

# Digital Preservation: A Brief Introduction

Indigenous Digital & Archival Training  
Sustainable Heritage Network Workshop at University of Manitoba  
February 2 - 4, 2017

Jeanine Nault, Lotus Norton-Wisla

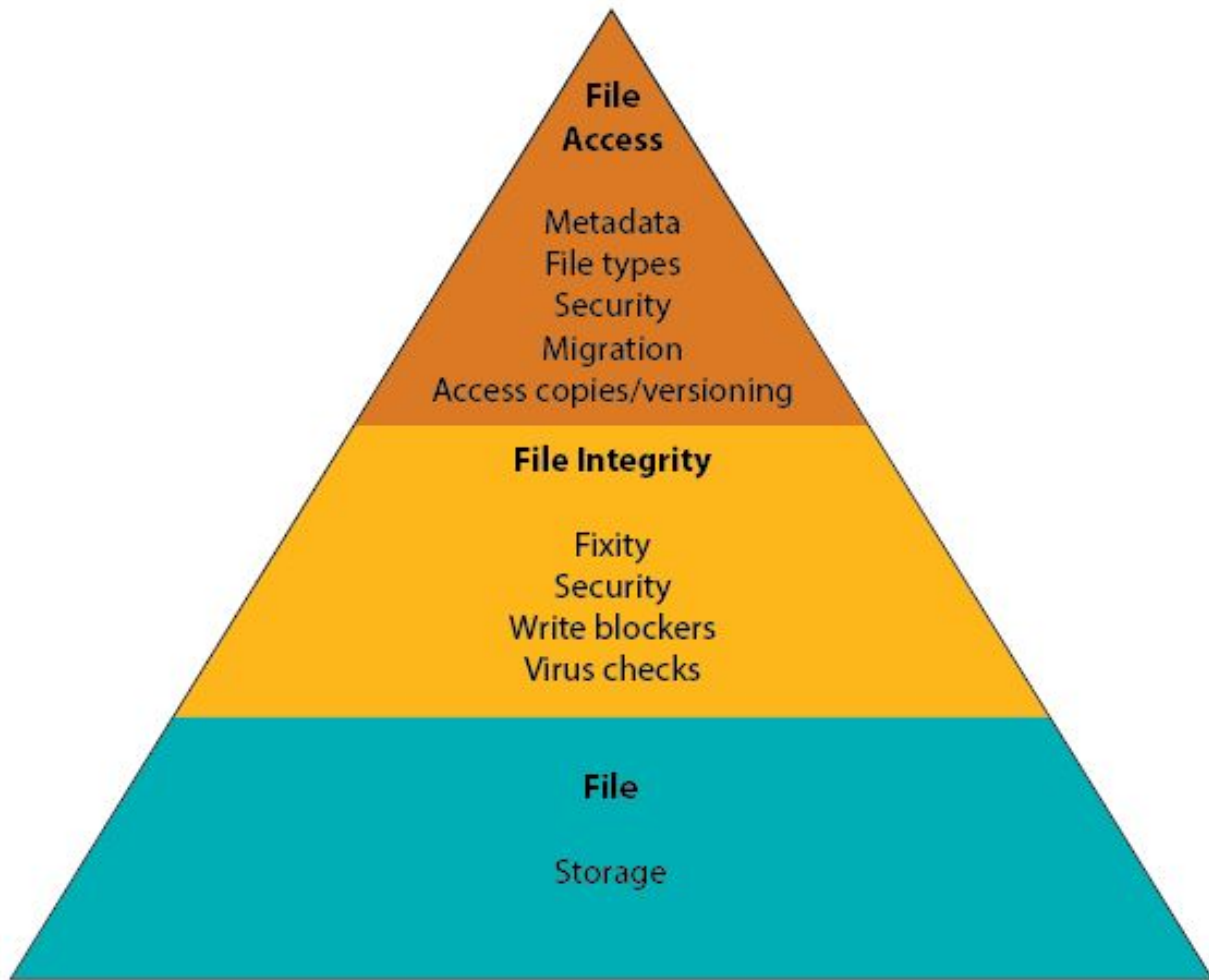
# Digital Preservation: Storage

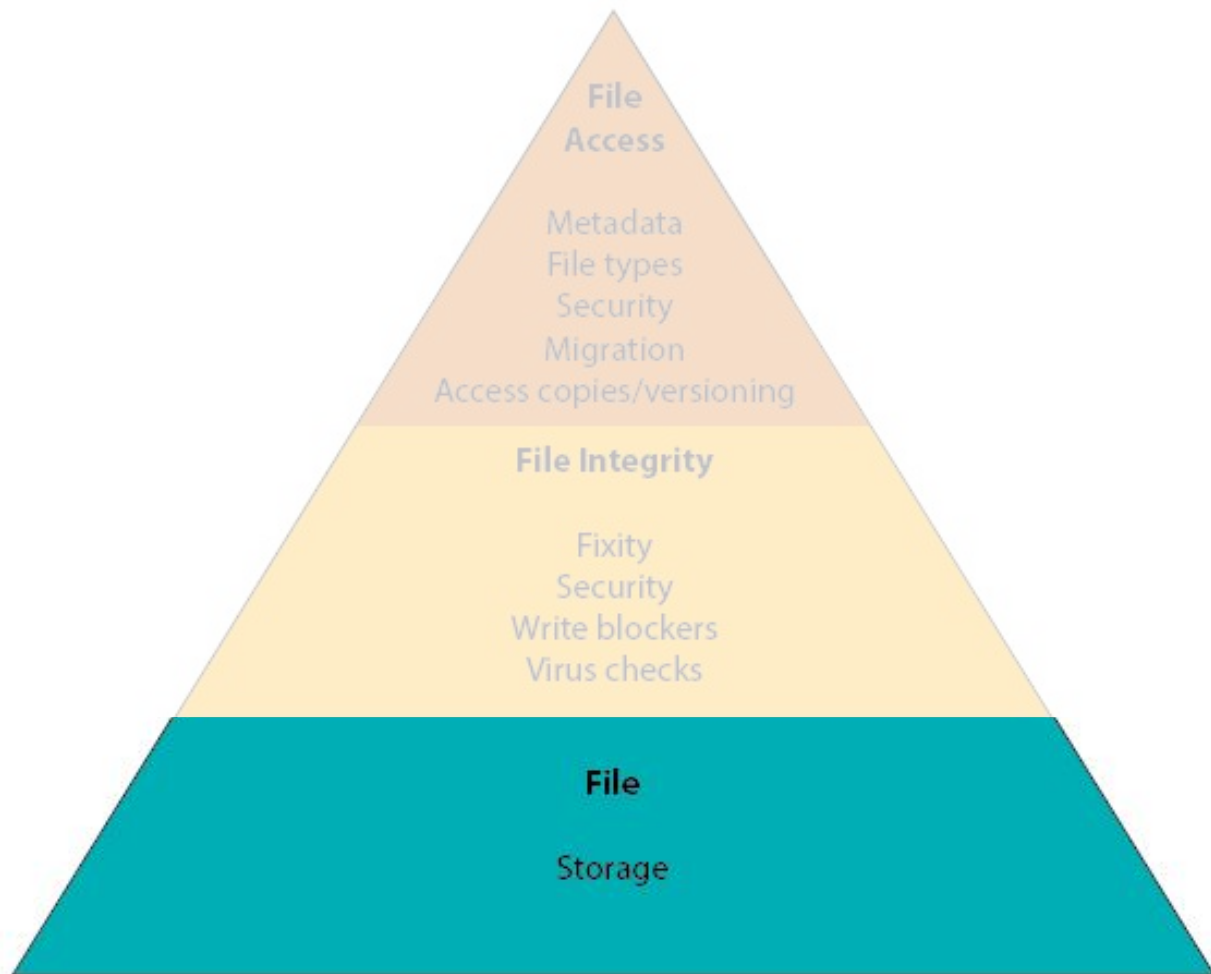
