

Sharing identifiable data: Lessons from Databrary

Rick O. Gilmore
Penn State



Databrary leadership



Karen Adolph



Rick Gilmore



David Millman

Databrary Team



Dylan Simon



Lisa Steiger



Mike Continues



Andrea Byrne



Lina Wictoren Roy



Eric Montellese



Jesse Lingeman



John Franchak

Agency Partners & Support

- ▶ Lisa Freund, NICHD, Science Officer
- ▶ Jim Griffin, NICHD, Program Officer
- ▶ Laura Namy, NSF, Director DLS

- ▶ NSF BCS-1238599
- ▶ NICHD U01-HD-076595-01



The Databrary project

- ▶ Data sharing library for video, related data
- ▶ Video scoring, analysis tools
- ▶ Data management tools
- ▶ Procedures and policies that enable sharing
- ▶ Changing scientific culture
- ▶ <http://github.com/databrary>





Databrary is an [open data library](#) for developmental science. Share video, audio, and related metadata. Discover more, faster.

[Learn more >>](#)

EXPLORE

Databrary is a repository of discoveries waiting to happen. Dig into the diverse source materials of developmental researchers worldwide. Spark your ideas, engage your students, and learn about behavior and development.

DISCOVER

Databrary offers tools for coding your videos, identifying patterns, and visualizing your data. Take advantage of our free tools for managing your lab and organizing your resources. Get the most out of your research.

SHARE

Advance developmental science. Increase transparency. Allow researchers to reuse your data for new purposes. Upload your materials with our lightning-fast importer and store it indefinitely. Never fear data-loss again.

- Databrary.org
- Private beta
early 2014
- Public release
spring 2014





Multiple OS support

Datavyu is a Java-based application that runs on Windows, Mac, and Linux operating systems. It requires no specialized hardware. Most labs can create Datavyu coding stations with computers already at hand. Datavyu builds on the [OpenSHAPA](#) coding tool.

Flexible and extensible

Datavyu accommodates your needs. You decide how you want to code your video, audio, physiology, motion, or eye tracking data. Power users can write scripts in the Ruby programming language and use the R statistical environment to extend Datavyu's functionality.

Free and open-source

Datavyu software and technical support is free, now and always. Datavyu is open-source. Anyone can improve Datavyu and share the new features with other researchers.

- Video coding tool
- 1.0 release Sept 2013
- Free
- 2.0 web-based



Why video?

- Rich, informative
- Transparent
- Can be re-used, re-analyzed, re-purposed



Sharing video is hard

- Identifiable
 - Faces
 - Voices
 - Locations
 - Others



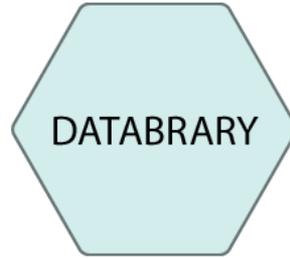
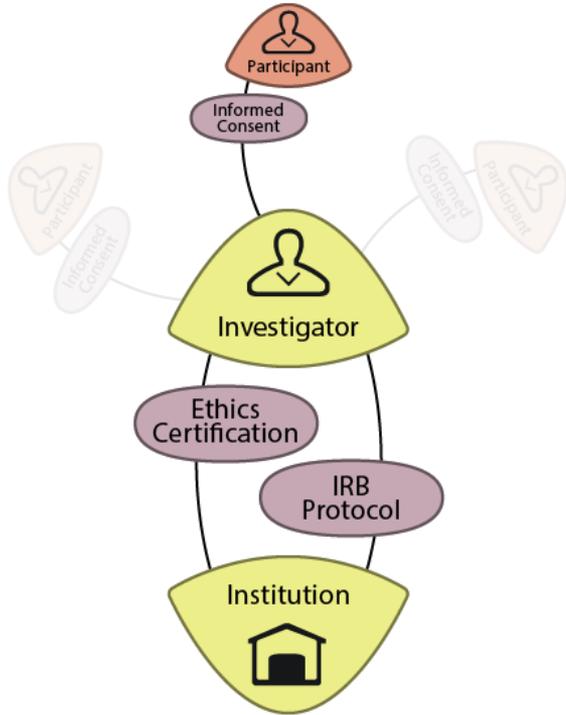
Data-sharing model



