



Training Archivists about Language Collections: A Modular Approach



AERI 2024
Abu Dhabi, UAE

Hugh J.Paterson III
Dr. Ana Roeschley
Sergio I.Coronado
Dr. Oksana L. Zavalina
University of North Texas

Language Archives

- ※ “any physical or digital item that is a product of language documentation, description, or development or is a tool that specifically supports the creation and use of such a product.” – Bird and Simons, 2022, p. 507
 - Authors conflate the concepts of “Archive”, “Memory Institution”, and “Collection”.
- ※ “representative of authentic language use” whether or not materials were created specifically for this purpose. – Burke and Zavalina, 2019, p. 365
- ※ Collections
 - frequently composed of resources created as part of linguistic or anthropological research work in collaboration with minority ethno-linguistic communities.
 - have a variety of modalities (Image, Video, Audio, Text) and carriers (digital and non-digital).

Bird, S., & Simons, G. F. (2022). The Open Language Archives Community: A 20-year update. *The Electronic Library*, 40(5), 507–524. <https://doi.org/10.1108/EL-08-2022-0192>

Burke, M., & Zavalina, O. L. (2019). Exploration of information organization in language archives. *Proceedings of the Association for Information Science and Technology*, 56(1), 364–367. <https://doi.org/10.1002/pra2.30>

The Context

- ※ **Insufficient training materials for language resource stewardship**
 - Oriented to specific language archives.
 - Oriented for depositors, both language scholars and language-community members.
 - Disconnected from information organization theory and are not relevant to collection management or institutional management.
- ※ **Lack of overt training for archivists (and other information professionals)**
 - No university-level curriculum in English specifically addressing language artifacts and their management.
 - The topic is rarely seen in archival science and library/information science training programs.

Project Goal and Team Composition

DIGITAL ARCHIVES

Community Memory and Language Archiving



With support from IMLS ([RE-254860-OLS-23](#)), project to develop a 4-module, Open Access curriculum introducing archivists and other information professionals to the contexts of language collections and issues encountered in their management.

Team composed of Professors at an iSchool (specialising in Archives, Information Organization, Digital Imaging, etc.), digital repository developer, a Linguistic Professor, other Linguists, and current and former archivists at language archives.

The 4 Modules

- 1. Planning, developing, and managing a community language archive.**
- 2. Ethical archival practices and digital curation in community language archives.**
- 3. Metadata, digital content management and web archiving for community language archives.**
- 4. Dissemination and use of community language archive content, evaluation of archival services.**

Module Content Use

Designed to be integrated or offered as standalone course.

- IU LING Corpus Linguistics (Module 1)
- IU LING Field Methods (Module 1)
- UNT INFO 5224 Advanced Metadata (Modules 2-3)
- UNT INFO 5960 Cultural Heritage Stewardship (Module 2)
- *UNT INFO 5742 Web Archiving (Module 3) - Planned Spring 2025*
- UNT INFO 5385/5680 Community Language Archiving and Curation for Information Professionals (all 4 modules)

Information Professionals – Relevant Roles and Impacts

- ※ Digital archivists and other informational professionals
 - responsible for **inventory** and **implementation of data management plans**
 - working with university faculty on **adding legacy-data collections to repositories**
 - whose responsibilities include **establishing a new archive** and/or development of existing digital collections
 - working on system/repository **migration tasks**.
- ※ Programming Librarians / Exhibit Production Specialists
- ※ Special Collections Archivists, Curators, and Collection Maintenance Specialists

Information Agency Management Topics

- ※ *Of-ness* in subject analysis and representation
- ※ Analysis and representation of **relations between works, their versions, formats**, etc.
- ※ **Part-whole relationships representation**
- ※ **Interactive modality and materiality representation** (including controlled vocabularies)
- ※ **Language identification** and representation
- ※ **Fiduciary responsibility**
- ※ **Evaluation**

Student feedback after first
standalone course offering:
5-week term of Summer 2024

The Training Effectiveness: Course Objectives 1-4

To what extent do you feel the course content and activities helped you meet the following course-level learning objective:	helpful + very helpful (% of responses, n=23)
<i>1 - Define community language archives and their function?</i>	26% + 74% = 100%
<i>2 - Describe procedures & considerations for community language archive planning & development?</i>	30% + 65% = 95%
<i>3 - Examine archival theory & practice and digital curation trends & perspectives relevant for community language archives?</i>	39% + 57% = 96%
<i>4 - Explain important access, preservation, and description issues and practical problems associated with community language archives?</i>	43% + 57% = 100%

