

# UNDERSTANDING AND IMPROVING DATE SEARCHING IN CONTENTOM



**Patrick Curtis** 



**Mateo Palos** 





Patrick Curtis
Metadata and Digital Collections Developer
Ball State University

### **PURPOSE**

- To share information regarding problems encountered when searching by date and methods/settings discovered to solve, and explain, such problems
  - Enhanced understanding of how CONTENTdm handles search function
  - Improved search results
- Focus: CONTENTdm Advanced Search's "Enter Date" interface



### DIGITAL MEDIA REPOSITORY

HTTPS://DMR.BSU.EDU

- Added first collection in February 2005
- As of January 2021, the DMR contains 311 digital collections
- Created from 764 digitized archival collections (full or partial)
  - Ball State University history and operations
  - Specific units within Ball State
    - Andrew Seager Archive for the Built Environment
    - Virginia B. Ball Center for Creative Inquiry
  - Local history of Muncie and Delaware County, Indiana
  - Student projects and significant coursework
- Over 250,000 records (digital objects)
  - Just shy of 1.2 million items (text, images, audio/visual)



#### **BACKGROUND – TRANSITION PERIOD**

Ball State used a local instance of CONTENTdm until August 2018

Transitioned to a hosted instance of CONTENTdm during Fall 2018 –
 Spring 2019

 Noticed that DMR records, known to exist, were not present in search results when searching by date in the "Enter Date" interface



#### **BACKGROUND – TRANSITION PERIOD**

- Deliberate testing across the spectrum of collections
  - Discovered that no results at all were being presented.

- Apart from use of an "ON" date search
  - Appears to operate like a keyword search
  - Search terms highlighted when found in any field
  - Not limited to specified date fields



#### CONTEXT — DATE FIELDS

 Metadata standards behind the DMR provide for three specified date fields and a date-related field

#### Original Date

- Most useful to user / researcher
- Provides context beyond a strict ISO standard format (text-based)
- Allows for inexact dates (e.g., circa, unknown, etc.)
- Informed by DPLA requirements for its harvesting of metadata

#### System Date

- As CONTENTdm developed, became a new and required field
- Date-based formatting standard
- Initially only YYYY, but eventually led to more specifically dates (YYYY-MM-DD) and date ranges (YYYY-YYYY)

#### CONTEXT — DATE FIELDS

#### Digital Date

- Administrative
- Useful for tracking collection history, metadata revisions, and other significant events

#### Time Period

- Locally developed to provide for broad date ranges (e.g., decades, centuries, etc.)
- Necessity for museum objects with BCE dates as the BCE/CE distinction is not handled in CONTENTdm



