

## Dynamic Systems for Humanities Audio Collections: The Theory and Rationale of Swallow

Jason Camlot, Tomasz Neugebauer, Francisco Berrizbeitia

This paper approaches a system that has been designed, and continues to be in development, for the aggregation of metadata surrounding collections of documentary literary sound recordings, as an object for theoretical and practical discussion of how information about diverse collections of time-based media should be managed, and what such schema and system development means for our engagement with the contents of such collections as artifacts of humanist inquiry. Swallow (Swallow Metadata Management System 2019), the interoperable spoken-audio metadata ingest system project that is the boundary object for this talk, emerged out of the goals of the SpokenWeb SSHRC Partnership Grant research network to digitize, process, describe, and aggregate the metadata of a diverse range of sound collections documenting literary and cultural activity in Canada since the 1950s. Our talk, collaboratively written and delivered by a literary scholar and critical theorist, a digital projects and systems development librarian, and a library developer / programmer, outlines 1) a theoretical rationale for the audiotext as a significant form of data in the humanities, 2) consequent modes of description deemed necessary to render such data useful for humanities scholars, and 3) a rationale for the development of a specific form of database system given the material and systems contexts that inform our national holdings of documentary literary sound recordings at the present time.

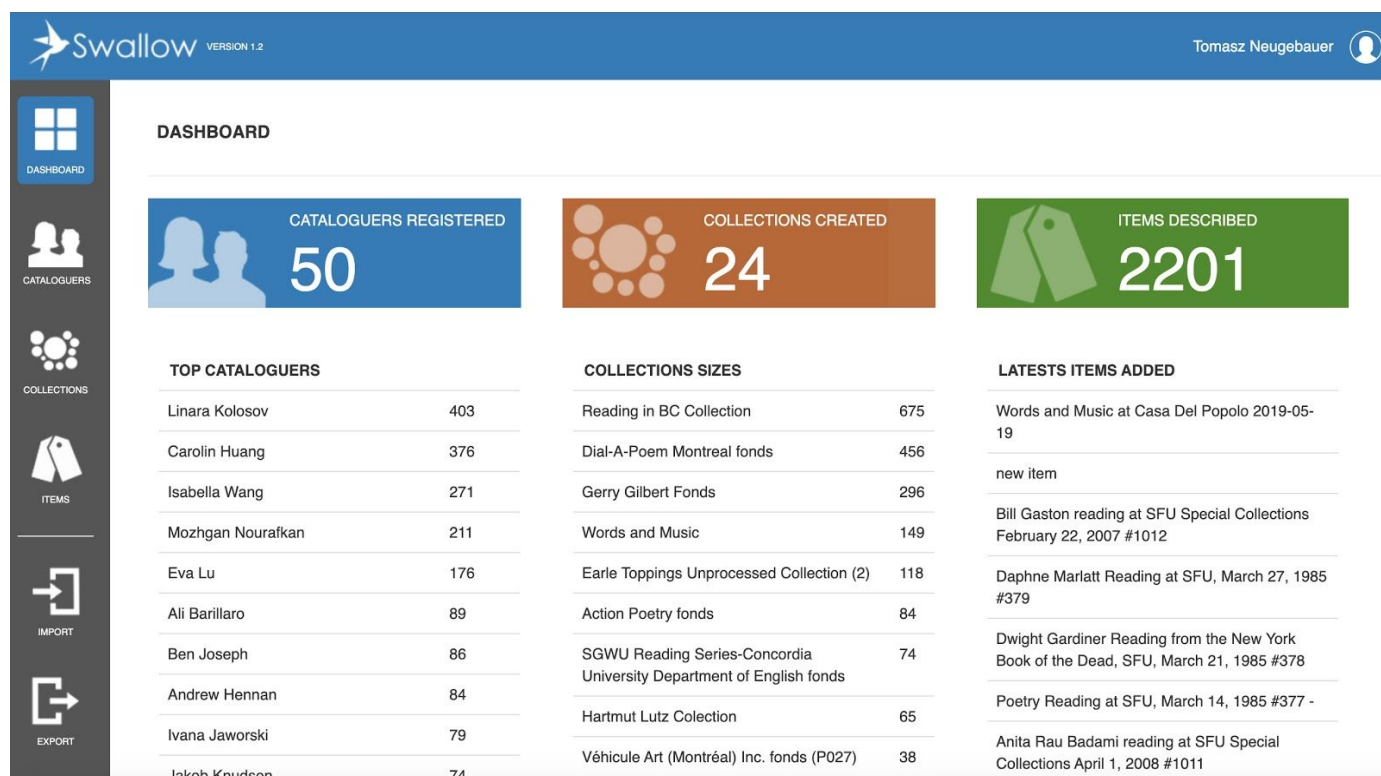


Figure 1. Screenshot of Swallow Dashboard.

## **1. Rationale of digitized audiotexts as humanities research data**

Why is the study of literary recordings important? With the introduction of viable sound recording technology in the last decade of the nineteenth century, a new form of material data for literary analysis emerged: *the audiotext*. This new literary artifact, based on the sound signal of the poet's acoustic performance, has been defined as "a semantically denser field of linguistic activity" that demands new methods of analysis (Bernstein). Once digitized, engagement with sound recordings move us further down the path of understanding "The Text" as "a methodological field" (Barthes), and towards increasingly elaborate "rationales of audio text" (Clement) realized in relation to information systems for the purposes of software analysis, content modeling, and cataloguing. The digitized archive of voices invites the production of thorough metadata, and thus expands the ways in which we