

# Contributor Roles in a Mixed Archive

Hannah Tarver, Mark Phillips, Hugh Paterson III. 2024. Contributor Roles in a Mixed Archive.  
Paper presented at the Language Documentation and Archiving Conference.  
Berlin-Brandenburgische Akademie der Wissenschaften, Berlin, Germany. September 4th–6th.  
Presentation slides Copyright 2024 licensed: CC-SA-ND.

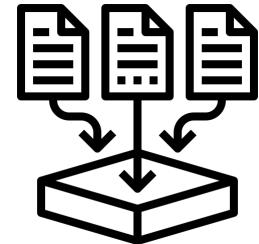
# Archive

- A Corpus
- Special Collections Unit within a Memory Institution
- A Building
- An Organization
- An Infrastructure marked by Stewardship of Static Resources
- An Infrastructure marked by Curation and Dynamic Handling of Resources

Paterson III, Hugh J. 2023. "Diversity and Identity: Categories for OAI Data-Providers in the Open Language Archives Network." Proceedings from North American Symposium on Knowledge Organization (NASKO) 9 (1): 18–31. <https://doi.org/10.7152/nasko.v9i1.16301>.

Paterson III, Hugh J. 2022. "Supporting Diversity Through a Typology of Data Providers: What is an Archive?". Paper presented at the 8th International Conference on Language Documentation and Conservation: Centering Justice in Language Work.  
<https://hughandbecky.us/Hugh-CV/talk/2023-supporting-diversity-through-a-typology-of-data-providers-what-is-an-archive>

# Mixed Archive



## Marked by:

- Broadly Scoped Accessions
- Single Administrative Structure
- Language “archive” exists as a *Special Collection* within the Memory Institution

## Implications:

- Single Metadata Schema
- Rigid and inflexible metadata schema evolution requirements; must make sense across many special collections
- The needs of one client are pitted against the demands for effectively serving many clients and maintenance costs impacting sustainability.

# Our Use Case

University of North Texas Digital Collections

CORSAL, a special collection within the broader  
Digital Library infrastructure.

UNTL is the overarching metadata schema for all  
collection items across various special collections  
within the Digital Library infrastructure.

# Our Team

**Hannah Tarver** - Supervisor, Digital Projects  
(UNTL Architect and Steward)

**Mark Phillips** - UNT Associate Dean for Digital  
Libraries

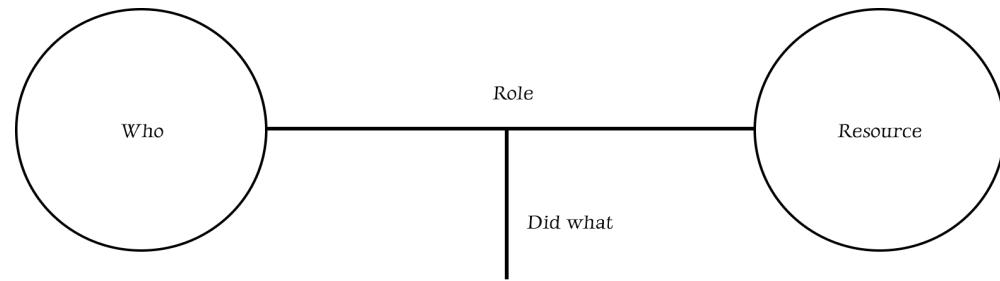
**Hugh Paterson III** - Analyst

Phillips, M., Burke, M., Tarver, H. and Zavalina, O.L. (2021), “Leveraging digital library infrastructure to build a language archive”, in Zavalina, O.L. and Chelliah, S.L. (Eds), Proceedings of the International Workshop on Digital Language Archives: LangArc 2021, doi: 10.12794/langarc1851182

Burke, M., Tarver, H., Phillips, M. E., & Zavalina, O. (2022). Using existing metadata standards and tools for a digital language archive: A balancing act. *The Electronic Library*, 40(5), 579–593. <https://doi.org/10.1108/EL-02-2022-0028>

