

9-22-2016

A Conversation about Digital Preservation

Tammy Troup

Bucknell University, tlt014@bucknell.edu

Follow this and additional works at: https://digitalcommons.bucknell.edu/lit_pubs



Part of the [Archival Science Commons](#), and the [Cataloging and Metadata Commons](#)

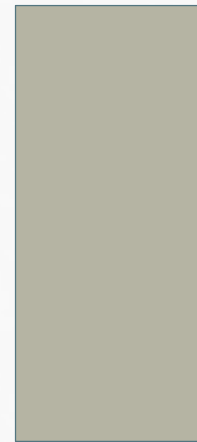
Recommended Citation

Troup, Tammy, "A Conversation about Digital Preservation" (2016). *Library and Information Technology Publications*. 8.
https://digitalcommons.bucknell.edu/lit_pubs/8

This Conference Proceeding is brought to you for free and open access by the Library and Information Technology at Bucknell Digital Commons. It has been accepted for inclusion in Library and Information Technology Publications by an authorized administrator of Bucknell Digital Commons. For more information, please contact dcadmin@bucknell.edu.

A CONVERSATION ABOUT DIGITAL PRESERVATION

Tammy Troup, MHS Digital Services Manager
"Just Scan It... Digitization Dos and Don'ts"
SHRAB Workshop, September 22, 2016



INFORMATION EXCHANGE



Oral

Written

Digital

Methods of Cultural Information & Knowledge Exchange



White Bull Interview. L.A. Huffman. Catalog #981-150. June 20 1901. Montana Historical Society Research Center Photo Archives. Helena, Montana.

Human



Human

Information

Language, Gesture,
Expression, Art,
Music, etc.

Information



White Bull Interview in regard to the Custer Battle. L.A. Huffman. Catalog #981-151. June 20 1901. Montana Historical Society Research Center Photo Archives. Helena, Montana.

Human



Human

Meaning

Emotion, passion,
persistence, organization,
stewardship,
standardization

Meaning



First Associated Press Wirephoto transmission, small plane in Adirondacks, 1935.
<http://norman.hrc.utexas.edu/nyjadc/ItemDetails.cfm?id=266>



Beatles arrive at JFK, 7 February 1964, By 1965 the Associated Press approved the first photo transmission standard

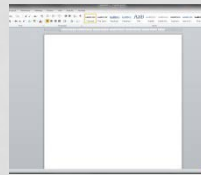


DIGITAL FILE MANAGEMENT

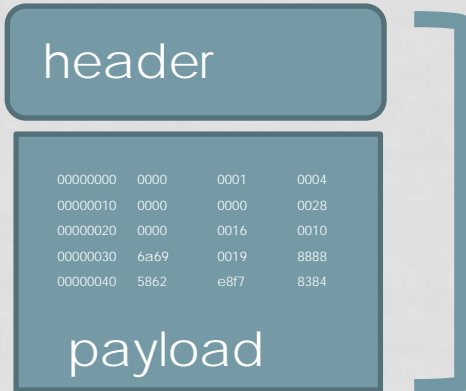
MOVING TOWARD STANDARDIZATION



TFW YOU REALIZE ROBOTS WILL STEAL YOUR JOB



```
000000 0000 0001 0001 1010 0010 0001 0004 0128
00000010 0000 0016 0000 0028 0000 0010 0000 0020
00000020 0000 0001 0004 0000 0000 0000 0000 0000
00000030 0000 0000 0000 0010 0000 0000 0000 0204
00000040 0004 8384 0084 c7c8 00c8 4748 0048 e8e9
00000050 00e9 6a69 0069 a8a9 00a9 2828 0028 fdfc
00000060 00fc 1819 0019 9898 0098 d9d8 00d8 5857
00000070 0057 7b7a 007a bab9 00b9 3a3c 003c 8888
00000080 8888 8888 8888 8888 288e be88 8888 8888
00000090 3b83 5788 8888 8888 7667 778e 8828 8888
000000a0 d61f 7abd 8818 8888 467c 585f 8814 8188
000000b0 8b06 e8f7 88aa 8388 8b3b 88f3 88bd e988
000000c0 8a18 880c e841 c988 b328 6871 688e 958b
000000d0 a948 5862 5884 7e81 3788 1ab4 5a84 3eec
000000e0 3d86 dc88 5cbb 8888 8888 8888 8888 8888
000000f0 8888 8888 8888 8888 8888 8888 8888 0000
00001000 0000 0000 0000 0000 0000 0000 0000
*
00001030 0000 0000 0000 0000 0000 0000 0000
0000103e
```



