

Washington University in St. Louis  
**Washington University Open Scholarship**

---

IASSIST & DCN - Data Curation Workshop

Workshop Schedule

---

Dec 11th, 3:00 PM - 4:15 PM

## Request and Augment Presentation

Wendy Kozlowski

*Cornell University Library*, wak57@cornell.edu

Jake R. Carlson

*University of Michigan - Ann Arbor*, jakecar@umich.edu

Follow this and additional works at: <https://openscholarship.wustl.edu/data-curation-workshop-2017>



Part of the [Library and Information Science Commons](#)

---

Kozlowski, Wendy and Carlson, Jake R., "Request and Augment Presentation" (2017). *IASSIST & DCN - Data Curation Workshop*. 9.  
<https://openscholarship.wustl.edu/data-curation-workshop-2017/schedule/Schedule/9>

This Presentation is brought to you for free and open access by the Conferences and Symposia at Washington University Open Scholarship. It has been accepted for inclusion in IASSIST & DCN - Data Curation Workshop by an authorized administrator of Washington University Open Scholarship. For more information, please contact [digital@wumail.wustl.edu](mailto:digital@wumail.wustl.edu).

# R and A

Request and Augment

C ⇒ U ⇒ **R** ⇒ A ⇒ T ⇒ E



Request missing information

- Generate a list of questions for the data author to fix any errors or issues.

# Determine needed recommendations or requests

How much is too much?

What are your priorities?

How hard do you push if you don't get those priorities from the submitter?

Rule of Thumb #1:



**KEEP  
CALM  
AND  
KEEP IT  
SIMPLE**

## Rule of Thumb #2: Be specific



**an example  
would be handy  
right about now**

# Rule of Thumb #3: Provide Resources



# Request info

- Email template from University of Michigan:

Dear [name of the person Identified as the contact for the data set as stated in the DBD metadata],

Thank you for depositing your data set, [title of the data set] to the library's Deep Blue Data repository.

After we receive a data set, we review it to ensure that the data sets we host are as complete, accessible and understandable as possible. We have reviewed your data set and have the following recommendations for you:

- Recommendation #1
- Recommendation #2
- Recommendation #3
- Recommendation #4

We look forward to hearing your response to our questions and requests for additional information.

Please do let us know if you have any questions about or recommendations. We would be happy to talk with you over the phone or meet with you in person to discuss our review of your data should you wish to do so.

Sincerely,

[Name of Liaison]

Thank you

What you need from them, and why

How they should get you the info

Offer to help



# Request info

- Email template from University of Minnesota:

Subject: Your submission to the Data Repository for the U of M (DRUM)

Dear <submitter>,

Thank you for your submission (insert link) to the Data Repository for the U of M (DRUM). I am in the process of curating your submission, and I have a few questions for you:

- Most pressing question here
- So that people can best understand and make use of your data, I'd like to add more documentation about each of your files and the variables within. I've started a "readme.txt" file using DRUM's standard format based on the information contained within in the submission, and attached it to this email. Would you please add more information to this text file? Specifically the following:
  - Request specific information here
- More questions here

Please let me know if you have any questions or concerns. Thank you very much for your submission, and I look forward to hearing from you.

Sincerely,

Name  
DRUM Coordinator

Thank you

What you need from them, and why

Offer to help

# Request info

- Basic email outline from Cornell:

Dear <SUBMITTER(S)>,

My name is Wendy Kozlowski, and I'm a data curator and administrator for eCommons. Thank you for your request to submit a dataset.

We strive to have the datasets in eCommons meet the FAIR data principles (Findable, Accessible, Interoperable, Reusable). I've looked at your dataset and have the following suggestions and questions:

<KEEP AND MODIFY FROM BELOW ONLY THE APPROPRIATE SECTIONS:>

1. File name(s)...
2. File format(s)...
3. File content...
  - A. For All Data...
  - B. For Tabular Data...
4. Discovery (system) metadata...
5. Science metadata...

