

# Where Have All the Collections Gone?

HUGH PATERSON III

## Abstract:

The Open Language Archives Community (OLAC) aggregator currently compiles 443, 217 records from 65 providers. Participating archives each provide Dublin Core metadata via an OAI feed.

Based on the needs of both linguists and language community members, Wasson et al. (2016) note that usability requirements are not met by language-archive records. Burke and Zavalina (2019, 2020) established that record composition for the free-text description field is used in various ways across the three archives they evaluated. Some of these free-text description fields indicated that the “item” was in fact more like a “collection”. However, collection records should have a different composition from individual artifact records because they each have distinct scopes. With this in mind, different record types should have distinct evaluation criteria. Unexplored in the literature are how record providers are utilizing distinct collection records.

The current study explores the nature of linguistic and anthropological language-focused collection records in contrast to artifact records by investigating the use of the DCMIType attribute “collection”. Across the OLAC records, 850 use the DCMIType “collection”; only 7 providers even use the “collection” DCMIType. By using the DCMIType “collection” and relating artifact records with collection records via the Dublin Core “hasPart” property, more about the original context of the collection is transferred from the host institution to the OLAC aggregator. When properly displayed this can lead to increased utility in browsing environments.

## About the author:

*Hugh Paterson III* is an unaffiliated collaborative scholar who is interested in the production, discovery, and reuse cycle of language resources. Language resources are often created or used as part of activities in anthropological linguistics, language documentation, and language revitalization. He recently completed an M.A. in linguistics which looks at the reusability of metadata from language archives. He has both field experience as a linguist and has worked in a language archive.