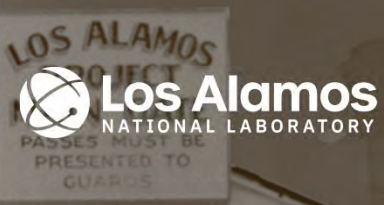


LOS ALAMOS
PROJECT
MAIN GATE
PASSES MUST BE
PRESENTED TO
GUARDS

NATIONAL SECURITY RESEARCH CENTER





Digitization Overview



Rizwan Ali
NSRC Director



Managed by Triad National Security, LLC for the U.S. Department of Energy's NNSA

About Us



- Born out of a need to make LANL Weapons Program archival material available to scientists and engineers
- Material is the result of **decades of consolidation** of mini-libraries and mini-archives at LANL
- Latest consolidation brought together LANL's **digital archives** and **physical archives**
- Houses **75+ years** of nuclear weapons research, designs, procedures, videos, photos, and other reports



The NSRC offers services similar to major university research libraries.

LOS ALAMOS
PROJECT
MAIN GATE

Powering Tomorrow's Innovations Today



Video has been removed to reduce file size.



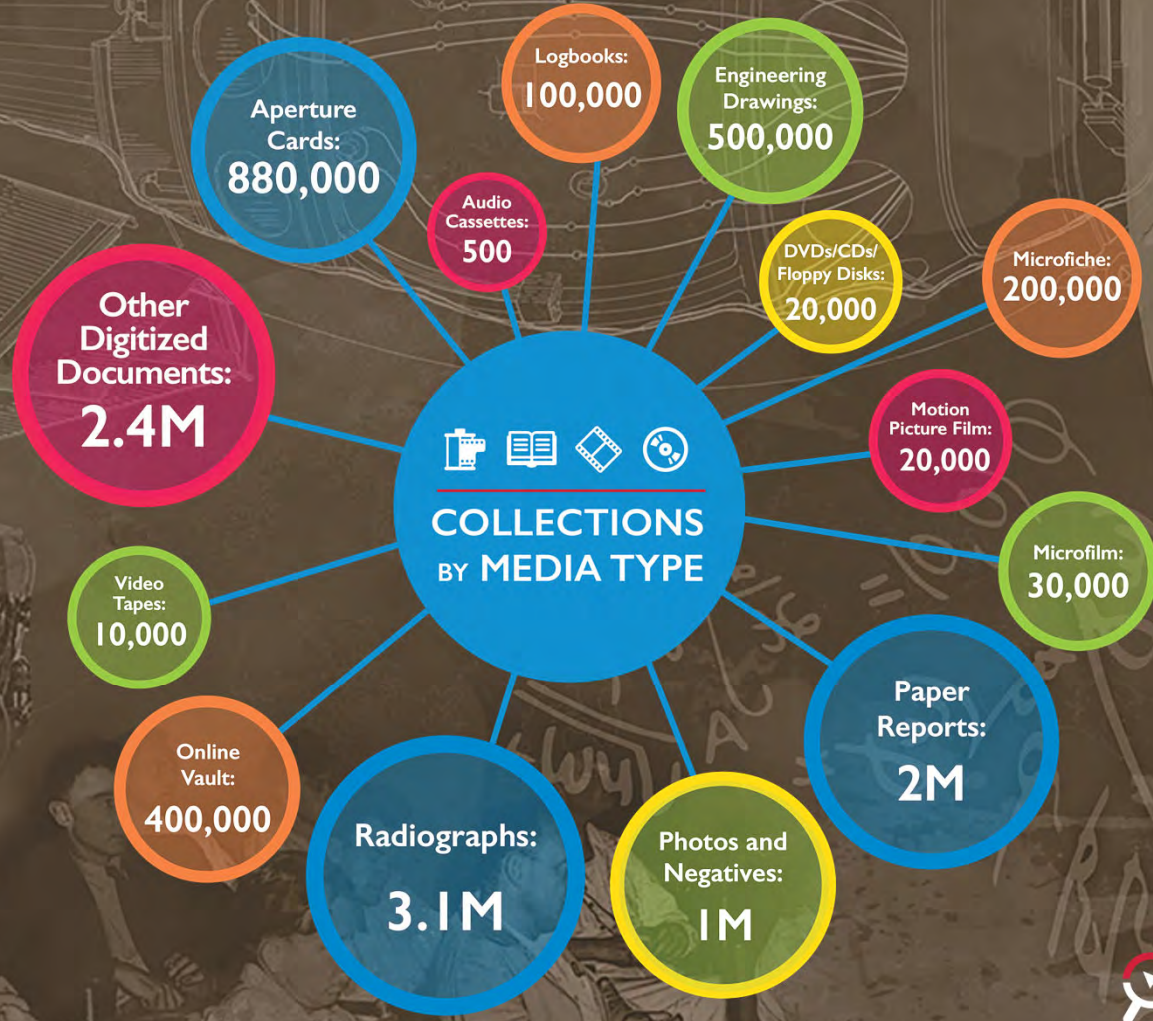
NSRC's Collections



- Not a traditional archive.
- NSRC's holdings are *active collections* used daily by researchers at LANL and throughout the NSE.
- Contains *millions* of unique and historically significant items documenting the U.S. nuclear weapons design, development, testing, certification, and assessment activities.

SAMPLE OF NSRC HOLDINGS
Nuclear weapons designs
Assembly drawings
