement in this field can also be the most rewarding because it yields increased user engagement and can add to staff time-savings in reference and product (exhibition, research, etc.) output.	This area of cleanup will take the longest time as it requires reviewing, editing, and creating narratives. But, it's also one of the most important areas of the record as it describes what's depicted in each photograph. This is considered a high priority, but one that should be approach in achievable batches.	Identify and prioritize chunks of data for cleanup. Specifics on how to break out this type of data cleanup into achievable chunks will be covered in the next section: A Plan for PastPerfect Online Data Cleanup. 1. Prioritize cleanup of the Description field for the 559 Online records. 2. Consider a policy for using a collection name or credit line.
4	Identifier	559	Low	An absent Accession Number may not hinder discoverability but it could impact collection management especially in situations where staff memory or institutional knowledge is no longer available. The missing Accession Numbers can also hinder data migration as it's often the data that helps tie related records (such as Conservation or Exhibition Loans) together.	If and when DCHS intends to participate in a collaborative portal this will become a higher priority. Until then, this is a low priority.	Check the PP 5 data to verify the correct accession numbers and add it.

Item #1	Item for Cleanup	Qty.	Priority	Assessment Determination	Priority Notes	Cleanup Instruction
5	Publisher	559	Low	While there's no immediate impact, this piece of data will be required for any future collaborative portal opportunity.	If neither Collection nor Home Location feel like a good fit for the repository name, consider using a customizable field in PP to capture the data. If and when DCHS intends to participate in a collaborative portal this will become a higher priority. Until then, this is a low priority.	PP 5 doesn't have a specific field carved out for capturing a repository name. In the DCHS data the Collection field is sometimes used to indicate the "Deschutes County Historical Society"; however, it's also used for capturing more granular collection names for collections at DCHS. The repository name can also be captured as part of the Home Location field in PP 5, but it will share space with more specific location information—information that may not be desirable for public consumption.
6	Title	559	High	Titles serve to increase discoverability of materials and their use. Titles are one of the first (and sometimes only) fields a user reads before they make an adjudication on the item's usefulness to their search and move on.	Given the immediate attention a title can command when reviewing items in search results, this is recommended as a high priority area to focus on.	It can be intimidating for some to create a title for an item. To help steer past any roadblocks consider the following prompts: •What's the most important thing about this photograph? •Who or what is the focus of the photograph? •What event is being captured in the photograph? Examples: •The Exterior of Pilot Butte Inn Featuring the Patio •Mrs. Jane Smith with Juniper Elementary School 1st Graders •A Harvesting Scene with Farm Equipment, Grain, and Horses
7	Creator	559	Medium	Capturing creator names help provide contextual information to both collection workers and collection users. Using the Creator field provides a tie across collection items that share a creator and can offer insight into the inter-relationship of the collection.	This is a helpful field but not quite as critical as Date or Subject. It's a medium priority field.	The creator is most often a controlled vocabulary, can be local, an authority, or both. In the case where a creator is not known, DCHS has elected to confirm this with the use of "Unknown" in the Creator field. Examples: •Unknown •Smith, Jane •The Bend Bulletin Remember: Use the "People" classification in PP to establish the correct names of Creators so that they can be searched for and applied consistently.

Item #1	Item for Cleanup	Qty.	Priority	Assessment Determination	Priority Notes	Cleanup Instruction
8	Date	559	High	Date is one of the top fields we (and our audiences) use to help limit materials we're searching for and intend to work with. With that in mind, any date provided, even an estimation based on contextual information, can be incredibly helpful.	It's recommended that when using an approximation, pair it with a year range. Not all systems are able to recognize "circa 1910" as an acceptable record for search results when someone wants an item created between 1905 and 1920. Given the potential helpfulness this data lends to item discoverability, this is a **high priority** area to focus on.	The creation date of the object. Many of the descriptions capture a date or era and can be used to populate the Date field. In PP there are two date capture fields: •Date: free text can capture circa, month names, etc. •Year Range: YYYY-YYYY Examples to Use in Date: •August 15, 2010 •August 2010 •2010 •Summer 1964 •circa 1990s Examples to Use in Year Range: •1890-1910 (if spanning multiple items; e.g. an archival
9	Format	559	Low	While format is an important field to understanding the item being viewed, it's doesn't immediately impact management or use of the item.	This is a **lower priority** item and one to reconsider if and when DCHS wishes to participate in a collaborative portal—as facet searching based on format is a commonly used tool by our audiences.	For Archives and Photograph collections, this is where we would typically expect to see what is currently in Object Name; such as "Print, Photographic". DCHS may also elect to use the Materials vocabulary. Dimensions can also be included as part of the Format for data entry. However, in PP there is a separate Dimensions field to use.
10	Language	559	Low	For heavily visual collections this is a "would be nice" requirement. It doesn't immediately impact discovery or use in a significant way and can wait until other high priority data areas have been addressed.	This is a **low priority** area.	When applicable: This field is most used in the Archives collection module; however, it can increase accessibility when used to indicate if there is written language present within the image.
11	Subject	559	High	Subject and Search Terms are incredibly powerful tools to aid in item discover and use. If there are no subjects or inconsistent subjects in a CMS then it makes the items harder to be find—relying heavily on Title (if it exists) and Description fields. No or inconsistent Subject or Search Term use also hinder the ability to browse the collection as there's no thru-line in the material tied together by subject matter.	This is a **high priority** field that can yield immediate benefits to the discovery and use of the collection.	Subjects are used as search tools to indicate thematic relationships and meaning among items. There doesn't appear to be any Subject field use in the 559 PP Online records. In this case, it's suggested that DCHS refer to Nomenclature and utilize the People, Subjects, and Classification search terms to improve findability.

