

Supporting Diversity Through a Typology of Data Providers: *What is an Archive?*

Hugh Paterson III. 2023. Supporting Diversity Through a Typology of Data Providers: What is an Archive?. Paper presented at the 8th International Conference on Language Documentation and Conservation. University of Hawaii Manoa. Presentation slides Copyright 2022 licensed: CC-SA-ND.

Inspiration

Yi I, Lake A, Kim J, Haakman K, Jewell J, Babinski S, et al..

Accessibility, Discoverability, and Functionality: An Audit of and Recommendations for Digital Language Archives.

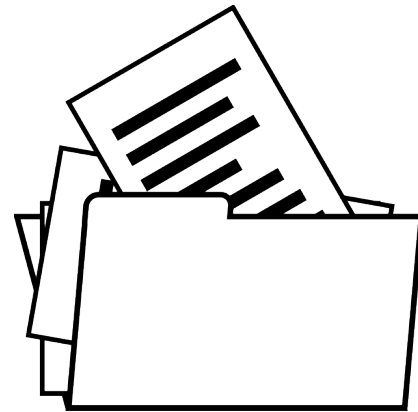
Journal of Open Humanities Data. 2022;8:10. DOI: <http://doi.org/10.5334/johd.59>

Definition of an archive

A Institution with a mandate for preservation.



A Collection of Resources (Assumed to be from the same institution), usually on matters related to the institution.



Definition of an archive

A Institution with a mandate for preservation.



A Collection of Resources (Assumed to be from the same institution), usually on matters related to the institution.



Definition of an archive

A Institution with a mandate for preservation



The screenshot shows the homepage of the Archive of the Indigenous Languages of Latin America (AILLA). The header is orange with a globe icon and the text "The Archive of the Indigenous Languages of Latin America". Below the header is a navigation bar with links: Collections, Languages, Countries, Announcements, Organizations, Persons, and a search bar. The main content area is yellow and features a large black and white photograph of several people in a boat. To the right of the photo is a "User" section with links for "My Account", "User Login", and "Log In". Below the photo is a "Main Menu" with links: Welcome, Browse, Depositor Info, Rights, Events, AILLA Papers, Contact, Contribute, News and Announcements, and FAQ.

Welcome to the Archive of the Indigenous Languages of Latin America (AILLA)

Indigenous Peoples' Day, October 9, 2017, marks the public launch of the newly migrated and updated Archive of the Indigenous Languages of Latin America. AILLA is a digital languages archive of recordings, texts, and other multimedia.

A Collection of Resources (Assumed to be from the same institution), usually on matters related to the institution.

Pahka'anil (Tübatulabal) Text Project

Michael Ahland
Department of Linguistics
California State University, Long Beach



Home Texts Credits Research

The Pahka'anil Language

Pahka'anil (Tübatulabal) is an indigenous Uto-Aztecan language spoken by the Pakanapul (also known by the exonym Tübatulabal) in the Kern River Valley of Kern County, California (see Golla's map). Today, the language is spoken by a growing community of second-language learners, teachers, and community members. The language's status is best represented as "Reawakening." The Pakanapul Language Team, which includes tribal leadership and language teachers, has been working for years to develop pedagogical and literacy materials as well as run language classes in the Lake Isabella and Bakersfield areas.

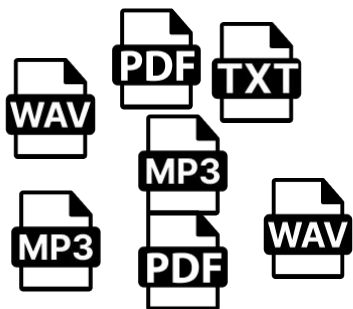


The Project

Characteristics of the different senses of “archive”

