



Digitization Deconstructed:

A Survey and Census
of Digital Repositories
Being Built by Federal Libraries

June 9, 2010



Who Are We

- Federal Library and Information Center Committee
 - Advocacy for Federal libraries
 - Training opportunities
- Working group brings together librarians interested in preservation and digitization issues



June 9, 2010



June 9, 2010



Purpose

- Collect information about digitization efforts going on in Federal libraries
 - What is being digitized
 - How are libraries going about it
- Facilitate sharing of information collected
 - Share procedures and best practices
 - Identify barriers and strategies to overcome them
- Assess training needs of library personnel about digitization



Goals

We want...

- 100% response rate
- A single, authoritative response from each library

This requires us to...

- Identify all Federal libraries
- Identify a contact within each



What We're Asking

- Identification
- What
- How (technology)
- How (workflows)
- How (human & financial resources)
- Why
- Challenges
- How can FLICC help?



Process



- Develop family tree
- Survey, Instructions, and Definitions
- Answer questions
- Follow-up with those who don't reply
- Repeat as necessary



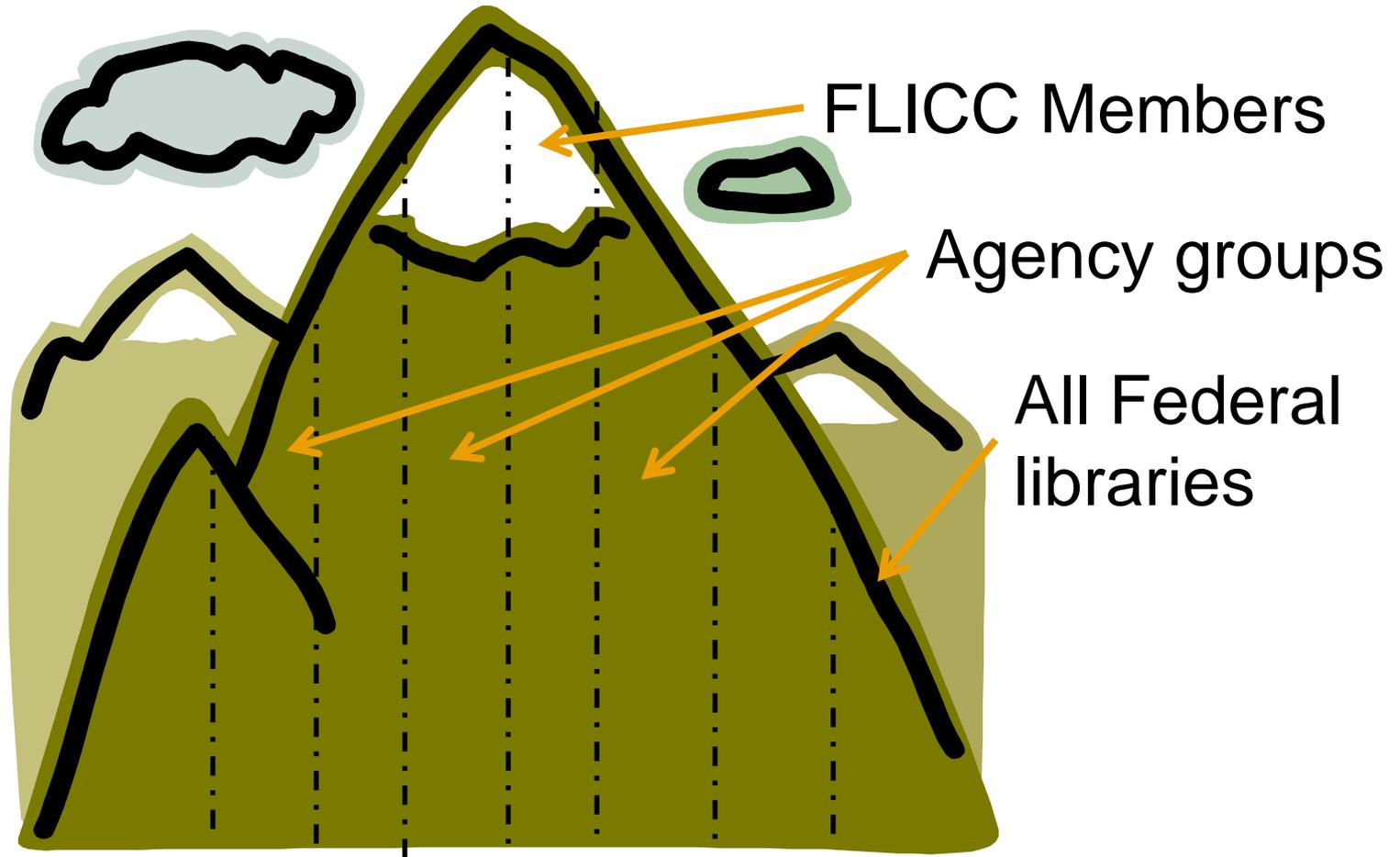
Survey Iteratively

Survey Federal library community in parts

1. Survey cross-sectional sample of federal library community
2. Identify discrete communities
3. Map community and identify individual libraries
4. Send survey
5. Repeat steps 2-4 until all federal libraries have been surveyed



Until You've Surveyed the Whole Mountain



June 9, 2010



Pilot Study FLICC Members

- Permanent members represent all cabinet-level agencies
 - Representative often comes from largest library or program office
 - Large agencies may have multiple representatives
 - Tends to be Washington DC centric
- Rotating members represent library community at large
 - Smaller libraries
 - Interests outside of DC

**Not a random
sample**



Pilot Results

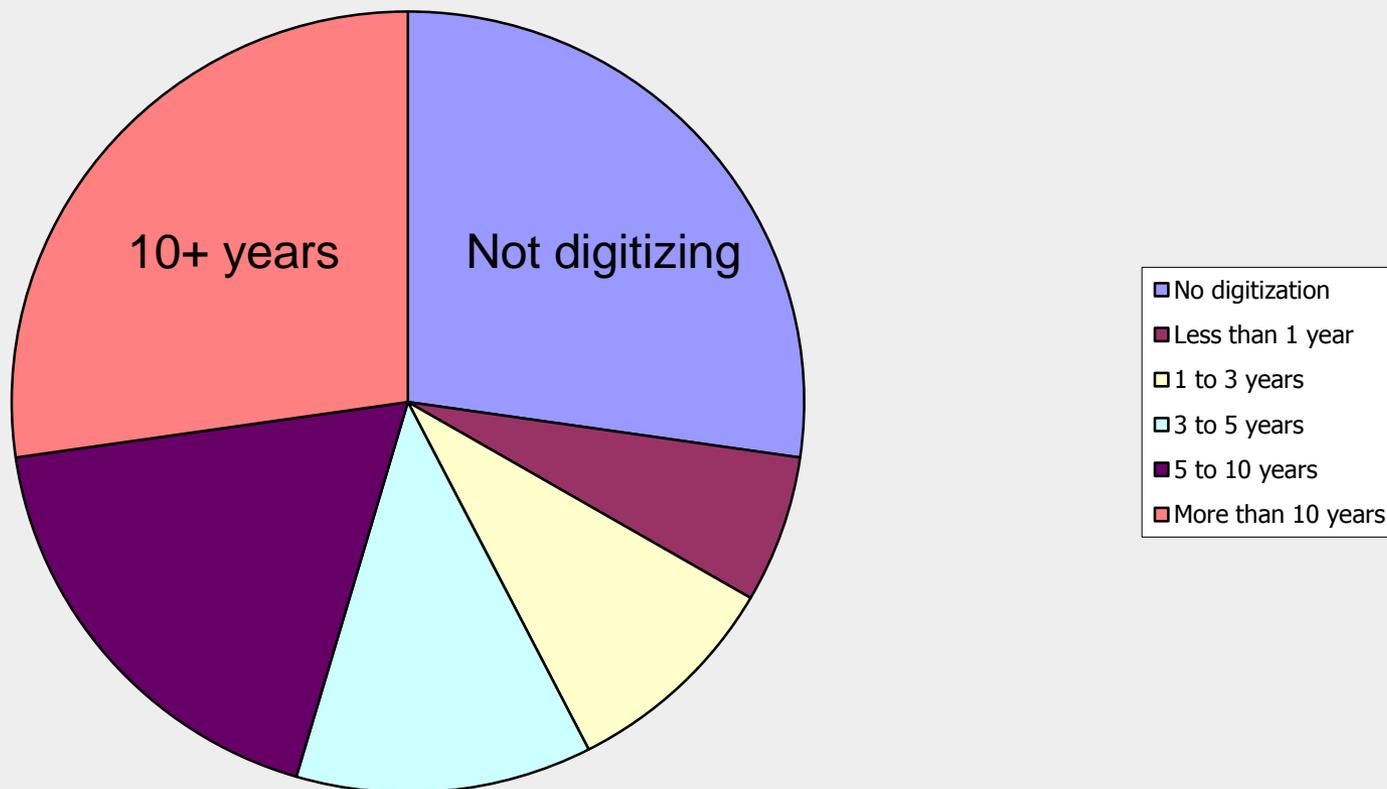
- 35/ 52 responded (67%)
- Wide range of subject materials being digitized
- 18/35 (51%) capturing born-digital materials
- Cross-section of formats being digitized
 - Archival materials
 - Technical reports
 - Monographs
 - Materials specific to their agency's work



