

Annex III -- Databrary Access Guide

Version date: April 29, 2020

This Databrary Access Guide provides information useful for Institutions and Authorized Investigators who wish to use Databrary's services.

For Institutions

Databrary is a restricted access data library specialized for storing, streaming, and sharing of video and audio recordings. Databrary's mission is to support the sharing of video and audio recordings and related data collected as part of research in the social, behavioral, educational, and neural sciences.

The responsibility for protecting research data and research participants is shared among Authorized Investigators and their Affiliates, Institutions, and Databrary.

Databrary has a multi-pronged approach to fostering the sharing of video and related data while upholding ethical research principles. Access is restricted to Authorized Investigators whose access must be approved by their Institution. Institutions must provide ethics oversight of Authorized Investigators through a research ethics board or Institutional Review Board (IRB) with standards equivalent to that required to secure Federalwide Assurance (FWA) status from the U.S. Government (see: <https://www.hhs.gov/ohrp/register-irbs-and-obtain-fwaf/index.html>). Authorized Investigators who share data with others must secure permission to share from research participants and record that permission level for each file stored on Databrary. In turn, Authorized Investigators who access Databrary and use data shared by other Authorized Investigators must follow ethical research principles, including respect for the sharing permissions granted by research participants. Taken together, these features aim to protect the rights of research participants while reducing barriers to sharing data.

The Databrary Access Agreement permits data use and data contribution

Databrary restricts access to Authorized Investigators whose Institutions sign a formal Databrary Access Agreement.

The Databrary Access Agreement combines provisions for data use and data contribution. Unlike data use agreements that allow specific individuals access to a particular dataset for a discrete research purpose, the Databrary Access Agreement allows any Authorized Investigator to access all shared datasets on Databrary and use the datasets for many different Institutionally-approved purposes. In addition, the Databrary Access Agreement permits Authorized Investigators to upload data and to share it with other Authorized Investigators in accord with the sharing permissions granted by participants and any required Institutional approvals. Databrary believes that combining data use and data contribution in a single agreement reduces barriers to data sharing, meets emerging funder mandates to share research data, and accelerates progress in research.

Video's use cases extend beyond research

In addition to its many uses in research, video may be used for a variety of pre-research and non-research uses, such as the documentation of methodological details, the demonstration of computer-based tasks, or the showing of video clips for teaching or scientific presentations. Databrary encourages the storage and sharing of video for all of these purposes, in a manner consistent with subject permissions and institutional approvals.

Institutions determine when research ethics board or IRB approval is necessary

Databrary does not require prior research ethics board or IRB approval for an Authorized Investigator to be granted access to the system or to use Databrary's shared resources for non-research, educational, or pre-research uses. However, in signing the Databrary Access Agreement, all Authorized Investigators promise to secure research ethics approval whenever their use of Databrary requires it under the Institution's policies.

Non-research, educational, or pre-research use cases may not require research ethics board or IRB review and approval at every institution. Institutions decide which use cases require research ethics board or IRB review and approval.

What data are shared

Databrary contains images, video, and audio recordings that may be identifiable, data and metadata that are not typically sensitive, and de-identified data. Exact birth dates may be stored for use in calculating exact ages, but exact

birth dates are only exposed to Authorized Investigators or Affiliates given specific access to data. Names, addresses, email addresses, financial information, government ID numbers, detailed geographic location information, and other personally identifiable data elements are **not** stored and shared on Databrary.

Securing permission to share

In order to contribute and share identifiable data with Databrary, Authorized Investigators must secure permission to store the recordings on Databrary from research participants and any other people (e.g., research staff) who are recorded.

Databrary has developed a Sharing Release Template that informs people depicted in recordings of the potential risks associated with sharing video and other data on Databrary. The Sharing Release Template is one of the resources that Databrary has made available to researchers. The Sharing Release Template may be adapted for use at any Institution and added to any research ethics or IRB protocol. However, an Institution may require other release language in order to provide equivalent protections.