# Contributor Roles

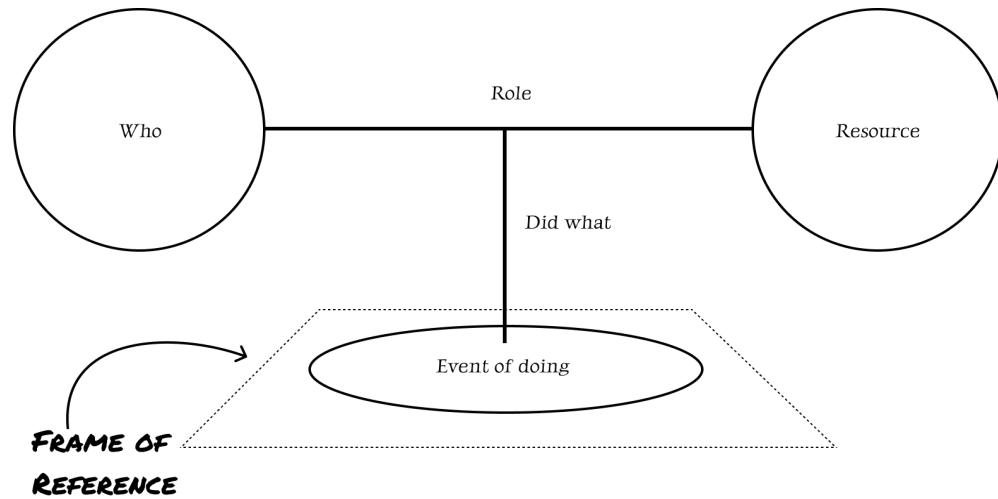
- Describe how people have impacted a resource.



Agent --- Relationship --- Item

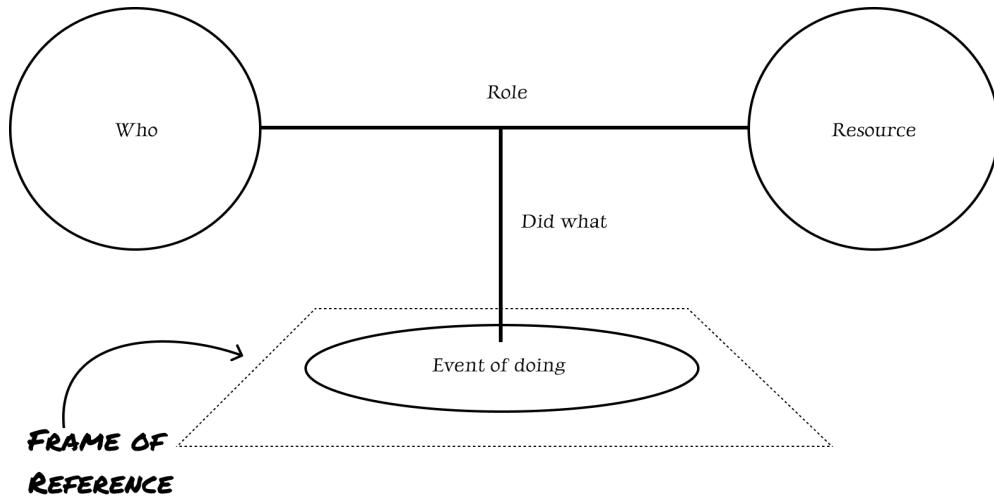
# Contributor Roles

- Describe how people have impacted a resource.
- Infer a frame of reference for the activity of impact.



# Contributor Roles

- Describe how people have impacted a resource.
- Infer a frame of reference for the activity of impact.
- It works in conjunction with other parts of the record such as **provenance** and **genre/form** identification.