may remodel our understanding of historical cultural artifacts (Camlot 2019, Camlot and Mitchell, Gitelman). Over the past two decades, due in part to the founding of online repositories of literary recordings like PennSound (PennSound center for programs in contemporary writing at the University of Pennsylvania 2020) and SpokenWeb (SpokenWeb 2020) the potential of the literary sound archive for research has become discernible. The implications of such research are momentous for future methodologies in the humanities, and yet the realization of such research is limited, at present, by the status and structure of our collections of literary sound recordings. In Canada, thousands of hours of literary readings, performance and related activities have been recorded on magnetic tape since the 1950s. The affordances of tape recording and subsequent media technologies enabled new literary uses of sound recording, including the documentation of literary events, readings and conversations, both public and private. This extended the reach of capturing literary forms and events, and consequently, transformed our understanding of what comprised the literary and cultural events. As valuable archival materials, these recordings represent a massive, largely untapped, undifferentiated, and often undiscovered resource for the study of literary art, culture and society. Research in literary sound studies depends upon a collaborative, interdisciplinary approach to the development of historical knowledge about these valuable cultural heritage materials, and of a networked structure of digital repositories and research-focused interfaces for the preservation and meaningful presentation of these materials to researchers, students, artists and the public. Swallow, the metadata system we launched in 2019 and intend to augment into a robust tool for sharing metadata about our national audio holdings across multiple institutions and a diverse range of metadata and access systems, is a coded software system that represents and works to realize the theoretical idea of audiotexts as significant digital humanities data.

## **2. A Poetics of Audio Description, or, a Metadata Schema for Literary Audio**

The diversity of recordings held at SpokenWeb partner institutions across Canada include recorded conversations, dictations, compilations, performances, interviews and lectures, among numerous other generic audiotextual categories. How do we begin to describe the features and contents of such recordings, and in what ways may we make collections of documentary sound recordings speak to each other, beyond their local venues of origin and siloed media formats and archives?

