

DataCite Implementations at the University of Massachusetts

Presentation for DataCite General Assembly

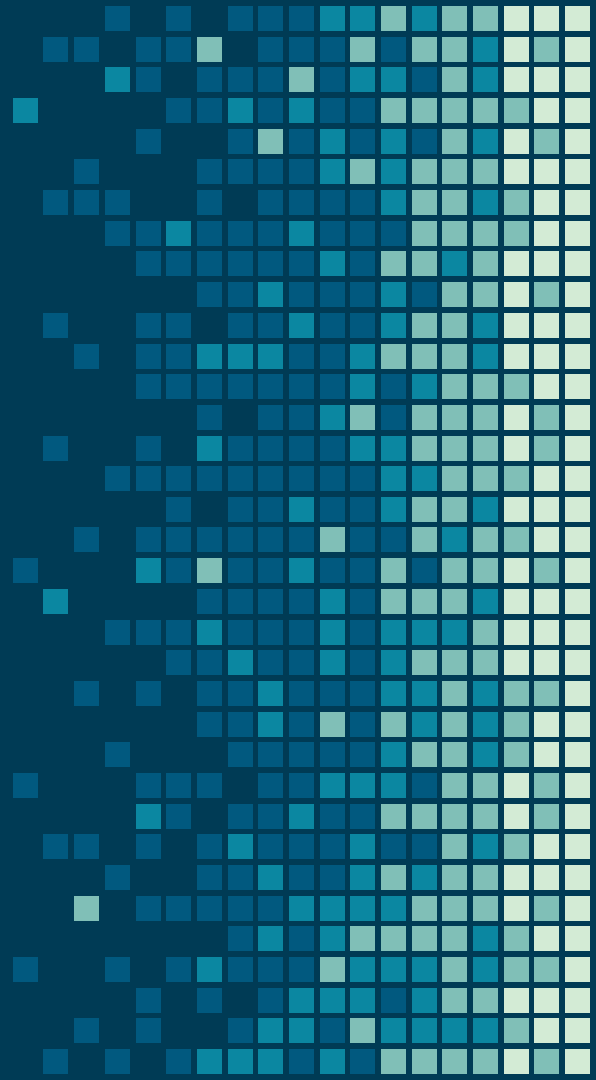
April 1, 2019

Tess Grynoch, @Te_Gryn

Erin Jerome, Libraries, UMass Amherst

Christian Haselgrove, CANDI, UMass Medical School

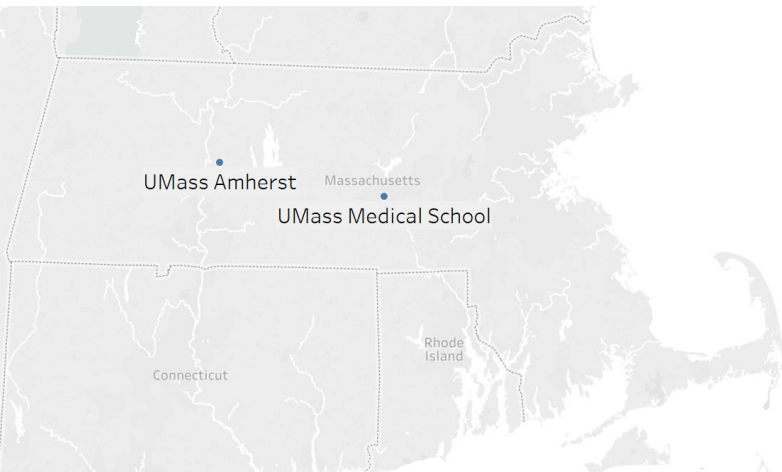
Lisa Palmer, Lamar Soutter Library, UMass Medical School





