

Digital Preservation

ATALM Post-Conference SHN Workshop October 2016

Alex Merrill, Lotus Norton-Wisla

Agenda

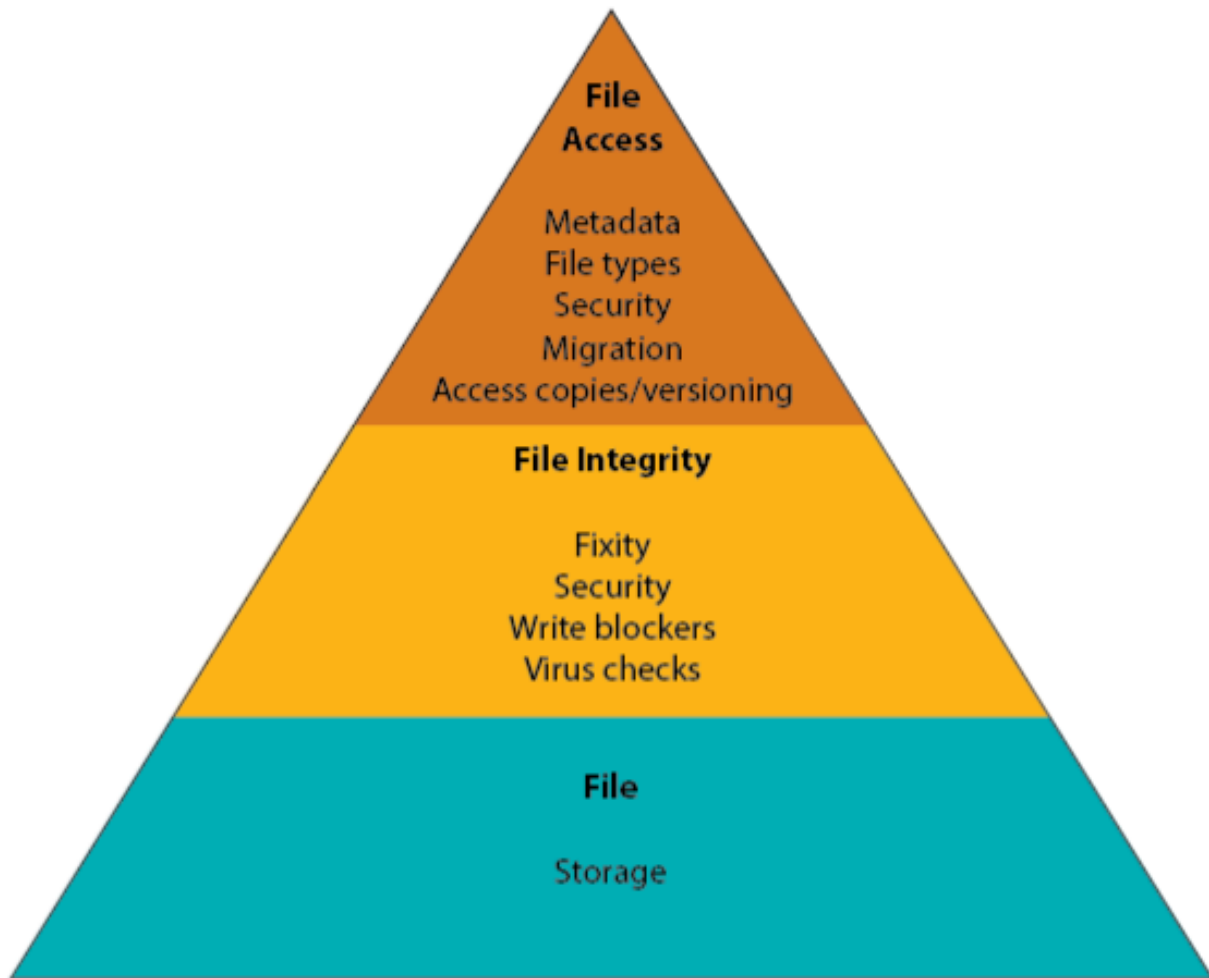
- 2:15-3:30 Digital Preservation Overview
- 3:30-4:00 Storage Activity
- 4:00-4:30 Digital Preservation Action Plan
- 4:30-5:00 SHN wrap up, evaluations