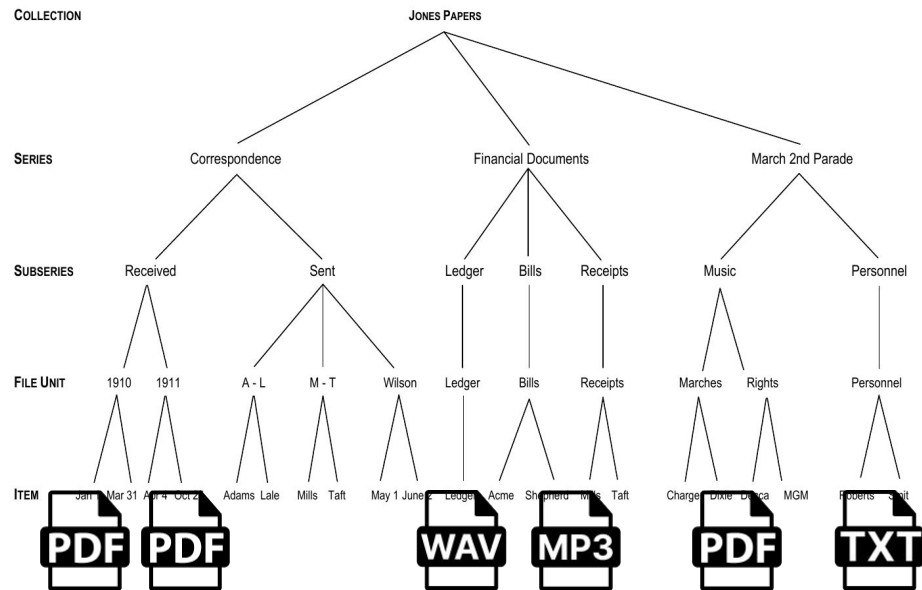
Document Server

- Independent documents



Archival Bond

- Associations between documents made explicit



Characteristics of the different senses of “archive”

Document Server



arXiv is a free distribution service and an open platform for the fields of physics, mathematics, computer science, electrical engineering and systems pre-reviewed by arXiv.

Subject search and browse:

News
Read about recent news and updates on arXiv's blog. (View the former "what's new" pages here).
Read robots beware before attempting any automated download.



Welcome to Cogprints, an electronic archive for self-archive papers in any area of Psychology, Neuroscience, and Linguistics, and many areas of Computer Science (e.g., artificial intelligence, robotics, vision, learning, speech, neural networks), Philosophy (e.g., mind, language, knowledge, science, logic), Biology (e.g., ethology, behavioral ecology, sociobiology, behaviour genetics, evolutionary theory), Medicine (e.g., Psychiatry, Neurology, human genetics, imaging), Anthropology (e.g., primatology, cognitive ethnology, archeology, paleontology), as well as any other portions of the physical, social and mathematical sciences that are pertinent to the study of cognition.

[Browse Repository](#)
Browse the items in the repository by subject.

[About this Repository](#)
More information about this site.

Physics

- Astrophysics ([astro-ph](#) new, recent, search)
 - includes Astrophysics of Galaxies; Cosmology and Nongalactic Astrophysics; Earth and Planetary Astrophysics; High Energy Astrophysical Phenomena; Instrumentation and Methods for Astrophysics; Solar and Stellar Astrophysics
- Condensed Matter ([cond-mat](#) new, recent, search)
 - includes Disordered Systems and Neural Networks; Materials Science; Mesoscale and Nanoscale Physics; Other Condensed Matter; Quantum Gases; Soft Condensed Matter; Statistical Mechanics; Strongly Correlated Electrons; Superconductivity
- General Relativity and Quantum Cosmology ([gr-qc](#) new, recent, search)
- High Energy Physics - Experiment ([hep-ex](#) new, recent, search)
- High Energy Physics - Lattice ([hep-lat](#) new, recent, search)
- High Energy Physics - Phenomenology ([hep-ph](#) new, recent, search)
- High Energy Physics - Theory ([hep-th](#) new, recent, search)
- Mathematical Physics ([math-ph](#) new, recent, search)
- Nonlinear Sciences ([nlin](#) new, recent, search)
 - includes Adaptation and Self-Organizing Systems; Cellular Automata and Lattice Gases; Chaotic Dynamics; Exactly Solvable and Integrable Systems; Pattern Formation and Solitons
- Nuclear Experiment ([nucl-ex](#) new, recent, search)
- Nuclear Theory ([nucl-th](#) new, recent, search)
- Physics ([physics](#) new, recent, search)
 - includes Accelerator Physics; Applied Physics; Atmospheric and Oceanic Physics; Atomic and Molecular Clusters; Atomic Physics; Biological Physics; Chemical Physics; Classical Physics; Computational Physics; Data Analysis; Statistics and Probability; Fluid Dynamics; General Physics; Geophysics; History and Philosophy of Physics; Instrumentation and Detectors; Medical Physics; Optics; Physics and Society; Physics Education; Plasma Physics; Popular Physics; Space Physics
- Quantum Physics ([quant-ph](#) new, recent, search)