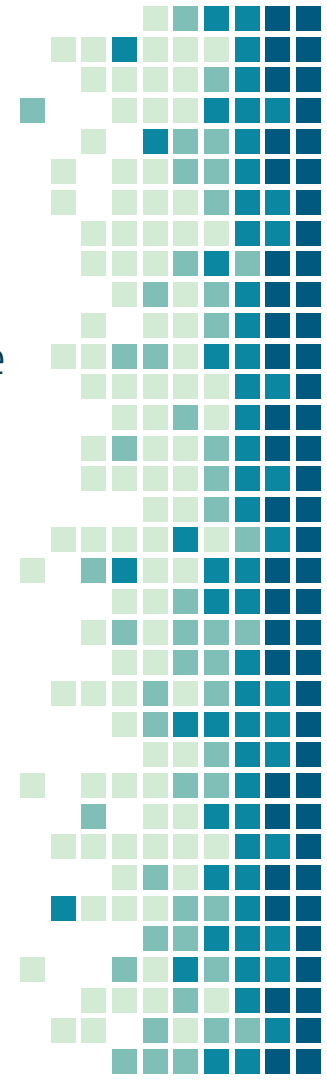
University of Massachusetts (UMass)

- 5 campuses, 3 groups on two of the campuses currently using DataCite



- UMass Amherst Libraries
- UMass Medical School Library
- CANDI (Child and Adolescent NeuroDevelopment Initiative), based at UMass Medical School

- Joined DataCite: March 2018 (migrated from EZID)



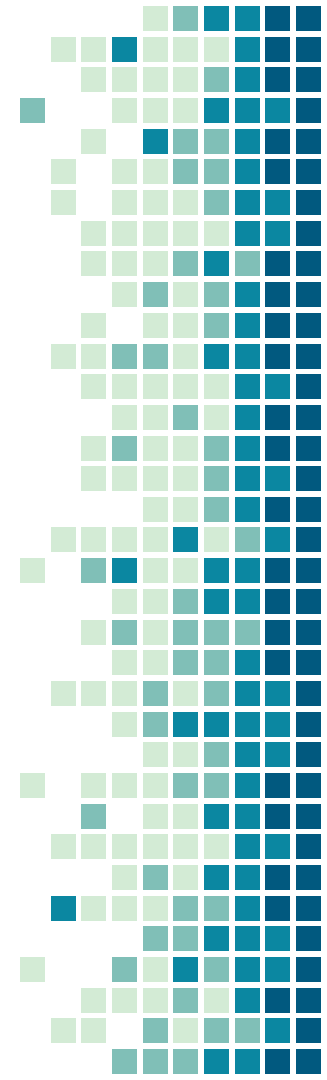
DataCite Implementation: UMass Medical School Library



The screenshot shows the homepage of the eScholarship@UMMS repository. At the top left is the UMass Medical School logo. The main header features the text "eScholarship@UMMS" and a background image of a person in a laboratory. Below the header is a navigation bar with links for Home, About, FAQ, My Account, and Submit Research. A search bar is located on the left side, with a dropdown menu for "in this repository" and a "Search" button. Below the search bar are links for "Advanced Search" and "Notify me via email or RSS". A "Links" section includes "Guide to eScholarship@UMMS", "Policies", and "Lamar Soutter Library". A "Browse" section lists "Collections", "Disciplines", and "Authors". The main content area is divided into two columns: "Research and Scholarship" with a "Follow" button and a list of categories (School, Department, or Program; Student Scholarship; Personal Researcher Pages; Books), and "Featured Collections" with a list of specific collections (Cancer Concepts: A Guidebook for the Non-Oncologist; UMass Center for Clinical and Translational Science Supported Publications).

■ [eScholarship@UMMS](#)

- Data
- Posters
- Presentations
- Theses & Dissertations
- Post-prints



DataCite Implementation: UMass Amherst Libraries

UMassAmherst Libraries

My Library Account Dashboard

You have...

Borrowed materials	Pending Requests	Enrolled Courses	Courses Taught
5	0	3	4

Earned Badges: 6

Logout

Cowabunga
You found the Teenage Mutant Ninja Turtles statue in the Du Bois Library

© 2019 University of Massachusetts Amherst • Site Policies • Site Contact

- [ScholarWorks@UMassAmherst](#)
 - Using API to retrospectively assign DOIs to dissertations and theses
- Earn badges in the library