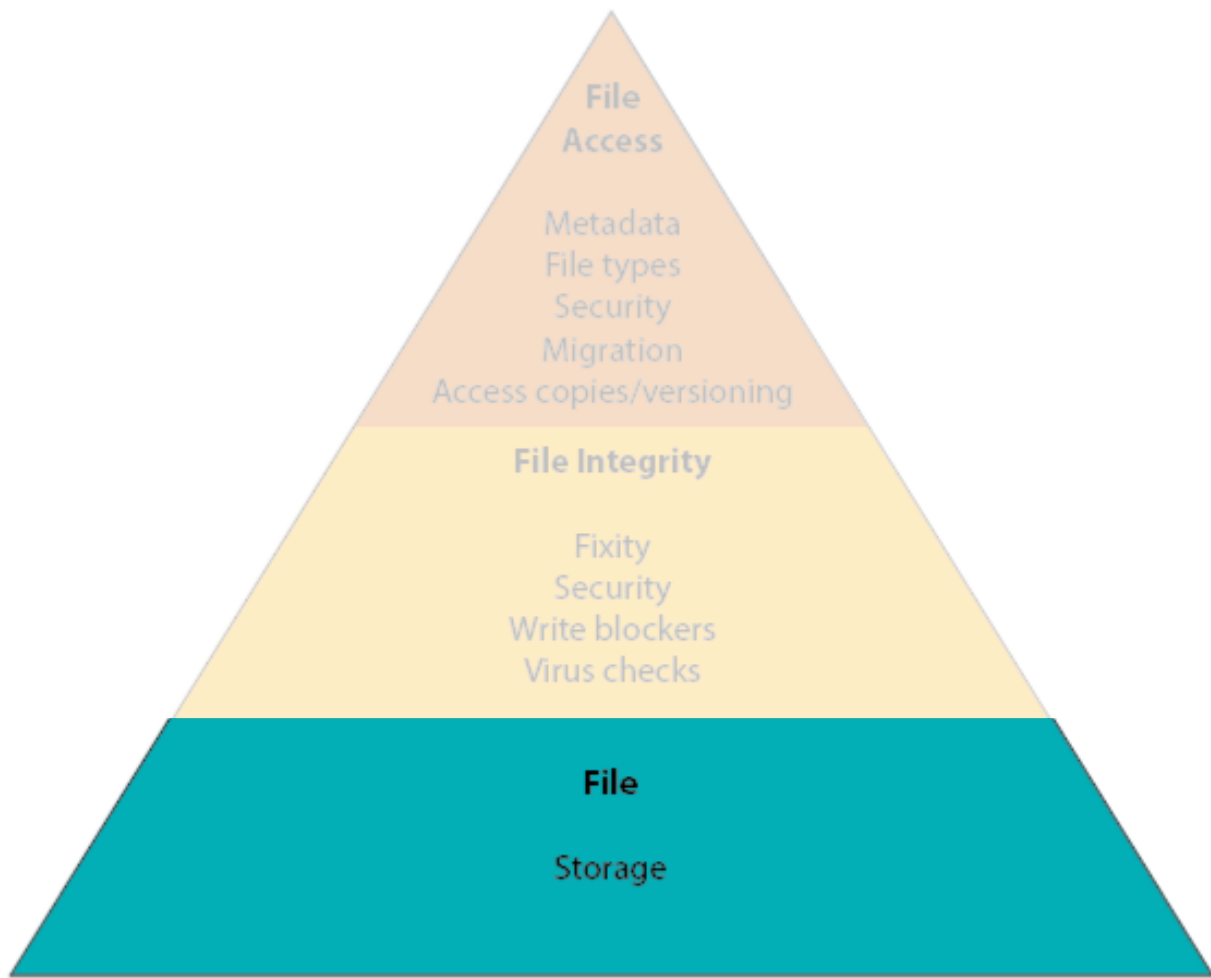
Digital Preservation

- Long term storage and preservation of your digital files
- Create a Digital Preservation Plan
- Runs through all of your digital projects
- Can't just “set it and forget it”
- Must be updating, researching, monitoring

Digital Preservation -

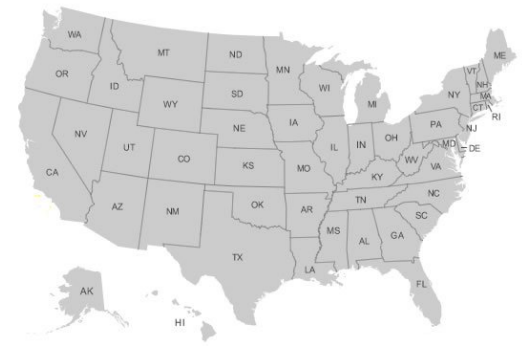
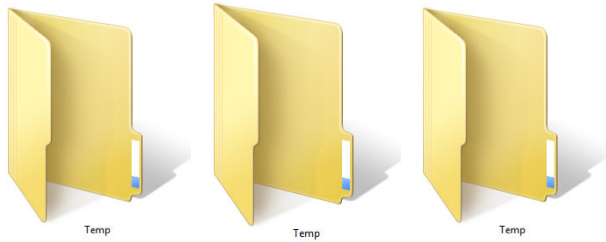
- Find others to bring into the conversation
- Fit your needs into what already exists
- Maintain communication
- Balance responsibilities