Mathematics

Archival Bond

Japanese Language and Literature

SEARCH THE JOURNAL

This title is part of a longer publication history. The full run of this journal will be searched.

TITLE HISTORY

2001-2021 • [Japanese Language and Literature](#)

1972-2000 • [The Journal of the Association of Teachers of Japanese](#)

1963-1972 • [The Journal-Newsletter of the Association of Teachers of Japanese](#)

Japanese Language and Literature publishes contributions in the areas of Japanese literary studies, Japanese linguistics, and Japanese language and literature pedagogy, as well as articles from other disciplines that help interpret or define the problems of Japanese literary history, literary or linguistic study, or classroom practice. Occasionally, an issue contains several articles on a single topic and is designated a "special issue."

All Issues

~ 2020s

2021 (Vol. 55)

[No. 2 OCTOBER 2021 pp. 407-510](#)

[No. 1 APRIL 2021 pp. 1-406](#)

2020 (Vol. 54)

[No. 2 OCTOBER 2020 pp. 167-606](#)

[No. 1 April 2020 pp. 1-366](#)

~ 2010s



PUBLISHED BY
[American Association of Teachers of Japanese](#)

[Journal Home Page](#)
[Subscribe](#)

COVERAGE

2001-2021 (Vol. 35, No. 1 - Vol. 55, No. 2)

MOVING WALL  [Learn more](#)

1 year

ISSN

15367827

EISSN

23264586

SUBJECTS

Language & Literature, Asian Studies,
Education, Social Sciences, Area Studies,
Humanities

Real world language usage I

Your laptop to the rescue: Using the Child **Language** Data Exchange System **archive** and CLAN utilities to improve child **language** sample analysis

[NB Ratner](#), [B MacWhinney](#) - ... in *Speech and Language*, 2016 - [thieme-connect.com](#)

... The second is a review of data obtained from the Child **Language** Data Exchange System (CHILDES) project **archive** that we use to evaluate the potential utility of certain LSA measures ...

☆ Save  Cite Cited by 39

Indian Sepoy Experience in Europe, 1914–18: **Archive, Language, and Feeling**

[S Das](#) - *Twentieth Century British History*, 2014 - [academic.oup.com](#)

... A photograph from the Imperial War Museum **archives** ... In such a context, it is essential to go beyond official **archives** ... in Hindi; in a small **archive** in Chandernagore, a former French ...

☆ Save  Cite

Issues in the creation of a digital **archive** of a signed **language**

[T Johnston](#), [A Schembri](#) - ... : *Proceedings of the conference held at ...*, 2006 - [books.google.com](#)



... Overall, the primary aim of the project is to create a set of digital **archive** of Auslan as used ... **language** remains possible in the face of rapidly decreasing numbers of signers. The **archived** ...

☆ Save  Cite Cited by 56 Related articles All 9 versions

Bringing user-centered design to the field of **language archives**

[C Wasson](#), [G Holton](#), [HS Roth](#) - 2016 - [scholarspace.manoa.hawaii.edu](#)

... (UCD) and **language archives**. One of the challenges facing **language archives** is the fact that ... UCD, informed by design anthropology, can help developers of **language archives** identify ...

☆ Save  Cite Cited by 37 Related articles All 6 versions 

Real world language usage II

Language documentation in the Americas

[B Franchetto, K Rice - 2014 - scholarspace.manoa.hawaii.edu](#)

... international researchers and institutions hosted in **the archives** to future generations • to bolster data exchange among **the archives**, creating mechanisms to facilitate mutual access ...

