

presented by:

Isaiah Beard

Digital Data Curator

Rutgers University Libraries

New Brunswick, NJ

Association for Recorded Sound Collections
2014 Pre-Conference Workshop
Chapel Hill, NC



FROM REELS TO BYTES

LESSONS LEARNED IN MAINTAINING
RUTGERS' JAZZ ORAL HISTORIES

TOPICS COVERED

- About myself / About Rutgers Libraries
- About the Jazz Oral History Collection
 - State of the recordings
 - Technical/Preservation challenges
 - Rights Concerns
 - What we've done
 - What we must do
- Lessons Learned

ABOUT RUTGERS UNIVERSITY LIBRARIES



- A system of 26+ libraries and centers on 4 major campuses serving Rutgers, the state University of New Jersey (58,000+ students, 2,900+ faculty)
- Over 10.5 million holdings, including 3,656,283 volumes, 4,600,902 microforms, 2,535,892 documents

ABOUT ME

- Previous life: worked in the corporate sector. managed IT for electronic document publishing companies, and some broadcasting outfits.
- Started at Rutgers in early 2000s as an IT tech.
- Moved on to a more specialized task: fleshing out best practices and digitization standards for various types of digital surrogates: Photos, books, large documents, and later, audio and video formats.

- Within our library system: The Institute of Jazz Studies



- Founded in 1952, Rutgers became its permanent home in 1966.
- The archival collections at IJS consist of personal papers as well as archives of record companies and jazz-related institutions and organizations spanning from 1920 to the present.



- “...the largest and most comprehensive library and archive of jazz and jazz-related materials in the world.”

JAZZ ORAL HISTORY PROJECT



- Collection of 120 oral histories of pre-Swing and Swing era jazz musicians, most 60 years and older at the time of their interviews.
- Recordings made 1972 - 1983
- 5 to 35 hours in length, each.

JAZZ ORAL HISTORY PROJECT



- Majority of the recordings are in open reel tape format of various brands, mostly AMPEX and Scotch/3M.
- Some later interviews are in audiocassette format, final 1983 interview is a video recording in Betacam.

JAZZ ORAL HISTORY PROJECT



- Recordings are very frequently requested and sought after by Jazz researchers and aficionados.
- The recordings contain a biographical history of popular Jazz artists, in their own words, in their own voices.

JAZZ ORAL HISTORY PROJECT



- Exemplar list:
 - Count Basie
 - Roy Elridge
 - Teddy Wilson
 - Charles Mingus
 - Mary Lou Williams

TRANSCRIPTS

All recordings are accompanied by a paper transcript of the interview.

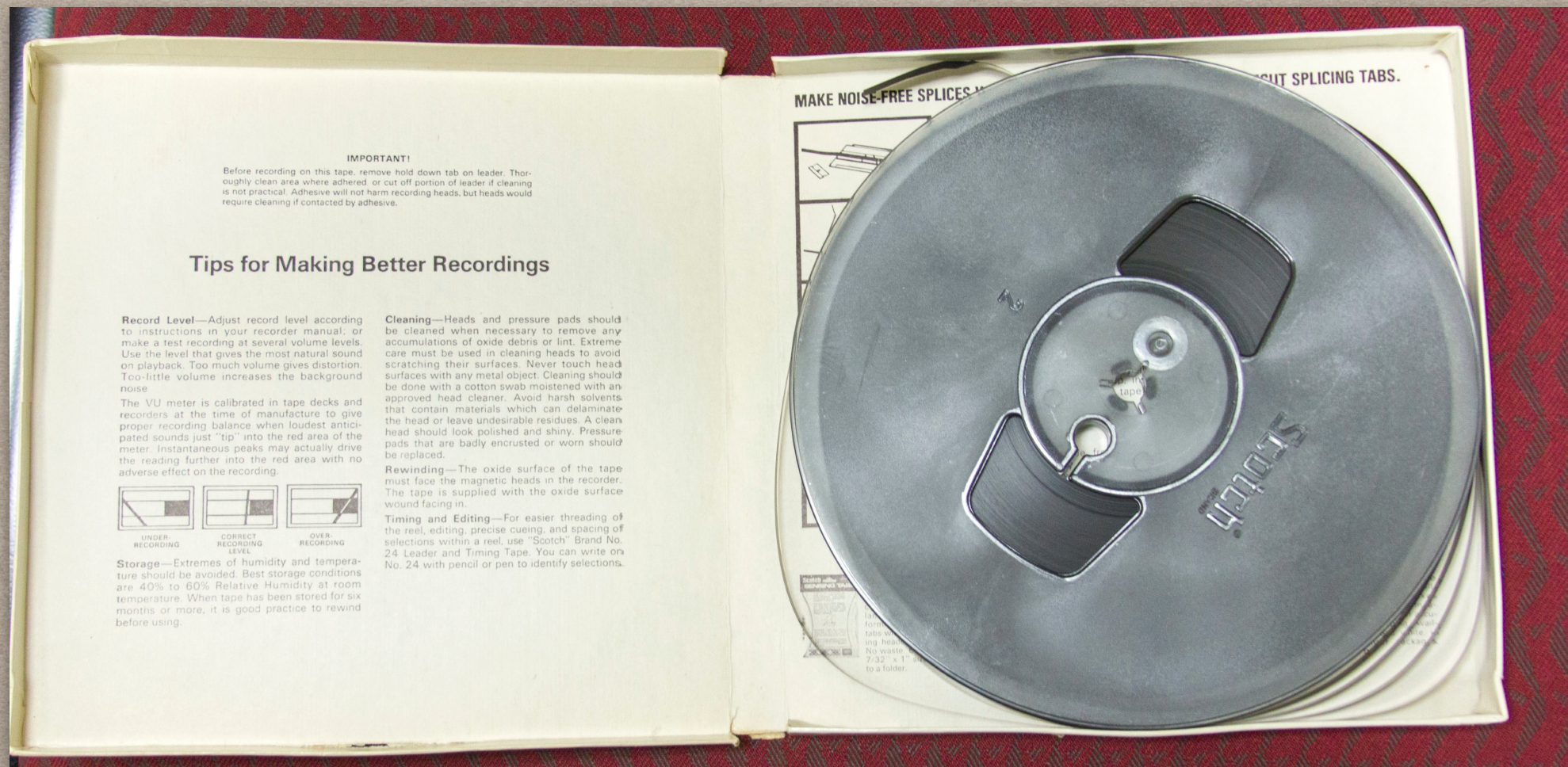


JAZZ ORAL HISTORY PROJECT



- Recordings have changed hands/locations over time
 - Jazz Interactions - NYC based non profit
 - Smithsonian Jazz Program
 - Administration turned over to IJS/Rutgers in 1979
 - Moved to new facilities in 1994

JAZZ ORAL HISTORY PROJECT



History and time have caught up with the original recordings.

JAZZ ORAL HISTORY PROJECT



History and time have caught up with the original recordings.