### **CONTEXT – DATE FIELD PROPERTIES**

#### Original Date

- Dublin Core mapped to Date-Created
- Data Type = Text
- Hidden = NO

#### System Date

- Dublin Core mapped to None
- Data Type = Date
- Hidden = YES

#### Digital Date

- Dublin Core mapped to Date-Available
- Data Type = Text
- Hidden = NO

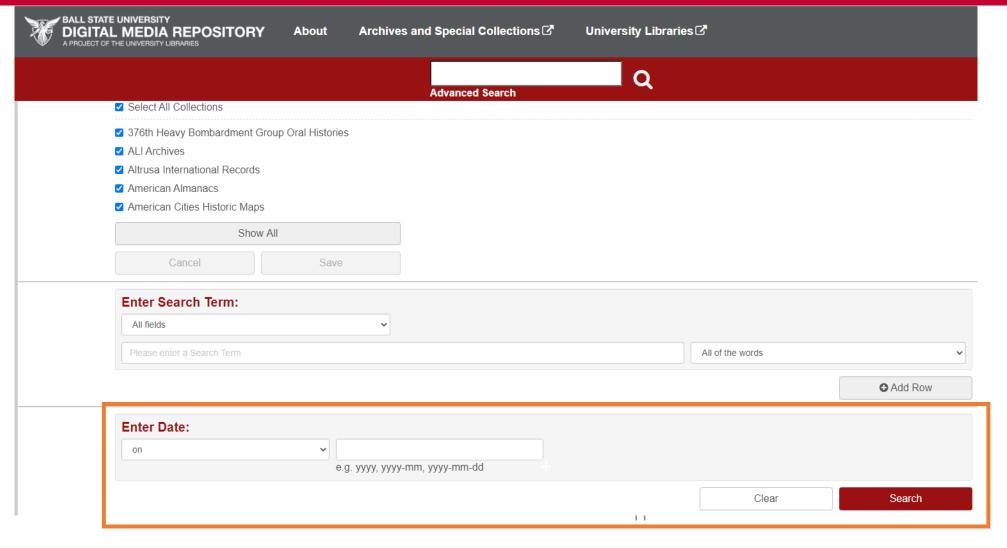






Mateo Palos
Library Data and Repository Analyst
Ball State University

### **Date Search Interface**





## What changed?

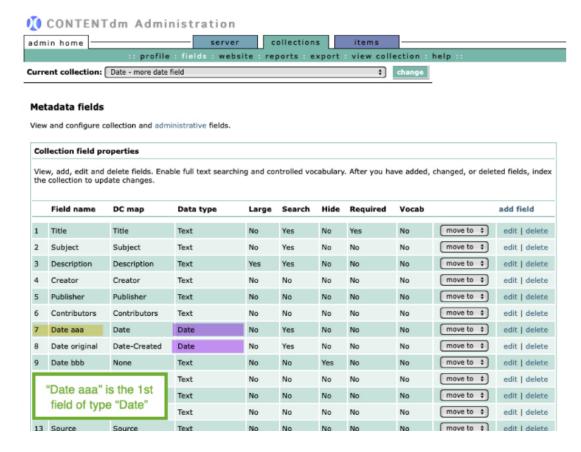
Why do only some date searches return results?



## June 2019 patch notes

#### Date search within a single collection uses the first field with the data type of "Date"

In Advanced Search, date search within a single collection now strictly uses the first field with the data type of "Date" to ensure seamless searching. The new rule for the Date search is simple and clear: in the fields list from CONTENTdm Administration, the first field of data type "Date" will be used for Date search. To ensure your Date search works and you are searching in the desired field, check your field settings in CONTENTdm Administration > collections > fields, as indicated in the image below:



https://help.oclc.org/Metadata\_Services/CONTENTdm/ Release\_notes/2019\_release\_notes/095\_CONTENTdm\_ release\_notes\_June\_2019



### What this means

There are two search contexts to design for:

- 1. Searching a single collection
- 2. Searching multiple collections

And these contexts look at very different things.



## What this means

	DC Map set to a Date element	DC Map not set to a Date element
Data Type set to Date	Single-collection searches and multi-collection searches return results	Single-collection searches return results (if and only if this is the first field with this data type)
Data Type not set to Date	Multi-collection searches return results	Does not return date results



## Searching a single collection

In Advanced Search, date search within a single collection now strictly uses the first field with the data type of "Date."

- Field must have "Date" data type
- Field must be the first field with "Date" data type
  - Others are ignored in this search context
- Field must be searchable ("Search" = Yes)
- Field must not be hidden ("Hide" = No)

Only the first field meeting these criteria is searched.



## Searching a single collection

	Field name	DC map	Data type	Large	Search	Hide	Required	Vocab		add field		
1	Title	Title	Text	No	Yes	No	Yes	No	move to ▼	edit   delete		
2	Creator	Creator	Text	No	Yes	No	No	No	move to ▼	edit   delete		
3	Subject - LCSH	Subject	Text	Yes	Yes	No	No	No	move to ▼	edit   delete		
4	Original Date	Date-Created	Text	No	Yes	No	No	No	move to ▼	edit   delete	←	No Date data type
5	Description	Description	Text	Yes	Yes	No	No	No	move to ▼	edit   delete		<b>,</b> ,
6	Alternative Title	Title-Alternative	Text	No	Yes	No	No	No	move to ▼	edit   delete		
7	Contributor(s)	Contributors	Text	No	Yes	No	No	No	move to ▼	edit   delete		
8	Publisher	Publisher	Text	No	Yes	No	No	No	move to ▼	edit   delete		
9	Subject - TGM	Subject	Text	No	Yes	No	No	Yes	move to ▼	edit   delete		
10	Subject - Other	Subject	Text	No	Yes	No	No	No	move to ▼	edit   delete		
11	Persons Pictured	Subject	Text	No	Yes	No	No	No	move to ▼	edit   delete		
12	Geographic Location	Coverage-Spatial	Text	No	Yes	No	No	No	move to ▼	edit   delete		
13	System Date	Date	Date	No	Yes	Yes	No	No	move to ▼	edit   delete	<b>←</b>	Hidden field
14	Time Period	Coverage-Temporal	Text	No	Yes	No	No	No	move to ▼	edit   delete		
15	Series	Relation	Text	No	Yes	No	No	No	move to ▼	edit   delete		
16	Notes	Description	Text	No	Yes	No	No	No	move to ▼	edit   delete		✓ "Date" data type,
17	Digital Date	Date-Available	Date	No	Yes	No	No	No	move to ▼	edit   delete	←	•••
18	Internal Notes	None	Text	No	Yes	Yes	No	No	move to ▼	edit   delete		Searchable, not Hidden



