

Leigh Phan

UCLA Library



UCLA LIBRARY
DATA SCIENCE
CENTER

San Diego Supercomputing Center - Summer Institute 2020

Applying machine learning & AI to collections and archival material in the Library as a way to make materials more accessible and engaging

- Oral History Research Center - transcripts and audio
- Text analysis/Natural Language Processing
- Detecting and identifying objects and people in historic photo collections
- Ethnomusicology archive
- UCLA Film & TV Archive

Computational challenges

- Still to be determined
- UCLA Film & TV Archive
 - Over 350,000 film and television programs
 - Over 100,000 News and Public Affairs programs
 - Over 2,000 radio programs
 - Processing and analyzing millions of moving images will take more advanced processing
- UCLA HPC cluster