Storage

- Storage space for content
- Integrate with IT
 - What does your IT department already have set up?
- Consider types of storage
- Multiple layers
- Buy new storage every few years



3

Copies

2

Types of
storage
media

1

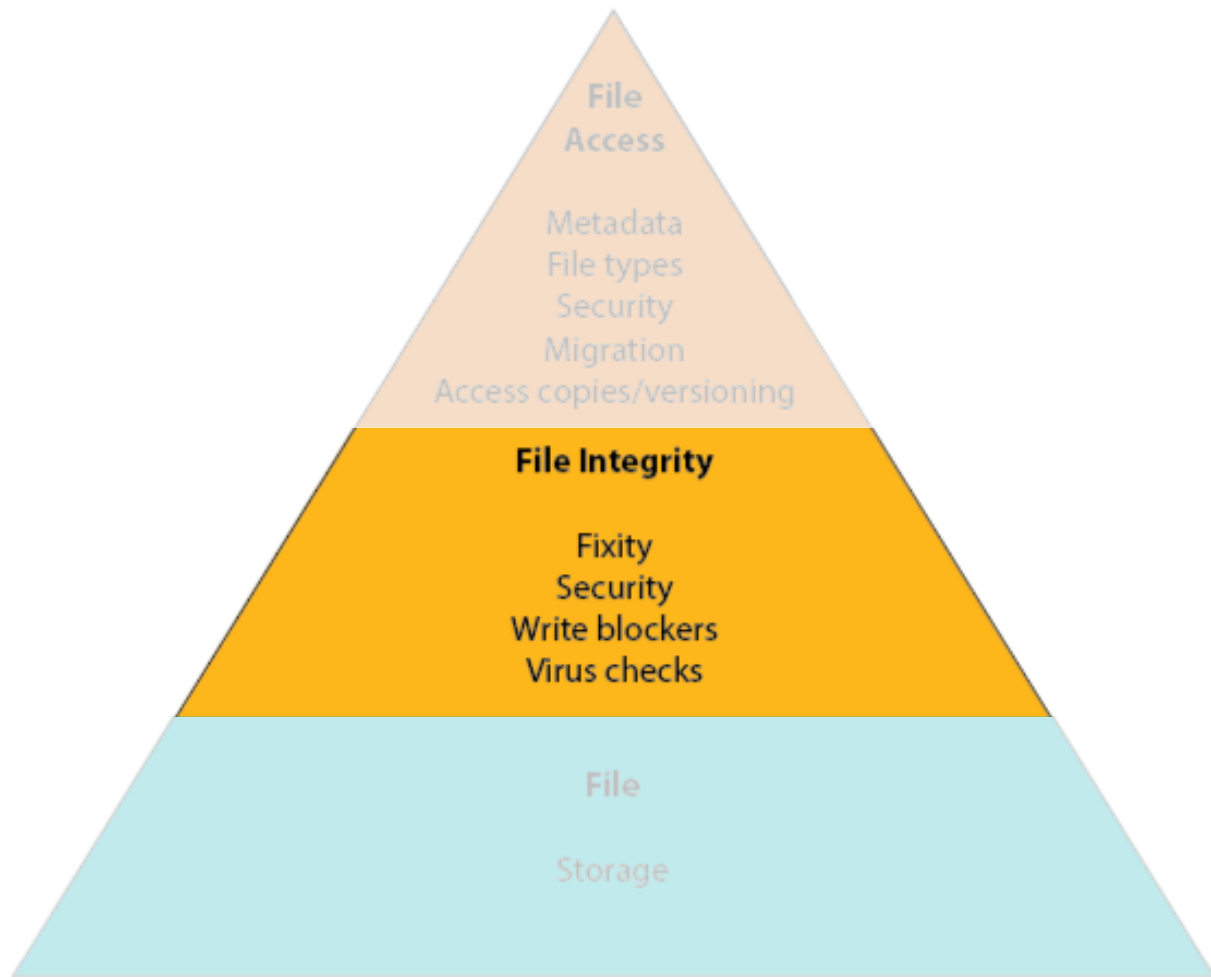
Different
geographic
location

Estimating Storage - Space Calculators

- Video - [Digital Rebellion](#)
- Audio - [The Audio Archive](#)
- Photograph - [NAU](#)
- Paper - [simple Cornell equation](#)

Storage Questions

- How often content is backed up?
- What types of storage devices are used, and how often storage devices are migrated?
- What risks are present in your region?