The Training Effectiveness: Course Objectives 5-8

To what extent do you feel the course content and activities helped you meet the following course-level learning objective:	helpful + very helpful (% of responses, n=23)
<i>5 - Identify metadata standards that can be utilized in representing community language archive materials to support information needs and user tasks of community language archive users?</i>	78% + 22% = 100%
<i>6 - Describe the problems related to selection of digital content management tools, depositing and web archiving for community language archive?</i>	30% + 61% = 91%
<i>7 - Identify efficient and ethical ways of disseminating the content of community language archives?</i>	26% + 65% = 91%
<i>8 - Discuss approaches for the evaluation of the services provided by community language archive?</i>	30% + 70% = 100%

The Overall Training Effectiveness: Student Feedback

- ※ “I learned so much about both community archives and community language archives. I was unaware that there was a community for either of these and now it’s a big, new world that I would love to be involved with.”
- ※ “I enjoyed this class because it tied together all of the classes I’ve taken while in the masters program.”
- ※ “I appreciated learning how to evaluate archives and use critical thinking skills and librarianship skills to find answers to problems.”
- ※ “I found areas each week that I wanted to research more. I also recommended the course to a friend in the MSLS program”
- ※ **“My concentration is [...] but after this course I'm wishing I had chosen Archives!”**
- ※ “I really did enjoy getting hands-on experience and being able to interact”
- ※ “Likely this was among the most interesting courses I have taken in this degree program. As such, it was definitely intellectually stimulating.”
- ※ “It was informative and interesting. I wish I could have done this in a 16 week semester”

Summary:

Increase in students' confidence across the key learning areas;
Growing awareness of language archives and motivating future information professionals to choose archival career path

Ongoing & Future Work

- Testing in longer semesters (10-16 weeks) and subsequent course revisions
- Open Access release of course content
- Modular textbook
- Sustainable curriculum development – collaborative revisions across teaching institutions

Questions:

How should training programs incorporate the language archive stewardship issues into more courses?

Please share your feedback with us:

hughpaterson@my.unt.edu

ana.roeschley@unt.edu

sergiocoronado@my.unt.edu

oksana.zavalina@unt.edu



Thank you!

DIGITAL ARCHIVES

Community Memory and Language Archiving



INSTITUTE of
Museum and Library
SERVICES



UNT Ψ

INDIANA UNIVERSITY

Please let us know
your thoughts on the
project and curriculum

Hugh J. Paterson III

Dr. Ana Roeschley

Sergio I. Coronado

Project PI: Dr. Oksana L.
Zavalina

Extra Slides if Discussion
Requires It

Module 1: Summer 2024 Student Feedback

How confident do you feel in your ability to ...	Pre-Course: <i>confident</i> + <i>very confident</i>	Post-Course: <i>confident</i> + <i>very confident</i>
Q1 - define digital language archives?	$15.6\% + 6.3\% = \mathbf{21.9\%}$	$64\% + 18\% = \mathbf{82\%}$
Q2 - explain the function of digital language archives?	$9.4\% + 9.4\% = \mathbf{18.8\%}$	$59\% + 32\% = \mathbf{91\%}$
Q3 - identify user needs of digital language archives?	$18.8\% + 12.5\% = \mathbf{31.3\%}$	$45\% + 27\% = \mathbf{72\%}$
Q4 - determine sources for budgeting and funding for digital language archives?	$6.3\% + 12.5\% = \mathbf{18.8\%}$	$45\% + 9\% = \mathbf{54\%}$ [+32% neutral]
Q5 - identify steps necessary for creating a digital language archive collection?	$9.4\% + 21.9\% = \mathbf{31.3\%}$	$55\% + 32\% = \mathbf{87\%}$