## Searching multiple collections

Data type and field order is irrelevant in this context; only DC map matters.

- Field must have "Date" or qualified Date DC map
- Field must be searchable ("Search" = Yes)
- Field must not be hidden ("Hide" = No)

<u>All</u> fields meeting these criteria are searched.



### What this means for collection metadata

Our fix: designating one field present in all collections as a "key date" that we can configure to be searchable in <u>all</u> possible searches.



### Our solution

#### 1. Field must be searchable (Search set to "Yes")

1 Tit	tle									
		Title	Text	No	Yes	No	Yes	No	move to ∨	edit   delete
2 Re	epository	Provenance	Text	No	Yes	No	No	No	move to $\vee$	edit   delete
3 Cre	reator	Creator	Text	No	Yes	No	No	No	move to $\vee$	edit   delete
4 Su	ubject - LCSH	Subject	Text	Yes	Yes	No	No	No	move to $\vee$	edit   delete
5 Sy	ystem Date	Date	Date	No	Yes	No	No	No	move to $\vee$	edit   delete

(Note that this is the first date field)



### Our solution

- 1. Field must be searchable (Search set to "Yes")
- 2. Field data type set to "Date"

Field name	DC map	Data type	Large	Search	Hide	Required	Vocab		add field
Title	Title	Text	No	Yes	No	Yes	No	move to ∨	edit   delete
Repository	Provenance	Text	No	Yes	No	No	No	move to $\vee$	edit   delete
Creator	Creator	Text	No	Yes	No	No	No	move to ∨	edit   delete
Subject - LCSH	Subject	Text	Yes	Yes	No	No	No	move to $\vee$	edit   delete
System Date	Date	Date	No	Yes	No	No	No	move to ∨	edit   delete
	Title Repository Creator Subject - LCSH	Title Title  Repository Provenance  Creator Creator  Subject - LCSH Subject	Title Title Text  Repository Provenance Text  Creator Creator Text  Subject - LCSH Subject Text	Title Title Text No  Repository Provenance Text No  Creator Creator Text No  Subject - LCSH Subject Text Yes	Title Title Text No Yes  Repository Provenance Text No Yes  Creator Creator Text No Yes  Subject - LCSH Subject Text Yes Yes	Title Title Text No Yes No Repository Provenance Text No Yes No Creator Creator Text No Yes No Subject - LCSH Subject Text Yes Yes No	Title Title Text No Yes No Yes  Repository Provenance Text No Yes No No  Creator Creator Text No Yes No No  Subject - LCSH Subject Text Yes Yes No No	Title Title Text No Yes No Yes No Repository Provenance Text No Yes No No No No Creator Creator Text No Yes No No No No Subject - LCSH Subject Text Yes Yes No No No No	Title Title Text No Yes No Yes No move to \to \text Repository Provenance Text No Yes No No No move to \to \text Creator Creator Text No Yes No No No move to \to \text Subject - LCSH Subject Text Yes Yes No No No move to \to \text No No No move to \to \text No No No Mo No move to \to \text No No No No move to \to \text No No No No Mo No move to \text No No No No Mo No Mo No Mo No Mo No No Mo N