Pilot Results

How long?

How long has the library/information center been digitizing? (Select the best option.)

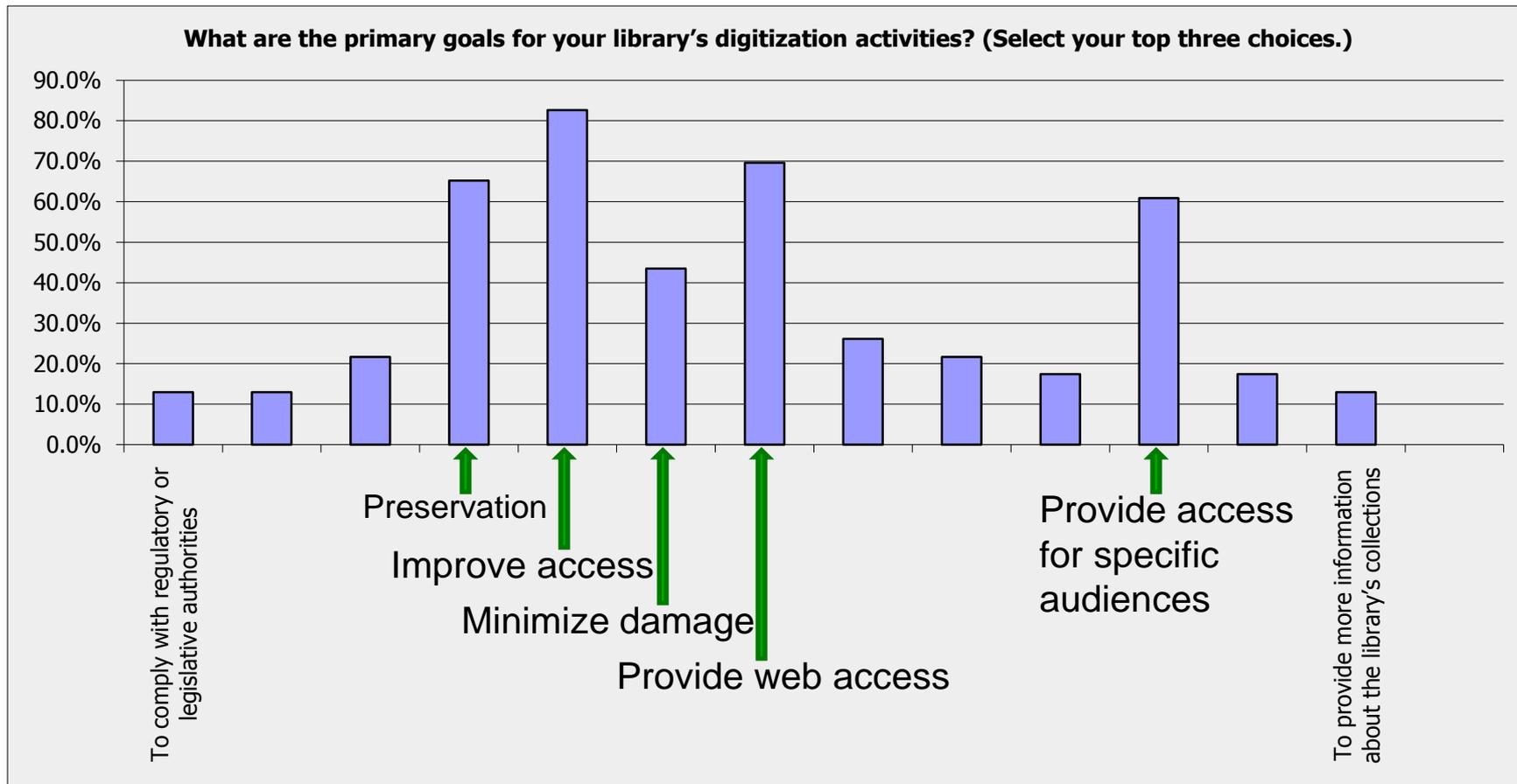


June 9, 2010



Pilot Results

Why?