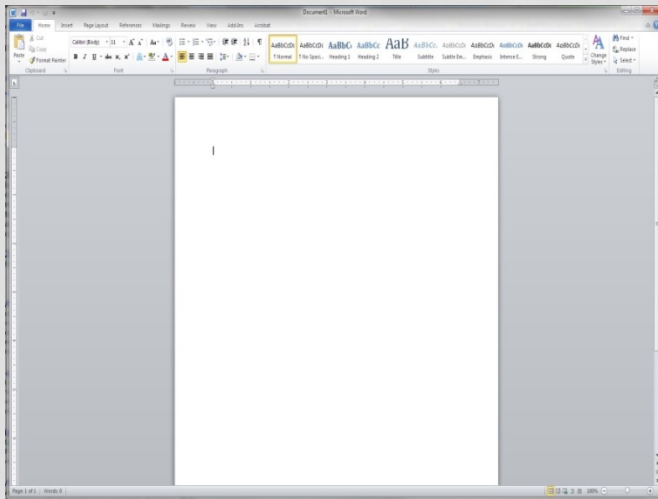
```
00000000 0000 0001 0001 1010 0010 0001 0004 0128
00000010 0000 0016 0000 0028 0000 0010 0000 0020
00000020 0000 0001 0004 0000 0000 0000 0000 0000
00000030 0000 0000 0000 0010 0000 0000 0000 0204
00000040 0004 8384 0084 c7c8 00c8 4748 0048 e8e9
00000050 00e9 6a69 0069 a8a9 00a9 2828 0028 fdfc
00000060 00fc 1819 0019 9898 0098 d9d8 00d8 5857
00000070 0057 7b7a 007a bab9 00b9 3a3c 003c 8888
00000080 8888 8888 8888 8888 288e be88 8888 8888
00000090 3b83 5788 8888 8888 7667 778e 8828 8888
000000a0 d61f 7abd 8818 8888 467c 585f 8814 8188
000000b0 8b06 e8f7 88aa 8388 8b3b 88f3 88bd e988
000000c0 8a18 880c e841 c988 b328 6871 688e 958b
000000d0 a948 5862 5884 7e81 3788 1ab4 5a84 3eec
000000e0 3d86 dc88 5cbb 8888 8888 8888 8888 8888
000000f0 8888 8888 8888 8888 8888 8888 8888 0000
00001000 0000 0000 0000 0000 0000 0000 0000
*
00001030 0000 0000 0000 0000 0000 0000 0000
0000103e
```

Digital File Container (.txt, .jpg, .tiff, .pdf, .xlsx, .docx, etc.)



Date
File Signatures
Owner

Rights
Structural Metadata
Styles Metadata



Header (XML)

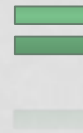
00000000	0000	0001	0004
00000010	0000	0000	0028
00000020	0000	0016	0010
00000030	6a69	0019	8888
00000040	5862	e8f7	8384

Payload (binaries)

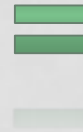
Digital File Container
(.docx –ZIP archive file with XML and binaries)