Whether de-identified or pseudonymized data may be shared with Databrary without seeking explicit permission from research participants depends on the Authorized Investigator's Institution. Some ethics boards or IRBs may allow de-identified or pseudonymized data to be shared without the explicit permission of the research participant.

Uploading data and assigning sharing permission levels

Authorized Investigators who upload data to Databrary must assign to each file uploaded the level of data sharing granted by research participants--from none to public access. The sharing level assigned to each file by default is *Private*, meaning that no person other than the Authorized Investigator who uploaded the file and any Affiliates granted access by the Authorized Investigator, may view the file.

Not only are data files assigned a *Private* sharing level by default, but all newly created datasets are initially made *Private*, accessible only to an Authorized Investigator, any Affiliates the Authorized Investigator selects, and any other individual Authorized Investigators granted specific access to the dataset. An Authorized Investigator may elect to create an overview of a dataset when it is created and make the overview publicly available. The overview can serve as the public face of the dataset while the dataset is being collected and analyzed. This can be useful in documenting progress on a research grant to sponsors. The dataset overview does not expose data unless the Authorized Investigator chooses to do so.

Sharing data

At a time when the Authorized Investigator chooses--e.g., when a paper goes to press or a grant period ends--the Authorized Investigator may choose to share the dataset with other Authorized Investigators on Databrary. This makes the data contained in the dataset available to other Authorized Investigators and possibly the public, but only in accordance with the sharing levels selected by Authorized Investigators and always subject to the permission granted by participants. That is, individual data files marked *Private* remain accessible only to people the Authorized Investigator specifically selects. Similarly, access to datasets labeled *Private* remains solely under the control of the Authorized Investigator.

Databrary and GDPR

Authorized Investigators may collect personal data from research participants who have rights under the General Data Protection Regulation (GDPR). GDPR governs the collection of personal data in the European Economic Area (EEA).

Databrary's servers are currently located in the United States, and data stored on Databrary may be accessed, downloaded, and re-used by Authorized Investigators and their Affiliates located outside of the EEA. There are more than 500 Institutions across the globe that have signed the Databrary Access Agreement. Databrary maintains a list of Authorized Investigators who have access to shared research data at https://nyu.databrary.org/search?volume=false&f.party_authorization=4&f.party_is_institution=false.

In seeking permission to collect data from research participants governed by GDPR or similar provisions, Authorized Investigators should communicate to research participants information about who may have access to shared data on Databrary and where the data are stored. Institutions and Authorized Investigators assume responsibility for

ensuring that research participants give sharing data permission that satisfies the provisions of GDPR and for abiding by other provisions of GDPR.

Databrary does not typically store information that links a person's identity to specific files unless there is a separate agreement between the Authorized Investigator, the Institution, and Databrary to store this type of data. Therefore, if a research participant requests a copy of data stored on Databrary or the modification or deletion of personal data that is stored on Databrary, the Authorized Investigator or Institution is responsible for responding to the request, not Databrary. Also, Databrary cannot guarantee that data previously shared with other Authorized Investigators via Databrary can be retrieved from all Authorized Investigators who accessed a participant's personal data prior to its modification or removal. Thus, Authorized Investigators should communicate to research participants that there are limits to a participant's right to modify or delete personal data that have already been shared on Databrary. The Databrary Release Template contains language for this purpose.

When accessing data shared by other Authorized Investigators, Institutions and their Authorized Investigators assume responsibility for ensuring that the sharing permission research participants have given in other contexts outside of the EEA meet relevant GDPR provisions.

When Authorized Investigators leave an Institution

Authorized Investigators are expected to inform Databrary when they leave an Institution. Institutions may also inform Databrary.

Access to non-public data shared with other Authorized Investigators will be suspended when an Authorized Investigator leaves an Institution. The suspension extends to an Authorized Investigators' Affiliates. Authorized Investigators will continue to have password-protected access to data they uploaded to Databrary once they leave an Institution. When a new Institution grants authorization, Authorized Investigators will regain access to non-public data shared by other Authorized Investigators.