You're welcome to contact me via email or phone, or we can meet in person if you prefer. I've also created a folder in Box for you to put your files: <URL TO BOX FOLDER>. Please don't hesitate to be in touch with questions or concerns.

Kind regards,  
<CURATOR NAME>

Thank you

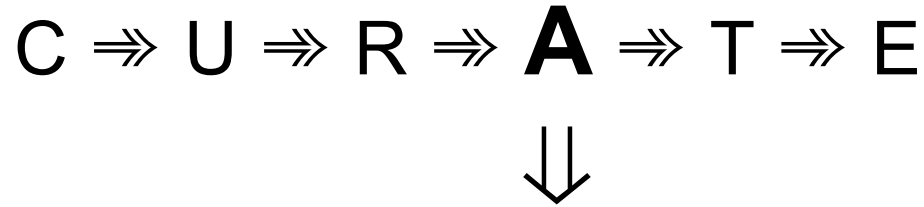
What you need from them, and why

How they should get you the info and offer to help

# Exercise: Request information

10 minutes: Write a letter to your dataset submitter

10 minutes: Trade with your neighbor and review



## Augment the submission with metadata for findability

- Enhance metadata to best facilitate discoverability.
- Create and apply metadata for the data record, including descriptive keywords.
- When appropriate, structure and present metadata in domain-specific schemas to facilitate interoperability with other systems.

# Augment the submission

- Discoverability Sufficient
  - Recommend (circle one) full-text index / file compression / file reorder / file descriptions / zip
  - Other \_\_\_\_\_
- Keywords Sufficient
  - Suggestions \_\_\_\_\_
- Linkages Sufficient
  - Link to Report/Paper
  - Link to related data sets
  - Link to source data
  - Link to other \_\_\_\_\_

# Exercise: Assess metadata

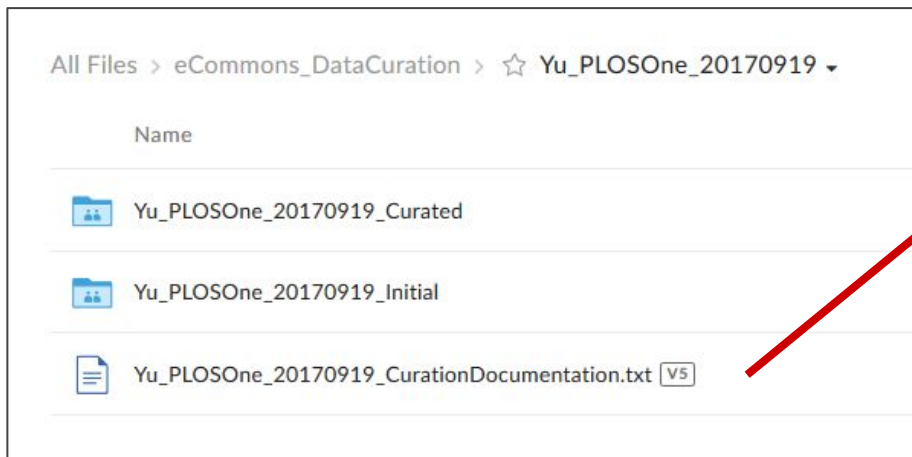
As a table:

10 minutes: Review metadata for your dataset

1 minutes each table: Report out - anything missing or that could be improved?

# Don't forget about the curation documentation

Example: Curation readme file,  
shared file storage space for files



-----  
Parent Folder:  
-----

Files in Initial Archive:  
-----

Date first received:  
-----

Curator:  
-----

Submission Contact Info:

Name:

Position:

Department:

Office Location:

URL:

e-mail/NetID:

phone:  
-----

Owner/Author Info (if different):  
-----

RDMSG Help Ticket #, if appropriate:  
-----

Repository:  
-----

Community/Collection URLs:  
-----

Curator Notes:

# Don't forget about the curation documentation

Example: Track using a service management software system (eg. OTRS, JIRA, RT, Remedy or others)

The screenshot shows a ticketing system interface for the University of Illinois at Urbana-Champaign. The main content is a table of articles related to a ticket. The table has columns for NO., TYPE, FROM, SUBJECT, and CREATED. The rows are color-coded: pink for internal notes, green for correspondence to the depositor, and blue for correspondence from the depositor. A 'Status Summary' sidebar on the right provides details about the ticket, including its age, creation date, status, and priority.