Example: Word Document

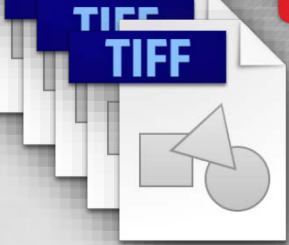
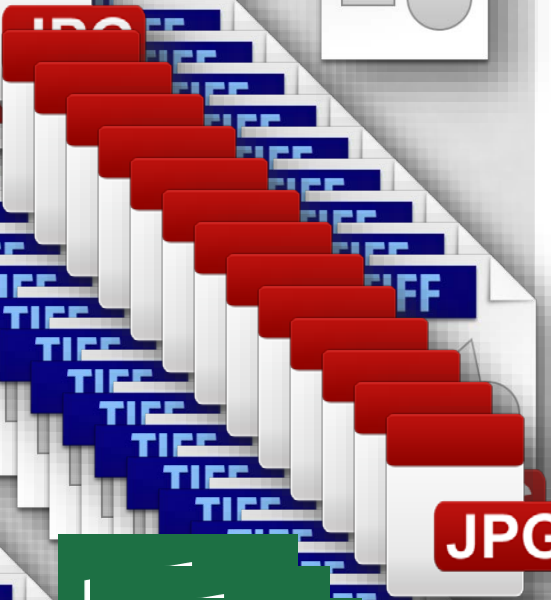
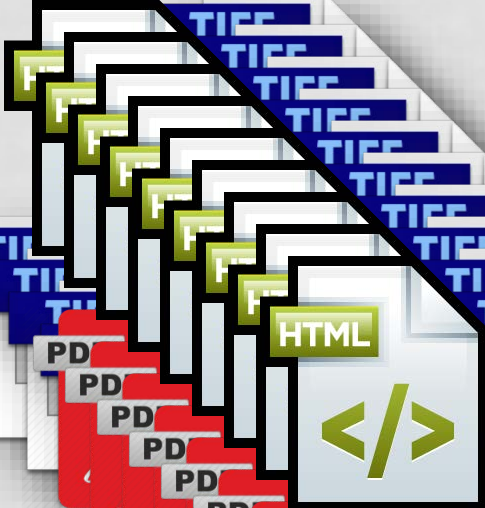
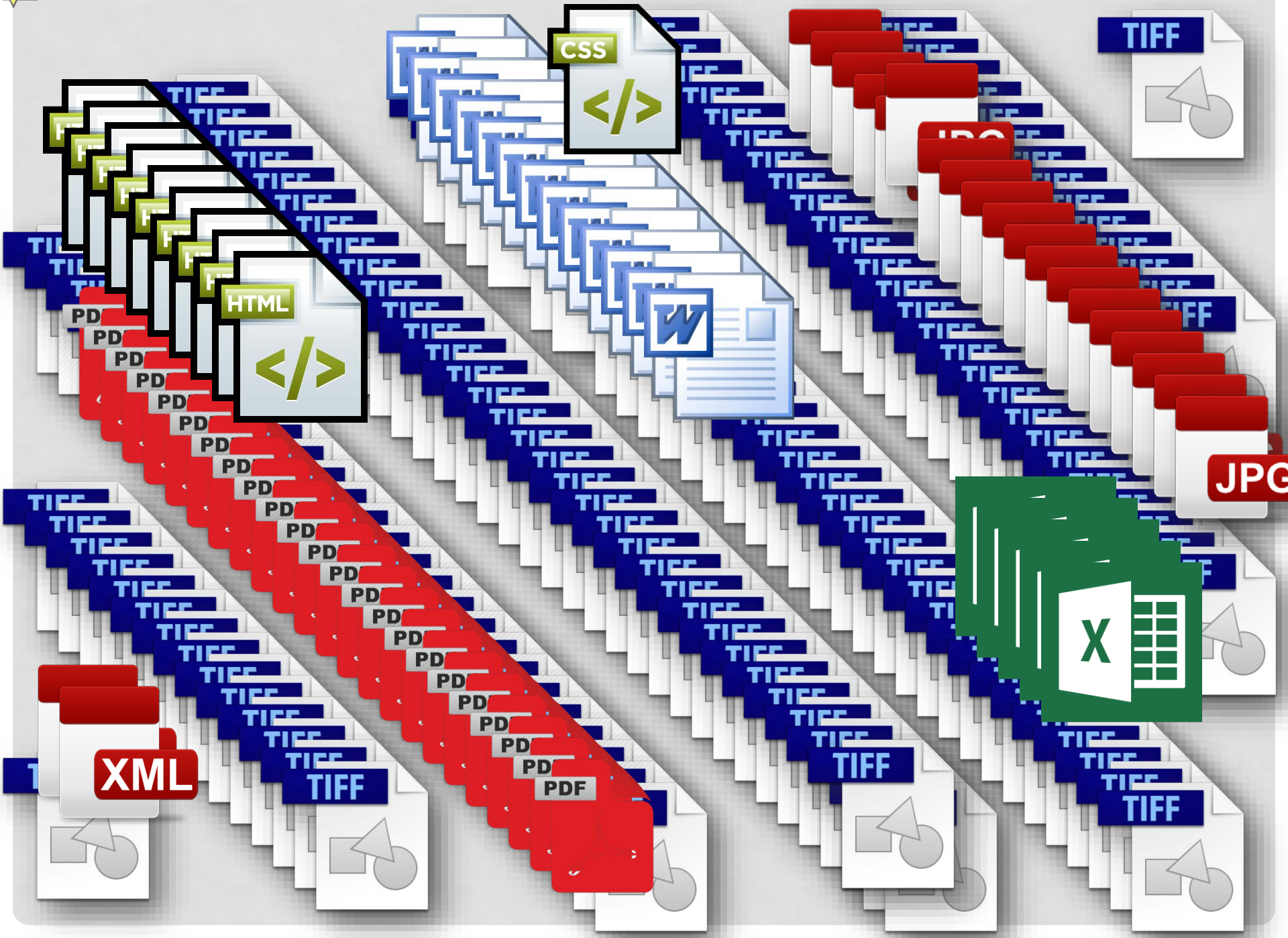




