Introduction

AVA (Archives Video Appraiser) leverages ML to generate recommendations on which born-digital video recordings should be permanently preserved or destroyed based on a set of pre-defined appraisal criteria.

AVA (Archives Video Appraiser)

is a ML and automation tool that generates archival selection recommendations with an accuracy of ~85% and has resulted in 1.5-person day savings per month.

What Does AVA Do?

• Scan designated network folders for new videos.

• Extract, generate, and store in DB: filename, images from video (1 frame per minute), text from images, object labels for images, and transcript using voice to text.

• Analyze video data. Identify empty and soundless recordings. Generate archival retention and destruction recommendations. AVA requires ~10 minutes to generate recommendations for 200 videos.

• Provide dashboard for archivists to confirm or correct each of AVA's recommendations.

• Transfer permanent videos to a digital preservation staging area. Delete videos identified for destruction. Audit reports are generated and captured in the World Bank Group’s EDRMS.

Lessons Learned

• AVA is effective and saves time!

• AVA needed 3000 sample appraisal decisions, previously generated manually by World Bank Group Archivists, in order to learn to differentiate between permanent and temporary content.

• Identifying the representative set of training data (those 3000 videos) required a high initial time investment for the archivists.

• Ongoing, human-driven, iterative training is required to increase accuracy.

• This use case required evaluation of straightforward criteria. Scaling AVA to tackle larger collections of mixed formats with more complex selection criteria will likely present greater challenges.

• We must remain vigilant to reduce bias included in our training data.

Technology used to develop AVA

M365: PowerApps, PowerApps API and Power Automate (Flow).

AWS: API Service, Python, Flask, Celery, ECS Lambda, Rekognition and RabbitMQ.

Ask us anything!

Pbeneitoarias@worldbankgroup.org
Jkramersmyth@worldbankgroup.org