Item #1	Item for Cleanup	Qty.	Priority	Assessment Determination	Priority Notes	Cleanup Instruction
12	Rights	559	Medium	This statement is a proactive risk mitigation tool to gently education audiences that there are different rights to using collection content. The inclusion of this statement is of particular importance if the collections are to be published, discoverable, and shared online. It makes it clear who the photograph belongs, if it's covered under copyright or CreativeCommons license, or can be used with the stated credit line. While the impact of this statement isn't immediate nor does it impact discoverability, it is an important piece of data to include for all collection items online.	Select a copyright, Creative Commons, or other statement approved by DCHS and apply it to all applicable items in the collection. This is a medium priority for collection items that are available online.	This field indicates to the external audience what rights are tied to this item and whether or not the item is eligible for use and what type of use. It also indicates to the user that there should be an appropriate citation if the item is used. This field can be found in the Notes, Legal, and Provenance sub-screen. The Web Rights field will most likely be the most common field to use; however, Legal can also be used if there are other legal or rights considerations beyond a DCHS copyright.
13	Relation	559	Medium	Taking the time to indicate relationships among items is an investment that can have tremendous benefit. While Creator or Subjects can assist by inferring a relationship, the use of the Relation field can help indicate relationships among items that aren't as obvious as Creator or Subject. Using this field can save future time searching for future exhibits and research.	Brainstorm how relationships may span the collection and which items should indicate the other. This field is a medium priority.	Given that several of these records refer to other related items, it makes sense to use this field to indicate relationships. The Relation field can also be used to capture a relationship to items at a different heritage organization. Note: There's not necessarily going to be a relationship to record for every item, so this field is indicated as required when applicable.
14	File Name	559	Low	This field may not have immediate impact, but it will if the data and related digital files ever needs to move to a different CMS platform, or be contributed to a collaborative portal. Knowing which record each digital file belongs to—and the actual file name as it exists outside of PP—can be critical to maintaining the pair.	It's not an immediate priority, but it is incredibly important as cataloging continues. Start capturing the file name for attached digital files moving forward. As time allows, retroactively populate the file names in a custom filed that you can call upon later when you need it. This is a low priority for cleanup.	In reviewing the data from PP 5 there does appear to be use of the IMAGEFILE field. A comparison of the PP 5 records with the 559 PP Online records may help with completing any data gaps for this field.