# Digital Preservation

- 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





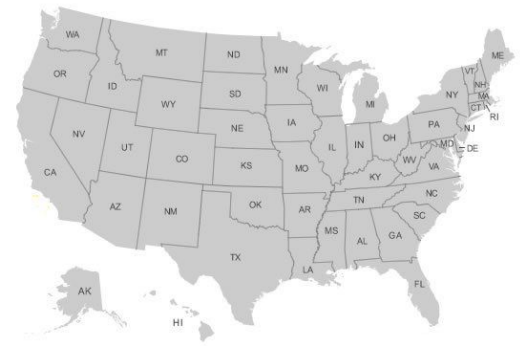
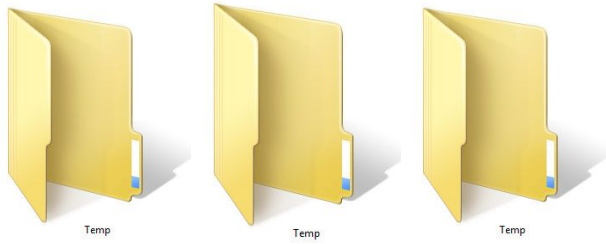
# Storage

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

# Estimating Storage - Space Calculators

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





3

Copies

2

Types of  
storage  
media

1

Different  
geographic  
location

# Types of Storage Media

- Hard disk drives
- ~~flash drives~~
- RAID hard drive
- cloud/hosted storage
- ~~CDs or DVDs~~
- SSD (solid state drives)
- LTO Tape
- Network Attached Storage



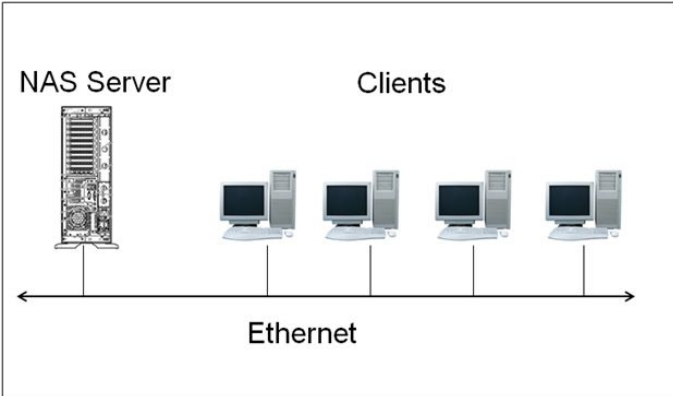
**There is no cloud**  
it's just someone else's computer











