



SpokenWeb Search Engine

Shreya Savant & Tomasz Neugebauer
Concordia University


Blacklight Summit (December 8th, 2025)




What Is SpokenWeb?

<https://search.spokenweb.ca/>



 SpokenWeb

Bookmarks History

All Fields Search... 

More options

Limit your search

- Institution >
- Collection >
- Series >
- Sub Series >
- Location >
- Production Context >
- Production Date >
- Publication Date >
- Performance Date >
- Material Designation >
- Recording Type >

SpokenWeb Search Engine



Collections of Literary Audio and Video Recordings in Canada

SpokenWeb is a research partnership that collaborates on the development of audiovisual collections documenting literary events and activities that have taken place in Canada since the 1950s. These collections are held at organizations and institutions across the country. Metadata about the recordings are held in the SpokenWeb Swallow Database. The contents of this database are made available to you through this SpokenWeb Search Engine which works as a directory to information about the recordings, and in some cases to the audiovisual materials themselves.

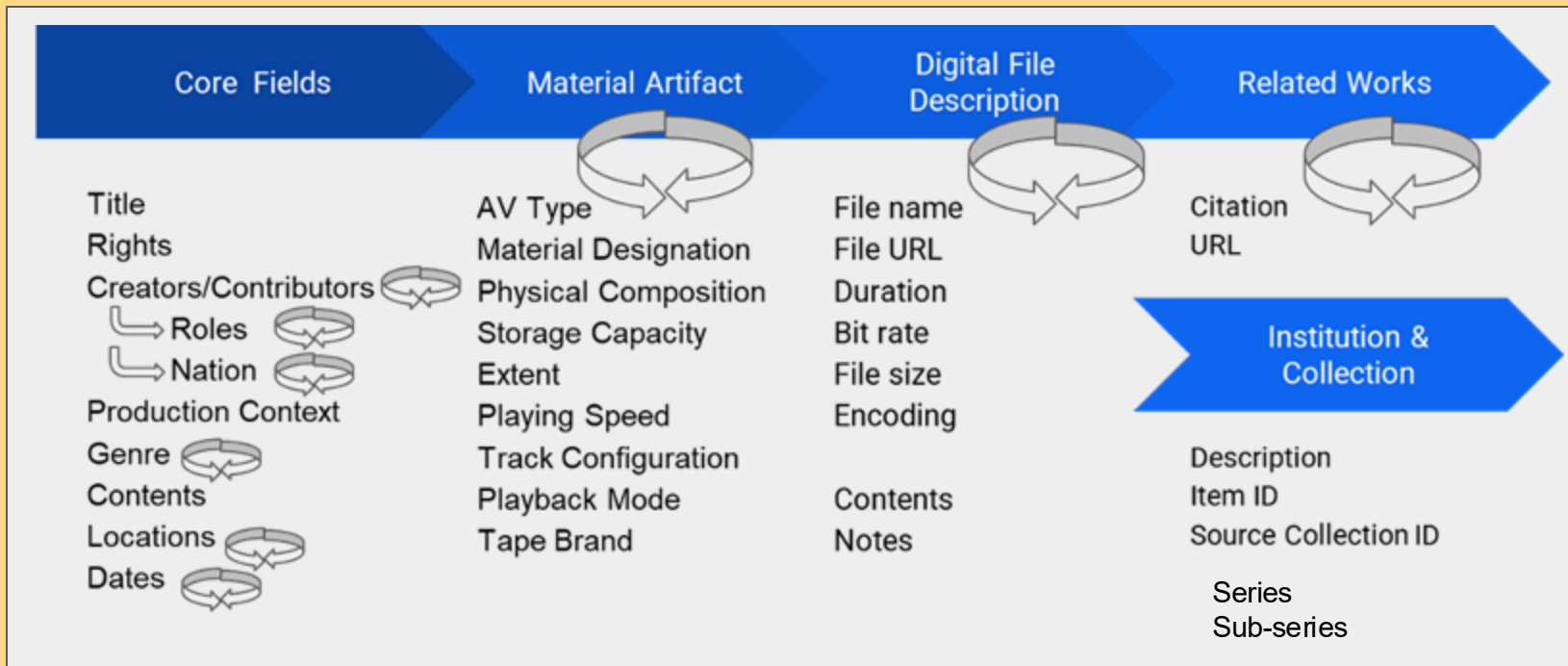
The SpokenWeb Search Engine allows users to search collections of literary sound recordings held at organizations and institutions across Canada. Metadata about the recordings are held in the SpokenWeb Swallow Database. The contents of this database are made available through this SpokenWeb Search Engine which works as a directory to information about the recordings, and in some cases to the audiovisual materials themselves.