JAZZ ORAL HISTORY PROJECT

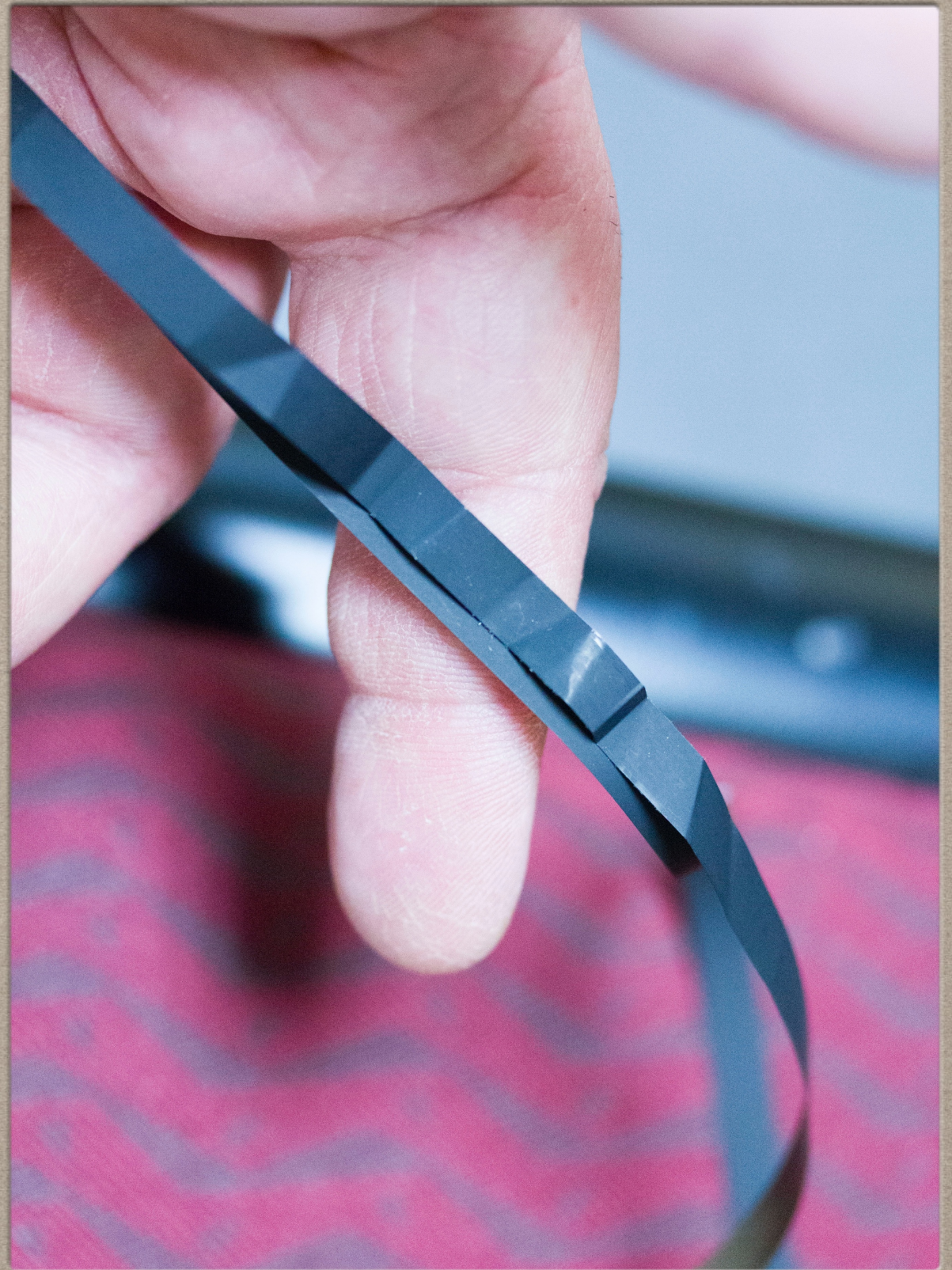


History and time have caught up with the original recordings.

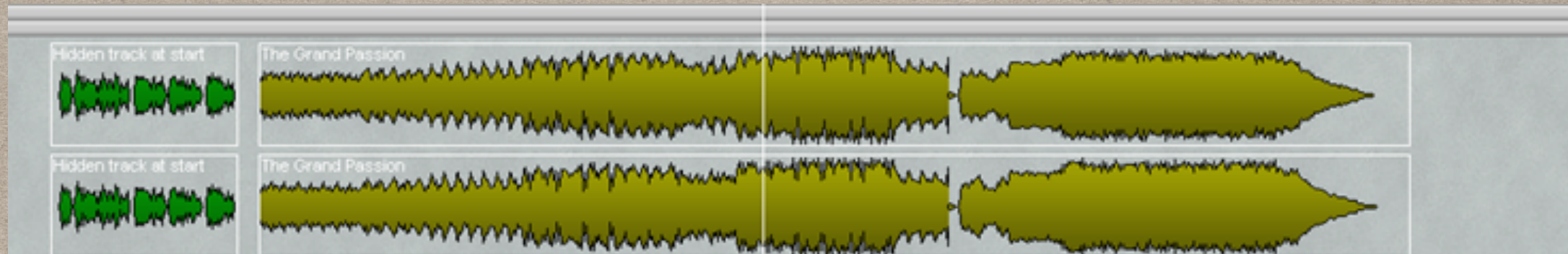
JAZZ ORAL HISTORY PROJECT

Sticky Shed Syndrome

"The condition of the original reel-to-reel and cassette tapes and some of the service copies had deteriorated to the point where the Institute could no longer offer access to large parts of the collection."

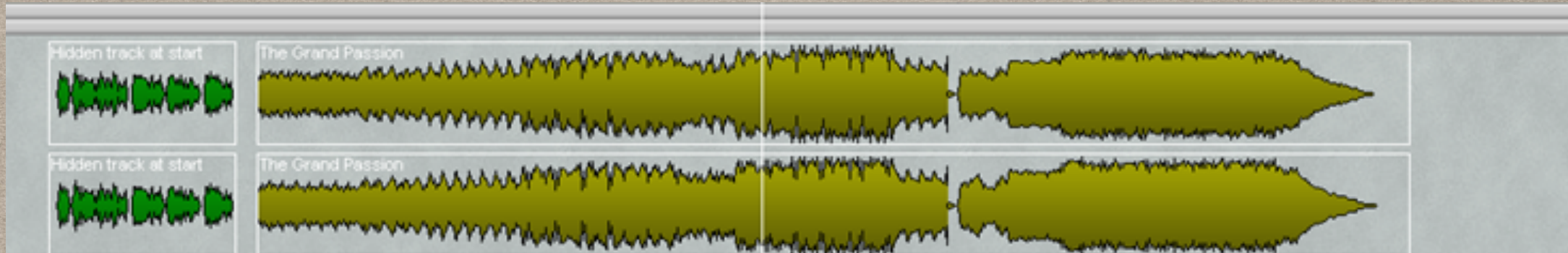


DIGITAL MASTERING



- Early 2000s, Rutgers applies for and receives a grant from the National Endowment for the Humanities to begin moving these recordings to digital format.

DIGITAL MASTERING



- Aims
 - Restore access to the public
 - Stabilize and retain the analog sources as best and as long as possible

THE FIRST PROBLEM: STICKY SHED MITIGATION

- Proposed solutions
 - Tape cleaning
 - Dehumidifying/dehydrating
 - Proprietary solutions (e.g. Rezerex)
 - Baking