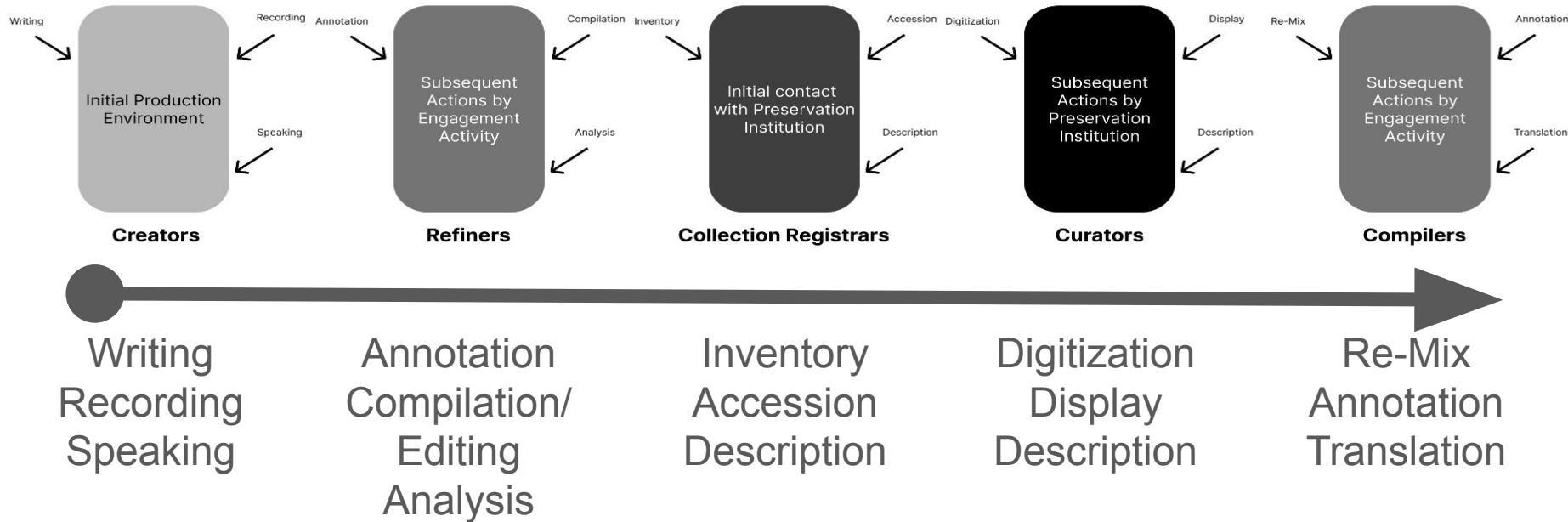
E.g., *Judge* vs. *Interviewee*

# Examples of Agent Paradigms (Frames of Reference)

- Legal
  - Intellectual Property
    - *Second party*
    - *First party*
    - *Licensee*
    - *Licensor*
    - *Patent applicant*
    - *Patent holder*
    - *Copyright claimant*
    - *Copyright holder*
  - Case Law
    - *Defendant*
    - *Plaintiff*
    - *Judge*
    - *Court Recorder*

# Broad Reference Frames in Resource Evolution

## - An Archival Perspective



# Management Questions

- What roles exist in which schemas and are they compatible?
- Which roles exist in which reference frames?
  - Are we missing critical reference frames?
- Do we need more or new roles to achieve the goals we have for the resources we steward?
- Is it cost effective?

# Research Questions

- Since UNTL roles are largely based on subset of MARC roles are there enough roles in UNTL to support CORSAL and OLAC aggregation?
- What are the overlaps in roles across the different schemas?
- Are there inaccessible frames of reference needed in the description of language resources?