Do you know of or hold a collection of literary audio or video that you can't find here? If you would like it to be added to the SpokenWeb directory please contact us at spokenwebcanada@gmail.com.

Please report any bugs that you may find, either navigating the site or in our data, via our [Online Ticketing Form!](#)



- Cataloging: Distributed cataloging by teams across Canada.
- SpokenWeb Schema



The Search Engine (Discovery Layer)

- Role: The public "Discovery Layer" for the Swallow database.
- Key Features:
 - Facet by Year, Genre, City, and Person.
 - Responsive Design.
 - Download batch/individual json/txt files for research
 - Link to available recordings
- Live URL: <https://search.spokenweb.ca/>



Technology Stack

- Frontend: Blacklight (v8) on Ruby on Rails
 - Industry standard for library discovery interfaces.
- Backend: Apache Solr (v9)
 - High-performance indexing and querying.
- Ingest Pipeline:
 - Python: ETL (Extract, Transform, Load) scripts to pull data from Swallow.
 - Traject: Ruby library for mapping metadata to Solr fields.
- Infrastructure: Docker & Docker Compose.

The SWSearchEngine Repository

- <https://github.com/spokenwebsites/SWSearchEngine>
- Containerized Development:
 - app: The Rails/Blacklight application.
 - solr: Pre-configured Solr instance.
- Easy Setup:
 - The use of Makefile gets the full stack system running locally with one command and simulates closest to the server behavior.
- Configuration:
 - Solr config files stored in version control for consistency.
 - Environment variables for secrets and API keys.



All Fields

Search...



More options

Limit your search

Institution >

Collection >

Series >

Sub Series >

Location >

Production Context >

Production Date >

Publication Date >

Performance Date >

Material Designation >

Recording Type >

AV Type >

Playback Mode >

SpokenWeb Search Engine

Collections of Literary Audio and Video Recordings in Canada

SpokenWeb is a research partnership that collaborates on the development of audiovisual collections documenting literary events and activities that have taken place in Canada since the 1950s. These collections are held at organizations and institutions across the country. Metadata about the recordings are held in the SpokenWeb Swallow Database. The contents of this database are made available to you through this SpokenWeb Search Engine which works as a directory to information about the recordings, and in some cases to the audiovisual materials themselves.

The SpokenWeb Search Engine allows users to search collections of literary sound recordings held at organizations and institutions across Canada. Metadata about the recordings are held in the SpokenWeb Swallow Database. The contents of this database are made available through this SpokenWeb Search Engine which works as a directory to information about the recordings, and in some cases to the audiovisual materials themselves.

Do you know of or hold a collection of literary audio or video that you can't find here? If you would like it to be added to the SpokenWeb directory please contact us at spokenwebcanada@gmail.com.

Please report any bugs that you may find, either navigating the site or in our data, via our [Online Ticketing Form](#)

**Action Poetry fonds**

(84)

**Adeena Karasick Fonds**

(18)

**Alan Lord collection**

(58)



SpokenWeb

Thank You!

<https://search.spokenweb.ca>

SpokenWeb.ca