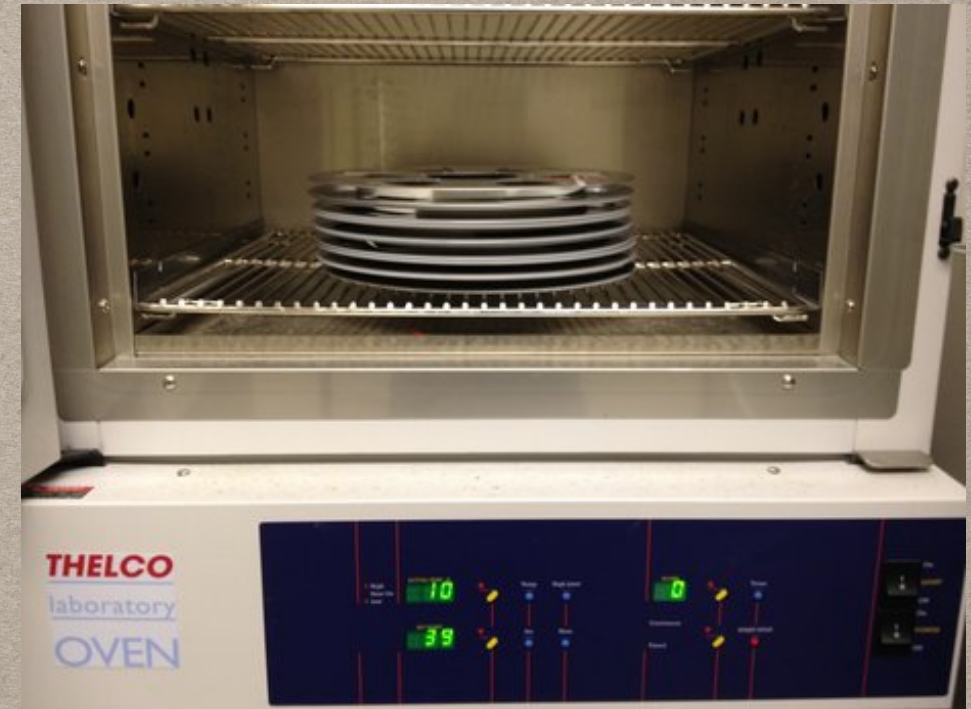


THE FIRST PROBLEM: STICKY SHED MITIGATION

Several remedial techniques enable playback of “sticky” tapes, but none stabilizes the material. Storage conditions may impact recurrence speed, but sticky tape damage is irreversible and reformatting should be given high priority. This does not negate arguments in favor of retaining damaged masters in stringently controlled low temperature and humidity storage, in anticipation of future improved reformatting options.

- U.S. Library of Congress

THE FIRST PROBLEM: STICKY SHED MITIGATION

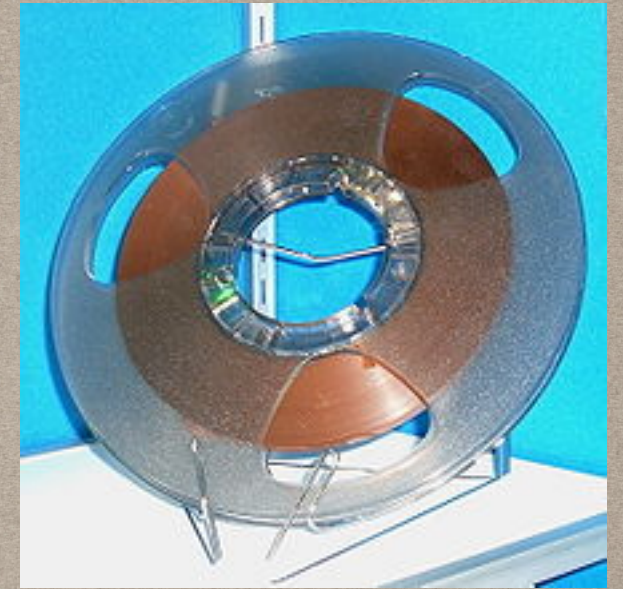


After much debate and examination by a committee of stakeholders in the project, the decision was made to “bake” the affected recordings, to enable digital mastering and retrieval of the recorded sound.

Photo credit: Lauren Sin/NPR

<http://www.npr.org/blogs/thisisnpr/2012/05/30/153917107/a-sticky-situation-baking-the-tapes>

STANDARDS AND BEST PRACTICES



A number of questions remained unresolved.

- * Which format is the right format?
- * What instructions do we give to the vendor for mastering:
 - * Sampling, bit rate, file formats?
- * Are we making the right choices?

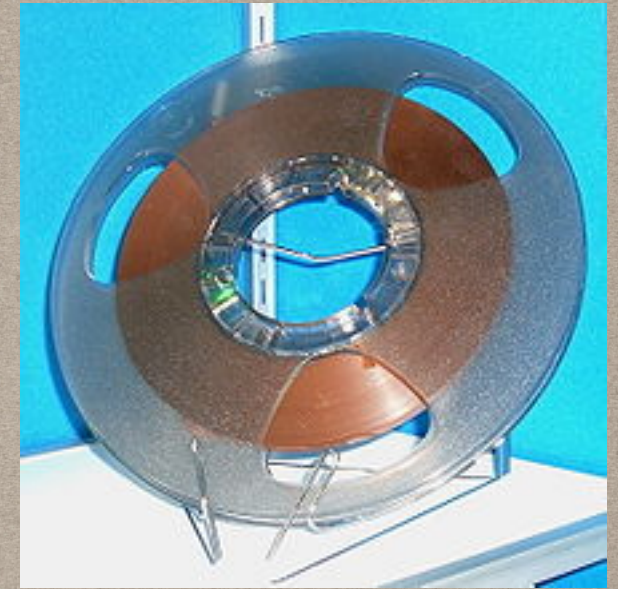
In the early 2000s, guidelines for archival audio mastering were still nascent, nothing formalized.

Photo credit: Arnold Reinhold

<http://en.wikipedia.org/wiki/File:Reel-to-ReelTapeBigHub.agr.jpg>

STANDARDS AND BEST PRACTICES

A committee for audio mastering and preservation was formed, and standards were developed.



Recommended Standards for NJDH and RU-CORE Audio Sampling

- **Minimum sampling rate: 44.1kHz 16-bit (CD Audio)**

This is the minimum acceptable rate to ensure a good preservation master. Most Compact Discs (CDs) are mastered at this rate. As such, all audio obtained from CDs will be archived at this rate.

Additionally, 44.1kHz is a suitable sampling rate for RU-Core partners when mastering recordings of spoken-word speech (i.e. interviews, speeches, press conferences and lectures), that are not accompanied by high-fidelity sound or music.

- **Recommended Sampling rate: 96kHz, 24-bit audio**

This is widely considered an ideal rate for high quality audio recordings, including DVD-Audio. For most audio formats, this sampling rate is the maximum sampling rate that also supports Quad (Dolby 4.0) and Surround (5.1) audio. When repository content partners are making a first generation sample of musical or high-fidelity recordings from an analog master, it is recommended that this sampling rate be used whenever technically possible.

- **High Level (Maximum) Sampling rate: 192kHz, 24-bit audio**

This sampling rate is often touted by audiophiles as one of the best sampling rates to work with in the editing of audio recordings and creating master samples. However, this format is generally not supported in current mass-produced formats for Quad or Surround sound. As such, recordings sampled at this rate should be limited to Mono or Stereo recordings. In general, this sampling rate, and higher rates, are recommended if there is a reasonable justification for using such a high sampling rate, and it is believed that the 96kHz rate will not be sufficient for accurate reproduction of the original sound.

DIGITAL MASTERING



Using NEH funds, the tapes were digitally mastered around 2004 by a 3rd party vendor and the digital audio stored on "Archival Grade" Gold CD-R media.