How it
works today



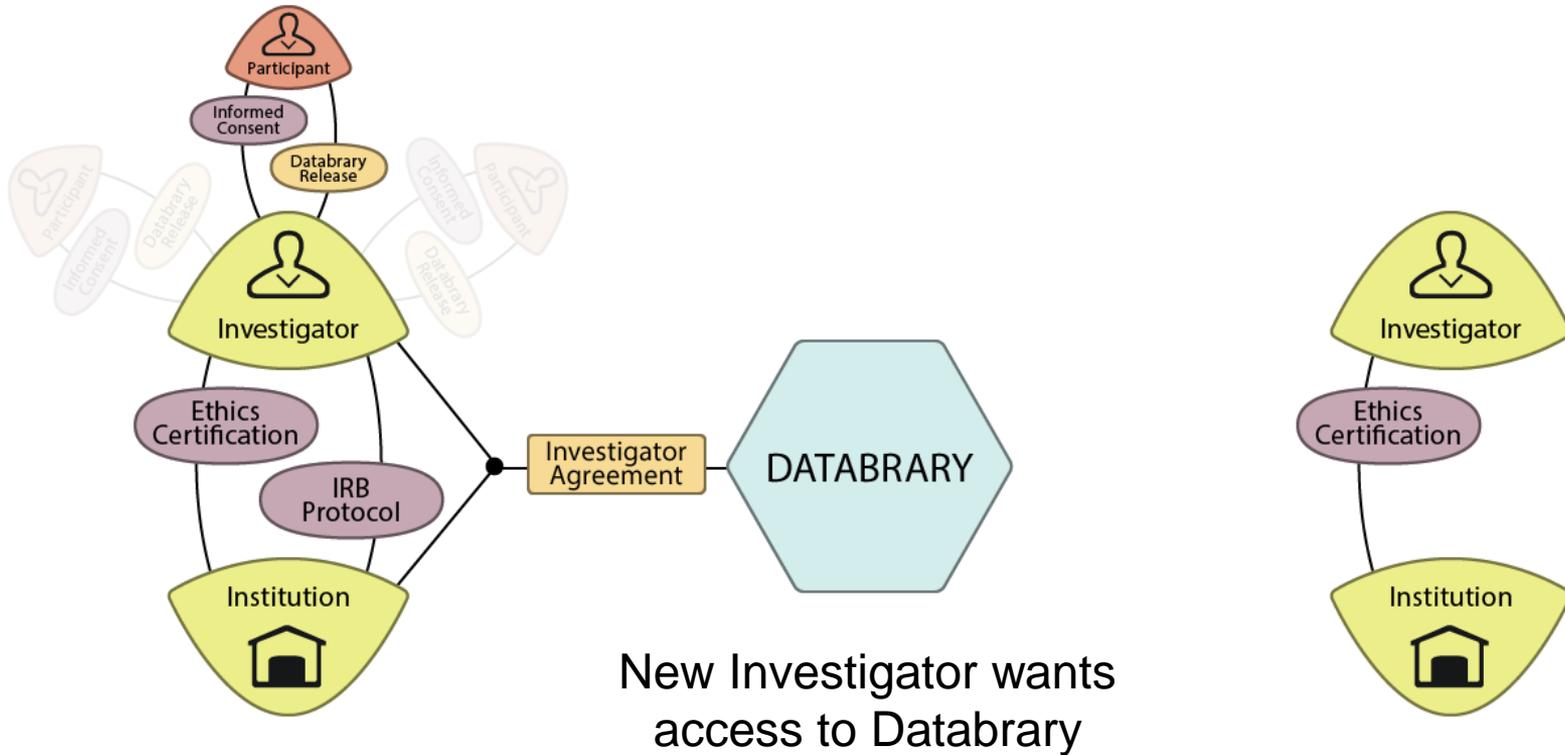
Data-sharing model



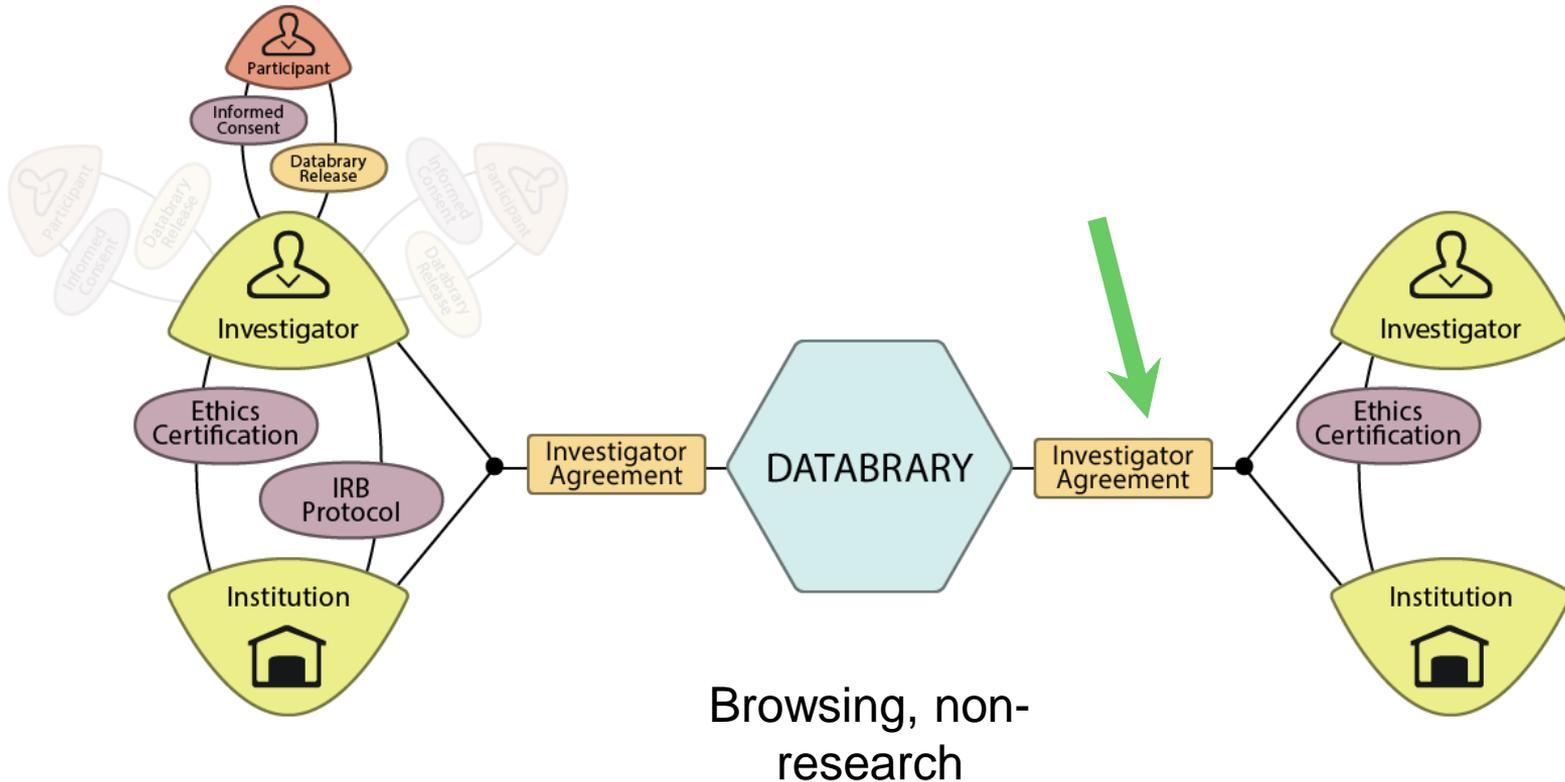
Enter Databrary



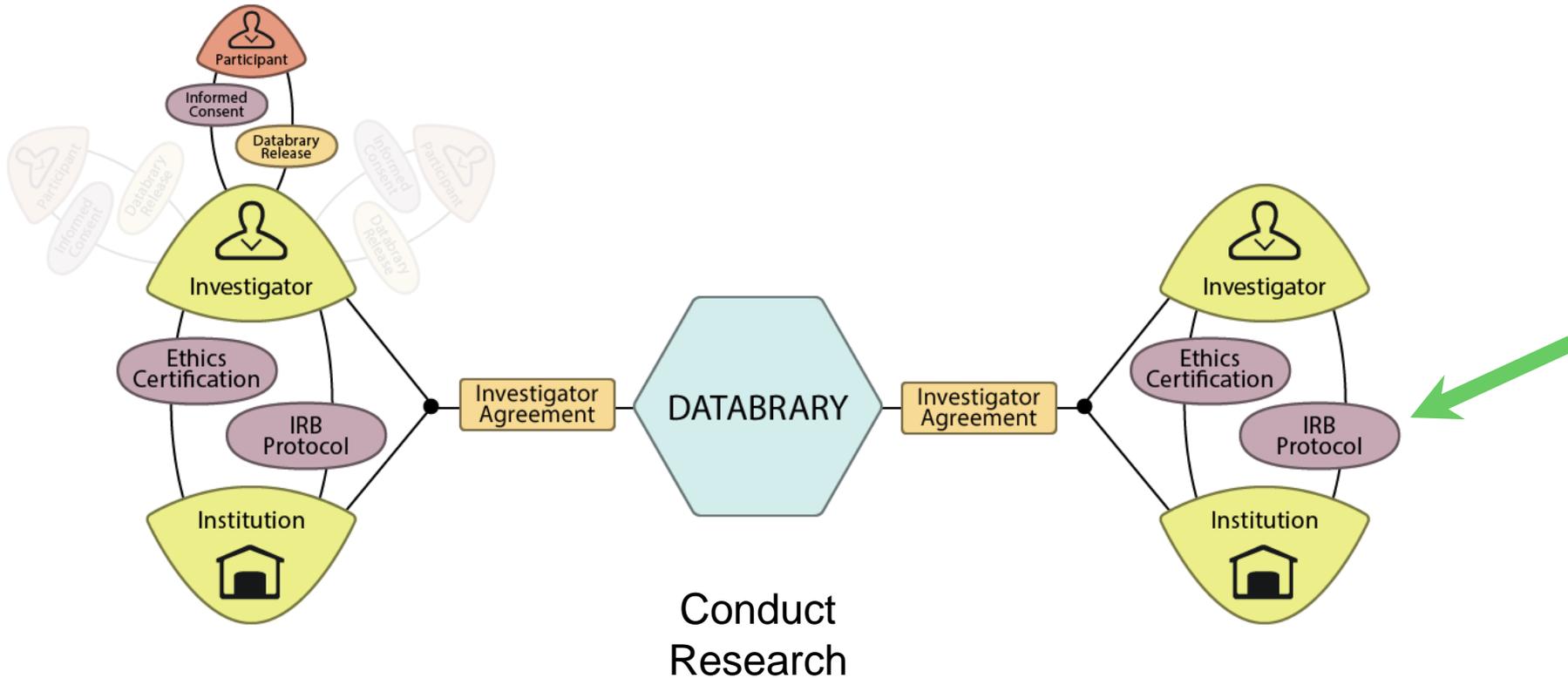
Data-sharing model



Data-sharing model



Data-sharing model



Sharing identifiable data

- Extending informed consent
- Permission to share \neq Consent to participate
- Clarifying risk
 - Public Excerpts vs. Sharing with researchers
 - [Personal Genome Project](#)
- Investigators must behave ethically



Policies that enable sharing

- ▶ Databrary Release Template
- ▶ [Investigator Agreement](#)
- ▶ [Definitions of terms](#)
- ▶ [Data Sharing Manifesto](#)
- ▶ [Bill of Rights](#)
- ▶ [Best Practices in Data Security](#)