### Our solution

- 1. Field must be searchable (Search set to "Yes")
- 2. Field data type set to "Date"
- 3. DC map set to a Date value

	Field name	DC map	Data type	Large	Search	Hide	Required	Vocab		add field
1	Title	Title	Text	No	Yes	No	Yes	No	move to ∨	edit   delete
2	Repository	Provenance	Text	No	Yes	No	No	No	move to $\vee$	edit   delete
3	Creator	Creator	Text	No	Yes	No	No	No	move to $\vee$	edit   delete
4	Subject - LCSH	Subject	Text	Yes	Yes	No	No	No	move to $\vee$	edit   delete
5	System Date	Date	Date	No	Yes	No	No	No	move to $\vee$	edit   delete
5	System Date	Date	Date	INO	Yes	NO	INO	INO	move to $\checkmark$	edit

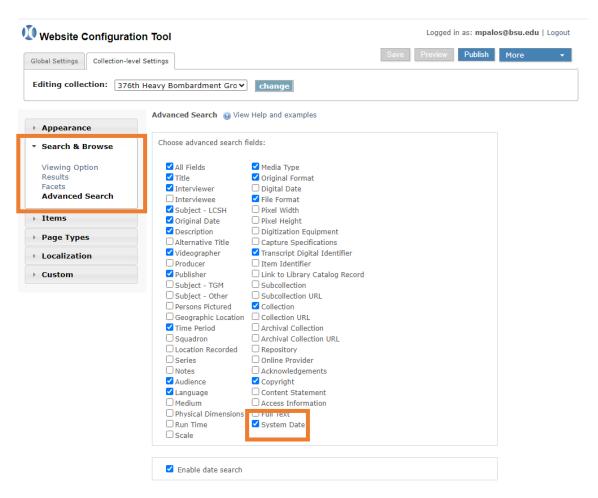


## One last thing

- Field must be marked as searchable in website collection-level advanced search configuration
- DC map must be marked as searchable in the website global-level advanced search configuration

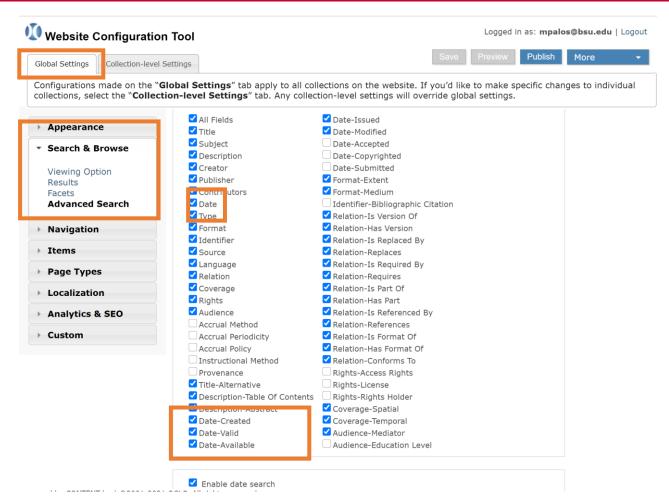


## **Collection-level settings**





## **Global settings**





### **Lessons learned**

- Successful metadata design requires keeping in mind two search contexts
- Collections should be designed with a "key date" field in mind:
  - The field should be set to the Date data type *and* be the first field in the collection with this data type
  - The field should be mapped to a DC.Date element.
  - The field should <u>not</u> be hidden
  - And has both the field and DC map exposed to the advanced search interface in the website settings

Field name	DC map	Data type	Large	Search	Hide	Required	Vocab	add field
System Date	Date	Date	No	Yes	No	No	No	move to V edit   delete



## Q & A

	DC Map set to a Date element	DC Map not set to a Date element
Data Type set to Date	Single-collection searches and multi-collection searches return results	Single-collection searches return results (if and only if this is the first field with this data type)
Data Type not set to Date	Multi-collection searches return results	Does not return date results



## Questions?

Please send all questions via the Chat panel.

- Click on the Chat box to make it the active window.
- To submit a question, please select All Panelists in the To box.
- Clicking the "Enter" key will send your message.
- We will answer as many questions as time allows.