Nuclear test diagnostic data
Pre-shot reports
Post-shot reports and analyses
Supporting and confirmatory experiments
Subcritical experiments
Stockpile stewardship experiments
Milestone reports
Classified research journals
US/UK collaboration documents
Internal memoranda
Working group notes
Histories
Code manuals
Test problems

The vast majority of the material is not duplicated anywhere else in the NSE.





- Less than 10% of the physical collection is digitized.
- Decades-long project to digitize the following media types:

---> *Paper*

---> *Video*

---> *Engineering drawings*

---> *Motion picture film*

---> *Microfilm*

---> *Audio*

---> *Microfiche*

---> *Notebooks*

---> *Photos*

---> *Negatives*

---> *Aperture cards*

---> *Card catalog*

The backlog in digitizing cannot be addressed with current processes.

Digitizing Efforts – Historical Rates



- ~1,300 videos per year
 - *Total collection: ~10,000 (6 videos per day)*
 - *~8 years to digitize*
- ~15 microfilm rolls per year
 - *Total collection: ~30,000 (1,500 – 2,000 pages per film roll)*
 - *~2,000 years to digitize*
- ~2,200 microfiche sheets per year
 - *Total collection: ~200K (10 – 20 pages per sheet)*
 - *~93 years to digitize*
- ~52 “banker boxes” of paper digitized per year
 - *Total collection: ~40,000 (~2,000 pages per box)*
 - *~760 years to digitize*