# File Copies - Master and Access

- Preservation Master
- Access Copy
- Web-ready derivative



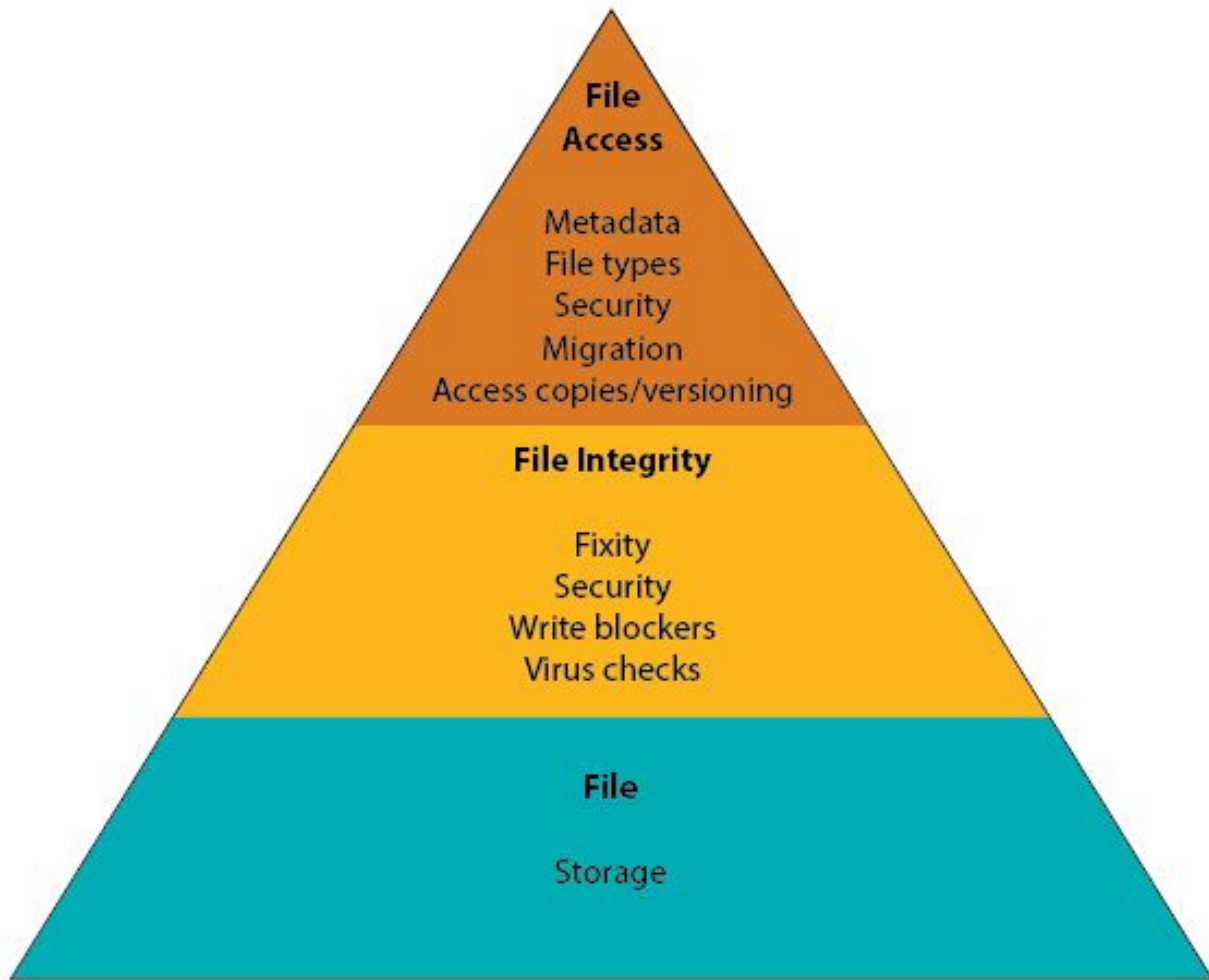
# Storage Questions

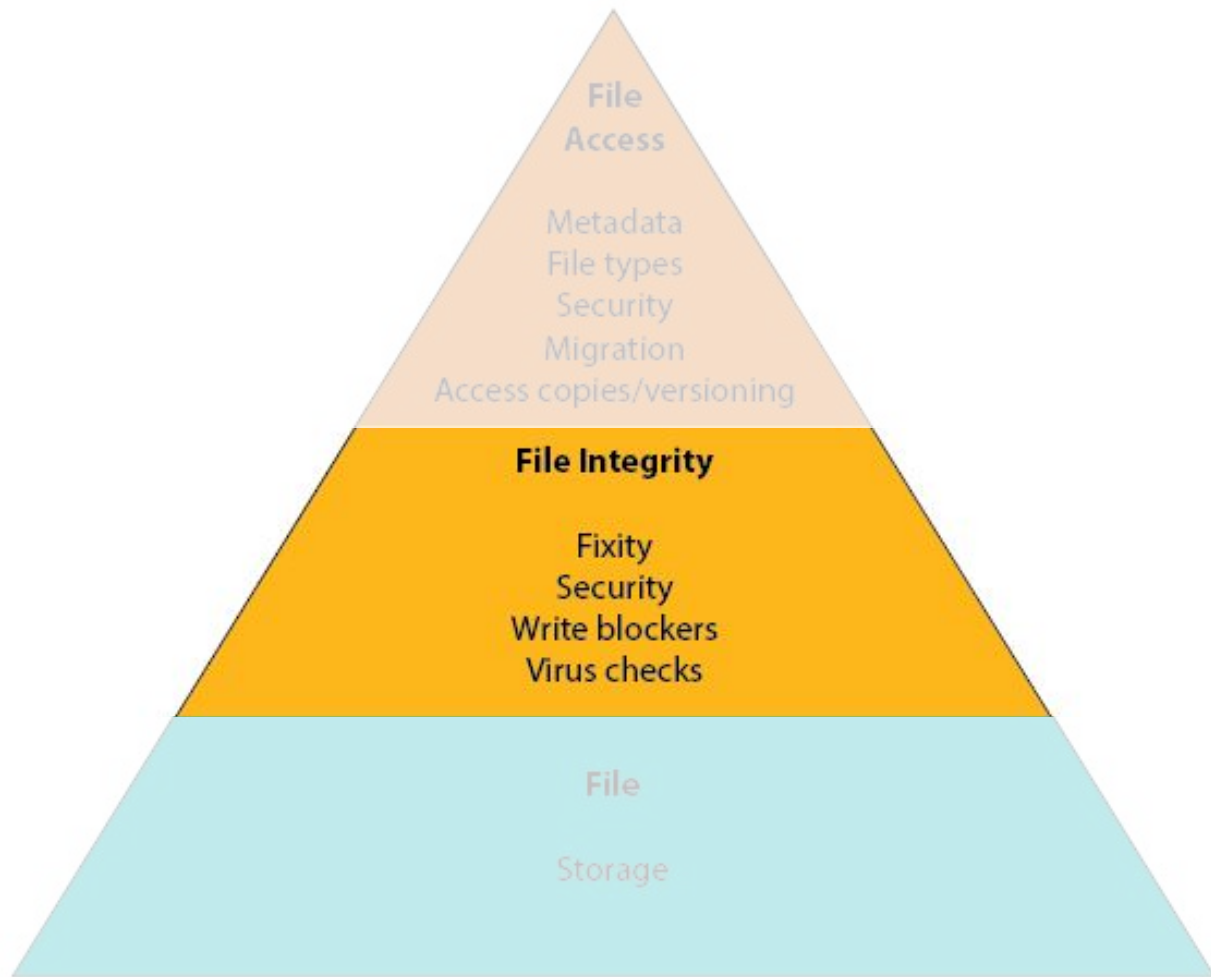
- What does IT already have in place?
- How often content is backed up?
- What types of storage devices are used, and how often storage devices are migrated?
- What risks are present in your region?