Levels of sharing

- ▶ **Private:** No sharing
- ▶ **Shared:** Sharing only with authorized researchers
- ▶ **Excerptable:** Sharing + excerpts may be created and shown by authorized researchers to the public
- ▶ **Open:** Sharing with the public



Databrary Release Template

Databrary Release for [STUDY NAME]

Version 2013-08-01

What we are asking

This form requests your permission to share material from this session in a secure, web-based data library called Databrary (Databrary.org). Material from this session may include video/audio recordings, photographs, and images; other information such as age, birth date, sex, ethnicity, education, and self-reported health information; and information supplied by researchers who analyze data from this session. The library allows only authorized researchers to have access to shared information. Data sharing will lead to faster progress in research on human development and behavior.

If you agree to share your data, this form also requests permission to allow authorized Databrary researchers to show selected video excerpts and images from recordings of this session for scientific presentations and/or informational and educational purposes, but never for commercial purposes.

Giving permission to share your information in Databrary is separate from consenting to participate in a research study. You do not have to give permission to share your information in the library. You can agree to share your data, but not agree to allow researchers to show excerpts and images. Your choices will not affect receipt of payment if offered for participation in this or future studies.

Will the information be confidential (kept private)?

No information will be included in the data library about how to contact you or your family. No data files will be identified by anyone's name. However, images and/or voices of people in the recordings may be seen or heard. First and/or last names of people may be spoken aloud. If the study takes place in your home, aspects of your home may be seen or heard. In addition, other information may have been collected about you, your child, or your family in conjunction with this session. Thus, someone might be able to identify you, your child, or other people unintentionally. But, the authorized researchers with access to the library have agreed not to reveal the identities of people in the library.

Who has access to information in the library?

- ▶ Sharing ≠ research participation
- ▶ Data privacy
- ▶ Who has access?
- ▶ How long?
- ▶ No compensation
- ▶ Minor assent
- ▶ Levels of sharing

Recording sharing permission

Date: _____

Databrary Release for [STUDY NAME]
Signature Page

Page 4 of 4

Permission to share with authorized researchers

I give permission to share the material from this session with authorized researchers in a secure data library called Databrary.

Permission to show excerpts and images

I give permission for authorized Databrary researchers to show selected video excerpts and images from recordings of this session for scientific presentations and/or informational and educational purposes, but never for commercial purposes. I trust that authorized researchers will use professional judgment and uphold ethical principles in determining which excerpts and images to share.

By signing below, I indicate that I understand the Databrary Release document and have recorded my own choices for sharing the material from this session in Databrary and allowing authorized Databrary researchers to show excerpts.

ADULTS					
Name	I give permission to share with authorized researchers	I give permission to show excerpts and images	Adult's Signature:	Date	
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
CHILDREN DECIDING FOR THEMSELVES (ALSO REQUIRES PARENTAL CONSENT)					
Researcher must read through Assent to Share Script with those that can provide oral or written assent. Parent or guardian must also decide whether to allow their child to share, and then sign on their behalf.					
Name	I give permission to share with authorized researchers	I give permission to show excerpts and images	Participant's Signature:	Parent/Guardian Signature:	Date
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
YOUNG CHILDREN WITH PARENT OR GUARDIAN PROVIDING CONSENT					
Parent or guardian must decide whether to allow their child to share and sign on their behalf.					
Name	I give permission to share with authorized researchers	I give permission to show excerpts and images	Parent/Guardian Signature:	Date	
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
TO BE COMPLETED BY RESEARCHER OBTAINING PERMISSION					
Researcher's Name	I certify that all participants present have consented or assented above and agreed to release this session to share in Databrary.		Researcher's Signature:	Date	
	<input type="checkbox"/> Yes <input type="checkbox"/> No				

- ▶ All depicted individuals
- ▶ Explicit yes/no boxes
- ▶ Adults and minors



Permission to share with authorized researchers

I give permission to share the material from this session with authorized researchers in a secure data library called Databrary.