If an Authorized Investigator retires, is incapacitated due to illness, or dies, the Institution should inform Databrary so that the Authorized Investigator's account may be suspended. An Institution may choose to designate a data steward for the Authorized Investigator's data and authorize that person to manage access to the data. If the Institution does not identify and authorize a data steward, Databrary staff will serve that role. However, the Databrary staff serving as data steward will not change the data sharing status of any specific dataset without explicit written guidance from the Institution.

Selecting Affiliates

If permitted by their Institution, an Authorized Investigator may grant a person outside their Institution Affiliate status. Institutions may establish policies governing the selection and management of Affiliates by their Authorized Investigators.

If a person is eligible to be an Authorized Investigator at their home Institution, that person must apply for access to Databrary through their home Institution. Except in unusual circumstances, researchers eligible to be Authorized Investigators at their home Institution should not be granted access to Databrary as the Affiliate of another Authorized Investigator either from the same Institution or a different one. Databrary welcomes questions about these policies. Please email contact@databrary.org¹.

For Researchers

This section explains how to apply for authorization, select and manage affiliates, collect data that can be shared, upload and manage datasets and studies, and explore, re-use, and extend the resources shared with Databrary.

Getting Authorized

Some materials on Databrary are available to the public, but access to most data requires that a person secure formal written authorization from an Institution. The Databrary Access Agreement is the legally binding document that formalizes the authorization process.

Individuals who are eligible to conduct independent research, submit funding proposals as the lead investigator, have Principal Investigator (PI) status, or have tenure-track faculty positions at their Institution, may apply to their

¹<mailto:contact@databrary.org>

Institution for approval to be named an Authorized Investigator. To receive authorization a person must apply for and receive authorization from their Institution; Databrary does not grant authorization separate from an Institution. In the U.S., a college or university's grants and contracts or sponsored programs office usually takes responsibility for granting authorization by signing the Databrary Access Agreement. Databrary staff will work with a researcher and an Institution to secure signatures on the Databrary Access Agreement. The authorization process typically takes a few days to a few weeks.

Selecting and managing Affiliates

Once authorized by their Institution, Authorized Investigators may then authorize staff, undergraduate or graduate students, or postdoctoral trainees to access Databrary as Affiliates. In granting a person Affiliate status, an Authorized Investigator takes responsibility for that person's use of Databrary.

An Authorized Investigator may grant a person outside their Institution Affiliate status if permitted by their Institution. If a person is eligible to be an Authorized Investigator at their home Institution, that person should apply for access to Databrary through their home Institution. Except in unusual circumstances, researchers eligible to be Authorized Investigators should not be granted access to Databrary as the Affiliate of another Authorized Investigator.

Authorized Investigators are responsible for ensuring that Affiliates have human subjects and/or animal research training appropriate for the Institution and remain in good standing with the Institution.

Affiliate access to an Authorized Investigator's data

Authorized Investigators determine which of the Authorized Investigator's datasets an Affiliate has access to. Authorized Investigators can set each Affiliate's access to one of the following levels:

- *No access*: Affiliate cannot access any of the Authorized Investigator's data. This is the default.
- *Read-only access*: Affiliate has view access to all of the Authorized Investigator's data. (The affiliate may view and download videos and other identifying information, including data that is not released for sharing.)
- *Read/Write Access*: Affiliate can access and edit all of the Authorized Investigator's. (The affiliate may view, download, add, change, and remove data as the Authorized Investigator without restriction.)
- *Manager*: Affiliate can access and change permissions on all of the Authorized Investigator's data. (The Affiliate may additionally change the access controls and sharing levels of the Authorized Investigator's data. Manager-level permission should be granted with utmost care.)

These levels are analogous to granting an Affiliate different levels of access to the Authorized Investigator's laboratory data server. Accordingly, the Investigator must trust Affiliates granted this access and supervise and manage Affiliates granted these privileges. The Authorized Investigator must assign an expiration date to each Affiliate's access.

Affiliate access to Databrary

Affiliates can only gain access to data on Databrary that has been shared by *other* Authorized Investigators if specifically granted that access by their supervising Authorized Investigator. The Authorized Investigator must assign an expiration date to each Affiliate's access.