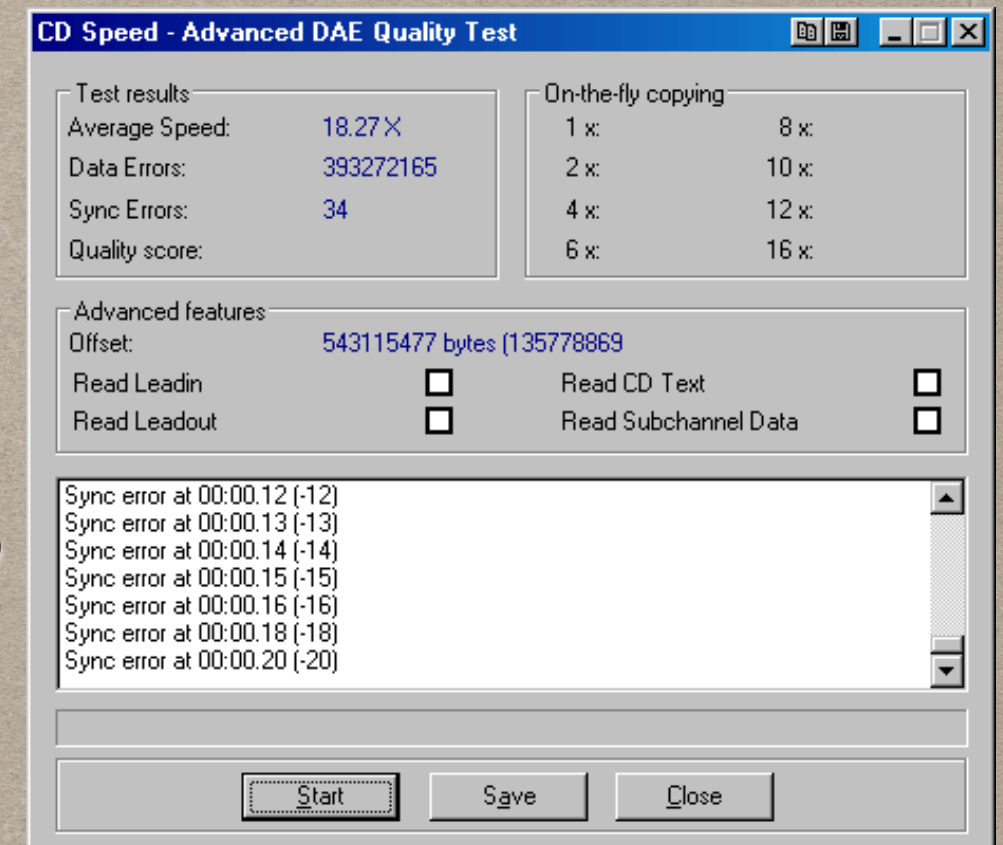
THE SECOND PROBLEM: CD-R NOT-SO-ARCHIVAL



CD/DVD experiential life expectancy is 2 to 5 years even though published life expectancies are often cited as 10 years, 25 years, or longer. However, a variety of factors discussed in the sources cited in FAQ 15, below, may result in a much shorter life span for CDs/DVDs. Life expectancies are statistically based; any specific medium may experience a critical failure before its life expectancy is reached.

THE SECOND PROBLEM: CD-R NOT-SO-ARCHIVAL

- New concerns
 - CD lifespan shorter than tape?
 - Proliferation of read errors (correctable, so far)
 - Future availability of CD players and readers



THE SECOND PROBLEM: CD-R NOT-SO-ARCHIVAL

- The CDs are not durable enough. Now what?
 - Digital files were made of all audio content in the collection.
 - Broadcast WAV, PCM audio 44.1kHz, 2 channels @ 1411.2Kbps
 - Stored on hard disks
 - Roughly 500GB total storage.

BWF files stored on a 1TB hard drive. Our "archival" copy.



Two copies on CDs. One "archival," one access copy.

End of story, everything's great. Right?



THE THIRD PROBLEM: HARD DISK FAILURES

*Hard drives have moving parts:
5,000-10,000 RPM spinning platters,
and read/write heads that fly over
these platters at widths of less than a
particle of smoke.*

*Eventually, they will fail, taking with
them the data stored on them.*



THE THIRD PROBLEM: HARD DISK FAILURES

The “bad batch” problem

A vendor released a line of low-cost, high capacity drives that had issues in their firmware that caused premature failure.

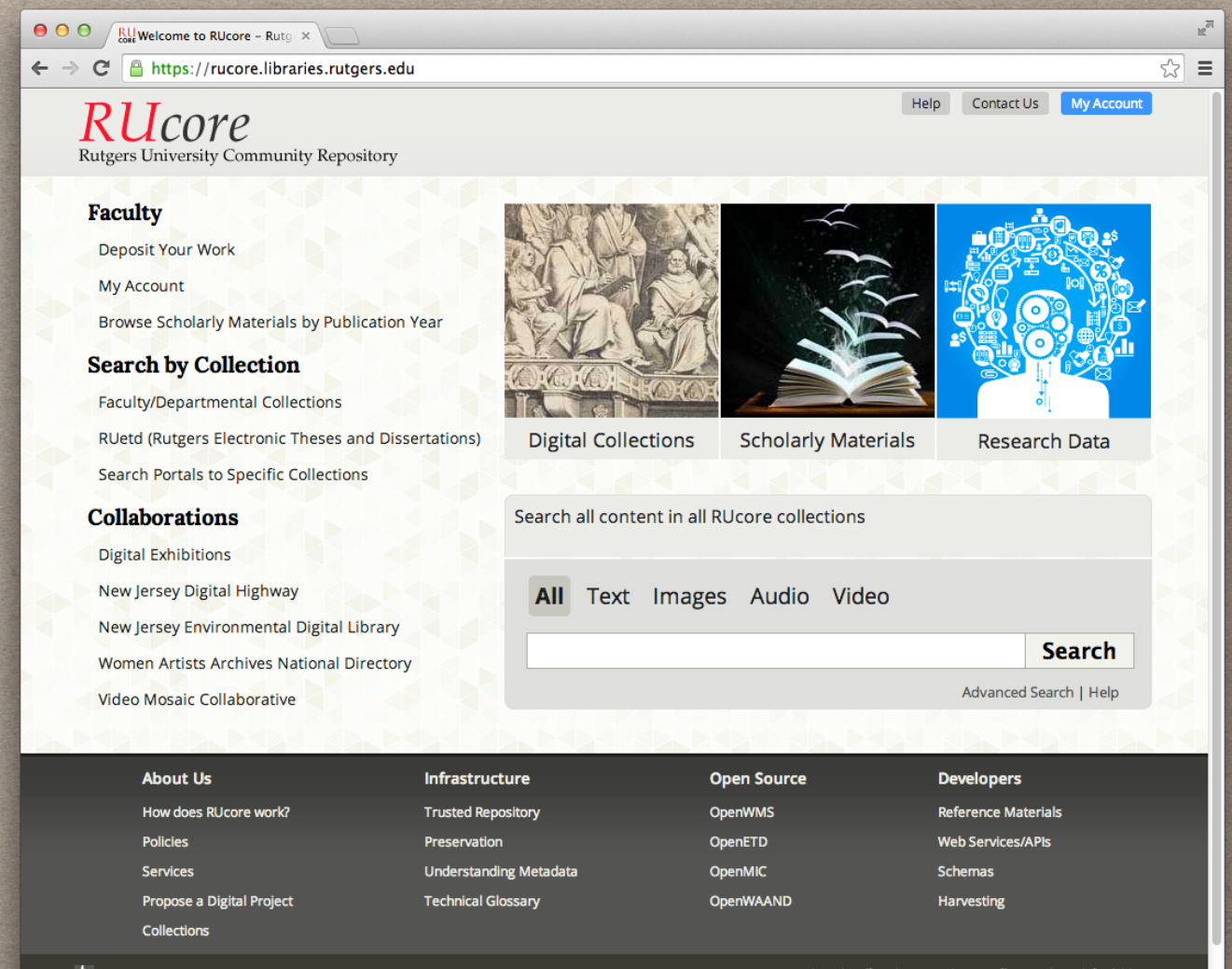
We unwittingly managed to buy a whole lot of the bad drives.