My Library Account Badges

Generation 1

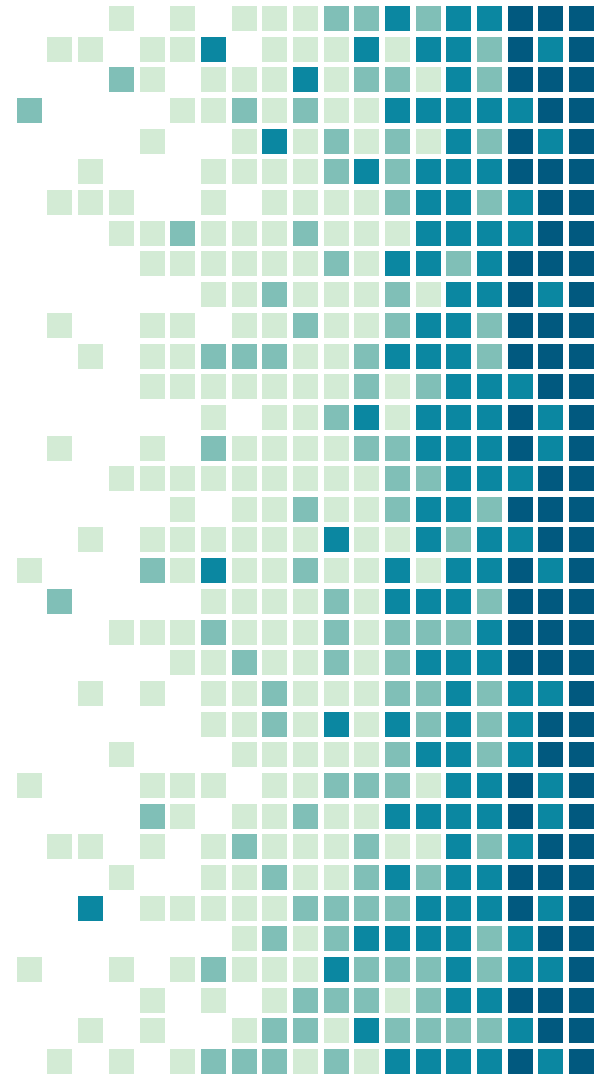
Early Adopter	?????	?????	Cowabunga	Adventurer	?????
?????	?????	?????	Teapot	?????	コナミコマンド

Logout

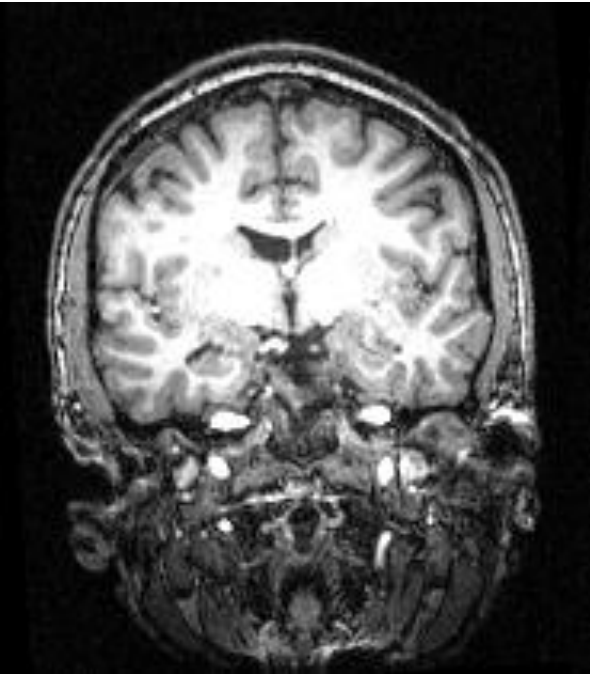


DataCite Implementation:

