WHEN APPRAISAL ISN’T ENOUGH:
SELECTING BORN-DIGITAL PHOTOGRAPHS AT SCALE

Elizabeth England, National Digital Stewardship Resident
@elizabeengland
eengland@jhu.edu
Sheridan Libraries, Johns Hopkins University

Society of American Archivists annual meeting
July 27, 2017, Portland, OR
COLLECTION BACKGROUND

◆ Campus photographers, Homewood Photography

◆ Since 2004 – digital workflow
  ◆ Accrual rate of 2-3 TB/year
  ◆ DVDs

◆ All jobs stored as JPEGs and NEF
  ◆ NEF – proprietary format; not ideal for preservation
  ◆ JPEG – open format; lossy; good for access but not preservation
WORKFLOW

1. "Incoming Transfer" Directory
2. "Processing" Directory
3. "Unsampled" Directory
4. "Sampled" Directory
5. Convert to .DNG

- CSV of job names
- OpenRefine
- CSV of updated job names

(delete)
PROCESSING

- "Incoming Transfer" Directory
- "Processing" Directory
- "Unsampled" Directory
- "Sampled" Directory
- Convert to .DNG

CSV of job names
- OpenRefine
- CSV of updated job names

Folders:
- 20071116_design_and_pubs_jump_rope
- 20071116_wse_mcveigh
- 20071117_Hope_Date_Auction_RAW_part_1
- 20071117_Hope_Date_Auction_RAW_part_2
- 20071118_sigma_phi_epsilon_banquet_part_1
- 20071127_JHU_Campus_Fall_Photos
SAMPLING

◆ Consulted archival sources:
  ◆ What percent to retain?
  ◆ Quantitative (sampling) or qualitative (selection)?

◆ Considered options:
  ◆ Retain more & sacrifice preservation file format?
  ◆ Preservation file format & retain less?
PERCEPTUAL HASHING
PROJECT DOCUMENTATION


or go to github and search ‘homewood photo’

THANK YOU!