**Pink: Internal Curation Notes**

**Green: Correspondence TO Depositor**

**Blue: Correspondence FROM Depositor**

**Status Summary**

NO.	TYPE	FROM	SUBJECT	CREATED
22	agent - note-internal	Heidi Imker	Note - Opening .czi files on a Mac	11/16/2017 16:41
21	agent - note-internal	Elise Dunham	Close	
20	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	
19	customer - email-external	Thapa, Sita	[Illinois Data Bank] Dataset succ...	
18	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	
17	customer - email-external	Thapa, Sita	[Illinois Data Bank] Dataset succ...	03/02/2017 09:50
16	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	03/01/2017 14:32
15	agent - note-internal	Elise Dunham	Note	02/28/2017 11:56
14	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	02/28/2017 11:20
13	customer - email-external	Thapa, Sita	[Illinois Data Bank] Dataset succ...	02/24/2017 15:40
12	customer - email-external	Thapa, Sita	[Illinois Data Bank] Dataset succ...	01/30/2017 08:50
11	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	01/30/2017 08:41
10	customer - email-external	Thapa, Sita	[Illinois Data Bank] Dataset succ...	01/27/2017 16:20
9	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	01/27/2017 16:14
8	customer - email-external	Thapa, Sita	[Illinois Data Bank] Dataset succ...	01/27/2017 15:40
7	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	01/27/2017 15:24
6	customer - email-external	Thapa, Sita	[Illinois Data Bank] Dataset succ...	01/25/2017 17:10
5	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	01/25/2017 17:00
4	customer - email-external	Thapa, Sita	[Illinois Data Bank] Dataset succ...	01/20/2017 11:10
3	agent - email-external	Illinois Databank	[Illinois Data Bank] Dataset succ...	01/19/2017 16:54
2	agent - note-internal	Elise Dunham	Note	01/12/2017 11:52

**Ticket Information**

- Age: 332 d 23 h
- Created: 01/06/2017 12:00
- Status: closed - answered
- Locked: lock
- Priority: 3 normal
- Queue: IDB
- CustomerID: databank@library.illinois.edu
- Owner: Heidi Imker

**Customer Information**

- none

**Linked Objects**

- none

**Article #22 - Note - Opening .czi files on a Mac** Created: 11/16/2017 16:41 by Heidi Imker

Reply to note | Mark | Print | Split

From: Heidi Imker  
To: Elise Dunham  
Subject: Note - Opening .czi files on a Mac



# Don't forget about the curation documentation

Example: Provenance records kept within your IR system

dc.description.provenance	Made available in DSpace on 2017-09-19T19:52:40Z (GMT). No. of bitstreams: 1 Place holder.docx: 13241 bytes, checksum: b58c982e49b9492d43d2f92545c3b547 (MD5) Previous issue date: 2017
dc.description.provenance	Item withdrawn by Wendy Kozłowski (wak57@cornell.edu) on 2017-09-19T19:53:29Z Item was in collections: Nutritional Sciences Publications, Manuscripts and Datasets (ID: 1209) No. of bitstreams: 1
dc.description.provenance	This item reinstated at the request of E. Yu, 2017-09-23 by W Kozłowski.
dc.description.provenance	Item reinstated by Wendy Kozłowski (wak57@cornell.edu) on 2017-09-23T16:09:29Z Item was in collections: Nutritional Sciences Publications, Manuscripts and Datasets (ID: 1209) No. of bitstreams: 3
dc.description.provenance	Readme file updated with publication citation information by W. Kozłowski, 2017-12-04, at request of E.Yu.

# Minute Paper

1. What was the most valuable thing you learned today?
2. What do you still have questions about?
3. Is there anything we could do to improve the workshop?