Apr 6th, 10:30 AM

Informationist: Informationist Breakout Session

Leah Honor

*University of Massachusetts Medical School*

Follow this and additional works at: [http://escholarship.umassmed.edu/escience_symposium](http://escholarship.umassmed.edu/escience_symposium)

Part of the [Scholarly Communication Commons](http://escholarship.umassmed.edu/escience_symposium)

This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 4.0 License](http://escholarship.umassmed.edu/escience_symposium).


[http://escholarship.umassmed.edu/escience_symposium/2016/program/7](http://escholarship.umassmed.edu/escience_symposium/2016/program/7)

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in University of Massachusetts and New England Area Librarian e-Science Symposium by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
Informationist Breakout Session
eScience Symposium
April 6, 2016
Hello!
I am Leah Honor

Informationist Liaison to the Child and Adolescent NeuroDevelopment Initiative (CANDI)

LSL Library Fellow
In this Session:

1. What we did, and why we did it
2. What I learned along the way
3. What we plan to do next
1. NIH Informationist Supplement Grant

Data Citation in Neuroimaging: Proposed Best Practices for Data Attribution and Citation
Data Sharing and Citation in Neuroimaging

Sharing and reusing data (images) is already accepted and common, but no standards exist on how to properly cite reused data, and how to maintain attribution to data creators.

Citations are often:

◎ Indirect - Cite the paper where the data was presented, but not the data itself
◎ Unofficial - Citations in acknowledgements or contribution sections
◎ Indefinite - descriptions of sources and datasets in methods sections are not machine readable
What ARE Citations?

At their most basic level, citations are just a way to identify your sources.

Citing data, which has no specific format or structure, led us to ask: how can data be identified in a way that is direct, official, and definitive (and hopefully machine readable)?

What granularity of identification will be needed to accurately cite reused data? What about new datasets that draw from many sources?
Data DOIs In NITRC

NITRC - Neuroimaging Informatics Tools and Resources Clearinghouse

Realized we needed nested levels of identifiers:

- **Project level** - assigned when new data is uploaded
- **Image level** - each image carries its own identifier, as well as a parent project ID
- **Functional level** - assigned to image sets created from existing data
To Create a New Functional Level DOI for an Image Set:

**Query**
Using the standard repository search functions, a new image set is defined using standard criteria such as age, gender, handedness, diagnosis, etc.

**Result/Selection**
Results can be reviewed and excluded, or additional images included, until the final collection of images has been resolved.

**Tag Preparation**
Basic metadata fields, such as authors, associated publication ID’s, funders, and project description fields must be completed in order to create an identifier for the newly defined image collection.
UMass/CANDI DOI Project

Use this form to start a search for data in the CANDI/UMMS DOI database, which can then be refined and tagged with a DOI.

Gender
- Female
- Male
- Either

Age
- Min
- Max

Handedness
- Left
- Right
- Either

Search

Functional DOI Process: Query
UMass/CANDI DOI Project

This page shows a collection of images starting with female, 5 <= age <= 18, right.

Actions:
- Tag: Tag this collection with a DOI.
- Download: Download the data in this collection (the download package may contain more data than requested).

Or refine the collection:

<table>
<thead>
<tr>
<th>Project</th>
<th>Subject</th>
<th>Gender</th>
<th>Age</th>
<th>Handedness</th>
<th>Scan</th>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_030</td>
<td>female</td>
<td>9</td>
<td>right</td>
<td>anat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_030</td>
<td>female</td>
<td>9</td>
<td>right</td>
<td>seg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_031</td>
<td>female</td>
<td>15</td>
<td>right</td>
<td>anat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_031</td>
<td>female</td>
<td>15</td>
<td>right</td>
<td>seg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_032</td>
<td>male</td>
<td>7</td>
<td>right</td>
<td>anat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_032</td>
<td>male</td>
<td>7</td>
<td>right</td>
<td>seg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_033</td>
<td>male</td>
<td>8</td>
<td>right</td>
<td>anat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_033</td>
<td>male</td>
<td>8</td>
<td>right</td>
<td>seg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cs_schizbull08</td>
<td>BPDwoPsy_034</td>
<td>male</td>
<td>10</td>
<td>right</td>
<td>anat</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

UMass/CANDI DOI Project

This page shows a collection of images starting with female, 5 ≤ age ≤ 18, right.
If you don't want to tag this collection, you can go back and refine or download the collection.

