

Data citation, attribution, and employability

Meagan Dailey & Ryan Henke | Department of Linguistics, University of Hawai'i at Mānoa

OUR QUESTION: Data management, archiving, and dissemination are the backbone of data citation and attribution. **How does data relate to employability?**

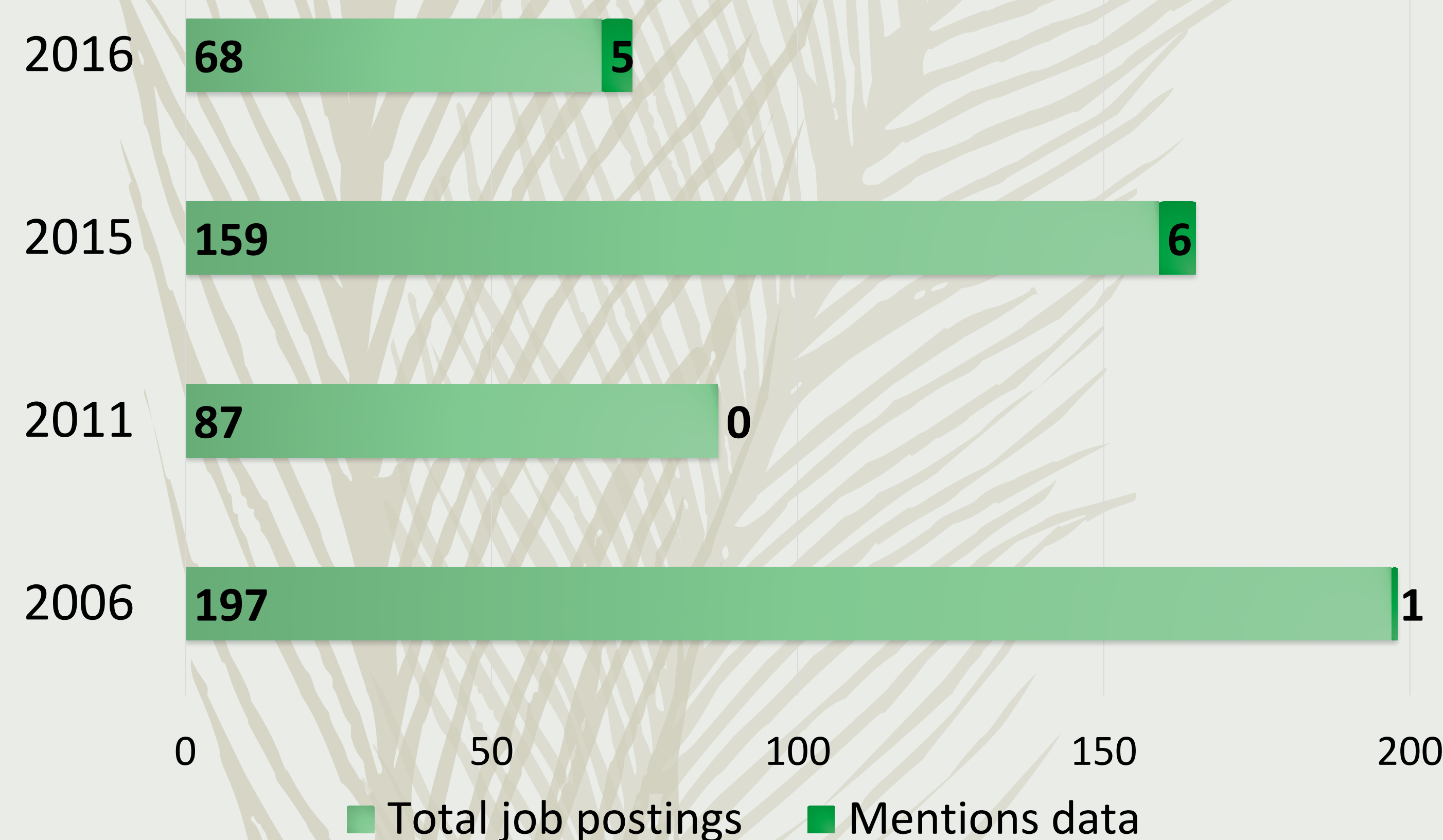
1. THE JOB MARKET

QUESTION: How does the linguistic job market approach *data*?