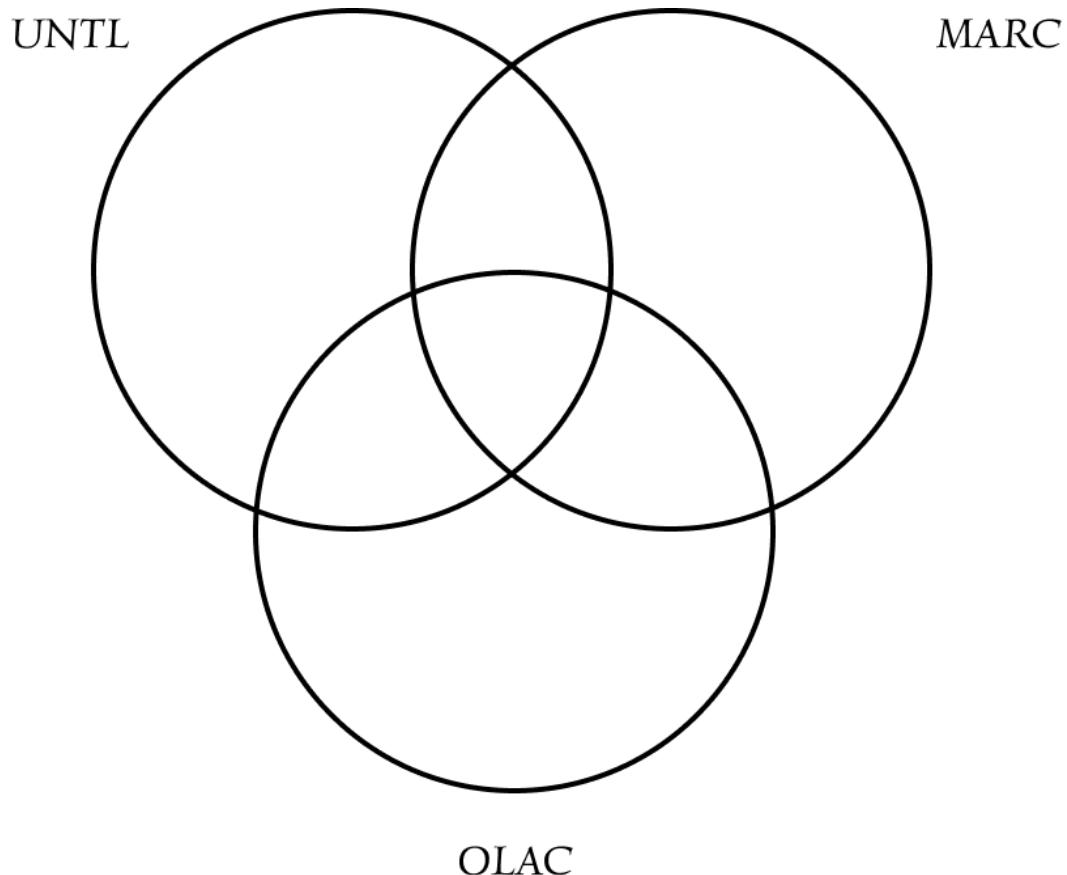
UNTL Role Vocabulary: <https://digital2.library.unt.edu/vocabularies/agent-qualifiers>