Fixity

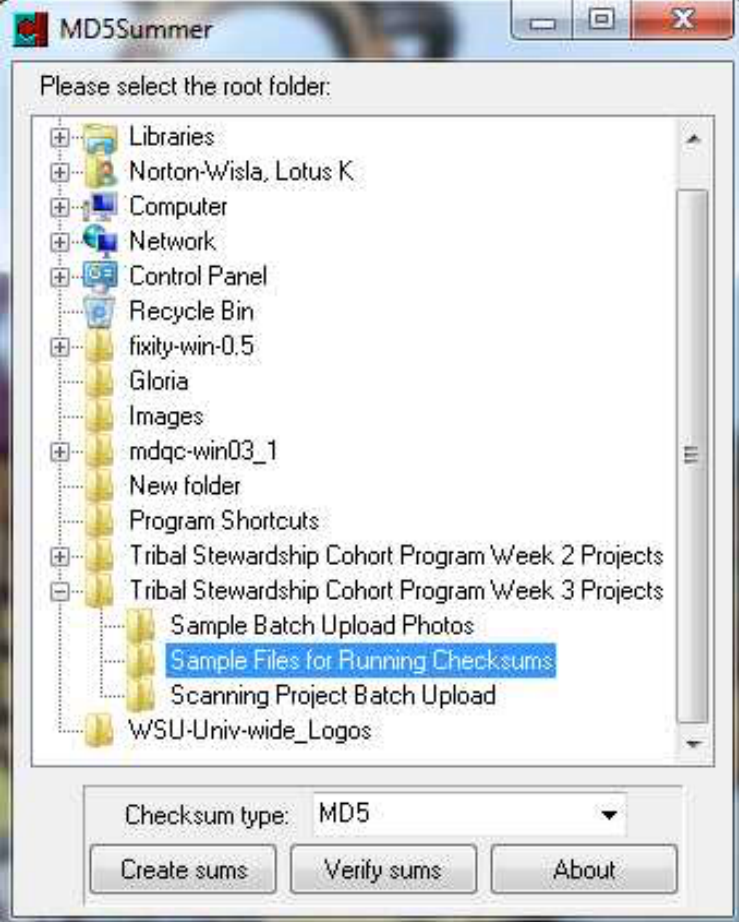
- Definition: Stability of a digital object
- Files should remain unchanged over time
- Digital files can degrade without warning

Fixity Levels

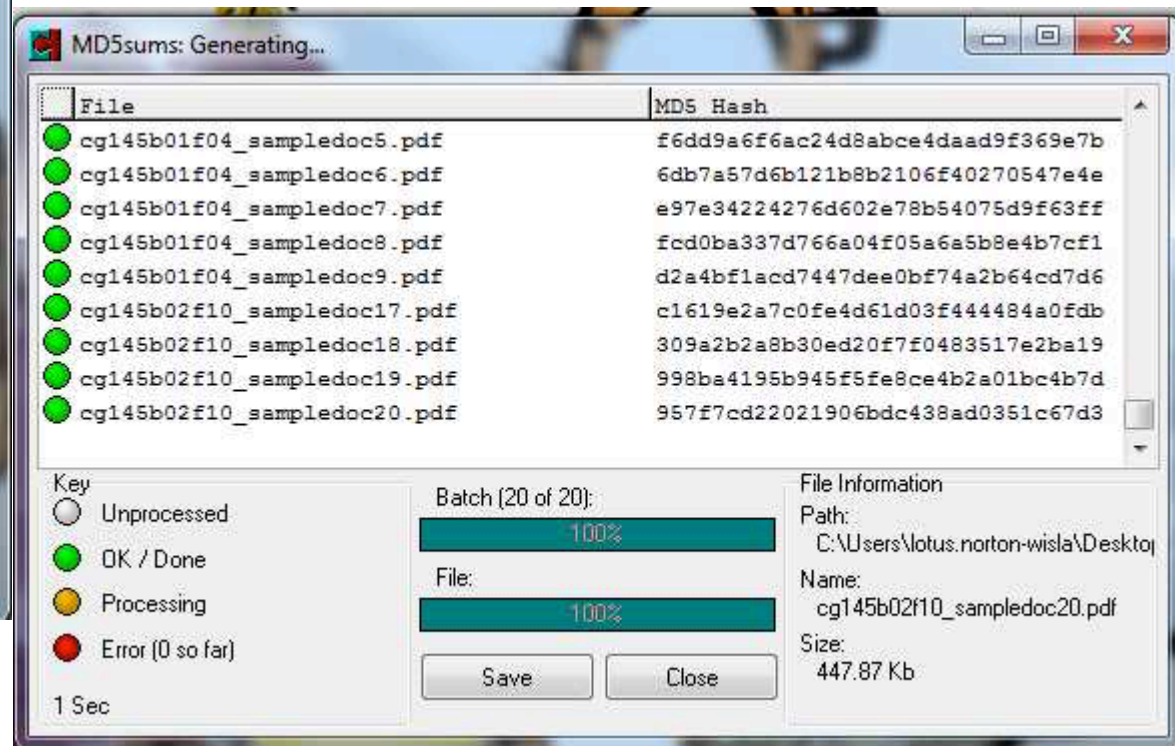
- Simple steps: Expected file size, expected file count
- Good/Better: Higher level of detail MD5 (hash algorithm)
- Best: Use a use tool that uses SHA1 (hash algorithm)