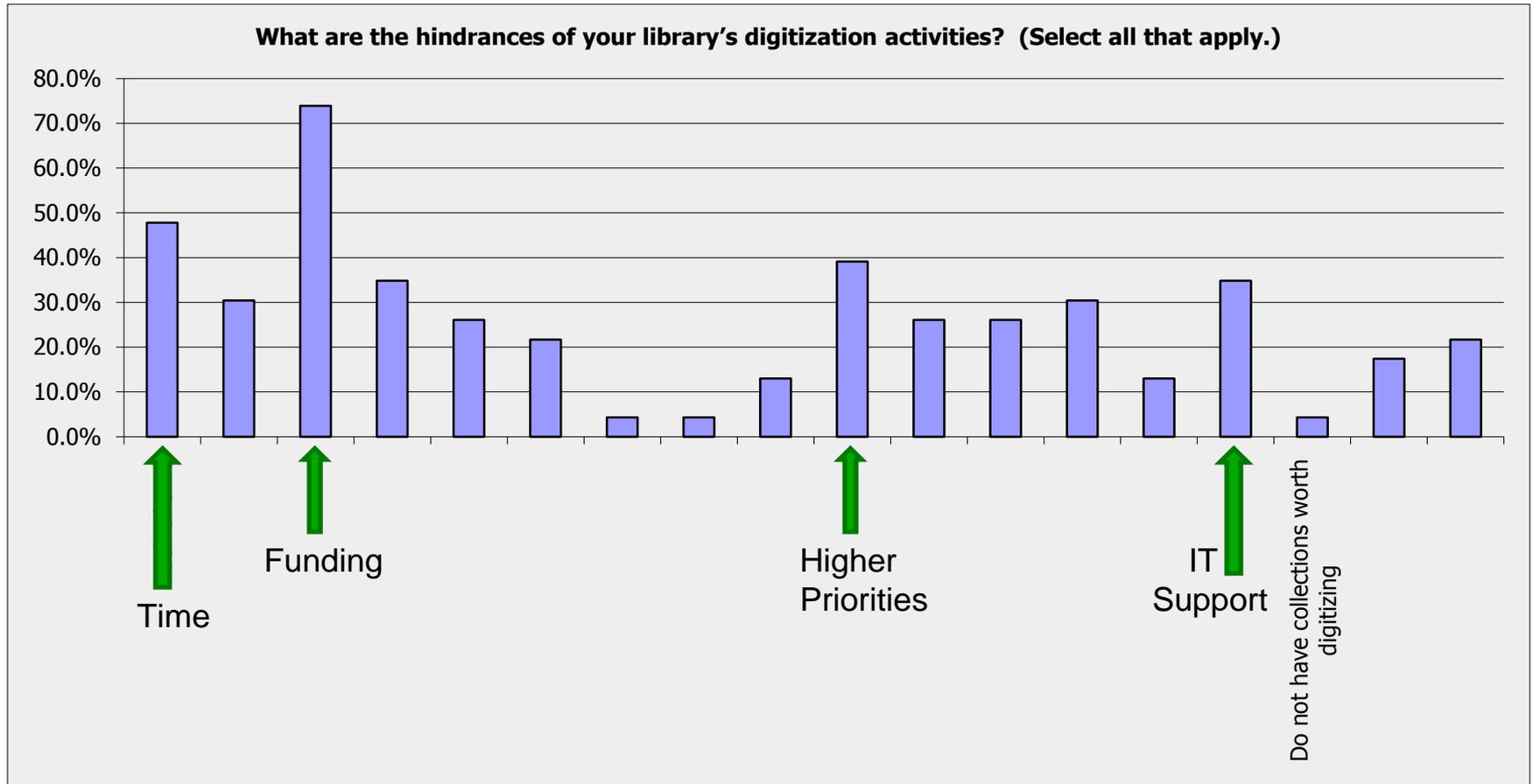


June 9, 2010



Sample Results

Challenges?



June 9, 2010



First Agency Groups U.S. Army and USDA

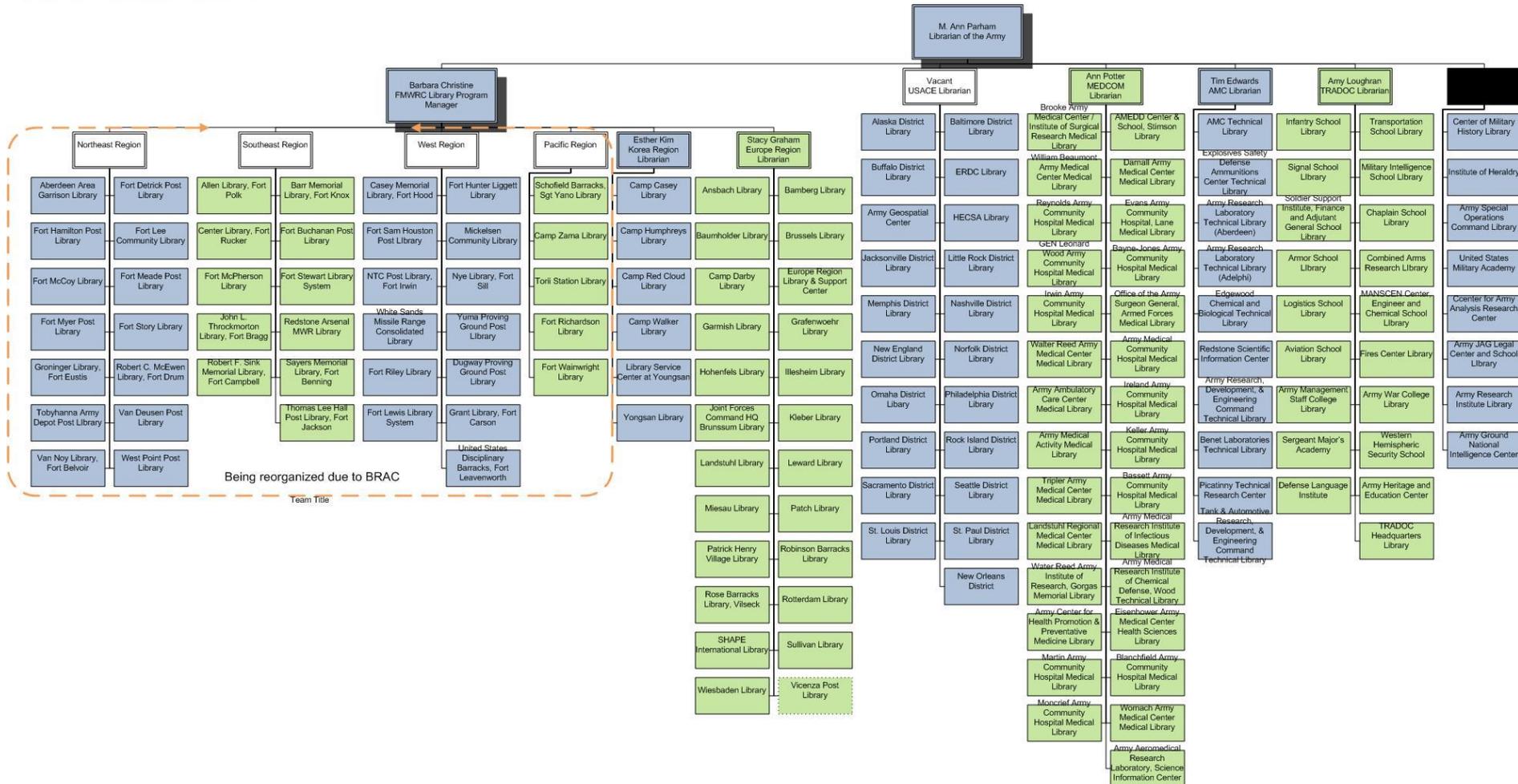
Why these agencies?

- Committee knowledge of agency library program structure
- Managable size group
- Volunteered