MARC Role Vocabulary: <https://www.loc.gov/marc/relators/relaterm.html>

OLAC Role Vocabulary: <http://www.language-archives.org/REC/role.html>

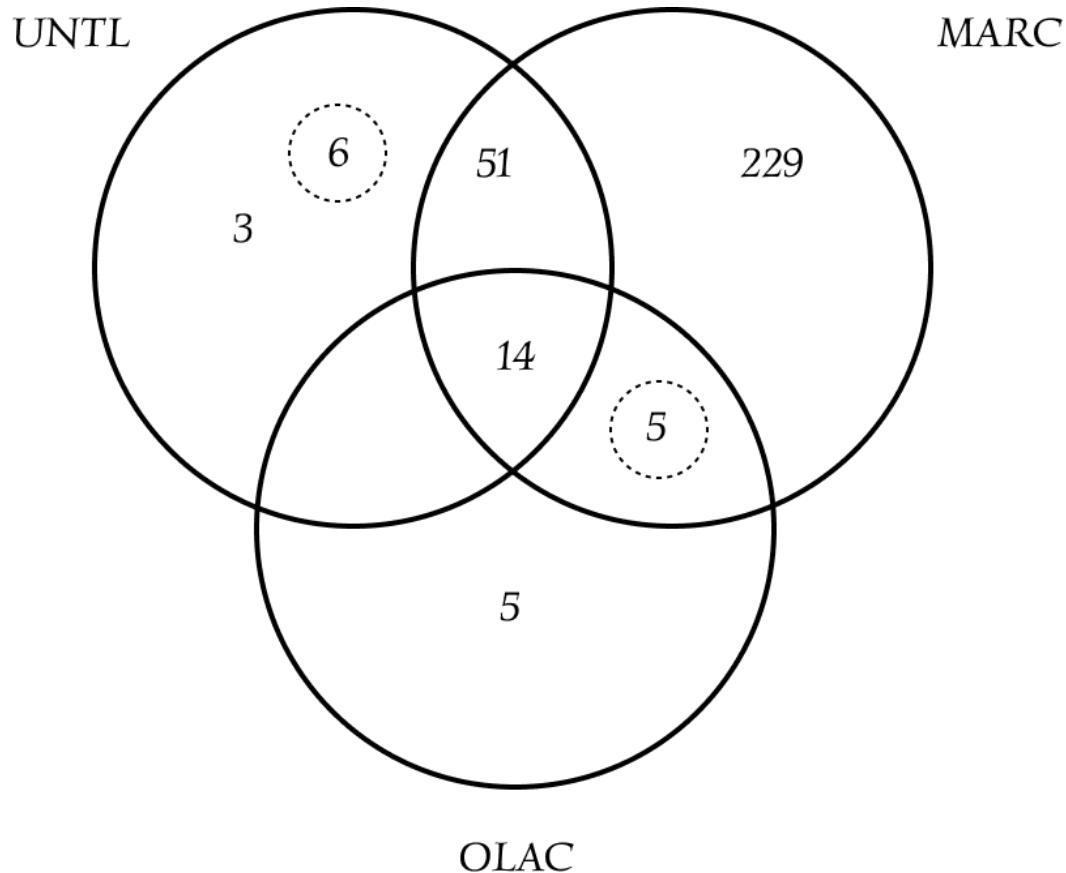
# The Comparison

- UNTL has 74 roles
- MARC has 299 roles
- OLAC has 24 roles



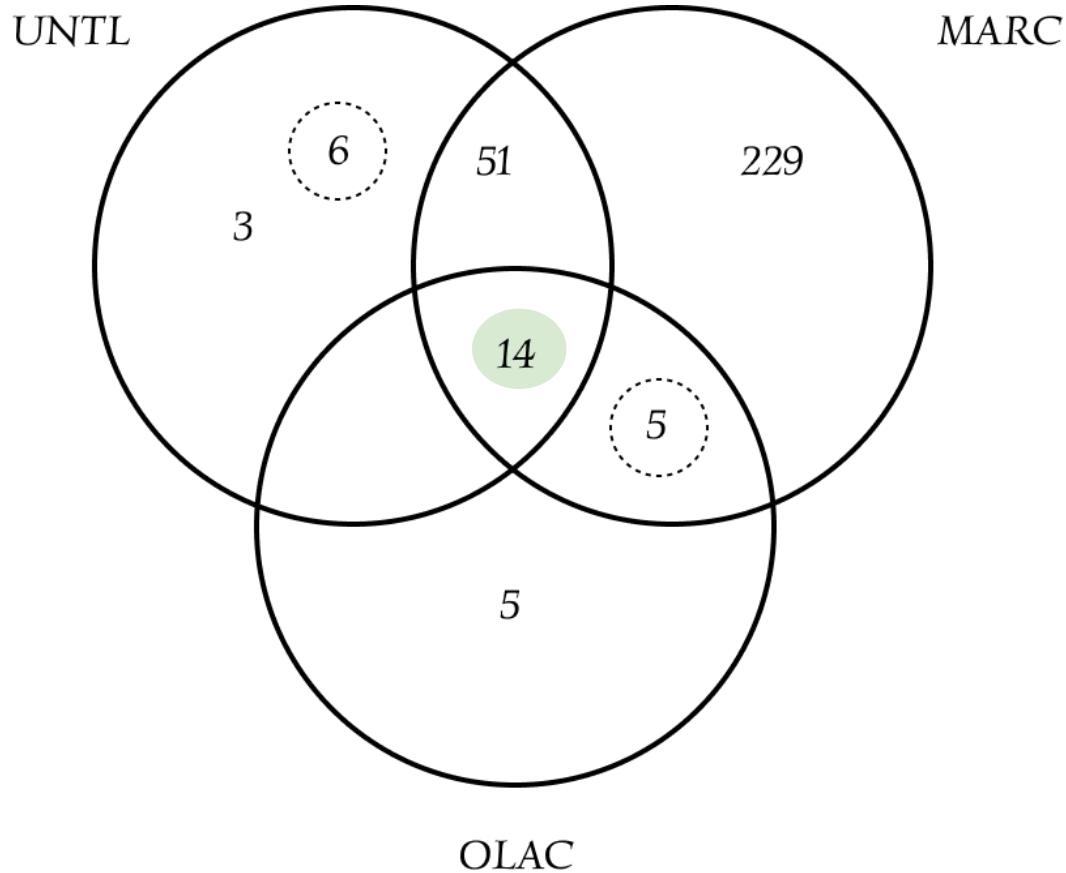
# The Comparison

- UNTL and OLAC share only 14 roles which are also common to MARC
- OLAC has 5 roles which are not found in MARC -- they are custom
- OLAC and MARC share 5 roles with a common name but different semantic meanings