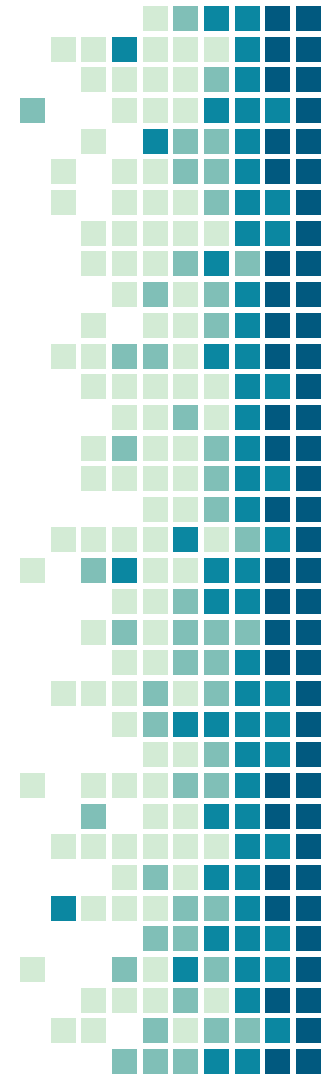
CANDI



The scene

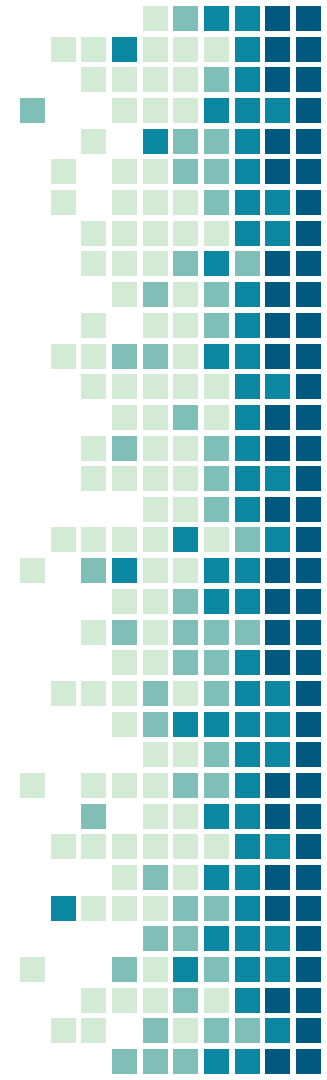


- Brain imaging studies:
 - Large numbers of subjects are needed, but scanning is expensive.
 - Data have become first-class research objects, and large data sets are collected and shared for reuse.



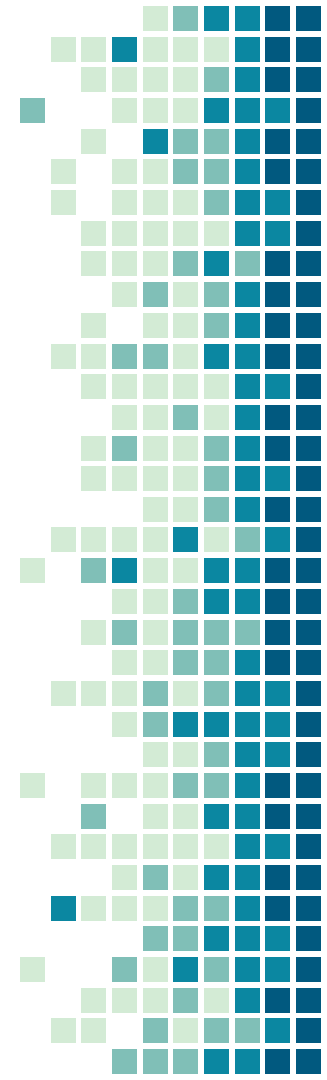
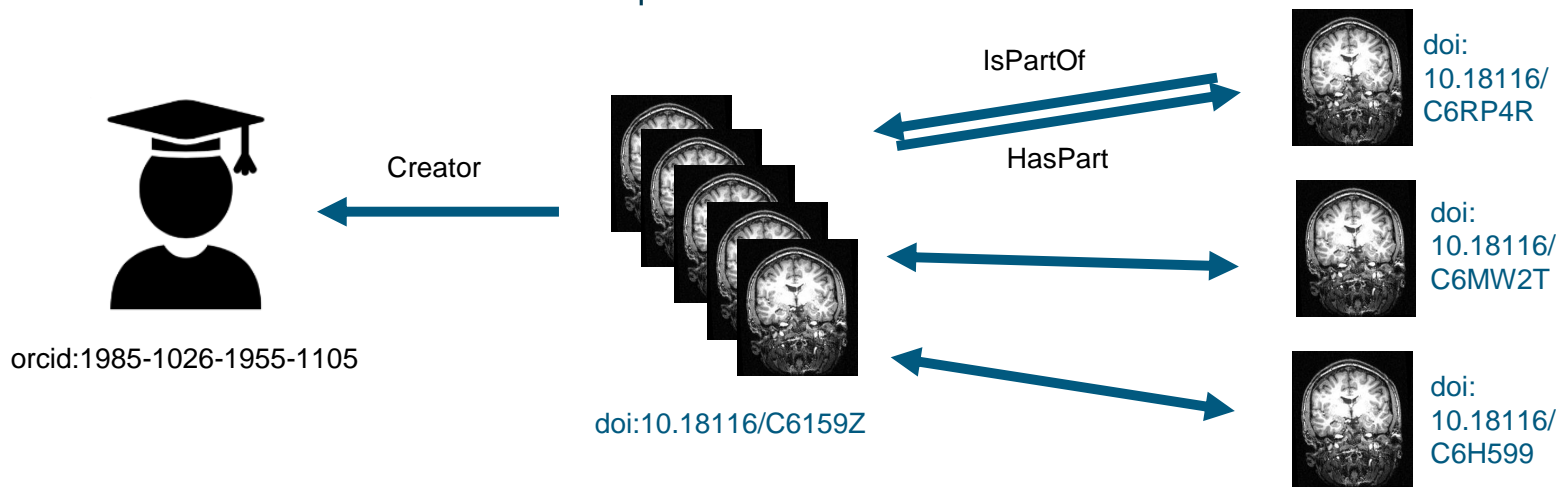
The problems