REPOSITORY-LEVEL PRESERVATION

A redundant, trustworthy repository solution was necessary to reliably store these valuable audio files, make adequate backups, and catalog and present these recordings to users.

Fortunately, one already existed.

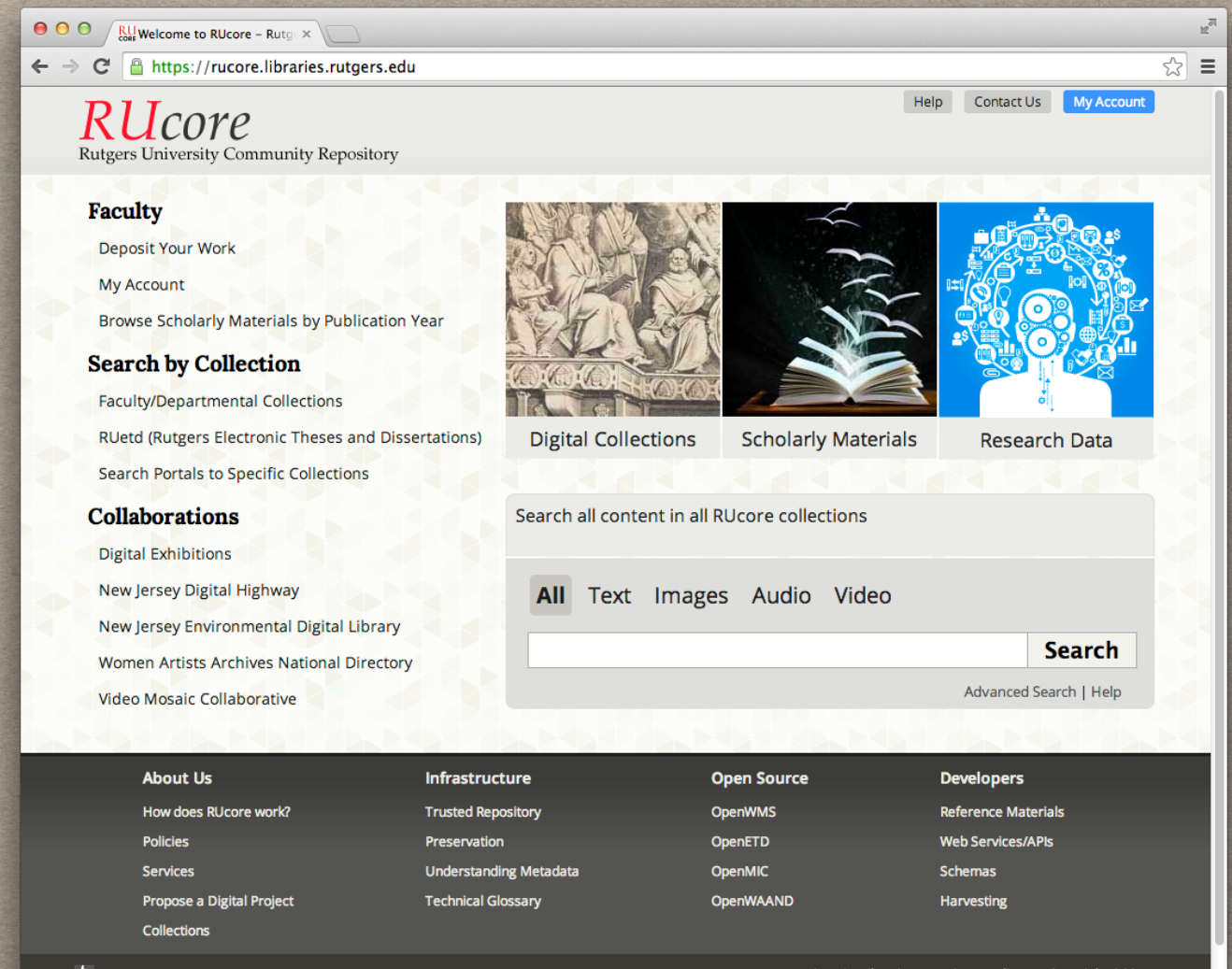


REPOSITORY-LEVEL PRESERVATION

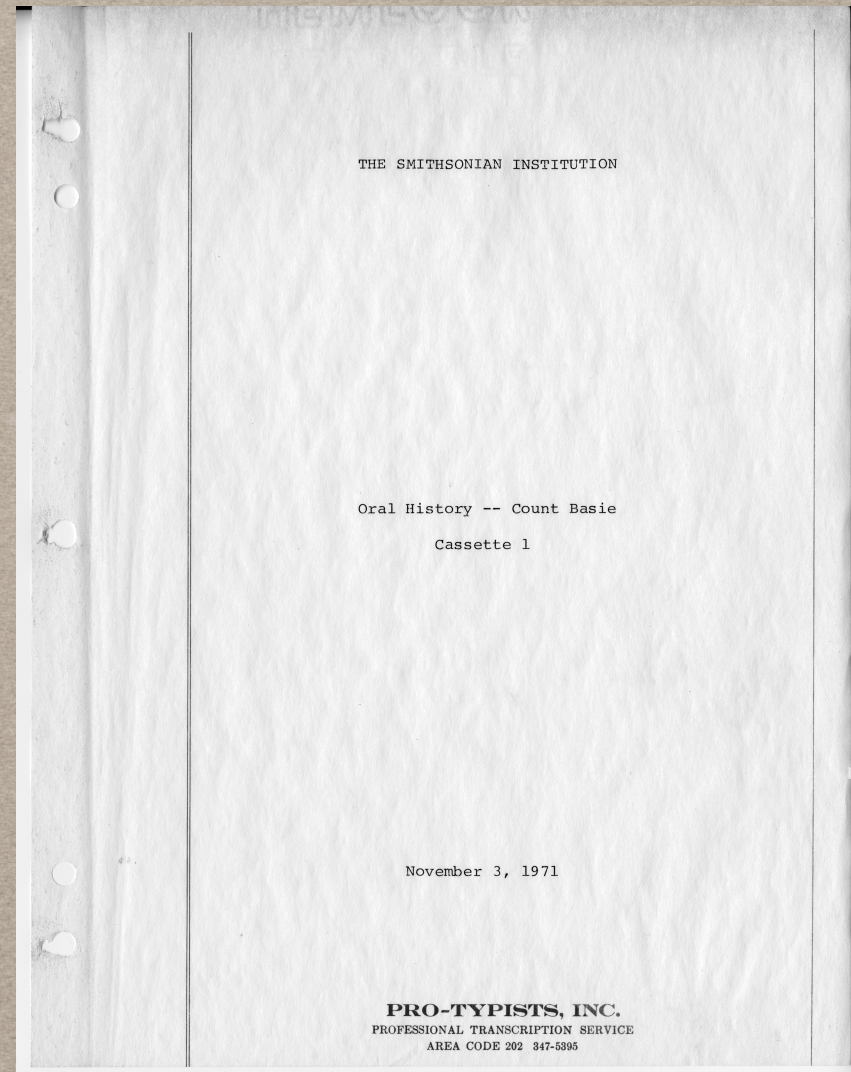
RUcore

The Rutgers Community Repository

- Large, redundant storage system plus backups (offsite, onsite)
- Digital signature checking of files to ensure integrity, alert us if something goes wrong.
- Online access: Electronic files, including sound recordings, can be accessed from the site.



REPOSITORY-LEVEL PRESERVATION



With RUcore, we could make available the audio recording AND the transcript of each interview, anywhere an internet connection existed.