Module 2: Summer 2024 Student Feedback

How confident do you feel in your ability to ...	Pre-Course: <i>confident + very confident</i>	Post-Course: <i>confident + very confident</i>
Q1 - define major features of archival theory?	$32\% + 4\% = \mathbf{36\%}$	$59\% + 18\% = \mathbf{77\%}$
Q2 - describe the description practices of language archives?	$16\% + 4\% = \mathbf{20\%}$	$50\% + 27\% = \mathbf{77\%}$
Q3 - discuss how closely language archives adhere to archival practices described in archival literature?	$12\% + 4\% = \mathbf{16\%}$	$49.5\% + 31.7\% = \mathbf{81.2\%}$
Q4 - analyze the differences and commonalities between the preservation and access for language archives and traditional community-based archives?	$24\% + 4\% = \mathbf{28\%}$	$45\% + 45\% = \mathbf{90\%}$

Module 3 (a): Summer 2024 Student Feedback

How confident do you feel in your ability to ...	Pre-Course: <i>confident</i> + <i>very confident</i>	Post-Course: <i>confident</i> + <i>very confident</i>
Q1 - identify how information needs & user tasks of community language archive users + other factors (e.g., ethical and cultural norms) affect organizing information?	$41\% + 5\% = \mathbf{46\%}$	$65\% + 22\% = \mathbf{87\%}$
Q2 - discuss entities and attributes of community language archive information objects and relationships between them as well as explain how representing these attributes and relationships supports user tasks?	$18\% + 5\% = \mathbf{23\%}$	$57\% + 17\% = \mathbf{74\%}$
Q3 - identify and contrast metadata schemes or application profiles used for representing community language archive materials to support user needs?	$14\% + 5\% = \mathbf{19\%}$	$48\% + 22\% = \mathbf{70\%}$

Module 3 (b): Summer 2024 Student Feedback

How confident do you feel in your ability to ...	Pre-Course: <i>confident</i> + <i>very confident</i>	Post-Course: <i>confident</i> + <i>very confident</i>
Q4 - compare, navigate & apply DACS and controlled vocabularies for representing community language archive materials to support user needs and?	$27\% + 5\% = \mathbf{32\%}$	$52\% + 26\% = \mathbf{78\%}$
Q5 - identify Digital Content Management (DCM) tools that could be used for organizing community language archives and DCM selection considerations?	$14\% + 0\% = \mathbf{14\%}$	$43\% + 9\% = \mathbf{52\%}$ <i>[+39% neutral]</i>
Q6 - discuss and compare community language archive deposit processes and approaches?	$14\% + 5\% = \mathbf{19\%}$	$61\% + 9\% = \mathbf{70\%}$
Q7 - identify web archiving opportunities for community language archives?	$9\% + 0\% = \mathbf{9\%}$	$48\% + 30\% = \mathbf{78\%}$

Module 4 (a): Summer 2024 Student Feedback

How confident do you feel in your ability to ...	Pre-Course: <i>confident + very confident</i>	Post-Course: <i>confident + very confident</i>
Q1 - discuss current practices and venues for effective and ethical dissemination of the content of community language archives?	$9\% + 4\% = \mathbf{13\%}$	$70\% + 13\% = \mathbf{83\%}$
Q2 - explain the role of finding aids and user guides in discovery of archival collections, with the focus on language archives?	$26\% + 9\% = \mathbf{35\%}$	$52\% + 39\% = \mathbf{91\%}$
Q3 - identify components of collection description in finding aids and user guides?	$22\% + 4\% = \mathbf{26\%}$	$48\% + 35\% = \mathbf{83\%}$
Q4 - describe the ways that are used to ensure effective access to digital archives and the related ethical implications?	$22\% + 4\% = \mathbf{26\%}$	$57\% + 22\% = \mathbf{79\%}$

Module 4 (b): Summer 2024 Student Feedback

How confident do you feel in your ability to ...	Pre-Course: <i>confident + very confident</i>	Post-Course: <i>confident + very confident</i>
Q5 - identify different services and user interaction features of a digital language archives that are commonly expected and evaluated?	$17\% + 4\% = 21\%$	$70\% + 22\% = \text{92\%}$
Q6 - identify the criteria in assessing discoverability of language resources and usability of language archives?	$9\% + 4\% = 13\%$	$61\% + 22\% = \text{83\%}$
Q7 - discuss the methods of collection and analysis of digital language archive use and audience engagement metrics?	$13\% + 4\% = 17\%$	$57\% + 35\% = \text{92\%}$