- Reproducibility
 - A study won't typically use all of the scans from a large database and won't report exactly which scans were used.
- Credit
 - If data is reused in a study, the data provider should be acknowledged for their contribution. Data providers will also want to track how their data is used.



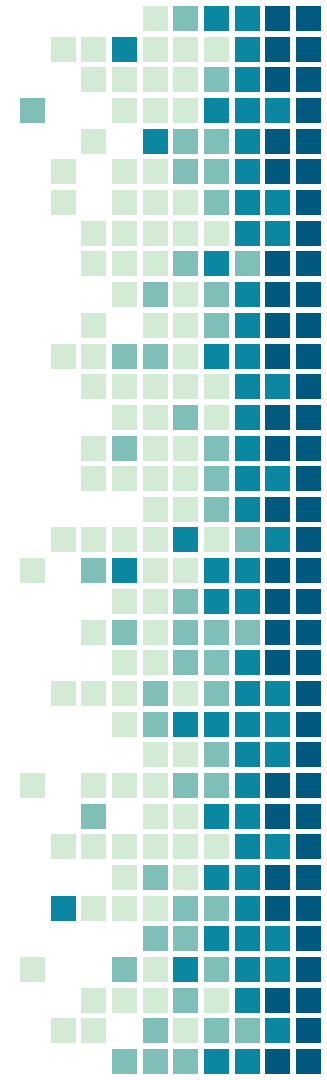
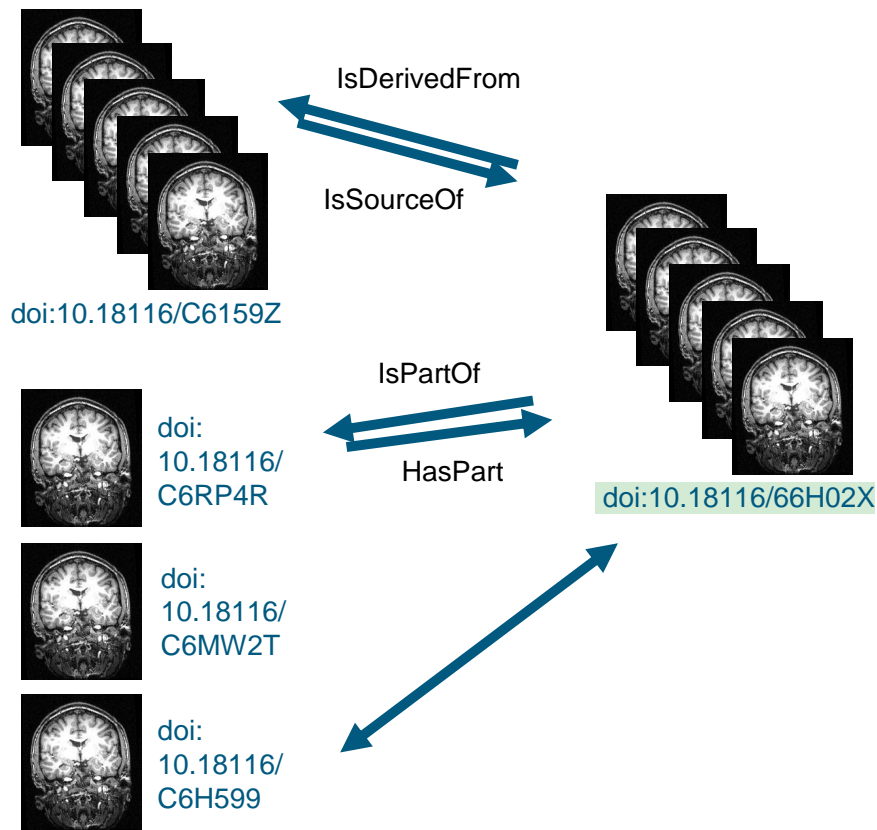
The approach