REPOSITORY-LEVEL PRESERVATION

Sample Recording

Herman Autrey

(December 4, 1904 - June 14, 1980)

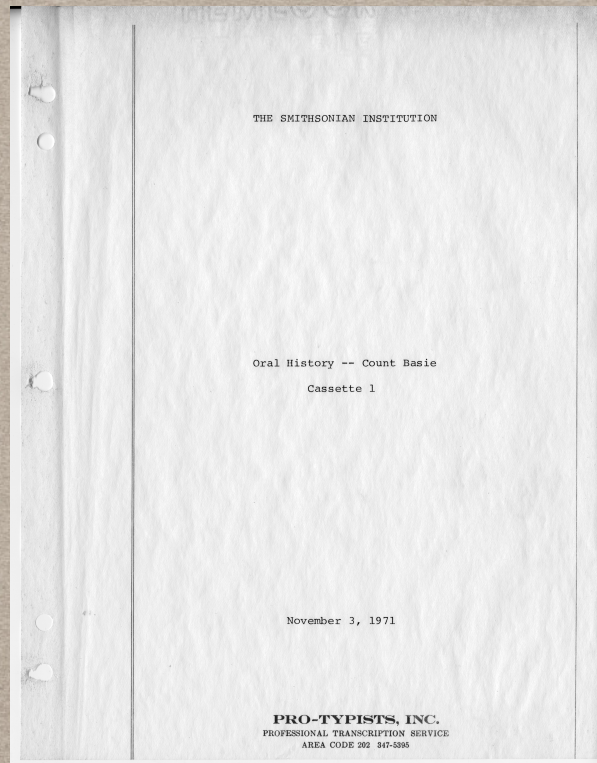
Jazz Trumpeter, and later vocalist,
active since the 1920s

Most closely associated with pianist,
vocalist, and bandleader Fats Waller

Interviewed April 23, 1975



THE FOURTH PROBLEM: PERMISSIONS AND COPYRIGHTS



The permissions and rights forms for these recordings were specifically designed with a limited, of-the-period use-case in mind.

- *One analog recording, in one place at one time*
- *No explicit permissions given for broadcasting or mass distribution*
- *None of the stakeholders had envisioned the internet, or anything like it!*

REQUIREMENTS FROM LEGAL

- A user viewing the content may not download any audio or transcript content to their computer.
- Users may not select, copy, cut or paste any audio or transcript content, and may not search any text.
- Users may not use our tools to remix, modify or create derivative works from the content
 - No playlists, annotations, reordering, etc.

WHERE WE STAND TODAY



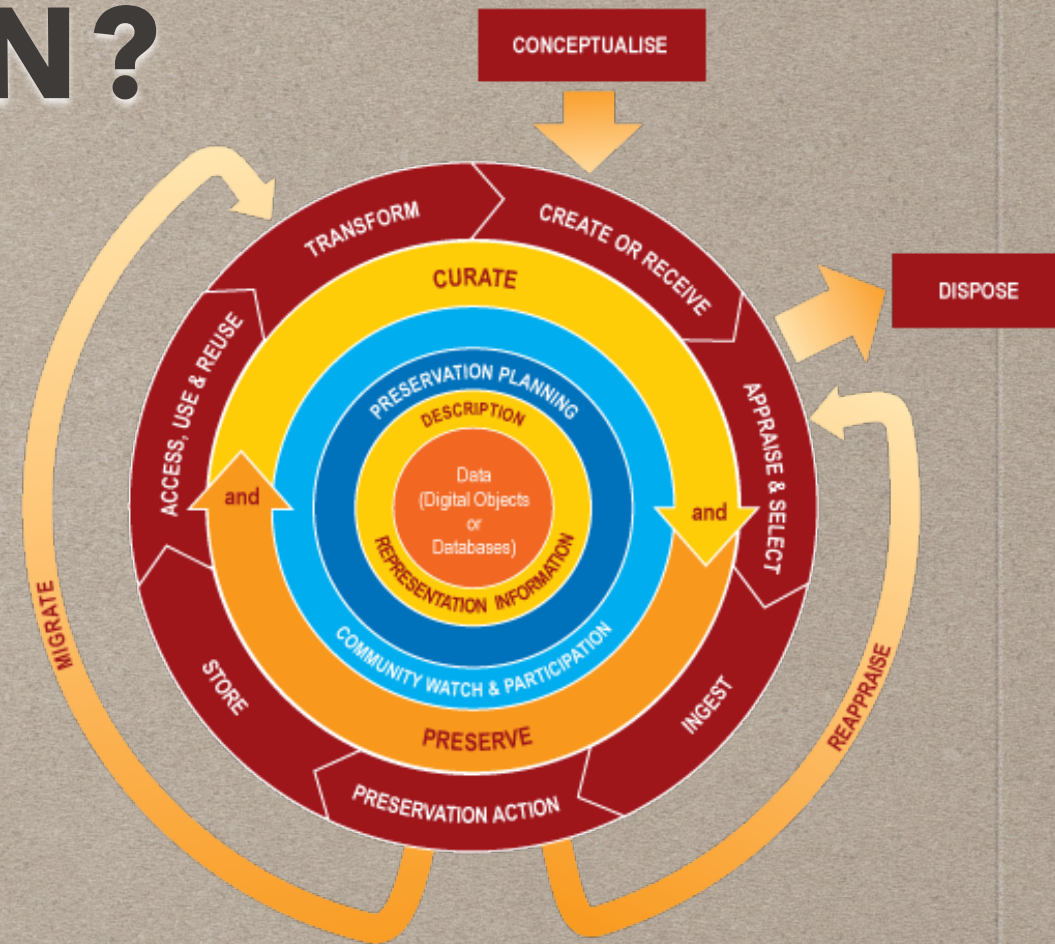
- Audio/transcripts not online yet.
- Stored on two redundant staging RAID storage systems, in a climate-controlled server room at the Digital Curation Lab in New Brunswick. Access restricted.
- Another digital HDD copy exists in Newark, also restricted.
- CD access copies and print copies still accessible to in-person IJS visitors.
- Tapes still in storage.