# Compatible Roles

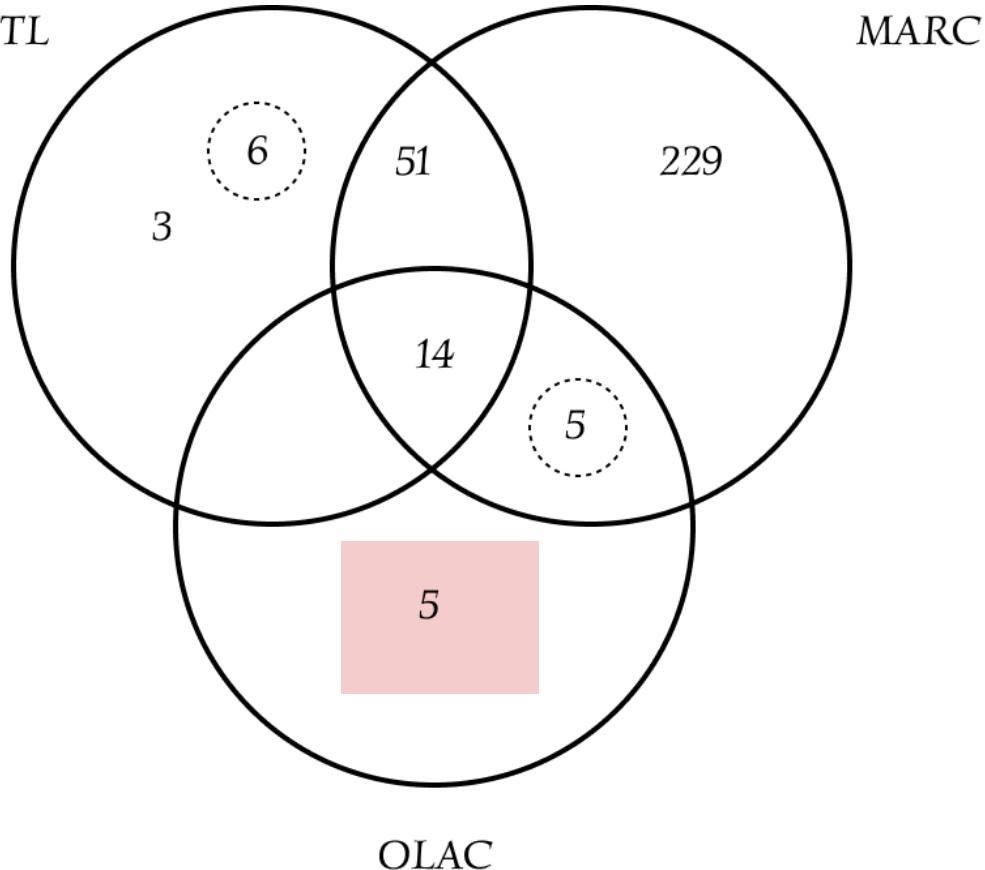
- Annotator
- Author
- Consultant
- Editor
- Illustrator
- Interviewer
- Performer
- Photographer
- Researcher
- Speaker
- Sponsor
- Transcriber
- Translator



# OLAC Exclusive Roles<sup>UNTL</sup>

- data\_inputter
- interpreter
- participant
- research\_participant
- developer

From a language resource perspective here is a significant difference between a *translator* and an *interpreter*.