Many Authorized Investigators grant Affiliates access only to datasets the Authorized Investigator controls, not to other datasets shared on Databrary. Some Authorized Investigators grant access to datasets shared on Databrary to senior graduate students, postdoctoral trainees, or trusted research staff.

Accessing data shared by others

Authorized Investigators automatically have access to data on Databrary that has been shared by other Authorized Investigators. Affiliates only have access to shared data on Databrary only if granted that access by the Authorized Investigator who supervises them.

Release levels

In order to make it clear to participants and researchers what level of access should be granted to specific files, Databrary has developed a set of standard release levels. These levels are indicated with consistent icons and language across the site. The levels are as follows:

- *Unknown*. Consent missing or not obtained. This content is not shared and access is restricted to specific people selected by an Authorized Investigator.
- *Private*. This content is not shared and access is restricted to specific people selected by an Authorized Investigator. *Note*: This is the default for all files and must be specifically changed by an Authorized Investigator (or an Affiliate with suitable privileges).
- *Authorized Users*. This content is restricted to authorized Databrary users (Authorized Investigators and Affiliates granted access to shared Databrary data) and may not be redistributed in any form.
- *Learning Audiences*. This content is restricted to authorized Databrary users (Authorized Investigators and Affiliates granted access to shared Databrary data), who may use clips or images from it in presentations for informational or educational purposes. Note that such presentations may be videotaped or recorded and those videos or recordings may then be made available to the public via the internet (e.g., YouTube).
- *Public*. This content is available openly to anyone.

All Authorized Investigators and Affiliates must follow the restrictions on uses indicated by these release levels. Furthermore, when using files released for access by *Learning Audiences* Authorized Investigators and Affiliates must exercise professional judgment in deciding what clips or images to show to what audiences and in what contexts.

Non-research uses

Many uses of Databrary do not constitute research, for example, browsing shared videos, downloading videos of displays or procedures, or showing excerpts in research or educational presentations. Authorized Investigators (and their Affiliates) must follow Institutional policies concerning whether specific non-research uses of Databrary require prior Institutional approval. Moreover, Authorized Investigators and Affiliates must always respect the release level associated with shared files. For example, data marked with an *Authorized Users* release level may only be viewed in research contexts (e.g., in the Authorized Investigator’s laboratory) by the Authorized Investigator or people under the supervision of the Authorized Investigator. Similarly, only those data with a release level of *Learning Audiences* or *Public* may be shown in research or educational presentations outside the Authorized Investigator’s laboratory.

Research uses

Authorized Investigators must secure Institutional approvals prior to conducting research with data shared on Databrary. In the U.S. context, conducting secondary analysis on shared video data may not qualify as human subjects research because there is no direct contact with research participants. Still, a research project involving video data reuse may require ethics board or IRB review and approval because the data are identifiable.

Contributing new data

Only Authorized Investigators may create new datasets, and only Authorized Investigators may share datasets with others.

Non-sensitive/non-identifiable data or materials

Authorized Investigators may upload to Databrary and share non-identifiable videos of research procedures, computer displays, research protocols, analysis code, or other research materials without any additional approvals, provided that sharing these elements with Databrary does not violate intellectual property rights of any person. De-identified or pseudonymized research data may be uploaded and shared subject to Institutional policies and approval. If an Authorized Investigator already has approval to share a de-identified or pseudonymized dataset with other researchers or with the public, then sharing that dataset on Databrary is permitted. All files uploaded to Databrary must be marked with a sharing release level. The default sharing release level is *Private*. While many non-sensitive/de-identified or pseudonymized data files could be made available to the public, it may be prudent to share de-identified or pseudonymized datasets only with Authorized Investigators.

Identifiable or sensitive data

Authorized Investigators must make preparations to share identifiable or sensitive data with Databrary. The next section describes some of these preparations.

Securing Institutional approval Authorized Investigators must secure approval from their Institutions prior to sharing data with Databrary. Institutional approval usually involves research ethics board or IRB review or its equivalent. Approval to share data may be included in a new research protocol application or as an amendment to an existing protocol.

