“We have the technology. We can make it better than it was.”: An Experiment in the Rapid Mass Digitization of Archival Material

Lindsay Dumas - Digital Projects Archivist
Harvard Law School Library Innovation Lab
BACKGROUND

1. 1949 Harvard receives a complete IMT and 12 NMT document records

2. Virtually inaccessible at Harvard--left unprocessed the collection is taken out of circulation 15 years ago due to its condition

3. HLSL decides in late 90’s to initiate project to preserve collection and offer access leveraging the power of digital technology and the Web

4. Funding intermittent--outside donors and Harvard-internal funds via university digital initiative and HLS Library funds

from Stadtarchiv Nürnberg
THE OPPORTUNITY

In the fall of 2014/winter of 2015 LIL and HSC had 18 weeks to digitize 360 boxes of case and evidence files using a high speed scanner.
GOAL

Add case and evidence files with corresponding metadata from the remaining trails at the document level into the existing database.
METHODOLOGY

Hypothesize (based on previous experience with the scanner and similar material) → test hypothesis in production → make adjustments as necessary
MEMO TO THE DEPUTY OF THE FUEHRER FOR CONSULTATION WITH THE FUEHRER

EXCHANGE OF NOTES WITH SWITZERLAND.
RESULTS

- 360 boxes
- 4,302 folders
- 413,743 document pages
  - case and evidence files of varying size and condition
  - primarily typescripts, but also photostats

... prepped and scanned by 2 (min) to 5 (max) FTE in 18 weeks

- under $.22 per document for digitization
  - yields = 3 files per page side
    a) 300 dpi lossless 24-bit JP2, b) Group 4 compressed TIFF, c) OCR-generated text

- averaged 5,600 documents per day over the 74 scan days for the project
  - about ¼ the rate of ‘sturdy’ modern pages
LESSONS LEARNED

• Be willing to revisit and update previous methods/guidelines to best fit the project
• Be agile and ready to make changes or exceptions to procedures
• Keep written procedures and visual aids clear and up to date
• Track material throughout process
• Take the time to walk through, in detail, the end to end process and any variations
• Proximity of team / building strong team
• Collaboration with other departments / resources