BAROMETER: LINGUIST List **job postings** in 2006, 2011, 2015, 2016
Any faculty positions in the U.S. | Any linguistic subfield

TARGET: Mentions data management, archiving, *or* dissemination

RESULTS: **12 total** postings over ten years mention data



KEY LANGUAGE FROM RELEVANT POSTINGS

- Research with "database design and management"
- Applicants for "interdisciplinary research in data science"
- Wanting individual to leverage "massive data repositories ... and/or their own data collection efforts"
- Specialization in "in very large data resources (so-called 'Big Data')"
- Someone to "utilize traditional and digital archival techniques to expand and maintain a public collection" of language data

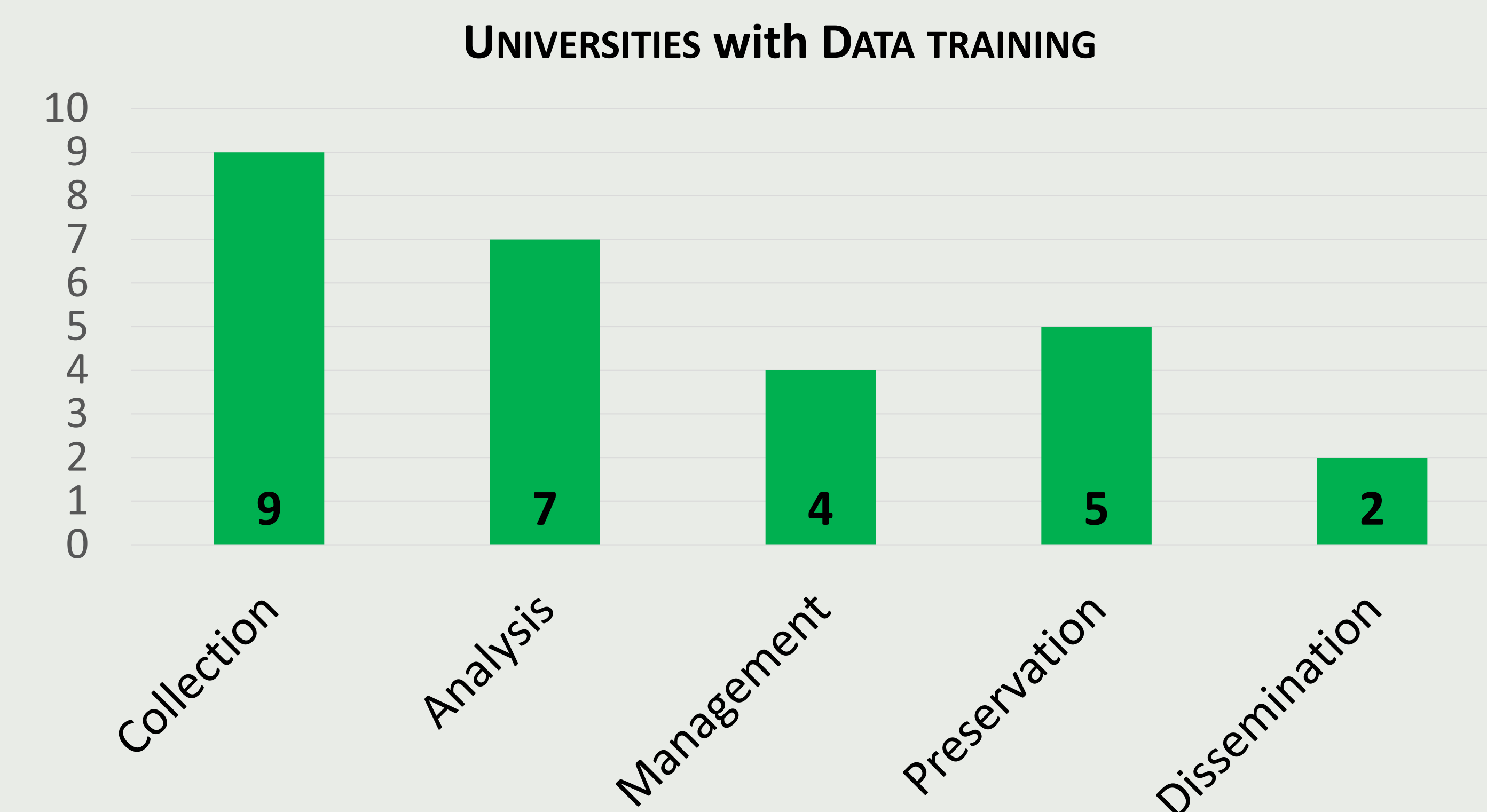
2. GRADUATE EDUCATION

QUESTION: How are we training future linguistics faculty in *data*?