Metadata and its structure has the potential to open and limit avenues of research. The aggregation of metadata from partner institutions requires either choosing an existing access system, such as Islandora (Islandora Open source digital asset management 2020) or Avalon (Avalon Media System 2020), with its own metadata schema, or to design one based on the evolving research needs of the project. We have chosen the latter because we do not wish to limit our research questions to any single digital repository platform. During the extensive metadata collection phase of our research program, we are seeking maximum usability. We are pursuing our work in describing the assets and contents of collections of documentary literary sound recordings using an agile, iterative development process. We have developed

our long-form metadata schema over the course of a year through the research, discussion, and regular meetings of a SpokenWeb Metadata Task Force consisting of librarians, literary researchers and students from across our partnership. We have provided the metadata schema (Camlot et al. 2020) and ingest software out to student cataloguers quickly for testing and feedback, so that the schema has been shaped by research use from the outset, and will continue to be shaped by researcher feedback throughout this seven-year cataloguing project. We based the SpokenWeb schema on standards such as IASA (International Association of Sound and Audiovisual Archives 1999), AACR2 (Joint Steering Committee for Revision of AACR 2005) and MODS (Library of Congress 2018), focusing on the most useful simplified instructions for the audiotext metadata context.

The initial result was a SpokenWeb Metadata Schema with standard core fields such as: Title, Rights, Creators/Contributors, Production Context, Genre, and Contents. The material artifact description recommendations in the schema include detailed sub-fields reflecting research interest, such as: AV Type, Material Designation, Physical Composition, Storage Capacity, Extent, Playing Speed, Track Configuration, Playback Mode, Tape Brand. The digital file description was a multiple field from the beginning, allowing for multiple files per material artifact. We later added Contents and Notes sub-fields that are capable of storing XML markup to Digital File Description, to make it possible to associate content annotations for each of the files in case of multi-file items. Following feedback from cataloguers who encountered event recording spanning multiple artifacts, it made sense to make Material Description into a multiple field as well. Other examples of changes include: the addition of “Performance Date” to Date Types and “Classroom recording” to Production Context. We recognized the research need to catalogue event/venue names and locations early on, and included location subfields for venue names, latitude and longitude, links to OpenStreetMap nodes, and notes for additional context.

For qualifying and describing authorship, we predefined a controlled vocabulary by selecting from the Library of Congress relators those roles relevant to audiotext, such as interviewer, author, presenter. In the second iteration of the schema, we made it possible to associate multiple roles for a contributor, and following feedback from one of our partner institutions working with First Nations content, we added the role of “Elder”. In addition, soon after the launch of Swallow, the First Nations Metis and Inuit Indigenous Ontology was released, and we added the Nation subfield and controlled vocabulary lookup based on this for Creators/Contributors.

The Contents field has been designed to facilitate navigation, searching and linking audio content of one recording to a wide collection of others. The inclusion of a Related Works field is one way to facilitate linking. The content information needs to be able to include time stamped structural and descriptive information about the works and creators on the recordings. There are many possible tools and formats to perform this type of time stamped metadata. Swallow allows for the storage of contents field information using text or XML markup. We also developed recommendations for how to annotate literary audio for the project and a Python/Flask script to convert text formatted annotations to Avalon XML (Neugebauer 2020). Avalon’s XML format for audio annotations was extended slightly to make it possible to package multiple digital files’ annotations into a single Avalon XML file.

### **3. Swallow and the rationale for an iterative database design**

The SpokenWeb collections are diverse and dispersed in archives and different repository systems. The networked approach we have pursued has allowed the research network to proceed with aggregating the metadata from partners without having to limit the infrastructure by choosing only one repository system with its associated metadata schema. However, we have also needed to move beyond spreadsheets, to improve the accuracy and usability of the cataloguing work. A spreadsheet of metadata quickly becomes

difficult to manage as the number of columns increases to accommodate multiple fields, to catalogue multiple creators/contributors, for example. In addition, the research requirements of the project call for a highly customizable and extendible set of metadata fields that will change over time as we analyze and catalogue the content and learn more about the collections. We needed to be able to rapidly iterate new versions of interfaces for evolving metadata schema. Our solution has been to avoid hardcoding the metadata schema into the relational database, and instead to store the metadata as unstructured JSON inside the database. A layer of abstraction then allows us to quickly modify configuration files that define the metadata fields and their associated controlled vocabularies and lookups. The resulting system's cataloguing interface may then be changed quickly, and is able to accommodate multiple schemas. Indeed, we have implemented numerous additions and changes to the metadata system over the course of our first year of cataloguing in response to the nature of specific collections we are describing.

