

Data Management Plan

Types of data:

This project will generate a diverse set of data, including:

- **Interview recordings and transcripts** performed with robotics researchers, which will be in video and textual formats. These will be stored on cassettes, hard drives, and DVDs in the locked offices of the PIs and participating subcontractor.
- **Open Source Software Code** which will be developed by the project's software developer. This data will be shared among collaborators using a version control site such as Git. A "beta" release, and a final "1.0" release will also be made available to open source developers on the web through Github, and/or through our Oral History of Robotics website.
- **Oral History of Robotics Website** will be initially hosted on a server at Indiana University for the prototyping, beta testing and development phase. Once complete, the archive will be permanently hosted on a site that is freely accessible to the public and supportable, either at Indiana University, IEEE, or another suitable site.

Data and metadata standards

Whenever possible, we will use standard, widely used file formats for storing the data and metadata produced throughout the course of the project. We plan to use MPEG-4 or WebM open source codecs for storing digital video, Word documents and Adobe Portable Document Format (PDF) for storing interview transcripts, and an OpenSQL database for storing bibliographic and other data.

These particular formats are chosen for the sake of familiarity of the research team and other possible users. All collected data will be tagged with the date, time, and place where it is collected.

Privacy and access policies

The oral history interview data that has been gathered and that will be collected in the future has gone through IRB approval and consent, in which the interviewees specify how we may use the videos in research and dissemination. A deed of gift must be signed by all the interviewed roboticists before any part of their interviews is made publicly available on IEEE website. The remaining proprietary datasets will not be made publicly available.

Data Access and dissemination

During the project, we will be disseminating research results through publications in professional meetings and journals, as well as through discussions with the media. We intend to make oral history interviews, transcripts of the interviews and database of full CVs publicly available. In cases where individuals who are part of the community under study may also have access to all the data and its results upon publication following the studies, we will be developing the norms of data use and identification with the input of the members of those communities. We intend to make as much of our data as possible freely available to external research groups upon request.

Research team access to data Aside from making selected data available to the public, the research team will also share more detailed data amongst the members through a secured database hosted on an Indiana University Bloomington server. Research team members (faculty and graduate students) will be given usernames and passwords, which they can use to access the data. This data will also be shared with external researchers upon request and approval of the external team's privacy protection plan.