Not all Families Look Alike

U.S. Army

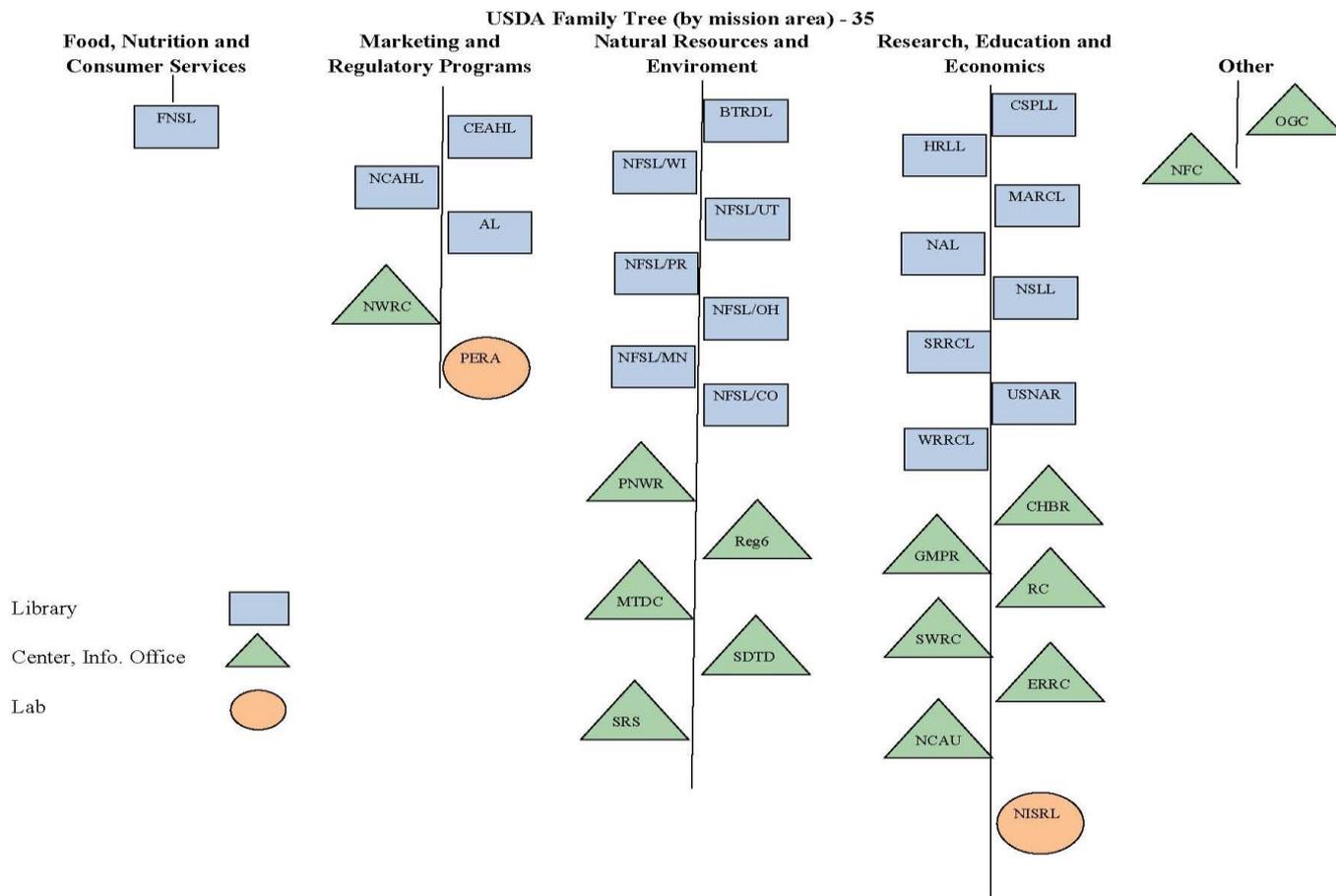


June 9, 2010



Not all Families Look Alike

USDA





Army/USDA Trends

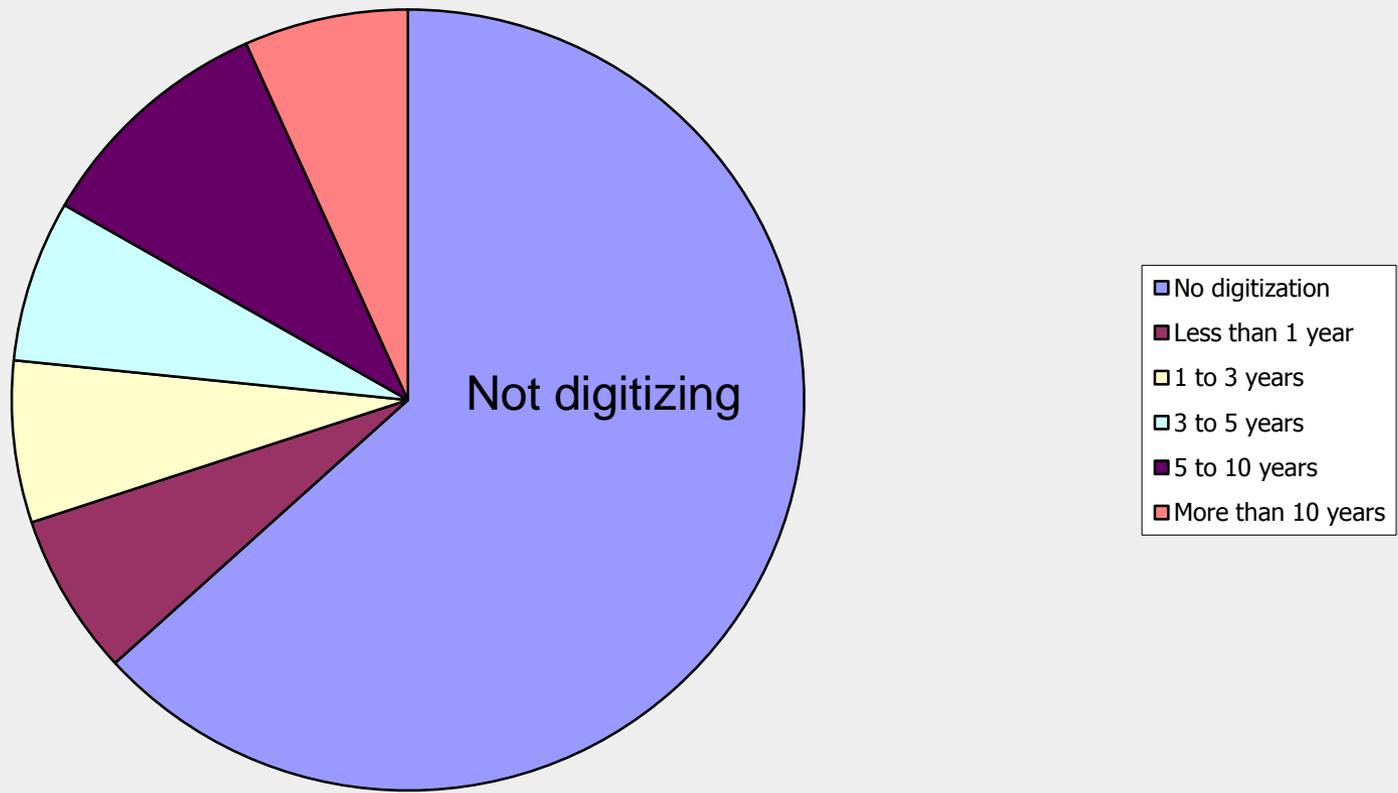
- 30/132 responded (23%)
- Collection strengths mirror agency mission
- 8/30 (26%) capturing born digital materials
- Cross-section of formats being digitized
 - Archival materials
 - Journal articles
 - Photographs
 - Rare books
 - Technical reports
 - Manuscripts
 - Maps



Army/USDA Trends

How long?

How long has the library/information center been digitizing? (Select the best option.)

