Description of Databrary

Databrary is a web-accessible, video data repository for researchers in the behavioural and learning sciences. It allows researchers to store, manage, and share the videos they collect through the course of their studies. It also allows them to attach metadata and supplementary documents to these videos to help with data management in their lab and for sharing with colleagues.

Databrary currently supports two types of adding data: 1) after-the-fact ingestion by Databrary staff and 2) allowing users to upload their videos and metadata as they conduct their study. The web application is largely geared for the latter, and is built with the intent of making this process more intiutive and similar to their existing local workflows.

Because video data of human subjects are sensitive, highly customizable access levels for the data users contribute are central to the functionality of Databrary. Users can choose from four different access levels (private, share with authorized Databrary users only, share with authorized Databrary users only adding the ability to use of video excerpts in educational settings, and public) to specify with whom and how they share their data. Because of this, what you can see and do on Databrary depends on a user's role.

Primary User Roles

There are two primary users of Databrary. The first are Principle Investigators (PIs), who are identified on Databrary as Authorized Investigators. The second are Lab members, who are identified on Databrary as Affiliates.

PIs (Authorized Investigators)

Authorized Investigators are the Databrary users that must be authorized by their institution to contribute and share data on Databrary. Often they are university faculty or lab directors at a research institution. They are generally considered to be the lead on a study and also to be the data owner for a volume (which is the representation of a study or dataset on Databrary). They have full read/write permissions for a volume and are the only users who can actually share the volume with the rest of the Databrary community once ready to do so.

While it is not a hard and fast rule, we believe that PIs should be creating and describing the volume and establishing the study design, which consists of selecting the metadata categories and fields that make up the structure for the description of participants and study sessions in their volume. Normally, they are not the parties that will be uploading videos and metadata to Databrary. Lab members who work for the PI will usually undertake these tasks.

The amount and nature of work a PI does on Databrary is potentially determined by where they are in their career. We've noticed that newly minted or more junior faculty tend to be more hands-on and therefore more fluent with new technical workflows. More experienced PIs may have less incentive in their day to day operations to learn new technical workflows, and often delegate these responsibilities to a lab manager. We have tried to make Databrary usable for both parties.

Lab members (Affiliates)

PIs sponsor members of their lab to participate and assist in contributing data to Databrary. We anticipate that much of the day to day work of uploading files and adding and managing metadata will be conducted by

Affiliates. Affiliates tend to be undergraduate and graduate students, postdoctoral researchers, or full-time staff or lab managers. Their traditional roles in the lab are to collect and organize data, use desktop software to annotate videos that will help them analyze the data they collect, and handle other routine operations of the lab. It will largely be their responsibility to take data and metadata from the labs local context and put it onto Databrary using the spreadsheet metadata viewer/editor and timeline interface, which depicts a single row of the spreadsheet view.

Other Users

Other users might be researchers who would like access to Databrary materials without requiring access to share or contribute data themselves. These users would interact mostly with search and browse capabilities in order to find datasets and highlights that they might use in their own work.

Key Terminology

Databrary has defined a set of terminology for navigating and using the service. Some of these terms may be found in our glossary and we do not assume that users will know the specific meaning of these terms ahead of time.

- **Volume:** A volume marks a complete dataset in databrary.
- **Session:** A session is a unit within a volume which consists of a set of videos and records. A session most often represents a single visit by a participant or set of particapants interacting together in a study.
- **Spreadsheet:** Refers to the user interface on a volume's homepage that allows users to quickly view and edit metadata across multiple volume sessions.
- **Timeline:** Refers to the user interface on a single session in a volume. It is designed to look like video editing software and allows users to upload and manage files associated with a session.
- **Metadata field:** A single metric that describes something about the participant or session such as participant birthdate or session recording date.
- **Metadata category:** A category is a grouping of fields such as participant, session tasks, or session context.
- **Highlights:** Assets or parts of an asset (video clip) highlighted by the investigator to feature on the volume page. Depending on the release level specified by the volume owner, some highlights may be used in talks and lectures.