# Digital Preservation: File Integrity and Access

# Storage - any remaining questions?

- What does IT already have in place?
- 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



**original image**



**one bit flipped**







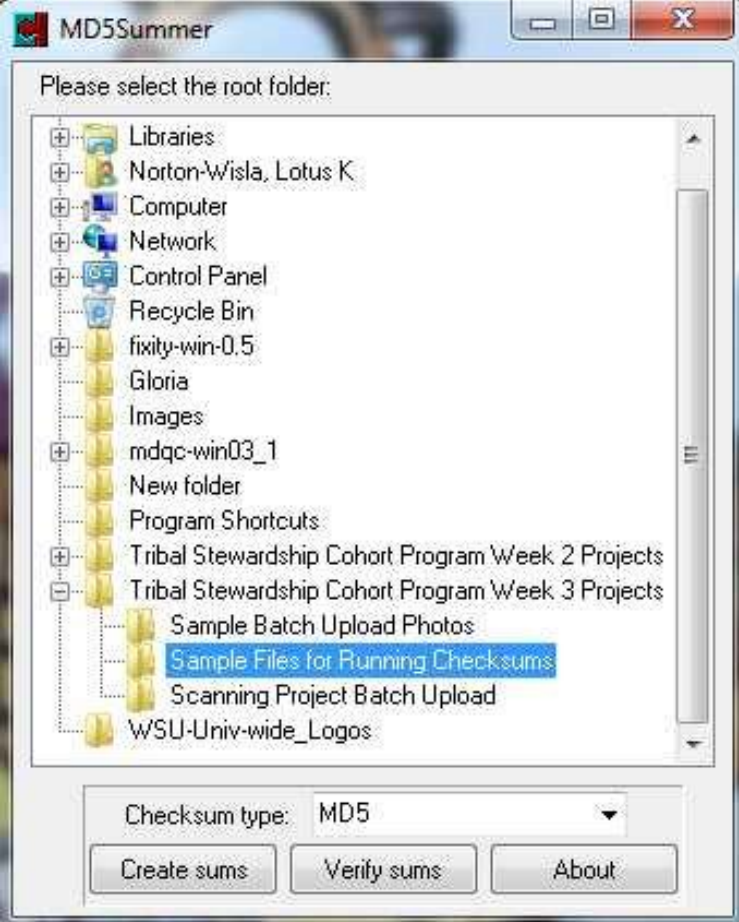


# Fixity Levels

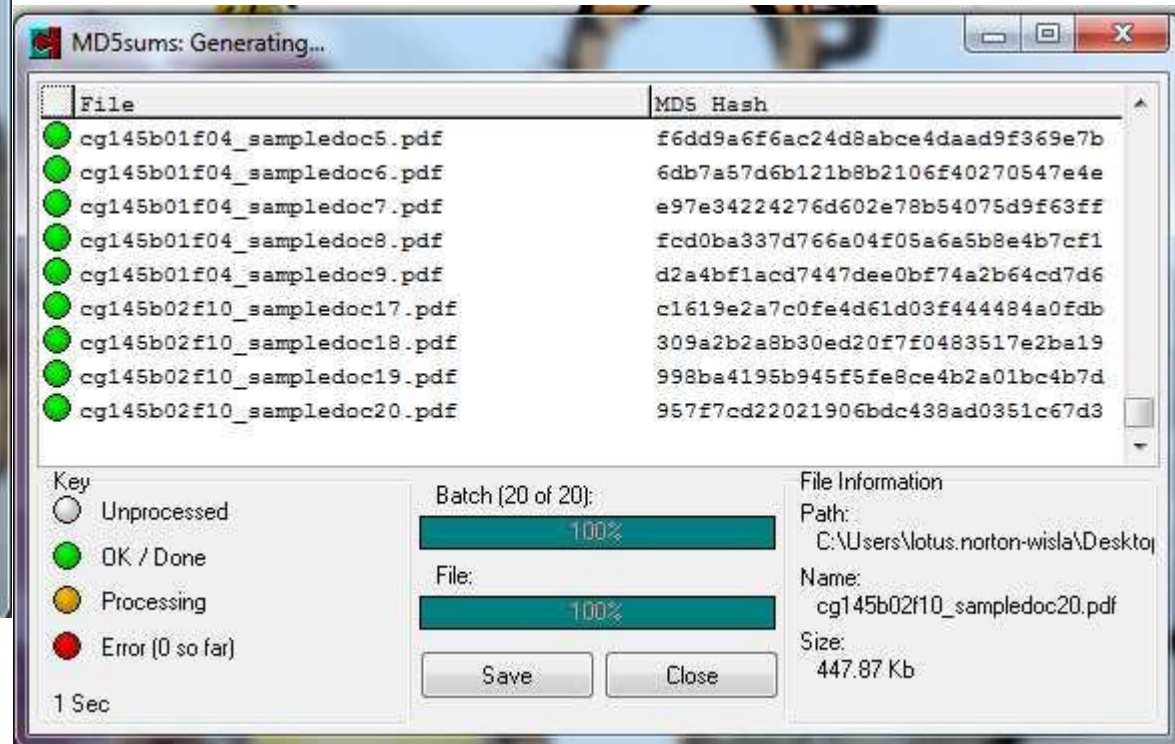
- Simple steps: Expected file size, expected file count
- Good/Better: Higher level of detail MD5 (hash algorithm)
- Best: Use a use tool that uses SHA1 (hash algorithm)