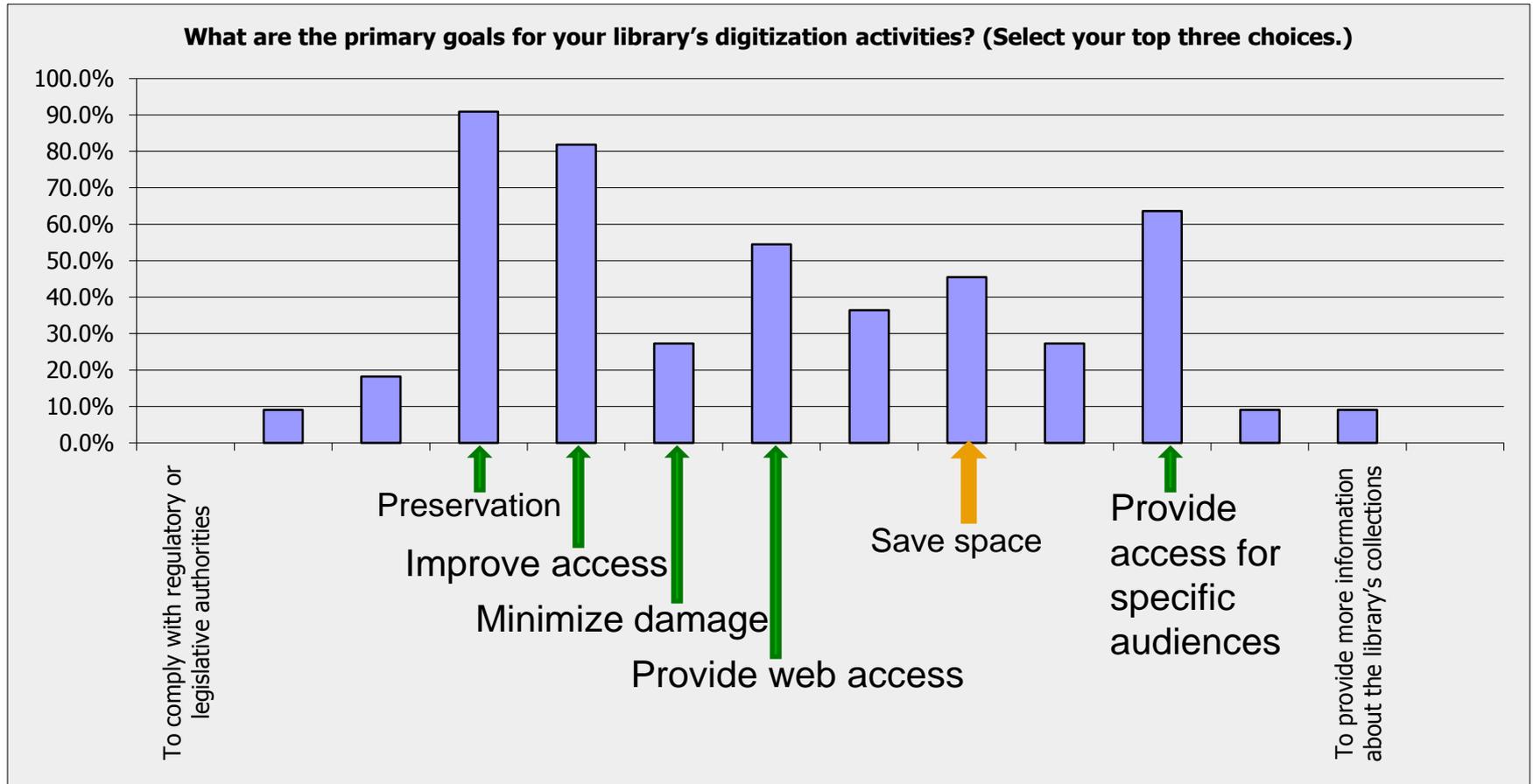


June 9, 2010



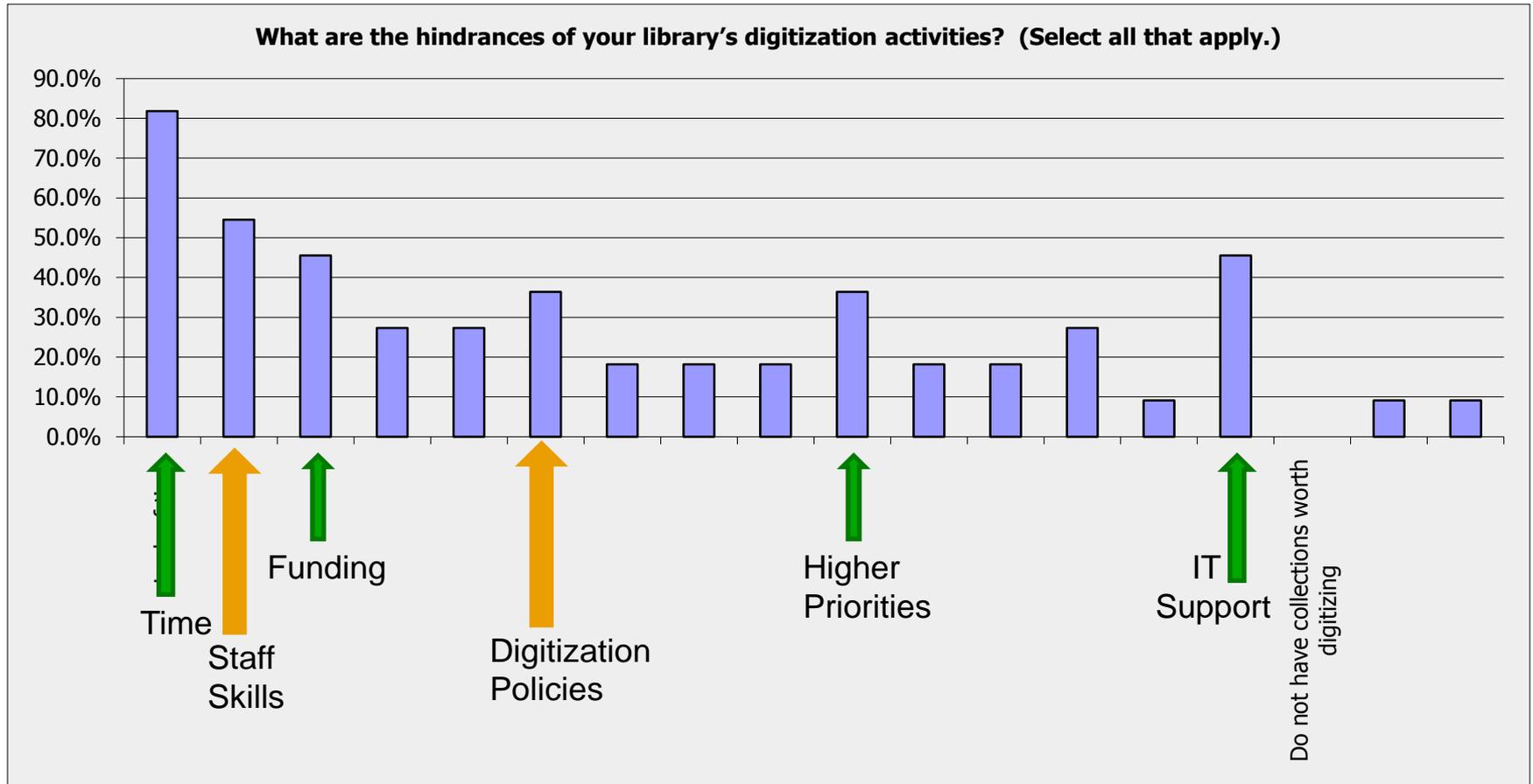
Army/USDA Trends

Why?





Army/USDA Trends Challenges?



June 9, 2010



Agency Similarity

U.S. Army

- 37/97 responded (38%)
- 57% serve customers outside the agency
- 57% not digitizing
- 56% capture born-digital
- 12% digitizing classified
- Digitization is both regular library operation and special project



June 9, 2010

USDA

- 23/34 responded (64%)
- 78% serve agency only
- 60% not digitizing
- 37% capture born-digital
- 25% digitizing classified
- Digitization is both regular library operation and special project





Can We Generalize?

Big Libraries

- Many have annual budgets of \$1M annually
- Spend 1-5% of budget on digitization activities
- Biggest challenge is **funding**



Small Libraries

- Many have annual budgets below \$100K annually
- Spend 10-20% of budget on digitization activities
- Biggest challenge is **time**