Publication PMID: 18003691
Publication DOI: 10.1093/schbul/sbm120
Authors (one per line):
Frazier JA
Hodge SM
Breeze JL
Giuliano AJ
Terry JE
Moore CM

Funder: National Institute of Mental Health

Description:
This data set of adolescent females was used to analyze hippocampal volumes in APP.

This collection has not been tagged with a DOI.
If you are test driving the system, please create a temporary test DOI. If you would like to create a permanent DOI for this data set, please indicate this by checking here: ☐

Tag

Functional DOI Process: Tag Creation
# UMass/CANDI DOI Project

**Image collection 10.5072/FK2H41Q83C**

UMass/CANDI DOI project, 2015  
DOI 10.5072/FK2H41Q83C  
[DataCite XML](#)  
[Refine/download](#)

<table>
<thead>
<tr>
<th>Description</th>
<th>PubMed ID</th>
<th>Publication DOI</th>
<th>Funder</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>This data set of adolescent females was used to analyze hippocampal volumes in BPD.</td>
<td>18003631</td>
<td>doi:10.1093/schbul/sbm120</td>
<td>National Institute of Mental Health</td>
<td>Frazier JA, Hodge SM, Breeze JL, Giuliano AJ, Terry JE, Moore CM, Kennedy DN, Lopez-Larson MP, Caviness VS, Seidman LJ, Zablotsky B, Makris N</td>
</tr>
</tbody>
</table>

**Source projects**: 10.18116/C6159Z  
**Source images**:  
- 10.18116/C6FP4R  
- 10.18116/C6MW2T  
- 10.18116/C6H999  
- 10.18116/C6CC7C  
- 10.18116/C67P43  
- 10.18116/C63W2S

---

**Functional DOI Landing Page**
Manual Segmentation for CANDI Share Schizophrenia Bulletin 2008 subject BPDwoPsy_041

Child and Adolescent NeuroDevelopment Initiative, 2008
DOI 10.18116/C6CC7C

Project 10.18116/C6159Z
Collections 10.5072/FK2H41Q83C
Sizes 1 files, 187197 bytes
Format NIFTI
Version 1.1
Rights Creative Commons Attribution 4.0 International License

UMass/CANDI DOI Project

Image collection 10.5072/FK2H41Q83C
UMass/CANDI DOI project, 2015
DOI 10.5072/FK2H41Q83C

Description
This data set of adolescent females was used to analyze hippocampal volumes in BPD.

PubMed ID 18003661
Publication DOI doi:10.1093/ajmhbr/mbr075
Funder National Institute of Mental Health
Authors

Source projects 10.18116/C6159Z
Source images
10.18116/C68P4R
10.18116/C68WZT
10.18116/C6HM58
10.18116/C6CQ7C
10.18116/C6TP43
10.18116/C6XZ06
2. Lessons Learned

Bridging the gap between library and research team
Know Your Strengths

How I succeeded as an Informationist:

Necessary

◎ Info. science skills
  ○ Metadata standards
  ○ Data management
  ○ Metrics and data usage reporting

Useful (but not critical)

◎ Neuroscience background
◎ Research experience
Main Takeaways:

There is no right answer! You are there to help find a better one than exists today.

When joining a research team you need to be willing to listen, to learn, and to be flexible in your ideas.

...but also remember you are not in this alone, and that research is a collaborative process. Have conversations, try different approaches, and keep the big picture in mind.
3. **Future Projects**

Machine readable mark-up and building reproducibility frameworks
Thanks!

Any questions?

You can contact me at:
Leah.Honor@umassmed.edu