# How to check/store Fixity information

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



# MD5Summer



# When to Check Fixity

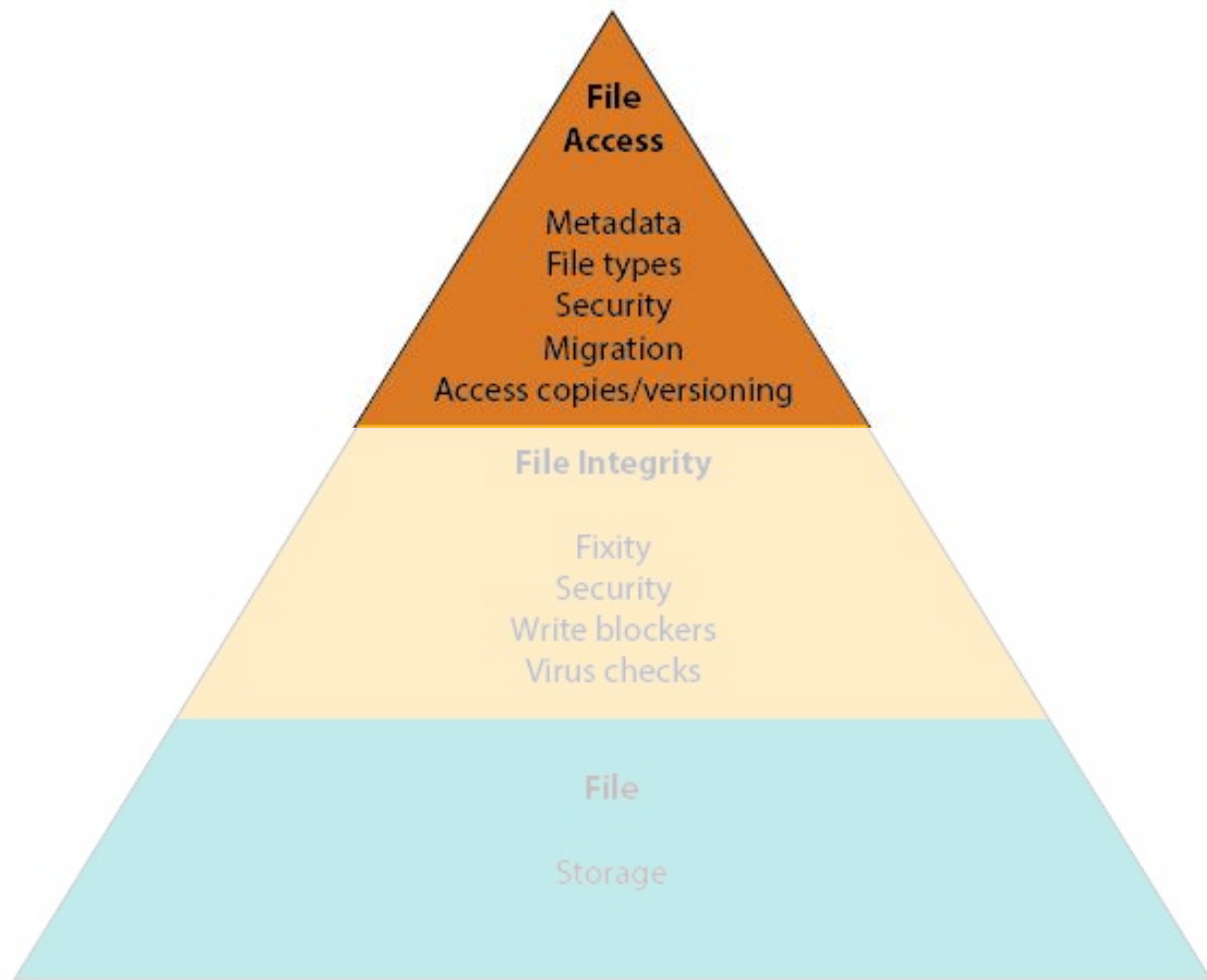
- Run a checksum tool
  - When created
  - At regular intervals (monthly, weekly)
  - During a change (transfer, recovery)
- What if you find a problem?

# 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)

# File 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 describing INFORMATION about  
your files
- For you, your staff, future staff

# Managing file types and metadata

- File Types
  - Document what formats you use
  - Avoid proprietary formats
  - Migrate formats
- Preservation Metadata
  - Inventory
  - Technical and Administrative metadata
  - Checksums
  - Representation information, migration, security

	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

# File 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?

## Next steps:

- People in your organization
- Taking stock of what you already know
- Things that you want to find out
- Create a Digital Preservation Plan, Policy

# People in your Organization

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

# 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](#)

# Additional Resources



# More Resources

- Sustainable Heritage Network
- Preservation Self Assessment Program
- FADGI
- Library of Congress
- Indigitization

Questions?

# Thank you!

[NaultJ@si.edu](mailto:NaultJ@si.edu)

[lotus.norton-wisla@wsu.edu](mailto:lotus.norton-wisla@wsu.edu)

[support@sustainableheritagenetwork.org](mailto:support@sustainableheritagenetwork.org)