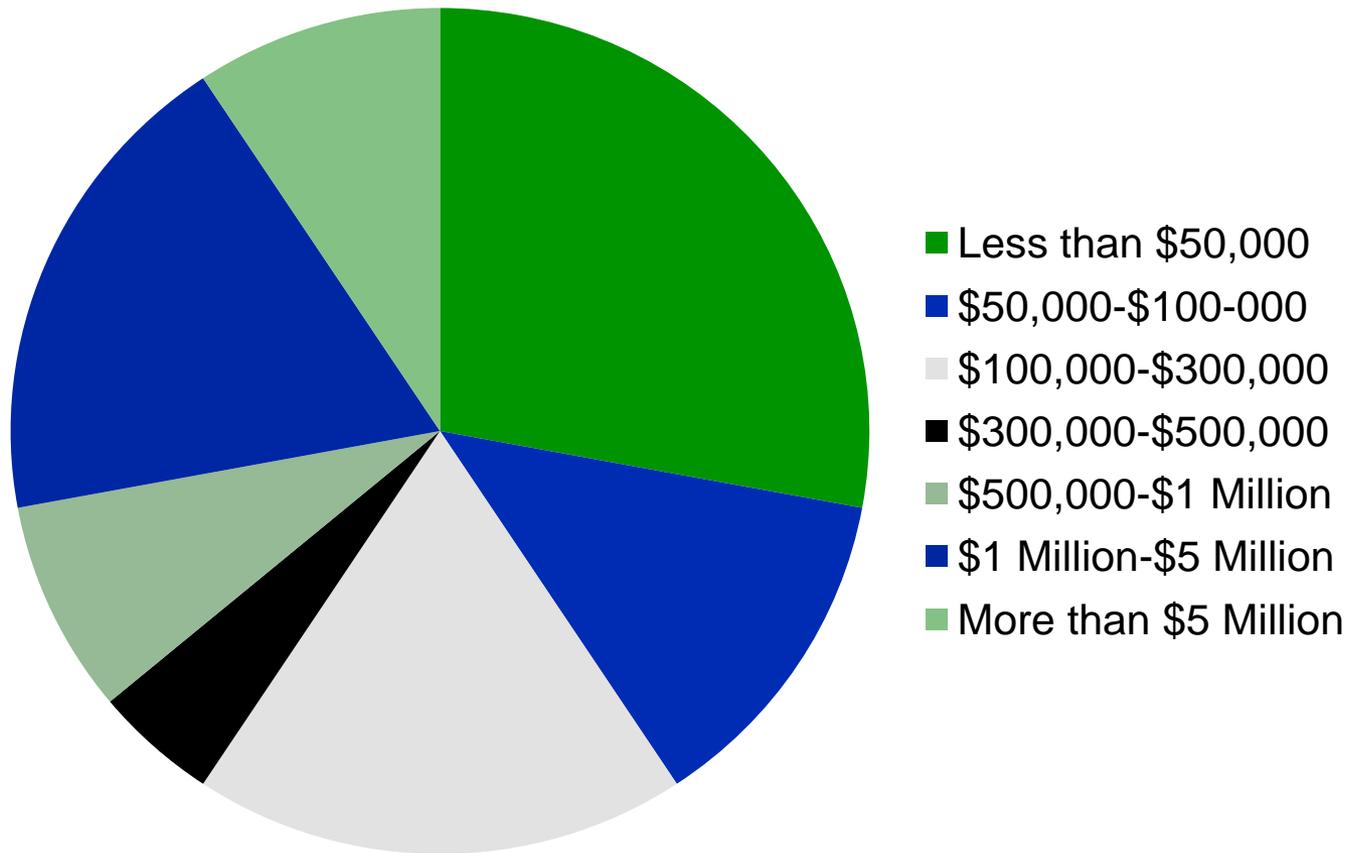


Overall Trends

- 86/187 responded (46%)
- 49% serve their agency only
- 72% are capturing born-digital for at least some formats
- 10% are digitizing classified materials



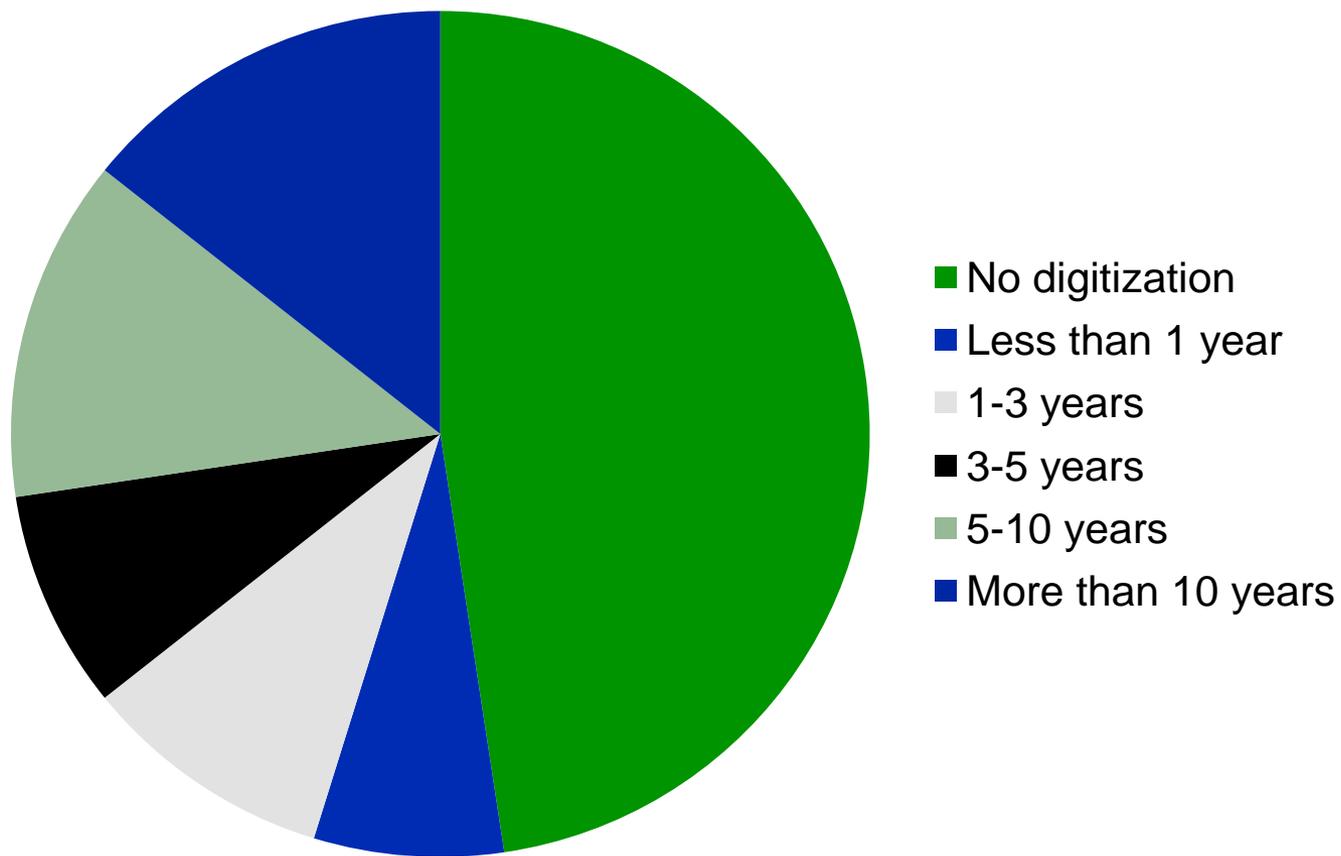
Budget Trends



June 9, 2010



Digitization Duration Trends



June 9, 2010



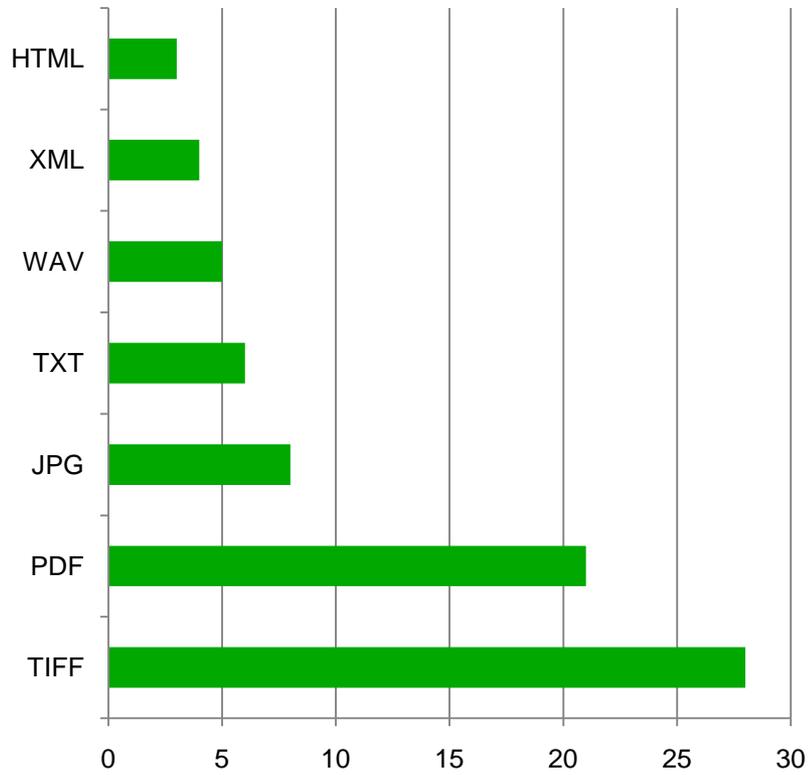
Digital Repository Trends

- 27% have more than one repository
- 18% have their repository hosted remotely
 - Lack technical expertise
 - Insufficient IT support
 - Less costly
 - System update more frequently

