

9. Data Management Plan

I. Legal Considerations for the Use of Podcast Metadata

There are no intellectual property restrictions that obstruct the proposed project. Whereas most of the audio files saved in the PodcastRE database are still protected by copyright, the basic metadata that describes these files are NOT subject to copyright protection. According to the Digital Public Library's Policy Statement on Metadata, "the vast majority of metadata is not subject to copyright protection because it either expresses only objective facts (which are not original) or constitutes expression so limited by the number of ways the underlying ideas can be expressed that such expression has merged with those ideas. To be protectable, a work must be original, which means that it must contain at least a 'modicum' of creativity in its creation, selection, or arrangement. Facts and ideas may not be copyrighted" (DPLA, 2013). Moreover, PodcastRE Analytics uses the metadata in a transformative manner; we either transform existing metadata fields to allow them to be analyzed or visualized in new ways, or we generate new metadata fields entirely through forms of sonic analysis. Ultimately, PodcastRE Analytics has two layers of copyright law on its side: we are making a transformative fair use of metadata, which is generally understood not to be copyrightable in the first place. All other data generated as part of the project will be authored by the project team, and the PodcastRE Analytics platform will be built exclusively using open source software and code.

II. Expected Data that Project will Generate

Type of data	When will they be shared?	How and under what conditions?
Metadata XML files that describe every podcast in the PodcastRE database.	January 2019 (the start of the second dissemination phase of the work plan).	XML files will be indexed in platform's Solr index and freely available for users to query. XML files will also be preserved and made accessible by Minds@UW.
Open source code for PodcastRE Analytics platform, built using Python, Javascript, and Solr.	January 2019 (the start of the second dissemination phase of the work plan).	The code will be installed and operational on a UW-Madison server as a web application, freely available to all users. The code will also be freely available on GitHub.
Two research journal articles, output as PDFs.	Upon their publication (most likely before the end of 2019).	The articles will be published in open access journals and accessible via the journal websites, PodcastRE website, and Minds@UW.
Blog posts reporting on the project's progress and early research results.	Immediately as they are posted online.	Blog posts will be available via the PodcastRE website and, by April 2019, Minds@UW.
Podcasts discussing project's progress and early research findings.	Immediately upon their completion.	Podcasts will be freely available via Apple iTunes, Stitcher, the PodcastRE website, and, by April 2019, Minds@UW.
White paper.	Immediately upon the completion of writing and editing.	White paper will be available via the PodcastRE website and Minds@UW before the end of 2019.

III. Period of Data Retention

All of the data described in Section II will be publicly available—in most cases, before the end of the funding period, but in some cases (such as the journal articles) up to a year after the funding period ends. The data will be retained during the grant period and for a minimum of 3 years afterwards on a secure Linux-based web server and XSAN local storage network, both of which are managed by the Instructional Media Center in UW-Madison's Department of Communication Arts. After the grant period ends, the data will be migrated to the University of Wisconsin-Madison Libraries and retained in Minds@UW, the Library's Fedora Commons-based Institutional Repository and preservation platform, for a minimum of 20 years.

IV. Data Formats and Dissemination

- Podcast metadata will be saved in XML files and indexed in PodcastRE Analytics's Solr index and for users to freely query. The XML files will also be preserved and made accessible in Minds@UW.
- The Python, Javascript, and Solr code of the PodcastRE Analytics platform will be installed and operational on a UW-Madison server as a free web application. The code will also be freely available for download on GitHub, where it will be maintained by Eric Hoyt.
- Research journal articles will be published open access in PDF format. They will be disseminated via the journal websites and PodcastRE websites and preserved in Minds@UW.
- Blog posts will be published and disseminated on the PodcastRE website, but also output as HTML files for preservation in Minds@UW.
- Podcasts will be output as MP3s and disseminated via Apple iTunes, Stitcher, and the PodcastRE website. They will also be deposited and preserved in Minds@UW.
- White paper will be disseminated on PodcastRE website and preserved in Minds@UW.

V. Data Storage and Preservation of Access

During the grant period, the work on PodcastRE Analytics will take place on the workstations, networks, and facilities managed by the Instructional Media Center, housed in UW-Madison's Department of Communication Arts. The specific Instructional Media Center resources we will be utilizing include the local XSAN storage network (16 TB), Linux web server (4 TB), and networked Mac Pro workstations (there are 38 different workstations available, any of which the project team can use with authorized network accounts). A letter of commitment from Erik Gunneson, director of the Instructional Media Center, is included as part of this proposal.

During the final three months of the grant period, Eric Hoyt (lead developer and one of the two project directors) will take responsibility for insuring the data generated by the project is migrated from the Instructional Media Center to Minds@UW, the University of Wisconsin-Madison Library's Fedora Commons-based preservation platform. Minds@UW will preserve the project data for a minimum of 20 years.

VI. Contingencies

The participation of two project directors and two established university units provides contingencies in the face of unexpected events that, while unlikely, could potentially befall any project. Eric Hoyt is the project director ultimately responsible for making sure this Data Management Plan is followed. However, if Hoyt were to die, become unable to work, or leave the University of Wisconsin-Madison, then Jeremy Morris will assume his responsibilities related to PodcastRE Analytics and the Data Management Plan. Similarly, if Morris leaves the institution, dies, or suffers serious injury, then Hoyt will take responsibility for the completion of PodcastRE Analytics.