How to check/store Fixity information

- Standalone software applications
 - MD5Summer, Fixity, Jacksum, BWF Metaedit
- Integrated into content management systems
 - Dspace, Islandora, Rosetta



MD5Summer



When to Check Fixity

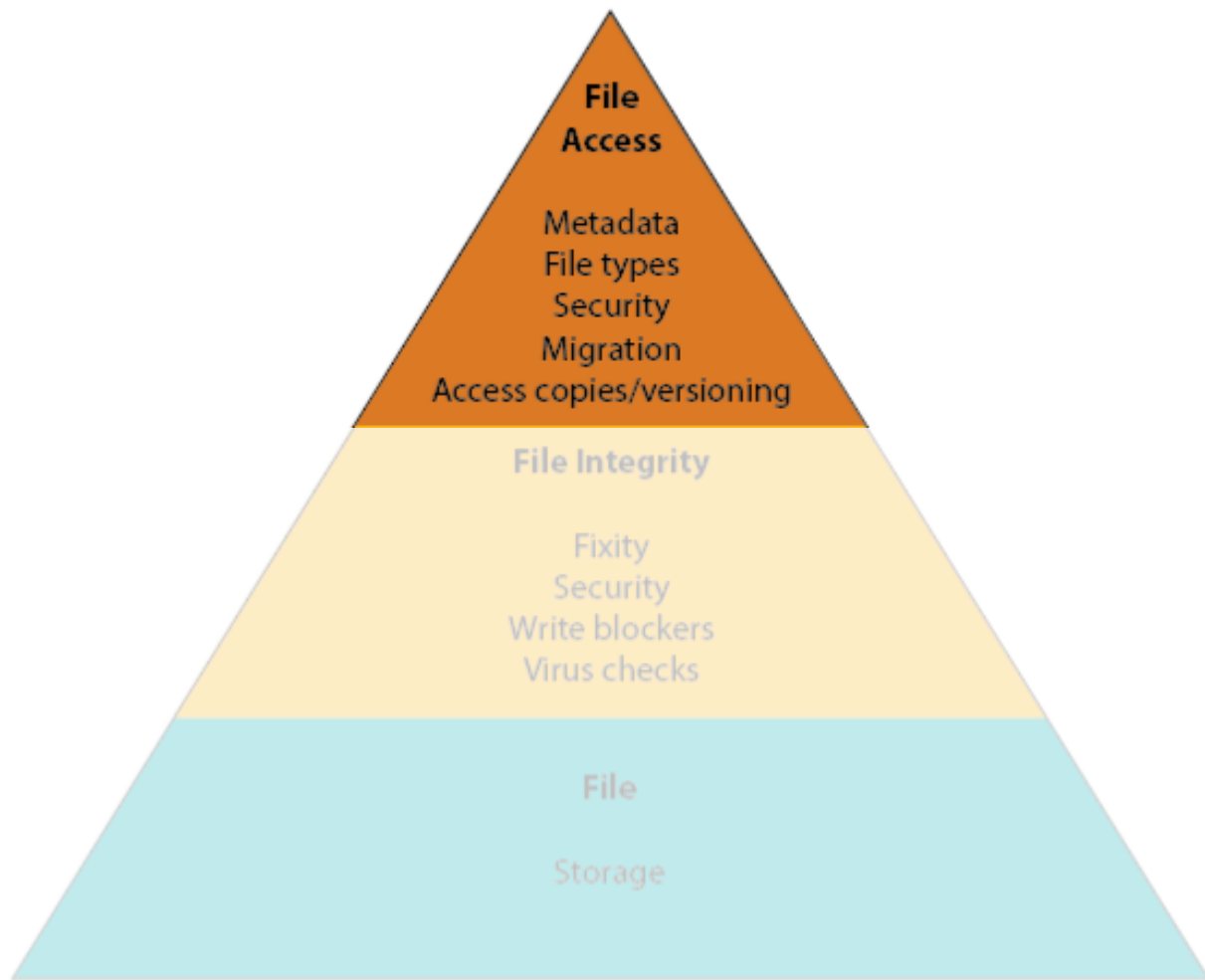
- Run a checksum tool
 - When created
 - At regular intervals (monthly, weekly)
 - During a change (transfer, recovery)