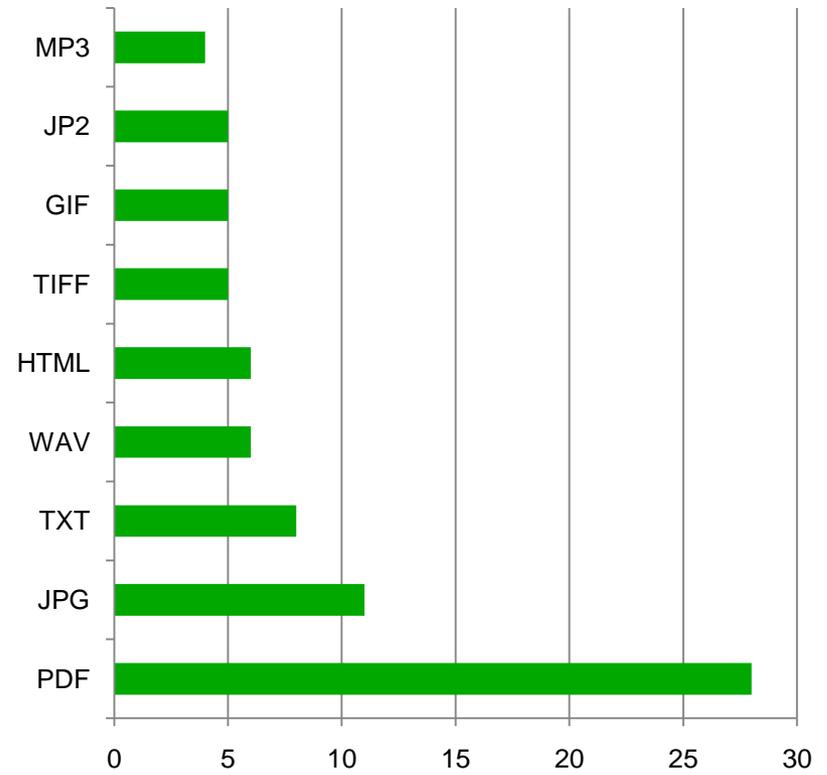


File Format Trends

Master File



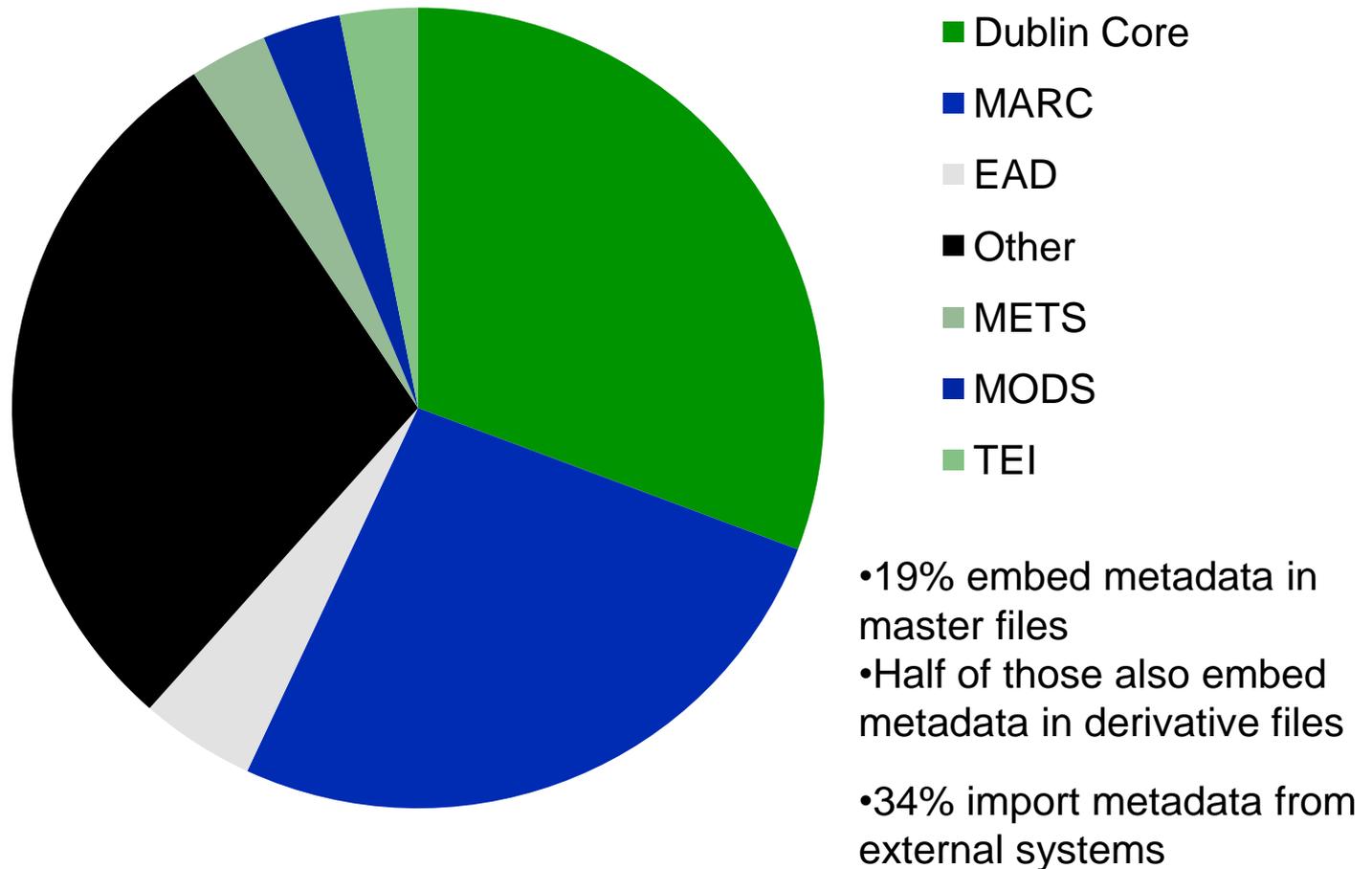
Derivative File



June 9, 2010

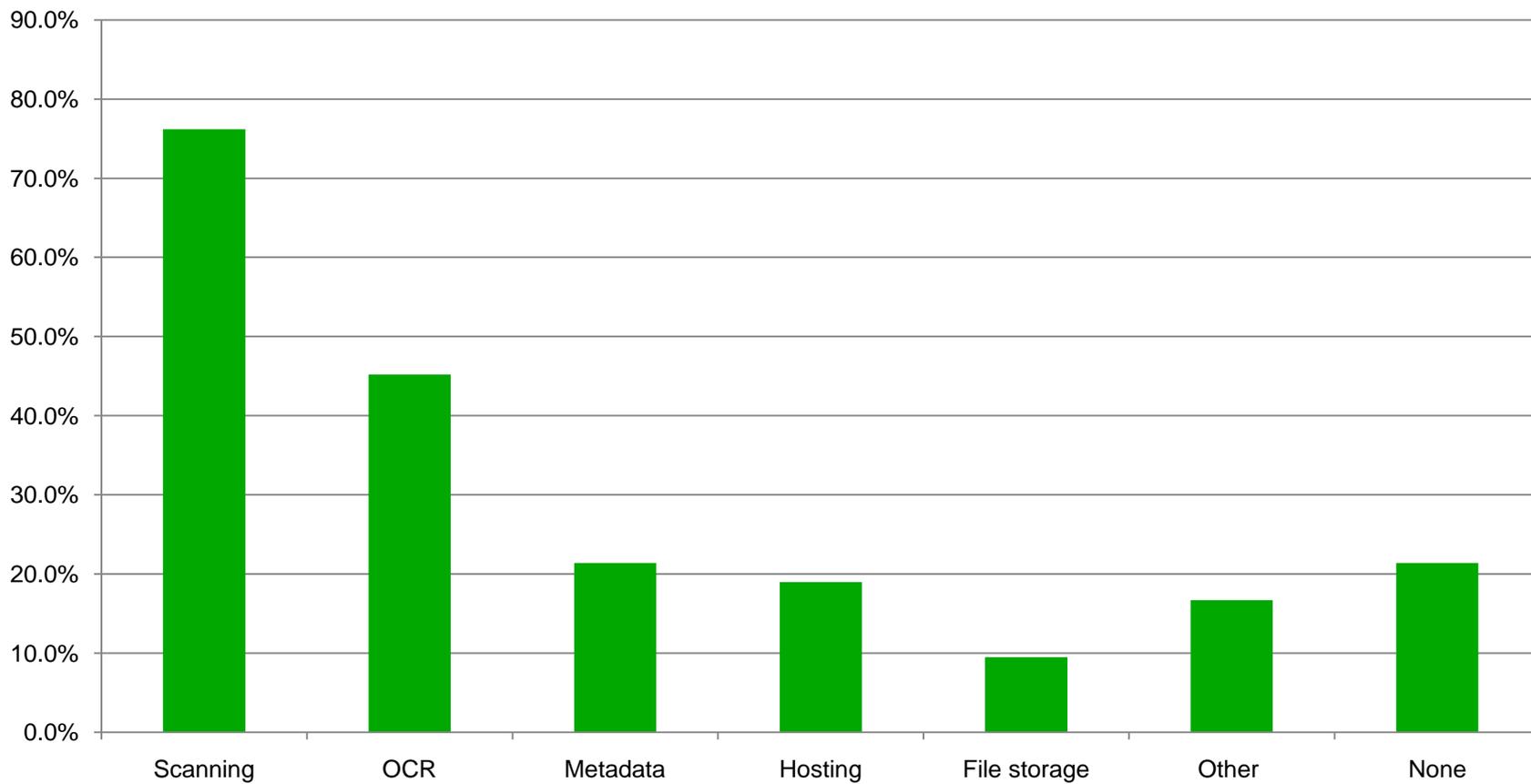


Metadata Scheme Trends





Outsourcing Trends



June 9, 2010



Written Policy Trends

57% have
written
digitization
policy
documents

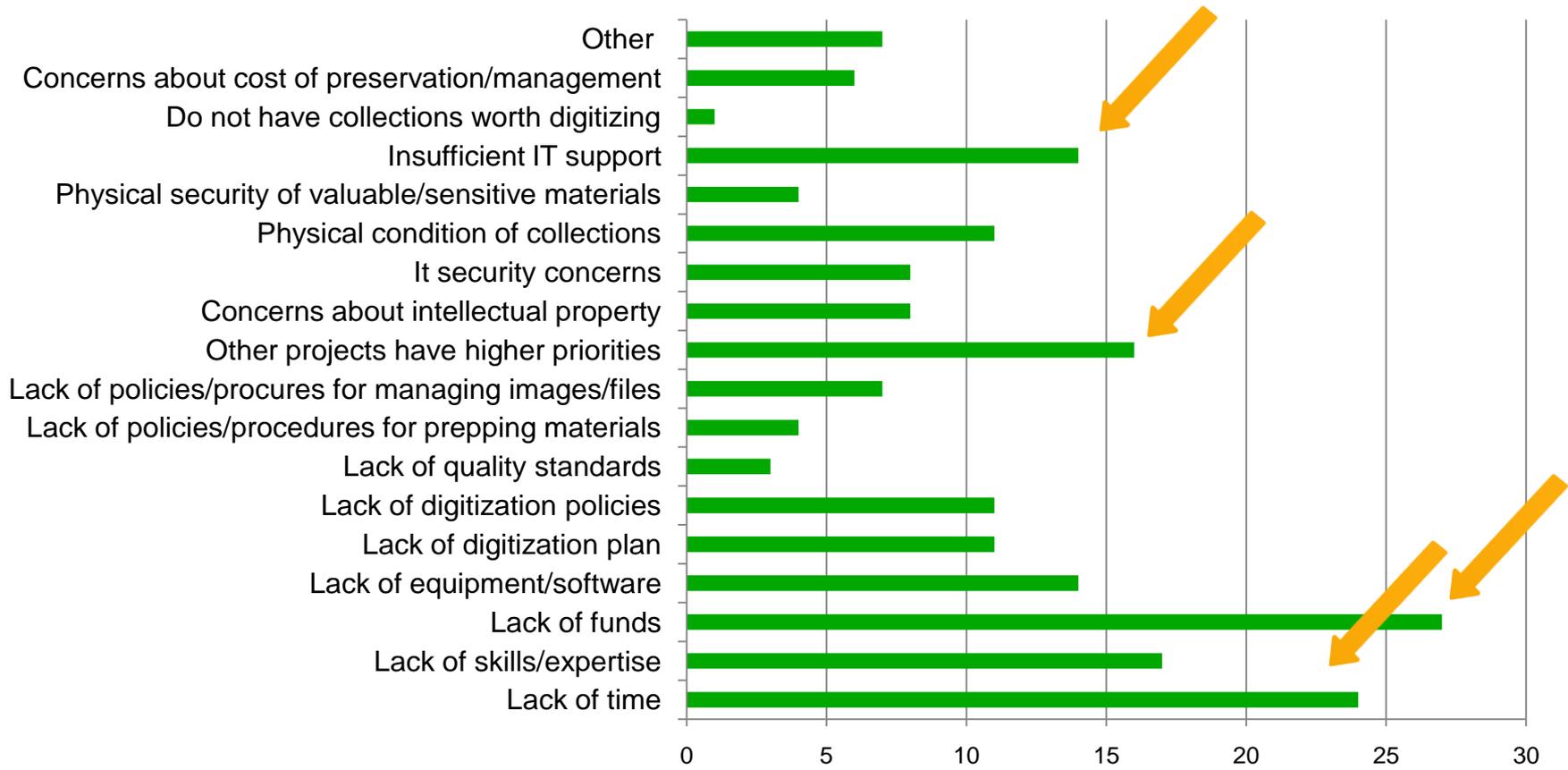


40% or more
of these
address...

- File naming conventions
- Metadata schemes & stds.
- Technical stds.
- Digitization stds. & best practices
- Quality control procedures
- File architecture
- Collection development/Digitization priorities
- Adding/Deleting files
- Who can access system
- Disposition/Preservation of originals



Library Challenges



June 9, 2010



Future Plans

- Survey rest of Federal libraries
 - Finish Army (Post Libraries)
 - Expand to DOD
 - Dept. of Interior
- Develop education programs to support library needs
- Continue outreach



We've Learned

- Digitization is taking place outside libraries in many agencies
 - How can we capture or survey this activity?
 - What do we want to find out about it?
 - Can we promote partnerships between libraries and other parts of our organizations?



Requesting Feedback

- Publishing
 - When do we have “enough” data to publish?
 - What are you interested in reading about this project?
- Best Practices
 - How do we define “best practices”?
 - Are these different from observable trends?



Questions?

June 9, 2010



Deborah E. B. Keller

Chair, Digitization Survey Committee,
FLICC Preservation & Digitization Working Group

U.S. Army Corps of Engineers
Humphreys Engineer Center Support Activity Library
deborah.eb.keller@us.army.mil

This work is the property of the U.S. government and is not eligible for copyright protection in the U.S.

This material reflects work in an ongoing project. Please contact the Federal Library and Information Center Committee to determine if more recent information on this project is available.

June 9, 2010