Data
Loss

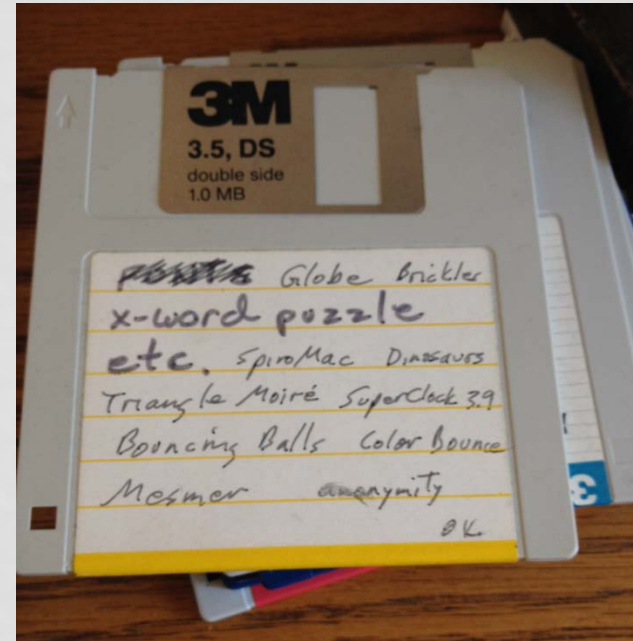


Data
Loss





Medium carries Information
and Meaning



Medium and Machines carry
Information and Meaning



STANDARDIZATION

PROCESSES FOR ACCESS & PRESERVATION

Human/Machine



Human/Machine

Administrative Metadata

Creator, Date, Title, Related Files, Rights, Contributor, Identifiers, File Names, etc.

Descriptive Metadata

Description of content, Keywords (tags), Subject-specific information, Coverage, Dates, etc.

Technical & Preservation Metadata

Software, Hardware, Structure, Format, Colorspace, Size, Checksum, Events/Actions, Dates, etc. Some of this metadata is automatically captured by originating software.

Header

00000000	0000	0001	0004
00000010	0000	0000	0028
00000020	0000	0016	0010
00000030	6a69	0019	8888
00000040	5862	e8f7	8384

payload

Digital File Container

(.txt, .doc, .jpg, .tiff, .png, .xlsx, .pdf)

Digital Object

Header

00000000	0000	0001	0004
00000010	0000	0000	0028
00000020	0000	0016	0010
00000030	6a69	0019	8888
00000040	5862	e8f7	8384

payload

XML declaration

```
<METS:fileGrp>  
<METS:file ID="MUSIC_FOLDER_DC" MIMETYPE="text/xml">  
<METS:FContent>  
<METS:xmlData>  
<oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"  
  xmlns:dc="http://purl.org/dc/elements/1.1/">  
  <dc:title>Music Folder</dc:title>  
  <dc:creator>Chris</dc:creator>  
  <dc:subject>music</dc:subject>  
  <dc:description>This folder contains my music files.</dc:description>  
</oai_dc:dc>  
</METS:xmlData>  
</METS:FContent>  
</METS:file>  
</METS:fileGrp>
```