Security

- Know who has access to your files
- Have policies and/or technology in place to restrict access to appropriate people
- File System Logging (Something to discuss with IT)

Integrity Questions

- Does your department /IT already check fixity?
- Who on your staff (or other departments) has permissions to view/edit/delete digital files?
- Are regular virus checks run on computers? What about when new digital collections come in?



Meaning of “Access”

- Providing access to your community
- **Digital preservation** access =
internal management of your files
and INFORMATION about your files

Meaning of “Access”

- Adequately describing the digital object
- Ensuring continued access to the information within the file (metadata, file types, representation information, migration, security)
- By you, your staff, and future staff

Access - Types of Copies

- Preservation Master files
- Access copies
- Web Ready copies

Access - Managing File Types

- **Good:** Decide which file formats to use across collections, open formats, document which formats
- **Better:** Validate files, keep up on threats to formats
- **Best:** Migrate formats, emulation, have plans to deal with obsolescence

Access - Preservation Metadata

- Documentation
- Long-term usability, understanding

Good: Inventory

Better: Descriptive, Administrative, and
Technical, metadata

Best: All preservation metadata

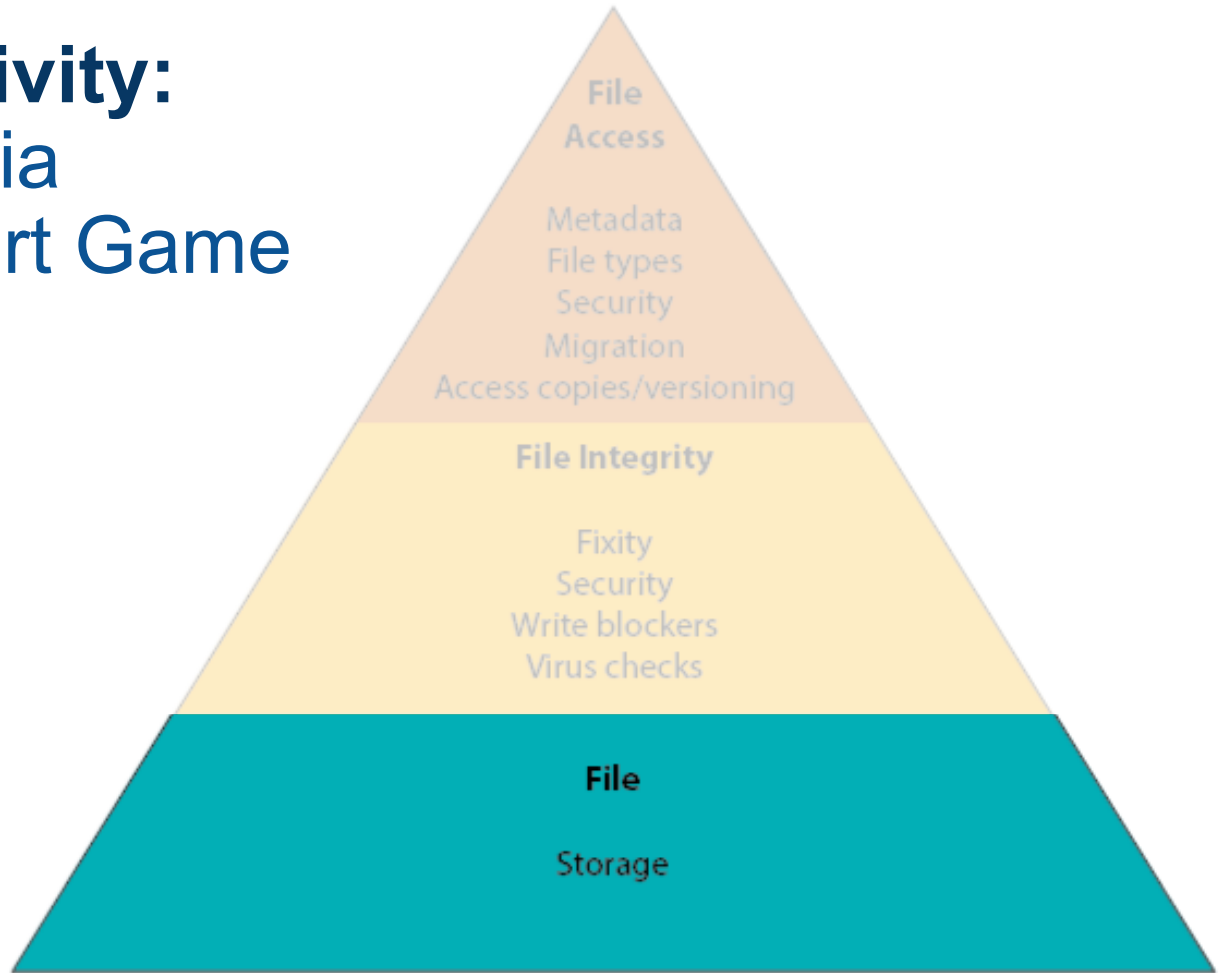
	Element name	Obligation	Suggested Value Control