WHERE WE STAND TODAY



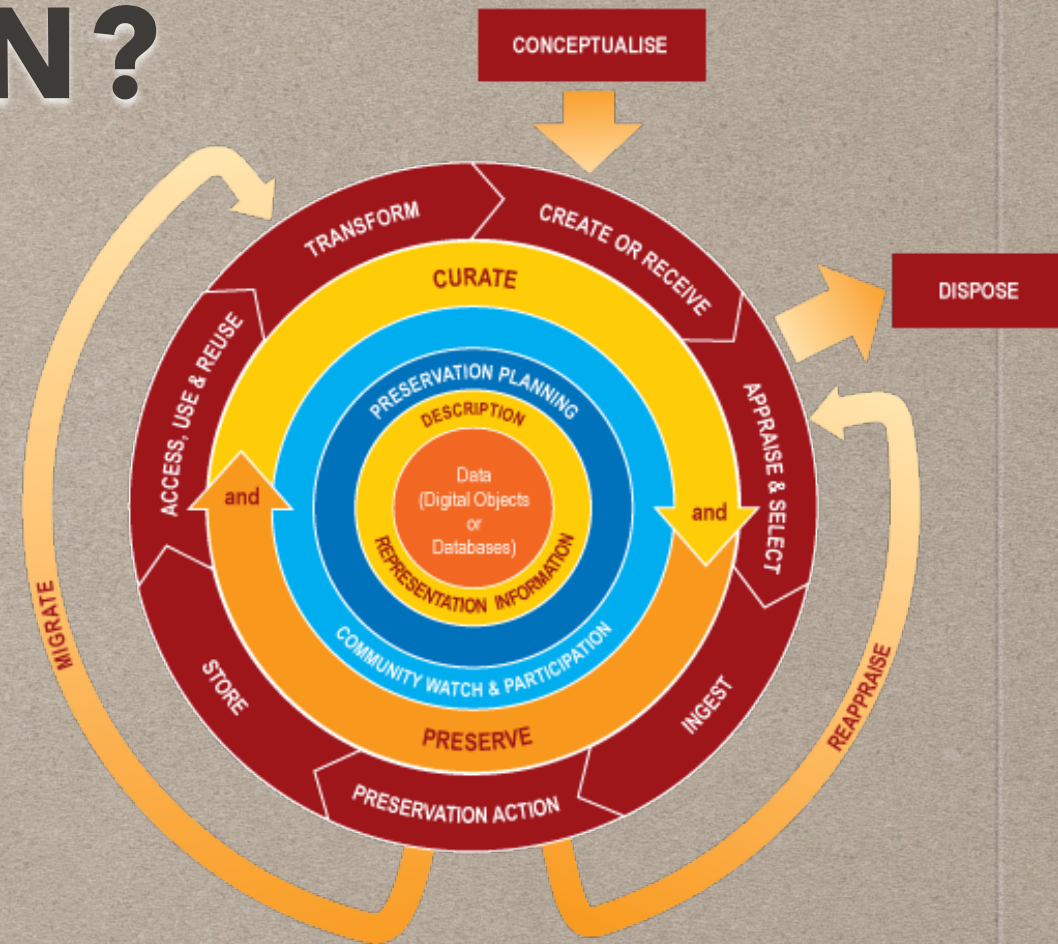
- The plan: software development is underway to enable all features requested by legal. Streaming-only, view-online-only copies of content will be the goal.
- Expected resolution: late 2014 or early 2015.
- In the meantime: Keep digital files verified and preserved on redundant storage.

WHAT DID WE LEARN?



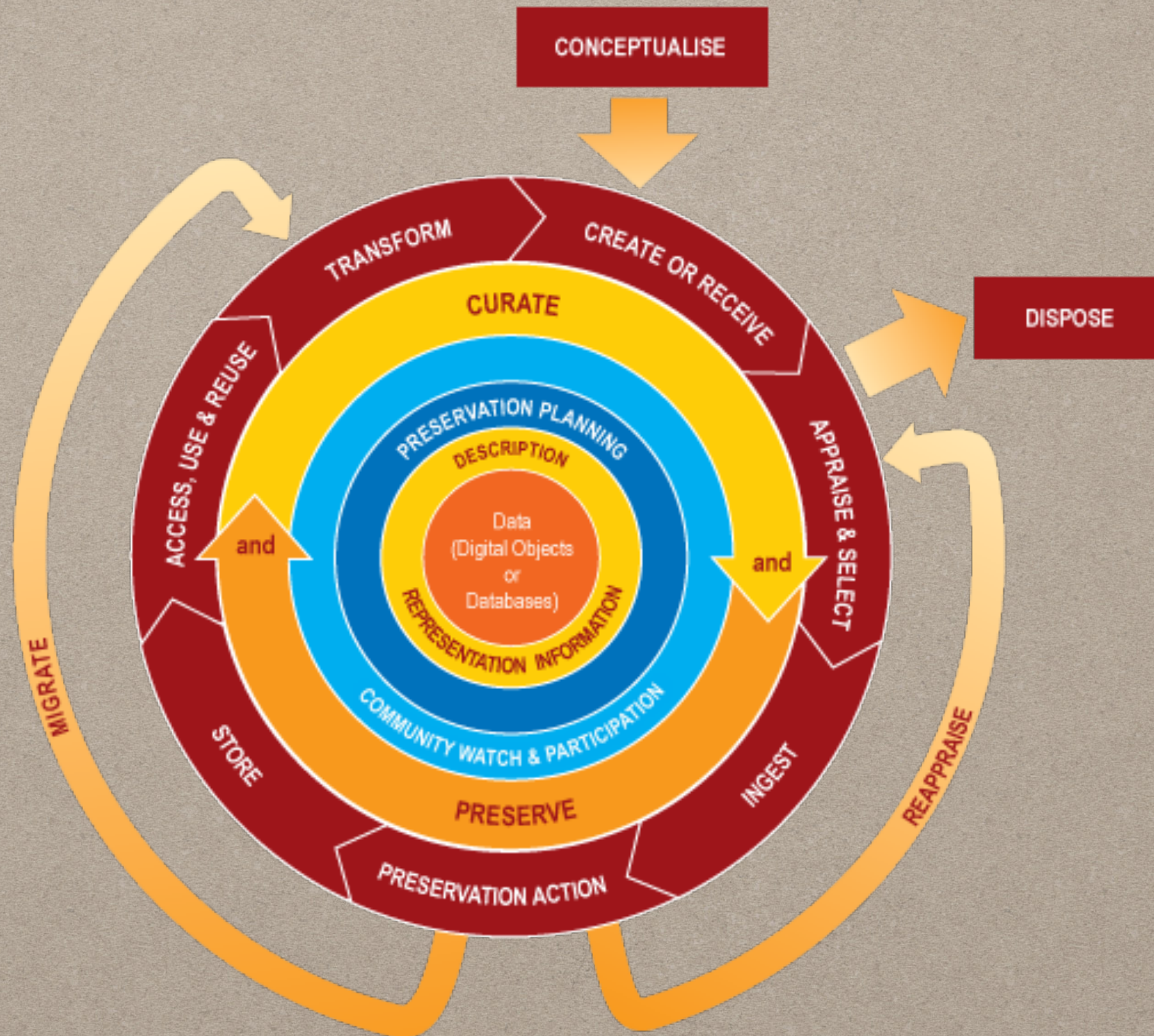
- There's more to digital mastering and preservation than just making a bunch of files and tossing them somewhere.

WHAT DID WE LEARN?



- A holistic methodology, discipline, approach to preserving digital content is necessary. This is the Digital Curation Lifecycle.

WHAT DID WE LEARN?



Source: <http://www.dcc.ac.uk/resources/curation-lifecycle-model>

WHAT DID WE LEARN?



Don't put all your eggs in one basket

Current storage/backup approach:

- At least 1 online *copy*
- At least 1 near line *copy*
- At least 1 dark/offline *copy* off-site

WHAT DID WE LEARN?



Don't operate in a vacuum

- Seek advice from other, more experienced digital archivists. Learn from their mistakes and successes. And share your own with others.

WHAT DID WE LEARN?