XML CONTENT

Digital file to be preserved and XML sidecar



txt--manifest



md5--checksum



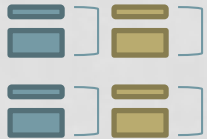
XML—collection level



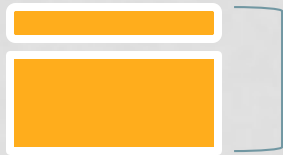
txt--collection description



pdf--rights documents



tiff & XML—digital objects



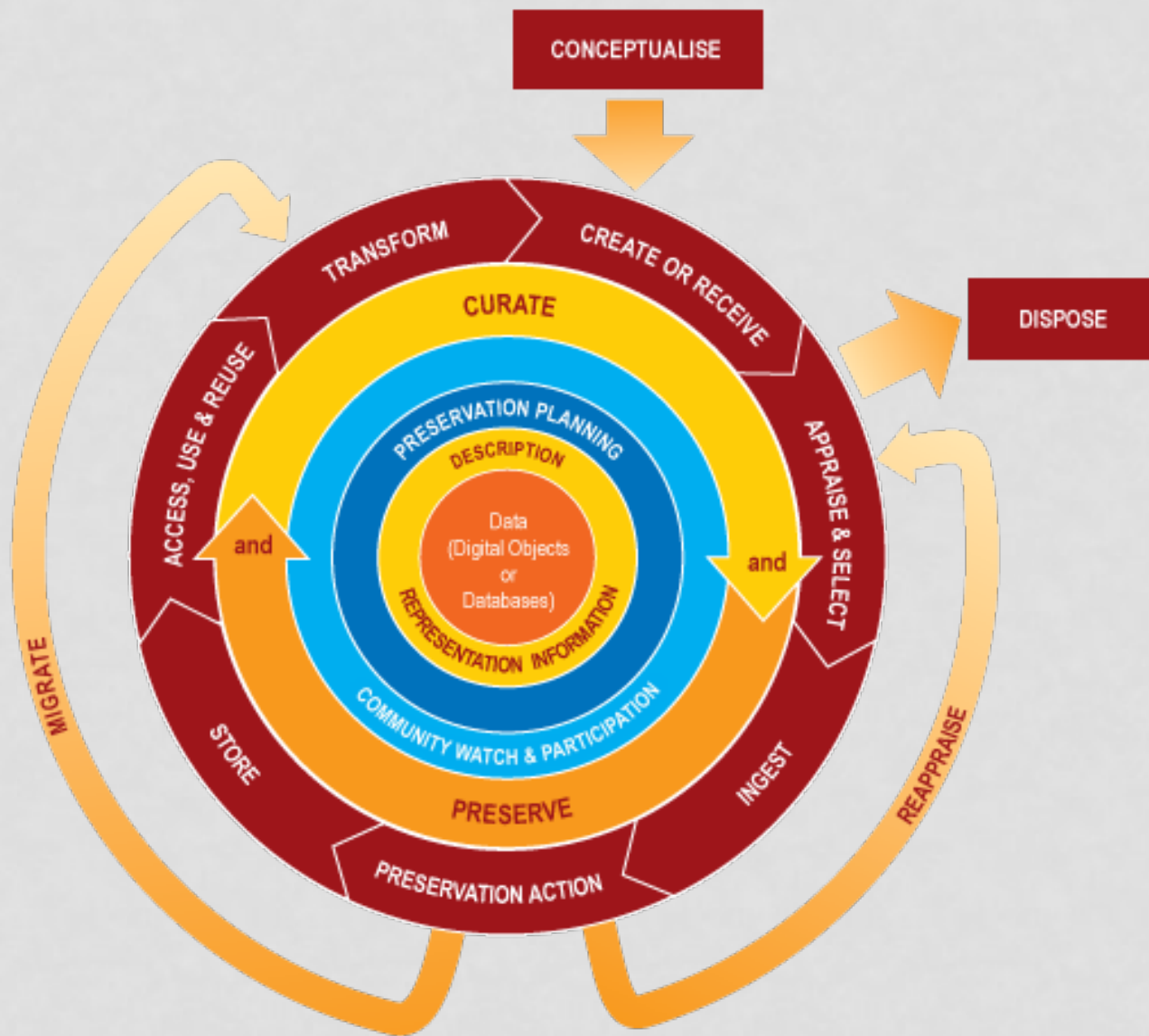
pst—communication files




Digital Object

Example
Submission Information
Package (SIP)

Ready for Preservation



Digital Curation Center Curation Lifecycle Model



Dissemination Information
Package (DIP)
(disposable)



Submission
Information
Package (SIP)