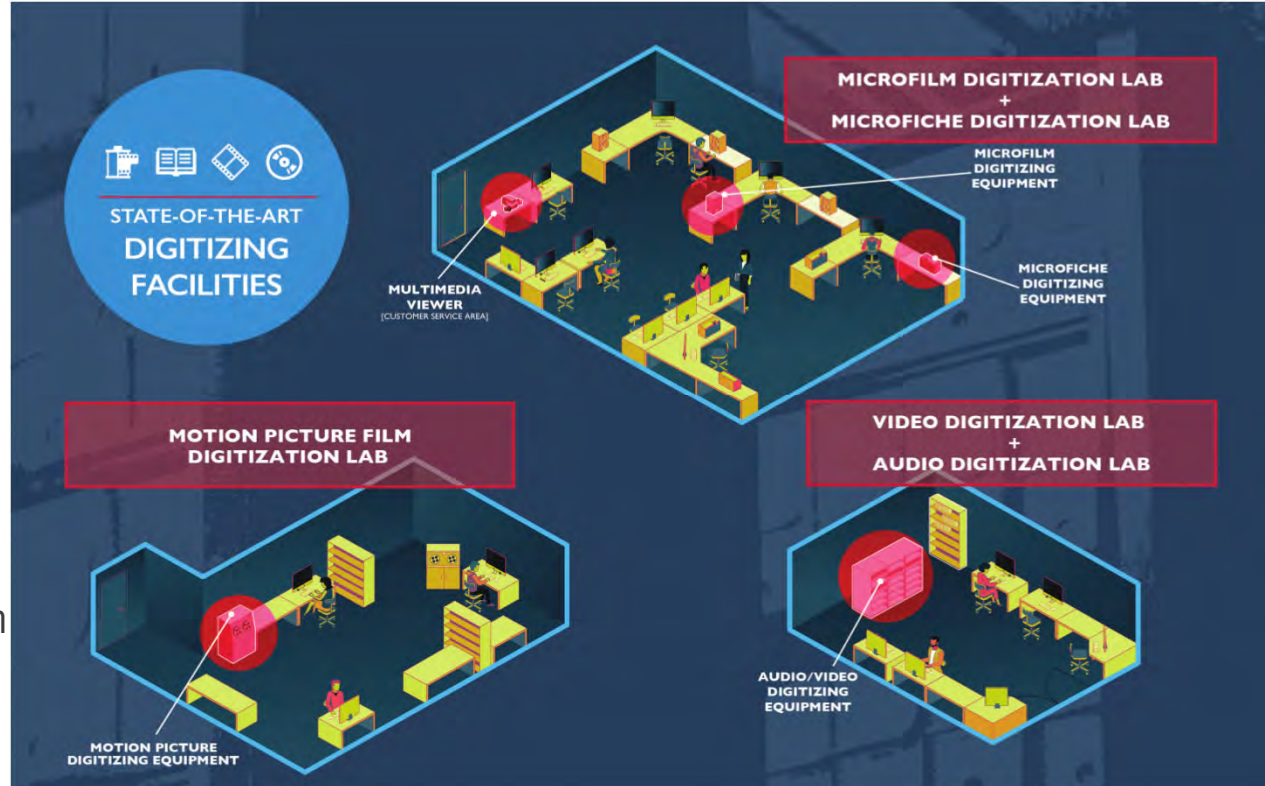
Digitization is limited by equipment, software, processes, training, and facilities.

Status of Digitization Services



A-1

- Seven new digitization labs in FY20 & FY21:
 - *Video Tape Digitization Lab*
 - *Audio Tape Digitization Lab*
 - *Microfilm Digitization Lab*
 - *Microfiche Digitization Lab*
 - *Motion Picture Film Digitization Lab*
 - *Rocky Flats Digitization Lab*
 - *Specialized Digitization Lab (paper, engineering drawings, photos, lab notebooks, aperture cards, etc.)*
- Established digitization equipment certification program
 - *All applicable staff to be certified on equipment and archival standards.*



Refocus Attention on Highest-Priority Media



- Video (VHS, Betamax, etc.)
 - Fastest *deteriorating media* in our collections
 - Rate of digitization may not address deteriorating media
 - Increased speed through number of digitization suites / FTEs
 - *Estimated completion: Reduced from 8 years to 5 years*

Video	Legacy	New
Quantity	10,000	
FTEs	1.0	1.5
Digitization suites	2	4
Rate per month	108	170
Years to complete	8	5

- Microfilm – Of immediate relevance to Weapons Program
 - New production-scale equipment and software
 - Revamped processes
 - Increase of personnel from 0.1 to 2
 - *Estimated completion: Reduced from 2,000 years to 16 years*

Microfilm	Legacy	New
Quantity	30,000	
FTEs	0.1	2
Digitization suites	1	3
Rate per month	1.3	160
Years to complete	2,000	16