☆ Save  Cite Cited by 11 Related articles All 5 versions >>

The Open **Language Archives** Community: An infrastructure for distributed archiving of language resources

[G Simons, S Bird - Literary and Linguistic Computing, 2003 - academic.oup.com](#)

... **language documentation** and description. After reaching consensus on the requirements for **language-resource archiving** (... OLAC, the Open **Language Archives** Community, is an inter...

☆ Save  Cite Cited by 69 Related articles All 9 versions

A brief history of archiving in **language documentation**, with an annotated bibliography

[R Henke, AL Berez-Kroecker - 2016 - scholarspace.manoa.hawaii.edu](#)

... poses of **language documentation** and endangered **language** ... , when increased attention to **language** endangerment and the ... In 2005, Hinton noted that **the archives** at Berkeley were "...

☆ Save  Cite Cited by 50 Related articles All 6 versions >>

Real world language usage III

What is “archiving”? Is it putting things in a collection or is it interacting with an institution?

The Open **Language Archives** Community: An infrastructure for distributed **archiving** of language resources

[G Simons, S Bird](#) - Literary and Linguistic Computing, 2003 - [academic.oup.com](#)

... **language documentation** and description. After reaching consensus on the requirements for **language-resource archiving** (... OLAC, the Open **Language Archives** Community, is an inter...

☆ Save ↗ Cite Cited by 69 Related articles All 9 versions

A brief history of **archiving** in language documentation, with an annotated bibliography

[R Henke, AL Berez-Kroecker](#) - 2016 - [scholarspace.manoa.hawaii.edu](#)

... , and trends in **archiving** for the purposes of **language** documentation and endangered **language** conservation. We identify four major periods in the history of such **archiving**. First, a ...

☆ Save ↗ Cite Cited by 50 Related articles All 6 versions ↗

Unanswered questions in **language documentation** and revitalization

[LA Grenoble](#) - ... directions in **language documentation** and **language** ..., 2013 - [torrossa.com](#)

... **language documentation** means that it provides little data about the very situations which have put these **languages** ... and the retrieval problem in **language** documentation and **archiving**. ...

☆ Save ↗ Cite Cited by 25 Related articles All 4 versions

Misconceptions

1. Access ≠ Archive

A website is a website is a website: Why trusted repositories are needed more than ever

V Ferreira, [L Lukschy](#), B Watyam... - ... Language **Archives** ..., 2021 - ideals.illinois.edu

... server does not mean they are **archived** and preserved for future generations. In this paper we discuss the differences between professional **archiving** systems and content management ...

☆ Save 📄 Cite Cited by 2 Related articles All 6 versions 🔗



2. The website is not the Archive

[PDF] The Formosan **language archive**: Linguistic analysis and **language processing**

E Zeitoun, CH Yu - ... Linguistics & Chinese **Language Processing** ..., 2005 - aclanthology.org

... users to determine the geographical distribution of each **language**/dialect. It is hoped that in ... to linguistics, **language** teaching, literature and music. The display of the **Archive** will not be ..

☆ Save 📄 Cite Cited by 36 Related articles All 9 versions 🔗

Language documentation in the aftermath of the 2015 Nepal earthquakes: A guide to two archives and a web exhibit

[KA Hildebrandt](#), T Burge-Beckley... - **Language Documentation** ..., 2019 - spark.siue.edu

... specifically, the **documentation** efforts and activities behind **the archives** and **web exhibit** that we describe in this paper arose from the two massive earthquakes that severely impacted ...

☆ Save 📄 Cite Cited by 7 Related articles All 7 versions 🔗

OLAC's current definition of a Data provider

OLAC archive description

- type="**institutional**" indicates that the repository is operated by an institution that is committed to maintaining it in the future, even after the individuals currently associated with it are no longer involved.
- type="**personal**" indicates that the repository is being operated by an individual (or a group of individuals) without the commitment of an institution for maintenance far into the future.