Swallow makes it possible for literary audio holding institutions to aggregate their metadata in a single system. The required metadata fields for access and preservation can change over time and vary by document type and source institution. Swallow allows for the aggregation of metadata in multiple metadata schemas in the same system. It generates a schema-specific cataloguing interface for each item stored, based on the item's respective schema configuration files that include a list of steps, metadata fields and controlled vocabularies. The Swallow schema specification functionality currently allows for the inclusion of URIs alongside metadata values that are a part of the SpokenWeb Schema, such as links to VIAF, OpenStreetMap, Wikidata, and Rights Statements.

### **Future development**

Figure 2 shows the proposed flow of information with Swallow at the center acting as a central metadata repository. The functionality that we are proposing to develop includes the capacity to batch ingest metadata from other systems currently used by SpokenWeb partners into the latest version of the SpokenWeb schema in Swallow. The functionality to access batch import mappings on the interface exists, and we will develop the mappings for Islandora, Avalon, and AtoM in collaboration with partner institutions, Concordia Spectrum Research Repository, SFU Radar, University of Alberta Dataverse, University of Toronto (Dataverse), UBC Brain Circuits (FRDR). The proposed functionality includes the development of batch editing that would allow cataloguers to modify the metadata of collections of items. We will also develop mapping specifications that will allow for the migration of items catalogued in one schema version into another. The proposed development would build into the interface the functionality to lookup and export metadata from Swallow to Wikidata. In addition, we will develop mappings for export of collections of metadata as data from Swallow to DataVerse and FRDR (Canadian Association of Research Libraries & Compute Canada 2020) .

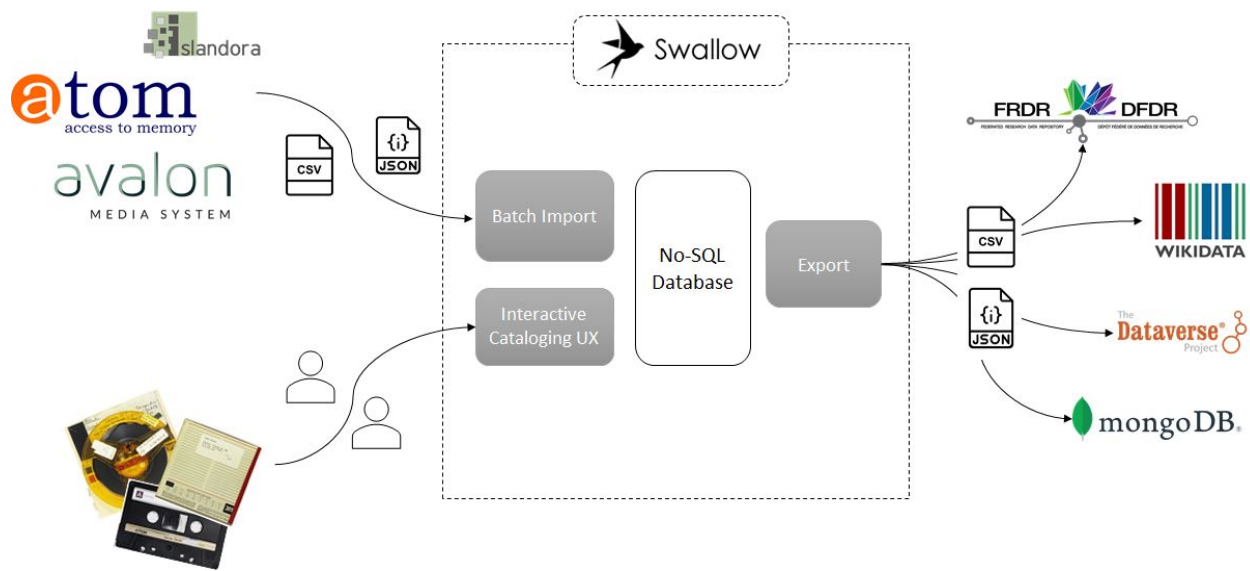


Figure 2. Swallow as central metadata repository.