1	<u>Bit Depth</u>	Strongly recommended	Controlled vocabulary
2	<u>Checksum</u>	Strongly recommended	None (free text)
3	<u>Collection Source</u>	Required	Controlled vocabulary
4	<u>Color Space</u>	Optional	Controlled vocabulary
5	<u>Compression Degree</u>	Recommended, if applicable	None (free text)
6	<u>Compression Type</u>	Strongly recommended, if applicable	Controlled vocabulary
7	<u>Creation Hardware</u>	Strongly recommended (digitized); Optional (born-digital)	Controlled vocabulary
8	<u>Creation Software</u>	Recommended (digitized); Strongly recommended (born-digital)	Controlled vocabulary
9	<u>Digital Creation Date</u>	Required	ISO 8601 Date-Time Format
10	<u>Digital Creator</u>	Required, when known	Controlled vocabulary
11	<u>Digital Object ID</u>	Required	None (free text)
12	<u>Extent</u>	Required	None (free text)
13	<u>File Format</u>	Required	Controlled vocabulary with Internet media types
14	<u>File Location</u>	Required	None (free text)
15	<u>Local Repository ID</u>	Strongly recommended	Controlled vocabulary
16	<u>Original Object ID</u>	Required, if applicable	None (free text)
17	<u>Resolution</u>	Required for static images	None (free text)
18	<u>Revision Date</u>	Strongly recommended, if applicable	ISO 8601 Date-Time Format
19	<u>Revision History</u>	Strongly recommended, if applicable	None (free text)
20	<u>Rights Statement</u>	Required	None (free text) or controlled vocabulary
21	<u>Security</u>	Optional	Controlled vocabulary