# Other MARC/OLAC Roles issues

<b>OLAC Role</b>	<b>in-Equivalence</b>	<b>MARC Relator Role</b>	<b>In UNTL</b>
consultant	Name	Consultant to a project	/ Consultant (MARC)
recorder	Name	Recordist	X
compiler	Semantics	Compiler	/ Compiler (MARC)
depositor (archive)	Semantics	Depositor (owner)	X
responder	Semantics	Addressee/witness	/ Addressee (MARC), witness
developer	Semantics/Name	Programmer/Redaktor/Author	X
signer	Semantics	Signer (paints signs)	X (= MARC speaker?)
data_inputter	eXistence		X
interpreter	eXistence		X
participant	eXistence		X
research_participant	eXistence		X

# OLAC Paradigms Compared with UNTL

Sign-language Paradigm	Archiving Paradigm	Lab Paradigm	Technology Operator Paradigm	Publishing Paradigm	Performing Arts Paradigm	Software Paradigm	Language Project Paradigm	Interview Paradigm	Audio to Text Paradigm
signer	depositor	researcher	photographer	author	performer	developer	consultant	interpreter	annotator
		research_participant	recorder	translator	singer			interviewer	transcriber
				editor				speaker	data_inputter
				compiler			sponsor	participant	
				illustrator				responder	

# Conclusion

- What are the overlaps in roles across the different schemas?
- Are there inaccessible frames of reference needed in the description of language resources? YES

Given the Red and Black highlighted roles, not only are there a significant number of OLAC roles inaccessible via the UNTL schema but the schema doesn't allow for description in key areas such as *Software development*, *Sign-language resources*, and *Archival Activities*.

## Conclusion II

# Examples



## OLAC Record

[oai:sil.org:40239](https://www.sil.org/resources/archives/40239)

### Metadata

<i>Title:</i>	Translating "Messiah," "Christ," and "Lamb of God"
<i>Contributor (author):</i>	King, Phil
<i>Date (W3CDTF):</i>	2005
<i>Description (URI):</i>	<a href="https://www.sil.org/resources/archives/40239">https://www.sil.org/resources/archives/40239</a>
<i>Extent:</i>	pages 1-27
<i>Format (IMT):</i>	application/pdf
<i>Identifier (URI):</i>	<a href="https://www.sil.org/resources/archives/40239">https://www.sil.org/resources/archives/40239</a> <a href="https://doi.org/10.54395/jot-tc4pm">https://doi.org/10.54395/jot-tc4pm</a>
<i>Is Part Of (URI):</i>	<a href="https://oai:sil.org:40276">oai:sil.org:40276</a>
<i>Language:</i>	English
<i>Language (ISO639):</i>	eng
<i>Publisher:</i>	SIL International
<i>Relation:</i>	Is Part of Series: Journal of Translation 1(3)
<i>Subject:</i>	Translation Relevance Theory
<i>Subject (OLAC):</i>	<a href="#">translating_and_interpreting</a>
<i>Type (DCMI):</i>	Text



## OLAC Record

[oai:www.mpi.nl:tla\\_1839\\_00\\_0000\\_0000\\_0016\\_4F8F\\_E](https://hdl.handle.net/1839/00-0000-0000-0016-4F8F-E)

### Metadata

<i>Title:</i>	Jornal de Minde no. 84 (1961)
<i>Contributor (recorder):</i>	Ramón Rodríguez
<i>Contributor (researcher):</i>	Vera Ferreira
<i>Coverage:</i>	Portugal Minde
<i>Description:</i>	Article published on Jornal de Minde no. 84 (1961). This article about some lexical and phonetic peculiarities of Minderico reproduces the text published by Reverend Mário Martins in the journal "Magnificat".
<i>Format:</i>	audio/x-wav text/x-eaf+xml application/pdf
<i>Identifier (URI):</i>	<a href="https://hdl.handle.net/1839/00-0000-0000-0016-4F8F-E">https://hdl.handle.net/1839/00-0000-0000-0016-4F8F-E</a>
<i>Is Part Of:</i>	DoBeS archive : Minderico
<i>Language:</i>	Minderico
<i>Language (ISO639):</i>	drc
<i>Publisher:</i>	The Language Archive, Max Planck Institute for Psycholinguistics
<i>Subject:</i>	Minderico language
<i>Subject (ISO639):</i>	drc
<i>Type (DCMI):</i>	Sound Text