Classes of data providers

Access Institutions

Exhibits and Standalone Collections

Reference Resources

Classes of data providers II

Access Institutions

Exhibits and Standalone Collections

Reference Resources

Archive: Preservation, Curation,
Community Re-engagement

Repository: Faithful Distribution

Library: Temporary Communal Access

Museum: Curator Driven Description
and Access

Gallery: Focused Sales

The business model of the institution will impact their approach to the web and artifact description.

Classes of data providers III

Access Institutions

Exhibits and Standalone Collections

Reference Resources

Special Collections: Limited scope by
topic or experience

Personal CV or Portfolio: Limited scope to
a single person's work

Lab or Department's Portfolio: An exhibit
of the work of a single organizational unit

Project Portfolio: Limited Scope to a single
research endeavor

Classes of data providers IIII

Access Institutions

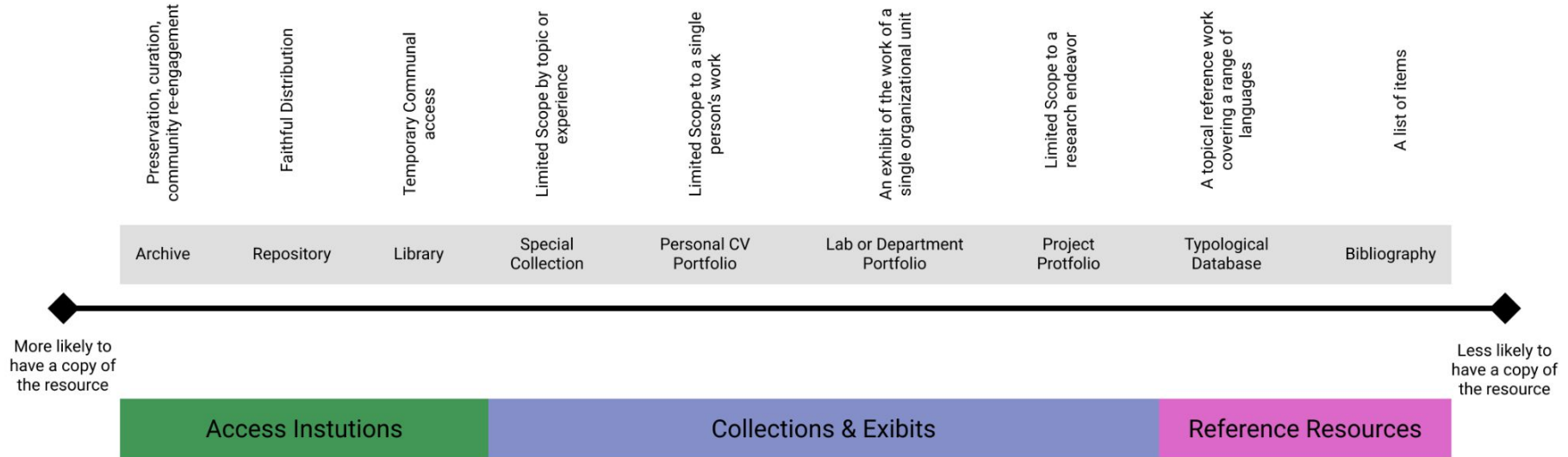
Exhibits and Standalone Collections

Reference Resources

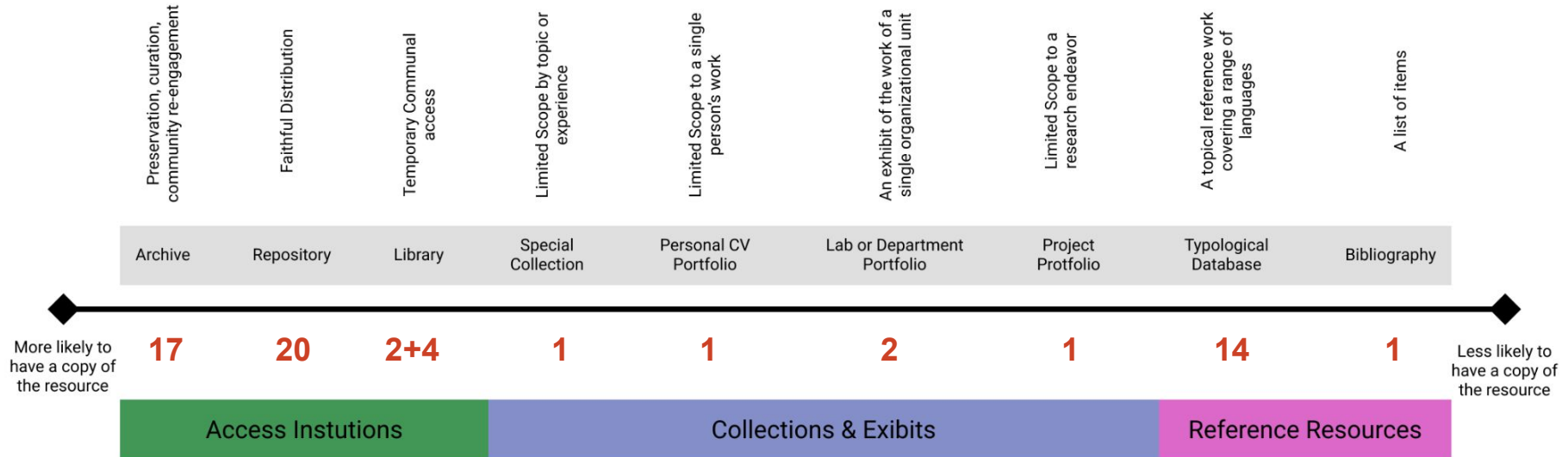
Typological Databases: A typological reference resource covering a range of languages

Bibliographies: A list of items

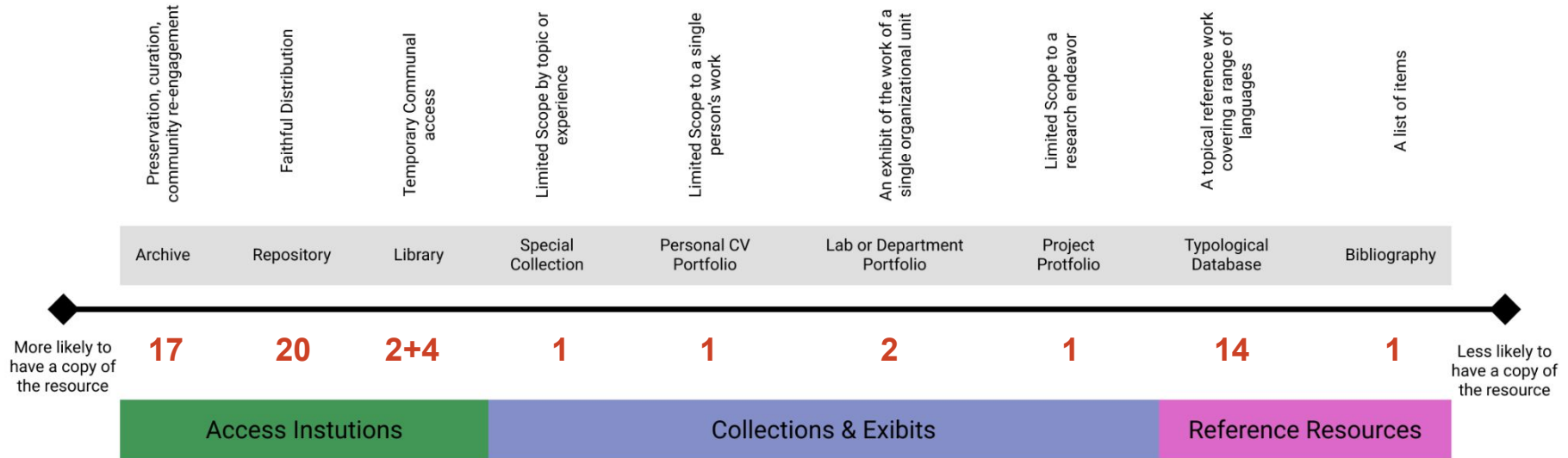
Is everything an archive?



What is the distribution we see?



Acknowledging Diversity Moderates Expectations



By recognizing the diversity in data providers it allows us to form interaction and quality based expectations around web-interactions.