Permission to show excerpts and images

I give permission for authorized Databrary researchers to show selected video excerpts and images from recordings of this session for scientific presentations and/or informational and educational purposes, but never for commercial purposes. I trust that authorized researchers will use professional judgment and uphold ethical principles in determining which excerpts and images to share.

By signing below, I indicate that I understand the Databrary Release document and have recorded my own choices for sharing the material from this session in Databrary and allowing authorized Databrary researchers to show excerpts.

ADULTS					
Name	I give permission to share with authorized researchers	I give permission to show excerpts and images	Adult's Signature:	Date	
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
CHILDREN DECIDING FOR THEMSELVES (ALSO REQUIRES PARENTAL CONSENT)					
Researcher must read through Assent to Share Script with those that can provide oral or written assent. Parent or guardian must also decide whether to allow their child to share, and then sign on their behalf.					
Name	I give permission to share with authorized researchers	I give permission to show excerpts and images	Participant's Signature:	Parent/Guardian Signature:	Date
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
YOUNG CHILDREN WITH PARENT OR GUARDIAN PROVIDING CONSENT					
Parent or guardian must decide whether to allow their child to share and sign on their behalf.					
Name	I give permission to share with authorized researchers	I give permission to show excerpts and images	Parent/Guardian Signature:	Date	
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No			
TO BE COMPLETED BY RESEARCHER OBTAINING PERMISSION					