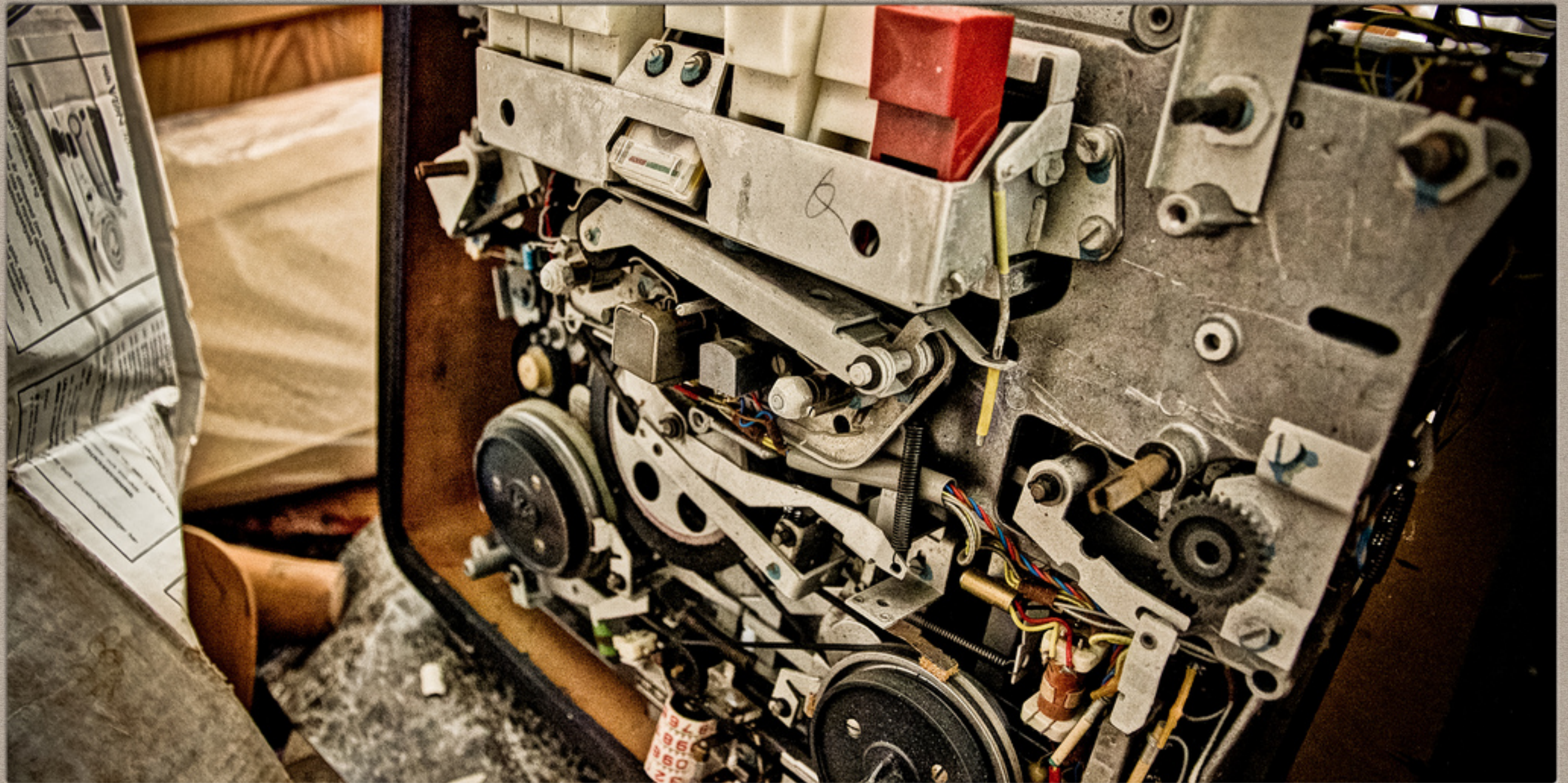
Diplomacy is a virtue, and a requirement for digital preservationists.

- There is extreme distrust prevalent in the “analog world” of recordings and collections, of digital technology and its applications.
- We’re not here to destroy the analog. We are here to provide an additional way to enhance access and *prolong the life of the original recording* for as long as we can in the process.



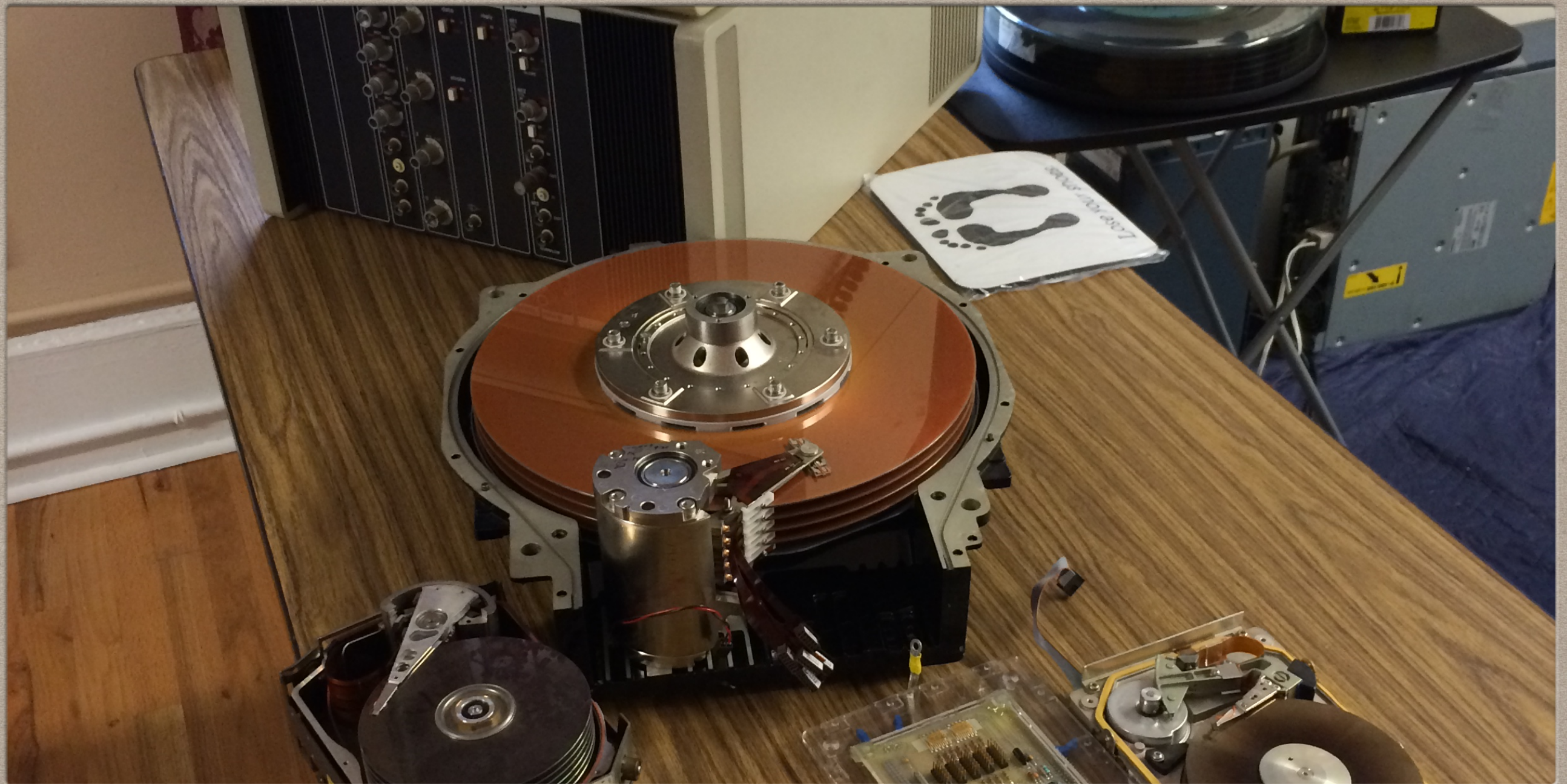
"We never go on vacation, because we're deathly afraid that *YOU GUYS* will descend on our offices like a band of ninjas and *digitize* everything!"

- *Anonymous colleague*



"Digitization is not preservation."

- Anonymous colleague



Digitization, for now, is our last, best hope of preserving the substance of a recording beyond the lifespan of the wax, or metal oxide, and the machines that play them.

Thank you for listening!

Isaiah Beard

Digital Data Curator

Rutgers University Libraries

New Brunswick, NJ

Association for Recorded Sound Collections
2014 Pre-Conference Workshop
Chapel Hill, NC



Digital Curation Blog: <http://page2pixel.rutgers.edu>

FROM REELS TO BYTES

LESSONS LEARNED IN MAINTAINING
RUTGERS' JAZZ ORAL HISTORIES