bConnected CASE STUDY

California Linguistics Archive

UC Berkeley's California Linguistics Archive houses historical materials — including field notebooks, video and audio recordings, and PDFs — that span the last 100 years. The archive, which focuses on American Indian languages, is an important depository for communities trying to maintain or revive their languages. When permissions allow, the public can view the materials and listen or watch the recordings.

Problem

The department has been digitizing and storing an increasing number of archives on its server. However, the server has its limitations; disk space is running out, and the archive needs more storage.

Secondly, there is only one IT person in the department managing the server, which creates a certain risk to the archived materials. "There is no backup in terms of personnel or the hardware itself," says Ronald Sprouse, Information Systems Analyst. "We don't have other servers to fall back on." In case of a server failure, this means that the materials won't be accessible. Another problem is that it has been challenging for both the department and users to consistently work with different browsers and operating systems to access materials.

Solution

The department has decided to move to Box for their storage needs. Currently, existing accessioned materials in CLA have been copied to Box. The next steps will be to deliver online content to the public via Box, and to eventually make Box the platform through which depositors contribute their materials.

Box, which offers unlimited storage, will accommodate larger video files and a growing archive, as it expands to take in a broader array of materials. Box also allows for easier control of permissions using a Special Purpose Account (SPA), which restricts access when necessary.

The idea is to make using the archive easy: Staff at linguistics helps with the metadata describing the materials, and encourages people to use the archive as they go along. This way, when the owners are ready, materials are easy to release to the public.

Currently, the archive has 3 terabytes digitized. "But we expect that to go up more quickly as more people start to contribute video materials," says Sprouse.

Berkeley
UNIVERSITY OF CALIFORNIA