For human subjects data, Databrary recommends adapting the Databrary Release Template although the exact Databrary Release Template language is not required. The Databrary Release Template serves as a guideline. The Institutional research ethics board or IRB must determine what specific language should be used for securing data sharing permission and whether requesting permission to share data should be part of the research consent process or separate from it. It is essential that whatever language and permission process the Institution requires enables the Authorized Investigator to determine clearly what level of sharing release participants have granted. Authorized Investigators must be able to assign a sharing release level to each file uploaded to Databrary that reflects the sharing permission granted by the research participant.

In some cases, research ethics boards and IRBs have deemed that certain release language in research consent forms from previously collected data permits sharing in Databrary. This has made archival data collected prior to the creation of Databrary to be stored and shared in the library. If an Authorized Investigator has archival data, Databrary staff are happy to discuss whether and how these might be stored and shared on the system.

Removing data destruction clauses There are often misconceptions about regulations pertaining to data destruction, especially where video or audio recordings are concerned. In the U.S., NIH and NSF do not require data destruction. Data destruction clauses inherently conflict with the ability to store and share data indefinitely via repositories like Databrary. Rather than promising to destroy data after some time period in a research protocol, Databrary recommends the following language: "Data will be stored for an indefinite period in a restricted data library on the internet (databrary.org)."

Obtaining participant release for sharing Once research protocol approval has been granted, Authorized Investigators may start securing permission from participants and recording those permissions.

Authorized Investigators must ensure that the process used for obtaining participant release for sharing meets relevant regulatory requirements, e.g. GDPR.

Why ask Video and audio recordings of participants are inherently identifiable. In order to share identifiable data on Databrary Authorized Investigators must secure permission from the people depicted in recordings or images. Authorized Investigators can share identifiable research recordings (or other identifiable data elements like birth dates) only when the Authorized Investigators (a) get approval from their Institutions to do so, (b) inform participants about what sharing these data elements means, and (c) secure participants' permission to share. Databrary calls this data sharing permission a "release" since it involves releasing data for use by other researchers.

When to ask Databrary recommends that Authorized Investigators ask participants for permission to share at the end of data collection after participants have acquired first-hand knowledge about the study procedures. This ensures that consent to participate is completely separate from permission to share and the risk that asking participants about sharing their data will affect their decision about whether or not to participate in research. However, some Institutions may prefer to have Authorized Investigators incorporate data sharing permission into the research consent process or even to have data sharing permission be part of an entirely separate protocol. The important and essential point is to secure sharing permission and record it faithfully.

How to ask To standardize the process of obtaining data sharing releases from participants---and make it as simple as possible for researchers---Databrary has developed the Databrary Release Template. The template can be added to or incorporated within human subjects research ethics protocols. The Databrary Release Template was developed and refined in close consultation with the IRB office at NYU. It gives researchers a standardized way to obtain the necessary permissions to share data in Databrary. The template makes it clear and unambiguous to participants about what agreeing to share data in Databrary will mean.

Uploading to Databrary

What to upload In addition to video, Databrary can store many types of data and supplemental files associated with a research project. These can be images, graphs, stimuli, protocols, blank questionnaires, analysis code, and

consent or release documents. Databrary permits Authorized Investigators to store any information permitted by the Institution. However, some information poses greater risks to participants than others, and these should be stored on Databrary only after special consideration. For example, first or last names, home location information more specific than state, and contact information are generally considered high risk for the disclosure of participants' identities. These data items should **not** be uploaded to Databrary.

Video and audio recordings may be uploaded to Databrary if the Authorized Investigator has secured participant (or parent) permission and the Institution approves. Under some circumstances, an Institution may determine that participant (or parent) permission is not necessary to store video and audio recordings on Databrary. For example, regarding recordings collected a long time in the past, participant (or parent) permission may be unavailable, and the Institution may determine that the risk to participants is low. Or, an Authorized Investigator may choose to store video and audio recordings on Databrary for sharing only among collaborators and Affiliates specifically chosen by the Authorized Investigator, but have no plans to share the videos with the broader Databrary community. This use of Databrary is permitted, but discouraged, as the purpose of Databrary is to foster data sharing beyond an individual research project or group.