Access Questions

- What metadata scheme does your department use?
Does it include preservation metadata?
- Do you have consistent file formats that you use for different types of files?
- Do you have preservation copy, access copy, edited copy for all files? How these are organized and saved?

Digital Preservation Storage Activity

Storage Activity: Storage Media Shopping Cart Game



Project Parameters

- Our collection: 10 terabytes
- Our budget: \$15,000
- Our grant period: 3 years

Types of Storage Media

- external hard drives
- flash drives
- RAID hard drive
- cloud/hosted storage
- SSD (solid state drives)
- LTO Tape
- Network Attached Storage



There is no cloud
it's just someone else's computer

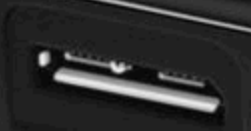




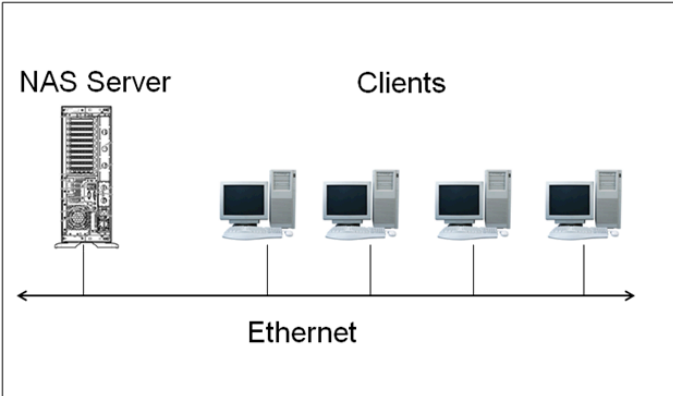




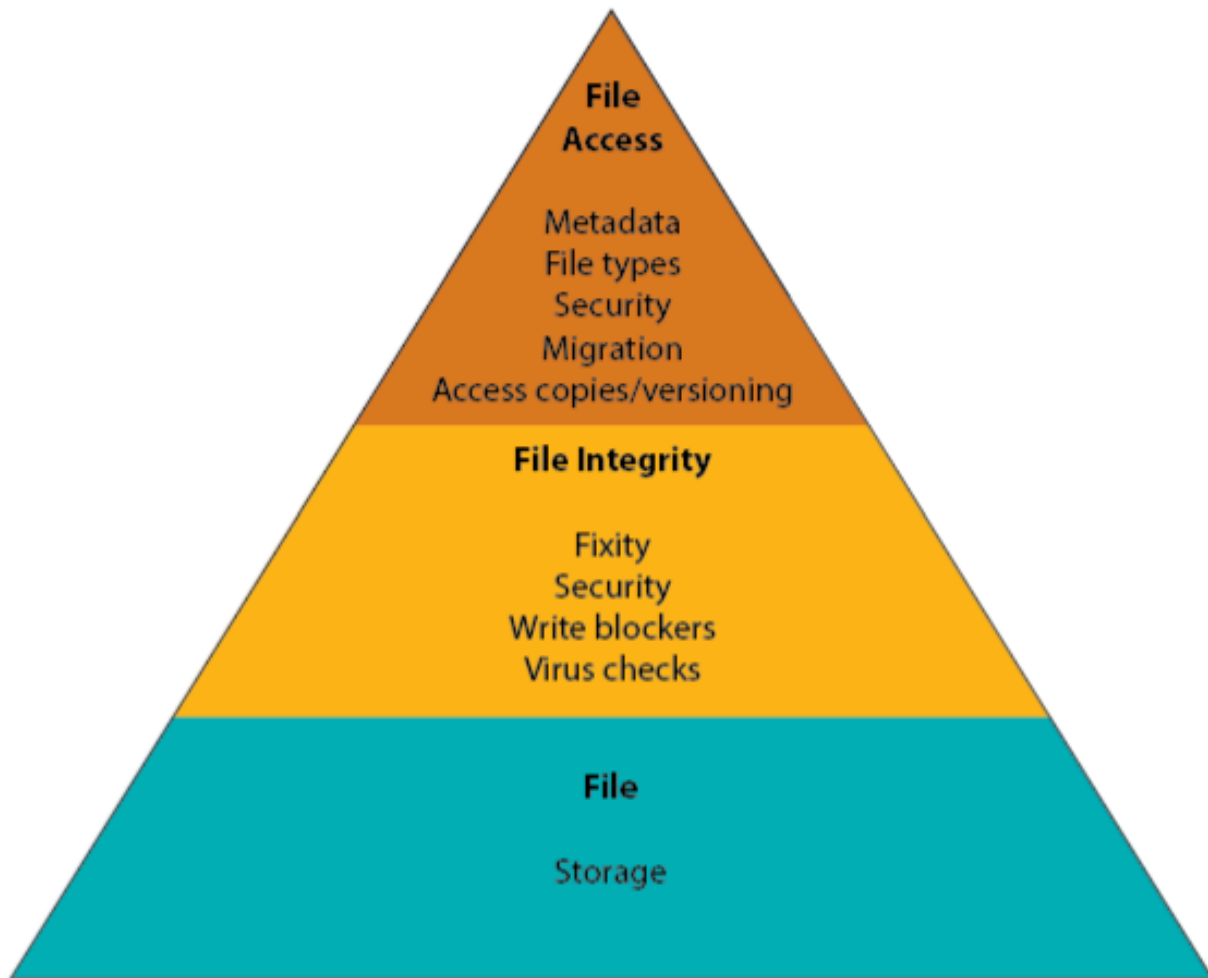
My Passport
Ultra







Digital Preservation Action Plan



What is an “Action Plan”

- People in your organization
- Taking stock of what you already know
- Things that you want to find out

People in your Organization

- Supervisors
- IT Department or Staff
- Staff
- Other Departments
- Funders
- Partners

What you know already:

- What is your current storage setup?
- Who is involved in digital preservation already?
- What digital content you have now?
- What digital projects do you have planned?

What you need to find out or accomplish?

- Inventory of digital files
- Who else do you need to talk to?
- Funding sources
- Questions from: Storage, Integrity, and Access
- Find out backup schedule (when and how they happen, when and how they check backups)

Fill out your Action Plan

- 3 people to meet with
- 3 things you know already
- 3 questions to ask (yourself or others)

Resources

- [The Signal](#) - blog from Library of Congress
- [Digital Preservation Q&A](#)
- [Digital Preservation Google Group](#)
- [COPTR Tool Grid](#)
- Resources on the [Sustainable Heritage Network](#)