- Data providers assign DOIs to large data releases and scans within those releases
- DOI metadata links the parts



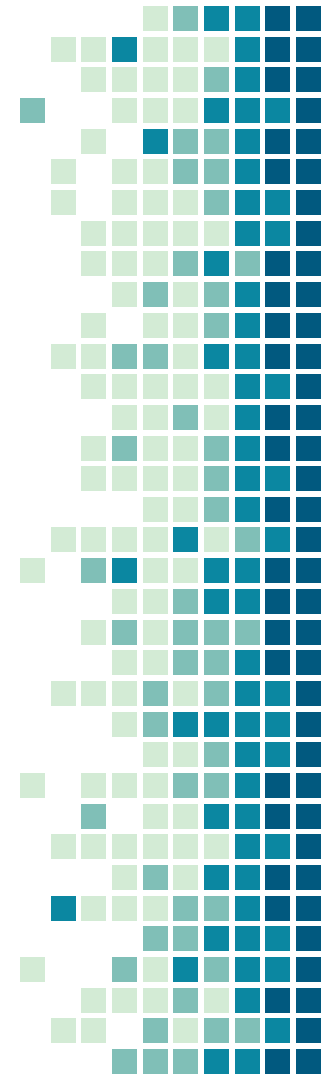
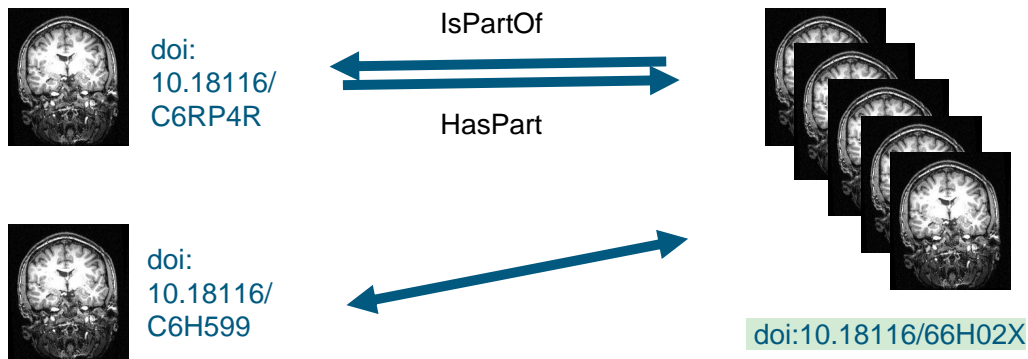
The approach (continued)

- New study takes a subset of the scans from multiple existing data releases.
- DOI is assigned to the collection of scans used in the new study.
- DOI metadata links the parts.