Exact birth dates or test dates may also be considered identifiable data elements. However, these dates can be highly valuable for future research uses, and may pose a lower level of risk to participants than some other data elements. Accordingly, Databrary permits the storage of exact dates, but treats exact dates with special care. The system only exposes exact dates to Authorized Investigators or their Affiliates who upload the data, and only exposes exact dates in shared datasets if the participant has agreed to share data (grant sharing release at the *Authorized User* level or above).

Other low risk data items like race, ethnicity, gender, and related personal characteristics may be stored on Databrary without specific permission.

Assignment of sharing release level The only essential item of information required for uploading files to Databrary is the sharing release level. Authorized Investigators are responsible for recording participant sharing release permissions and assigning them to files uploaded to Databrary. Databrary assigns all files a default release level of *Private*, so an Authorized Investigator must change the default *Private* release setting in order to share a file with people other than an Authorized Investigator's Affiliates or Authorized Investigators who are research collaborators on a specific dataset.

Who may upload Authorized Investigators may upload data, and Authorized Investigators determine who among their Affiliates may also upload data to Databrary.

When to upload Authorized Investigators need not wait until a research project ends to upload data to Databrary. Authorized Investigators may upload files to Databrary as soon as possible after a data collection. This is called active curation. It distributes the time required to curate a dataset for sharing and thereby reduces the overall burden of data sharing.

Non-human animal data

We are developing guidance for the use of Databrary to store non-human animal data.

Sharing data

Authorized Investigators determine when data stored on Databrary are ready to be shared with the broader community of Databrary Authorized Investigators. Typical 'triggers' for sharing a data collection are when a paper goes out for review or goes to press, when a grant report is prepared, or when a grant period ends. To share a dataset, the Authorized Investigator simply changes the access level on that dataset. This makes metadata about all data and materials visible so that other Authorized Investigators can access them, subject to the sharing release levels indicated for each file. Datasets receive a persistent identifier (DOI) when shared. Authorized Investigators may "unshare" a dataset, but except under unusual circumstances, this is not considered best practice.

Removing or altering participant data

Research participants in the European Economic Area (EEA) have certain rights under the General Data Protection Regulation (GDPR) pertaining to the use, sharing, and removal of their personal data.

Since Databrary does not typically store information that maps personal identities to the data stored on Databrary without a separate agreement, Authorized Investigators and Institutions are responsible for responding to any requests to provide copies of, change, restrict access to, or delete a participants' personal data stored on Databrary. Databrary cannot guarantee that personal data previously shared on Databrary can be retrieved from all Authorized Investigators who have accessed the data. Thus, in seeking permission from research participants to share personal data via Databrary, Authorized Investigators should communicate that there are limits to a participant's right to modify or delete personal data previously shared on Databrary.

Changing institutions

Authorized Investigators who change Institutions must notify Databrary of the change and its effective date. As soon as possible, the Authorized Investigator must secure authorization from their new Institution. Authorized Investigators who are awaiting authorization from their new Institution will have access to their own data but not to non-public data shared by other Authorized Investigators on Databrary. Once an Authorized Investigator has secured approval from the new Institution, full access privileges will be restored.

Definitions

Access Agreement: The Databrary Access Agreement between an Authorized Investigator and his or her Institution and New York University (NYU) on behalf of Databrary.

Affiliate: A person who conducts research under an Authorized Investigator's sponsorship and supervision and who has been granted access to Databrary by the Authorized Investigator. In most cases, Affiliates are students or research staff and are not eligible for Authorized Investigator status.

Authorized Investigator: A person who has permission from an Institution to conduct independent research and who has received authorization from an Institution and from Databrary to access Databrary.

Authorized Organizational Representative (AOR): A person granted authority by an Institution to sign the Databrary Access Agreement on behalf of the Institution and thereby bind the Institution to the terms of the Databrary Access Agreement.

Institution: An academic, not-for-profit, research, government, or health entity whose employees or students conduct scientific research or educational activities.