## Conclusion

Our work in metadata and systems development can be read as the development as a poetics of the audiotext, that is a poesis (a making) of the descriptive and digital infrastructure that allows digitized audio artifacts – artifacts that capture the sound of literature in its social contexts during the postwar era – to survive and take flight in the imaginations of contemporary critics, scholars, poets and listeners.

## Bibliography

*Avalon Media System 2020*, viewed June 15, 2020, <<https://www.avalonmediasystem.org/>>.

Bernstein, Charles, ed. *Close Listening: Poetry and the Performed Word*. New York, NY: Oxford UP, 1998.

Barthes, Roland. "From Work to Text." Trans. Richard Howard. *The Rustle of Language*. New York: Farrar, Straus and Giroux, 1986. 56-64.

Camlot, Jason. *Phonopoetics: The Making of Early Literary Recordings*. Stanford, CA: Stanford University Press, 2019.

Camlot, Jason and Christine Mitchell. "The Poetry Series." *Amodern* 4 (March 2015). <<http://amodern.net/article/editorial-amodern-4/>>. Web.

Camlot, J., Dowson, R. Ferrier, Fong D., Kail R., Luyk S., Meza A., Neugebauer T., Wiercinski J., Barillaro A., Hannigan L., Lu E., Mash C., Pickering H, Chandler N., Knudsen J, Kolosov L, Roberge, A.

*SpokenWeb Metadata Scheme and Cataloguing Process 2020*, viewed June 15, 2020, <<https://spokenweb-metadata-scheme.readthedocs.io/en/latest/>>.

Clement, Tanya. "Towards a Rationale of Audio-Text." *Digital Humanities Quarterly* 10.2 (2016). Web.

*Federated Research Data Repository (FRDR) 2020*, Canadian Association of Research Libraries & Compute Canada, viewed June 15, 2020, <<https://www.frdr-dfdr.ca/repo/>>.

Gitelman, Lisa. *Always Already New: Media, History and The Data of Culture*. Cambridge, MA: MIT Press, 2008.

International Association of Sound and Audiovisual Archives. "The IASA Cataloguing Rules (IASA 1999)". The International Association of Sound and Audiovisual Archives, viewed June 15, 2020, <<https://www.iasa-web.org/cataloguing-rules/>>.

*Islandora Open source digital asset management 2020*, Islandora Foundation, viewed June 15, 2020, <<https://islandora.ca/>>.

Joint Steering Committee for Revision of AACR (2005). *Anglo-American cataloguing rules*, 2002 revision. Chicago, Ill, American Library Association.

Library of Congress 2018. *Outline of Elements and Attributes in MODS Version 3.7*. viewed June 15, 2020. <<https://loc.gov/standards/mods/mods-outline-3-7.html#name.%20v.3>>.

Neugebauer, Tomasz. *Python Script to Convert Audio Annotation Files*. 2020, GitHub, viewed June 15, 2020, <<https://github.com/photomedia/AudioAnnotateConvert>>.

*PennSound center for programs in contemporary writing at the University of Pennsylvania 2020*, viewed June 15, 2020, <<http://writing.upenn.edu/pennsound/>>.

*RIGHTS STATEMENTS 2018*, viewed June 15, 2020, <<https://rightsstatements.org/page/1.0/?language=en>>.

*SpokenWeb 2020*, viewed June 15, 2020, <<https://spokenweb.ca/>>.

*Swallow Metadata Management System 2019*, GitHub, viewed June 15, 2020, <<https://github.com/spokenweb/swallow>>.

*VIAF: The Virtual International Authority File 2020*. viewed June 15, 2020, <<http://viaf.org/>>.