# Examples



## OLAC Record

[oai:scholarspace.manoa.hawaii.edu:10125/24768](http://oai:scholarspace.manoa.hawaii.edu:10125/24768)

### Metadata

<i>Title:</i>	A Guide to the Syuba (Kagate) Language Documentation Corpus
<i>Bibliographic Citation:</i>	Gawne, Lauren; 2018-04; Kaipuleohone University of Hawai'i Digital Language Archive; <a href="http://hdl.handle.net/10125/24768">http://hdl.handle.net/10125/24768</a> .
<i>Creator:</i>	Gawne, Lauren
<i>Date (W3CDTF):</i>	2018-04
<i>Description:</i>	This article provides an overview of the collection "Kagate (Syuba)", archived with both the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC) and the Endangered Language Archive (ELAR). It provides an overview of the materials that have been archived, as well as details of the workflow, conventions used, and structure of the collection. It also provides context for the content of the collection, including an overview of the language context, and some of the motivations behind the documentation project. This article thus provides an entry point to the collection. The future plans for the collection – from the perspectives of both the researcher and Syuba speakers – are also outlined, but with the overwhelming majority of items in the collection available to others, it is hoped that this article will encourage use of the materials by other researchers.
<i>Format:</i>	National Foreign Language Resource Center 31 pages
<i>Identifier:</i>	Gawne, Lauren. 2018. A Guide to the Syuba (Kagate) Language Documentation Corpus. Language Documentation & Conservation 12. 204-234. 1934-5275
<i>Identifier (URI):</i>	<a href="http://hdl.handle.net/10125/24768">http://hdl.handle.net/10125/24768</a>
<i>Publisher:</i>	University of Hawaii Press
<i>Rights:</i>	Creative Commons Attribution-NonCommercial 4.0 International
<i>Subject:</i>	Archives Nepal Tibeto-Burman Open Access Collections
<i>Table Of Contents:</i>	<a href="#">gawne.pdf</a>
<i>Type:</i>	Article
<i>Type (DCMI):</i>	<a href="#">Text</a>

Second Type Element

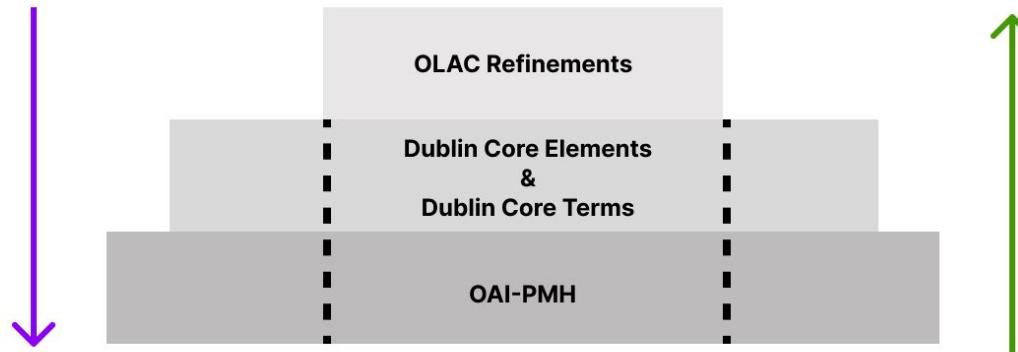
Bibliographic Citation element is confusing

Is the ISSN describing the Article or the Journal?

No Links to the collection described

# The Analytical Perspective

Austin, Peter K. 2013. "Language Documentation and Meta-Documentation." In *Keeping Languages Alive: Documentation, Pedagogy and Revitalization*, edited by Mari C. Jones and Sarah Ogilvie, 3–15. Cambridge, UK; New York: Cambridge University Press. doi:10.1017/CBO9781139245890.003.



Bird, Steven, and Gary F. Simons. 2004. "Building an Open Language Archives Community on the DC Foundation." In *Metadata in Practice*, edited by Diane Hillmann and Elaine L. Westbrooks. Chicago: American Library Association.

# Functional Requirements for Bibliographic Records