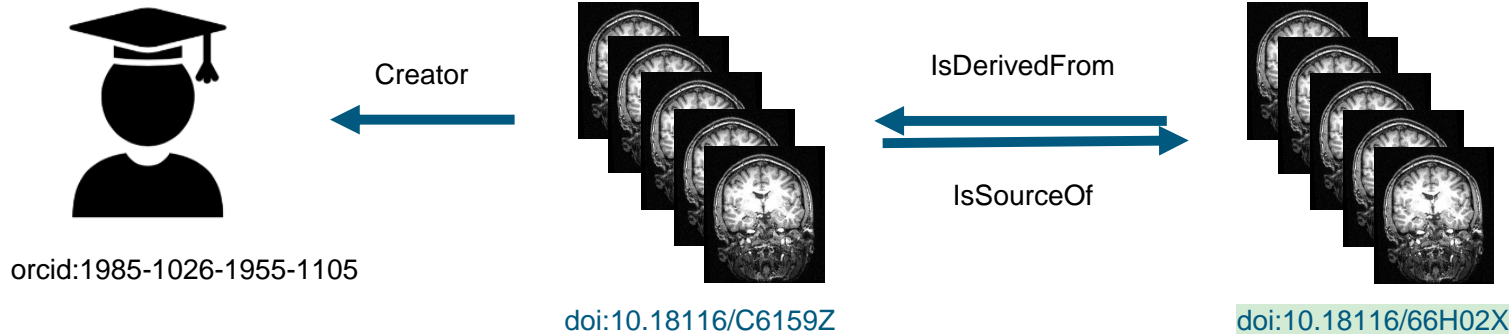
The solution

- The authors publish the DOI of the data set they used, and the individual scans can be identified.



The solution (continued)

- The source of the data can be tracked right back to the data providers, acknowledging their contribution and allowing them to track the use of the data they provided.



The solution (continued)

CANDI Share Schizophrenia Bulletin 2008 data

JA Frazier, SM Hodge, JL Breeze, AJ Giuliano, JE Terry, CM Moore, DN Kennedy, MP Lopez-Larson, VS Caviness, LJ Seidman, B Zablotzky & N Makris

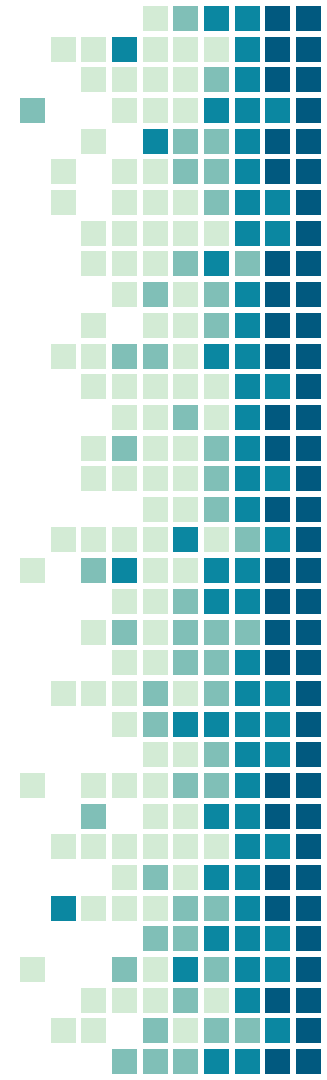
Imaging Data published 2008 via Child and Adolescent NeuroDevelopment Initiative

Data for CANDI Share Schizophrenia Bulletin 2008 (reference below) as part of the CANDI Neuroimaging Access Point.

This set includes preprocessed MRI images and segmentation results of all 4 diagnostic groups (Healthy Controls, N=29; Schizophrenia Spectrum, N=20; Bipolar Disorder with Psychosis, N=19; and Bipolar Disorder without Psychosis, N=35). Frazier JA, Hodge SM, Breeze JL, Giuliano AJ, Terry JE, Moore CM, Kennedy DN, Lopez-Larson MP, Caviness VS, Seidman LJ, Zablotzky B, Makris N. Diagnostic and sex...



<https://doi.org/10.18116/c6159z>  Cite





Questions?

Tess Grynoch

Tess.Grynoch@umassmed.edu

@Te_Gryn

Honor, L.B., Haselgrove, C. Frazier, J.A, and Kennedy, D.N. (2016). Data citation in neuroimaging: Proposed best practices for data identification and attribution. *Frontiers in Neuroinformatics* 10:34. doi: 10.3389/fninf.2016.00034

This work was funded in part by an administrative supplement award for Informationist Services in NIH-funded Research Projects, National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services, 3 R01 MH-083320-06S1.