- Microfiche – Of immediate relevance to Weapons Program
 - New production-scale equipment and software
 - Revamped processes
 - Increased personnel from 1 to 3
 - *Estimated completion: Reduced from 93 years to 19 years*

Microfiche	Legacy	New
Quantity	200,000	
FTEs	1	3
Digitization suites	1	4
Rate per month	180	900
Years to complete	93	19

Cataloging / Indexing Efforts



- Digital Collections

- Wide variety of metadata and indices ranging from nonexistence to rudimentary to robust
- *Search functions range from nonexistent to difficult-to-use*

- Physical Collections

- *Poor to nonexistent metadata*
- *Limited indices*
- *Classified Reports Collection has an old-fashioned card catalog*

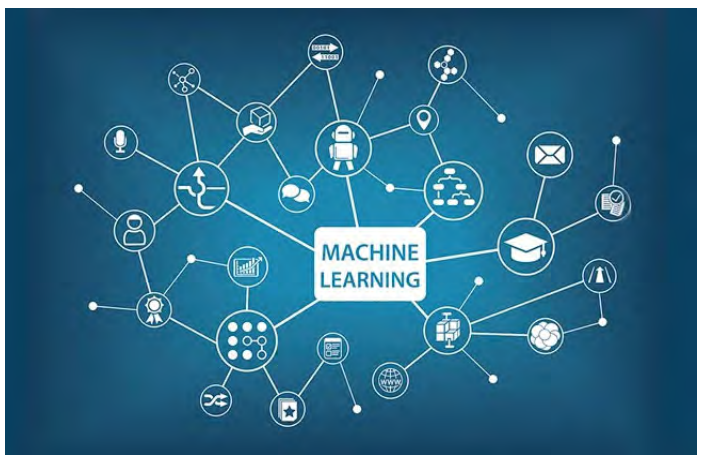
- Cataloging rate

- *Currently using **fully manual** process*
- *Process takes 10-30 minutes per document*
- *Total number of **digitized files** growing rapidly*

Cataloging snapshot of one digitized collection	
Quantity	2.4 million
FTEs	1.5
Rate per month	486
Years to complete	412

The backlog in cataloging/indexing cannot be met with a manual approach.

Titan on the Red Project



- **System must be able to:**
 - *Automatically extract metadata from digitized documents*
 - *Perform natural-language search*
 - *Enforce security classification and NTK protocols*
 - *Be expandable to ingest various data stores (Online Vault, PDMLink, shared drives, SharePoint, etc.)*
 - *Utilize commercial, public domain, and LANL ontologies*

- **Core Data Platform:**
 - *Company found through contacts in the US Intelligence Community*
 - *Company worked for several of the IC's three-letter agencies*
 - *Successfully funded (\$800K) and completed a 6-month demo on the unclassified network in FY20 and passed the Red Team evaluation*
 - *Brought capability on-contract for classified network in FY21*

- **FY21 efforts funded through combination for funds from NA-115 and LANL**

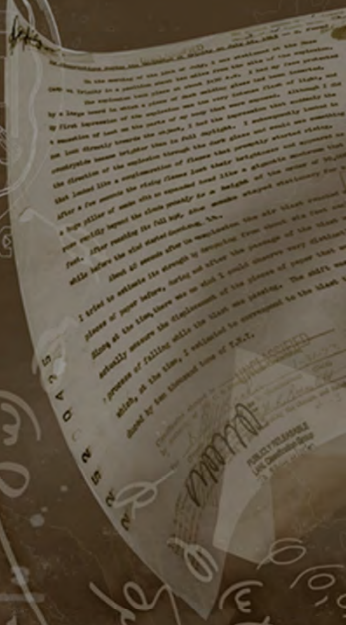
LOS ALAMOS
PROJECT
MAIN GATE
PASSES MUST BE
PRESENTED TO
GUARDS



QUESTIONS?



A-1



LOS ALAMOS
PROJECT
MAIN GATE
PASSES MUST BE
PRESENTED TO
GUARDS

NATIONAL SECURITY RESEARCH CENTER