BAROMETER: Top 25 Quacquarelli Symonds-ranked linguistics programs: 20 in U.S., 5 international
Course descriptions from any subfield

TARGET: Descriptions mentioning training in data management, archiving, *or* dissemination

RESULTS: 1) **19 schools** have course descriptions mentioning data training
2) 6 do not provide course descriptions, or their course catalogue is closed



WHICH KINDS OF COURSES FOCUS ON DATA TRAINING?

- Sociolinguistics
- Language documentation
- Field linguistics
- Computational linguistics

3. CONCLUSIONS

JOB MARKET

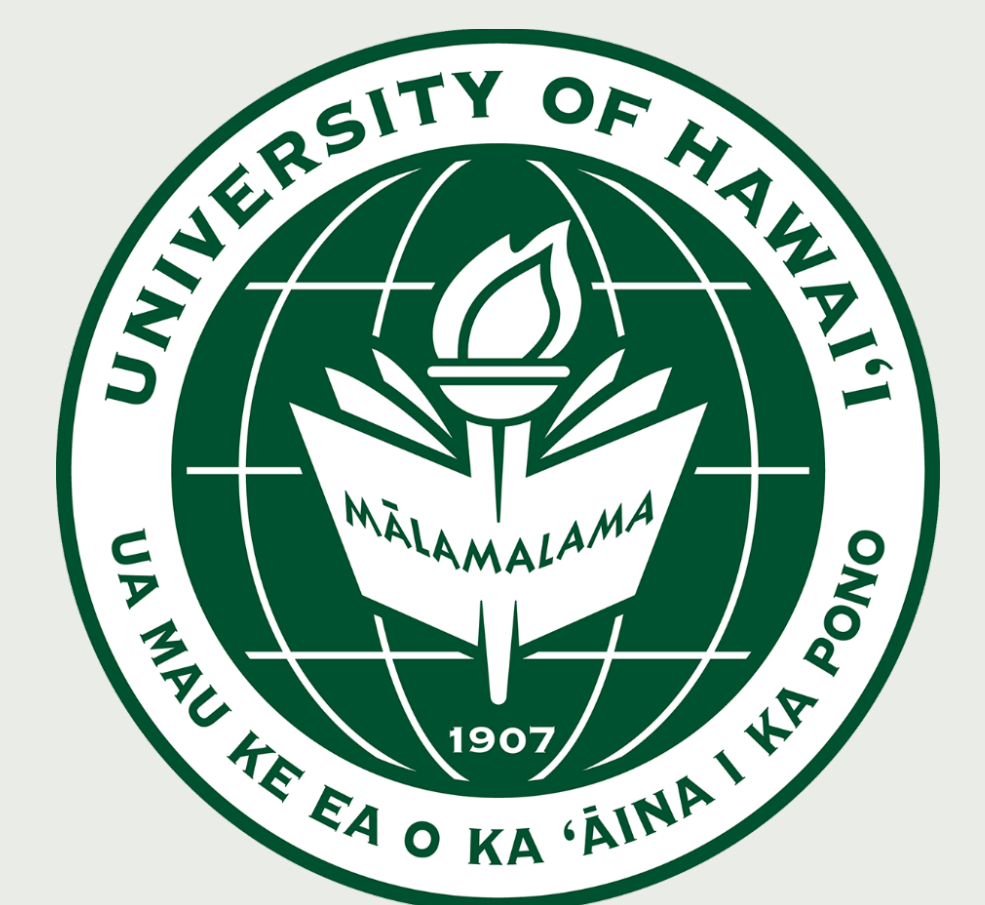
- Very few job postings talk about data management, archiving, or dissemination
- Postings mention little to nothing about reproducible research
- But the existing mentions of data are promising

GRADUATE EDUCATION

- Programs offering field methods courses addressed issues of data collection and analysis
- With few exceptions, only computational methods courses described long-term data management or dissemination

OVERALL

- Students enrolled in linguistics graduate programs have opportunities for training in data collection, management and preservation ...
- Especially if their department has a focus in **sociolinguistics**, **language documentation**, or **computational linguistics**
- Hiring committees are slowly becoming increasingly concerned that job candidates have experience with data beyond just doing research on primary language data



THANKS!

Meagan Dailey medailey@hawaii.edu

Ryan Henke rhenke@hawaii.edu