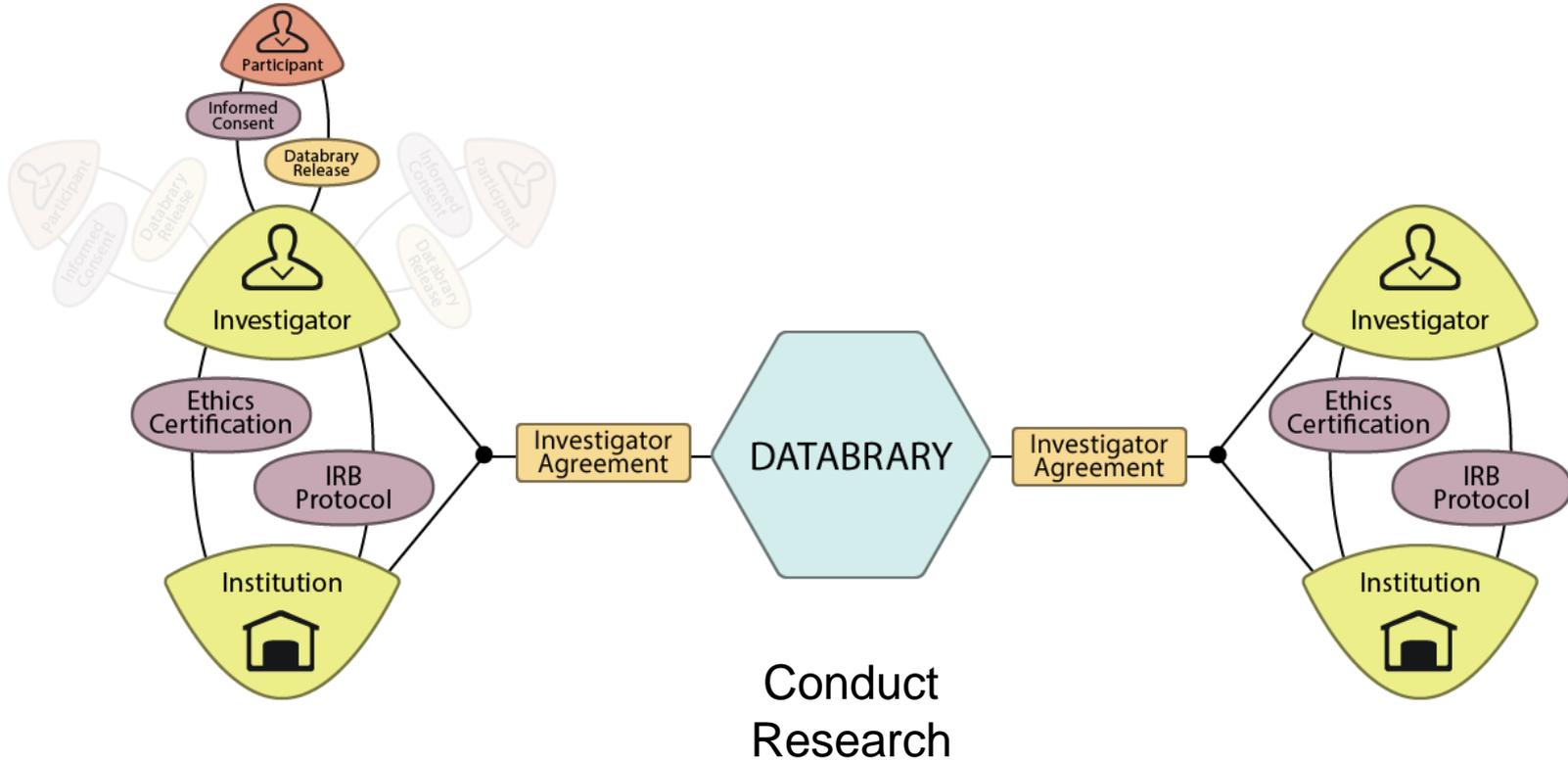


Getting permissions right

- ▶ Electronically recorded permissions
- ▶ Linked to session- and participant-level metadata
 - ▶ Avoid data entry errors
 - ▶ Honor participants' desired release level
- ▶ Spreadsheet template
- ▶ Web-based permission system



Data-sharing model



Investigator agreement

	Investigator	Institution	Databrary
Access to Data	Applies for access; Ethics training; notifies if change Institutions; supervise affiliates; not anonymous	Investigators are PIs, affiliates are associated with Institution; Certifies ethics training	Reviews, approves applications
Contributing data	Secures sharing permission from IRB, participants; transmits to Databrary; Removes PII from non-recordings	Authorizes Databrary to share	Maintains sharing permission; collects and hosts data, metadata

Who promises what

	Investigator	Institution	Databrary
Browsing, viewing data	Protect privacy; protect data; show excerpts only with permission	May receive information about usage	Keeps track of who views, downloads; citations; metadata
Using data for research	Secure IRB approval; communicate IRB to Databrary	Review and approve research protocols	Monitor research usage; store IRB protocol info
Monitoring and reporting	Report sharing or other violations	Reports sharing or other violations; Treats violations as violations of scientific integrity	Reports sharing or other violations; may deny access, remove data

The Databrary project

- ▶ Data sharing library for video
- ▶ Video scoring, analysis tools
- ▶ Data management tools
- ▶ Procedures and policies that enable sharing
- ▶ Changing scientific culture



Keeping it up

- ▶ Storage fee
- ▶ User fee
- ▶ Curation fee
- ▶ Consortial models
- ▶ Institutional support



Find & Analyze
Data

Membership in
ICPSR

Deposit Data

ICPSR Summer
Program

Resources for
Instructors

Data Management
& Curation

ICPSR Find & Analyze Data

Find Data Search/Compare Variables Find Publications Resources for Students Get Help

Sustaining Domain Repositories for Digital Data: A Call for Change from an Interdisciplinary Working Group of Domain Repositories

A pdf version of this article can be downloaded from the [ICPSR website](#) ↗.

June 24-25, 2013

Inter-university Consortium for Political and Social Research (ICPSR)

With support from the Alfred P. Sloan Foundation

Compiled by Cambridge Concord Associates

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/support/announcements/2013/09/sustaining-domain-repositories-for>

How agencies can help

- ▶ Seed domain-specific pilot projects
- ▶ Incentives for collaboration
- ▶ Incentives for sharing data, tools, expertise
- ▶ Find, empower, fund sharers
- ▶ Reduce right hand/left hand ignorance
- ▶ Promote value of public/shared data
 - ▶ NASA, National Weather Service, Census



Thank you



Databrary

Start your search...



Databrary is an **open data library** for developmental science. Share video, audio, and related metadata. Discover more, faster.

[Learn more >>](#)