APPENDIX 9 - RESOURCES TO REFERENCE

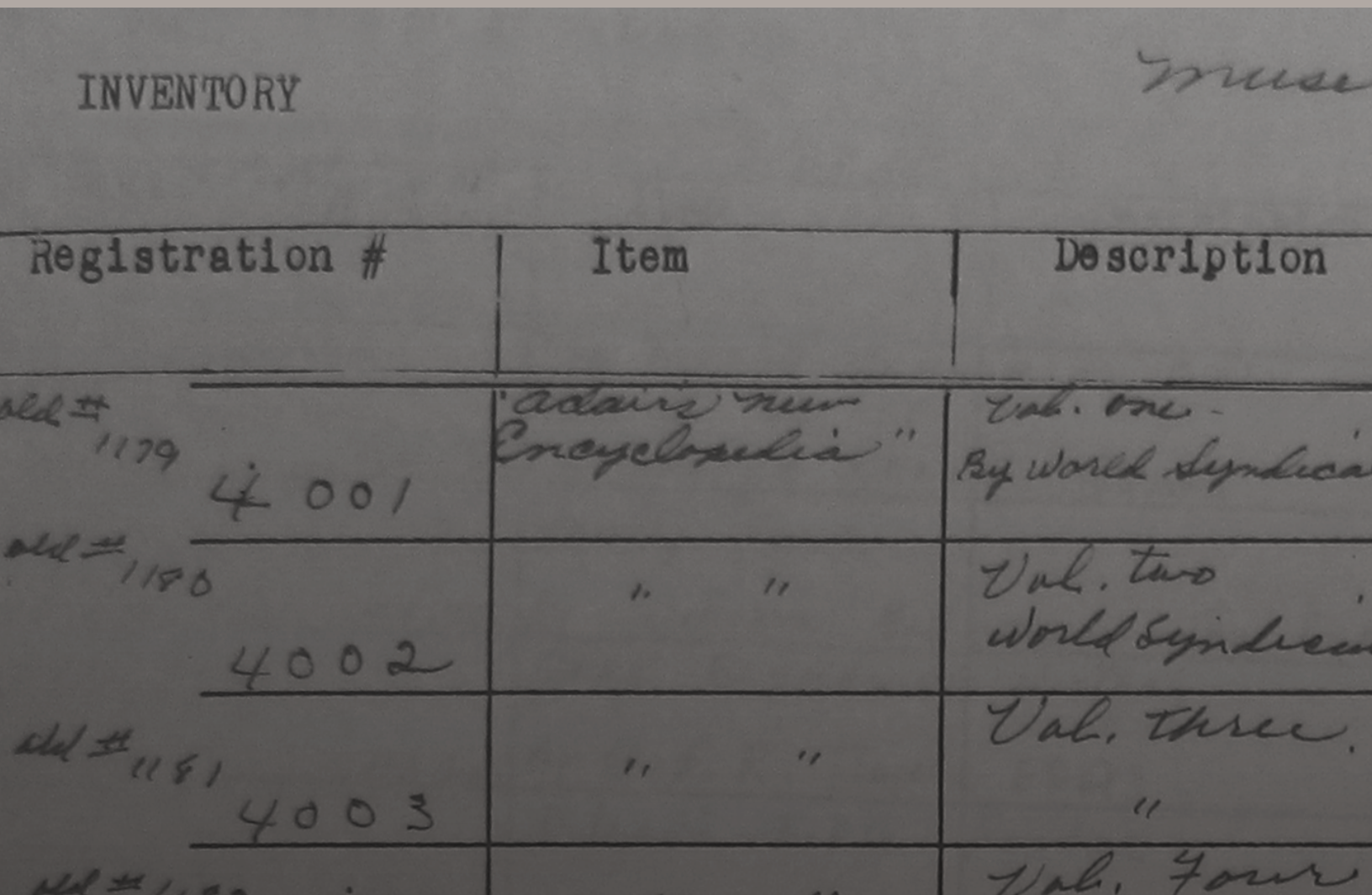
Northwest Digital Heritage Metadata Requirements, Version 1.0, July 2021. https://docs.google.com/document/d/1kiJGttioM7c4o7rvtIZhs_1Wns9Z24dBAIyYTrg74b0/edit

Northwest Digital Heritage Readiness Checklist and Content Policy – Washington State Library, State of Library of Oregon, and Oregon Heritage Commission, Version 1.0, adopted March 25, 2021 by the NWDH Council. https://drive.google.com/file/d/13Dd0fJeYWSbXSGdfSPz6Y4k_Hdqn9_iB/view

Orbis Cascade Alliance, Dublin Core Best Practices Guidelines, Version 2.3, January 29. 2018. https://drive.google.com/file/d/1ySAwESY1kOSsmH1d2H3aO493996yBCIi/view?usp=sharing.

Washington Rural Heritage Metadata Guidelines, Version 3.1, August 2018. Guidelines prepared 2018 by Nikki Chiampa, Digital Projects Librarian, Washington State Library and Evan Robb, Digital Repository Librarian, Washington State Library.https://www.washingtonruralheritage.org/digital/collection/wrh/id/266 .

OREGON PARKS AND RECREATION DEPARTMENT
OREGON HERITAGE/STATE HISTORIC PRESERVATION OFFICE
725 Summer Street NE, Suite C
Salem, OR 97301
(503) 986-0690



INVENTORY

Muse

Registration #	Item	Description
old # 1179 4001	"Adairs New Encyclopedia"	Vol. one - By World Syndica
old # 1180 4002	" "	Vol. two World Syndica
old # 1181 4003	" "	Vol. three. "
		Vol. four