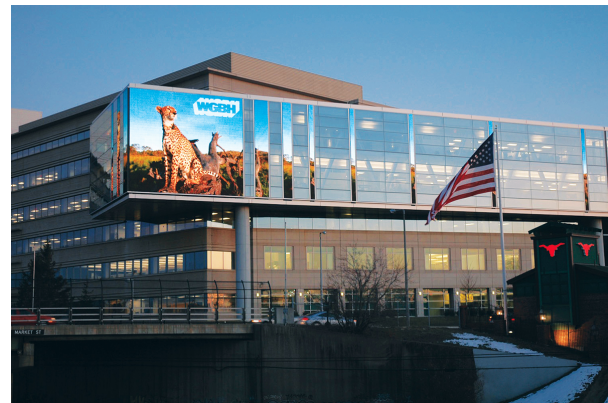# Transitioning Long-Term Storage

at WGBH Media Library and Archives

WGBH

The Archive

- ~1 PB of born-digital and digitized material

- Annual growth rate: ~350 TB
  - legacy digital material on outdated media
  - new material ingested from productions

- Each file currently stored on two geographically separated LTO 6 tapes (WGBH Vault and offsite storage)

- Proxies for audio and video stored on AWS & viewable through preservation database

The Setup

- 4 LTO 6 stations with 7 (aging) standalone decks

- 2 staff dedicating 30-50% of their time to digital preservation & storage workflows

- Media delivered from productions on drives or over shared network

NEH Requirements

- Migrate forward current collections from LTO 6 to a contemporary storage format (~1.3 PB)

- Digitize and ingest 83,000 at-risk media assets (~2.5 PB)

- Improve born-digital ingest from productions onsite

NATIONAL ENDOWMENT
FOR THE HUMANITIES

3-Year Goals

- Digital media stored in three ways and three different locations
  - 1 local, online copy on Cloudian Enterprise Object Storage
  - 1 local, offline copy on LTO 8
  - 1 cloud-based, offline copy on Amazon Glacier

- Metadata travels with media
  - Required metadata fields for uploading content to Cloudian
  - Metadata stored as key value pairs in Cloudian and on Glacier, and exported to digital preservation database as PBCore

- Proxies immediately available for all digital media
  - Automatically generated on ingest to Cloudian
  - Stored in Amazon S3 (tiered with Cloudian)
  - Accessible through MLA's Avalon Media System and through institutional Cloudian front end

Processes (in development)

- Migrate forward current collections from LTO 6 to a contemporary storage format (~1.3 PB)
  - vendor or in-house?

- Digitize and ingest 83,000 at-risk media assets (~2.5 PB)
  - files delivered and QCed on physical media before ingest, or directly uploaded to object storage?

- Improve born-digital ingest from productions onsite
  - workflows for verifying and documenting file integrity from creation on?
  - minimum metadata requirements for Cloudian front end?

Ideas and discussion welcome!

rebecca_fraimow@wgbh.org