Archival
Information
Package (AIP)





```

0000000 0000 0001 0001 1010 0010 0001 0004 0128
0000010 0000 0016 0000 0000 0000 0010 0000 0000
0000020 0000 0001 0004 0000 0000 0000 0000 0000
0000030 0000 0000 0000 0010 0000 0000 0000 0004
0000040 0004 0004 0004 0718 0000 4740 0040 0000
0000050 0000 0000 0000 0000 0000 2028 0028 ffff
0000060 001c 1010 0010 0000 0000 0000 0000 0000
0000070 0007 707a 007a 0000 0000 0000 0000 0000
0000080 8888 8888 8888 8888 200a 0000 8888 8888
0000090 0000 0700 0000 0000 7007 770a 0020 0000
00000a0 001f 70bd 0010 0000 407c 505f 0014 0100
00000b0 000c 0000 0000 0000 0000 0000 0000 0000
00000c0 0010 0000 0000 0000 0000 0000 0000 0000
00000d0 0000 0000 0000 0000 0000 0000 0000 0000
00000e0 0000 0000 0000 0000 0000 0000 0000 0000
00000f0 0000 0000 0000 0000 0000 0000 0000 0000
0000100 0000 0000 0000 0000 0000 0000 0000 0000
0000110 0000 0000 0000 0000 0000 0000 0000 0000
000011a

```

Information added/changed

Files transferred

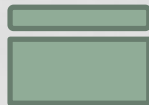


```

0000000 0000 0001 0001 1010 0010 0001 0004 0128
0000010 0000 0016 0000 0000 0000 0010 0000 0000
0000020 0000 0001 0004 0000 0000 0000 0000 0000
0000030 0000 0000 0000 0010 0000 0000 0000 0004
0000040 0004 0004 0004 0718 0000 4740 0040 0000
0000050 0000 0000 0000 0000 0000 2028 0028 ffff
0000060 001c 1010 0010 0000 0000 0000 0000 0000
0000070 0007 707a 007a 0000 0000 0000 0000 0000
0000080 8888 8888 8888 8888 200a 0000 8888 8888
0000090 0000 0700 0000 0000 7007 770a 0020 0000
00000a0 001f 70bd 0010 0000 407c 505f 0014 0100
00000b0 000c 0000 0000 0000 0000 0000 0000 0000
00000c0 0010 0000 0000 0000 0000 0000 0000 0000
00000d0 0000 0000 0000 0000 0000 0000 0000 0000
00000e0 0000 0000 0000 0000 0000 0000 0000 0000
00000f0 0000 0000 0000 0000 0000 0000 0000 0000
0000100 0000 0000 0000 0000 0000 0000 0000 0000
0000110 0000 0000 0000 0000 0000 0000 0000 0000
000011a

```

Formats migrated



```

0000000 0000 0001 0001 1010 0010 0001 0004 0128
0000010 0000 0016 0000 0000 0000 0010 0000 0000
0000020 0000 0001 0004 0000 0000 0000 0000 0000
0000030 0000 0000 0000 0010 0000 0000 0000 0004
0000040 0004 0004 0004 0718 0000 4740 0040 0000
0000050 0000 0000 0000 0000 0000 2028 0028 ffff
0000060 001c 1010 0010 0000 0000 0000 0000 0000
0000070 0007 707a 007a 0000 0000 0000 0000 0000
0000080 8888 8888 8888 8888 200a 0000 8888 8888
0000090 0000 0700 0000 0000 7007 770a 0020 0000
00000a0 001f 70bd 0010 0000 407c 505f 0014 0100
00000b0 000c 0000 0000 0000 0000 0000 0000 0000
00000c0 0010 0000 0000 0000 0000 0000 0000 0000
00000d0 0000 0000 0000 0000 0000 0000 0000 0000
00000e0 0000 0000 0000 0000 0000 0000 0000 0000
00000f0 0000 0000 0000 0000 0000 0000 0000 0000
0000100 0000 0000 0000 0000 0000 0000 0000 0000
0000110 0000 0000 0000 0000 0000 0000 0000 0000
000011a

```




DIGITAL PRESERVATION SYSTEM

BUILDING A FUNCTIONAL SYSTEM



Functional system – self-organizing, non-linear structure of conduct composed of synchronized distributed elements.

--V.G.Red'ko, D.V.Prokhorov, M.B.Burtsev, Theory of Functional Systems, Adaptive Critics and Neural Networks, in Proceedings of International Joint Conference on Neural Networks, Budapest, 2004, pp.1787-1792. Retrieved 15 September 2016 https://en.wikipedia.org/wiki/Theory_of_functional_systems



Standards

Policies

Hardware &
Software

Skilled Humans

Communication

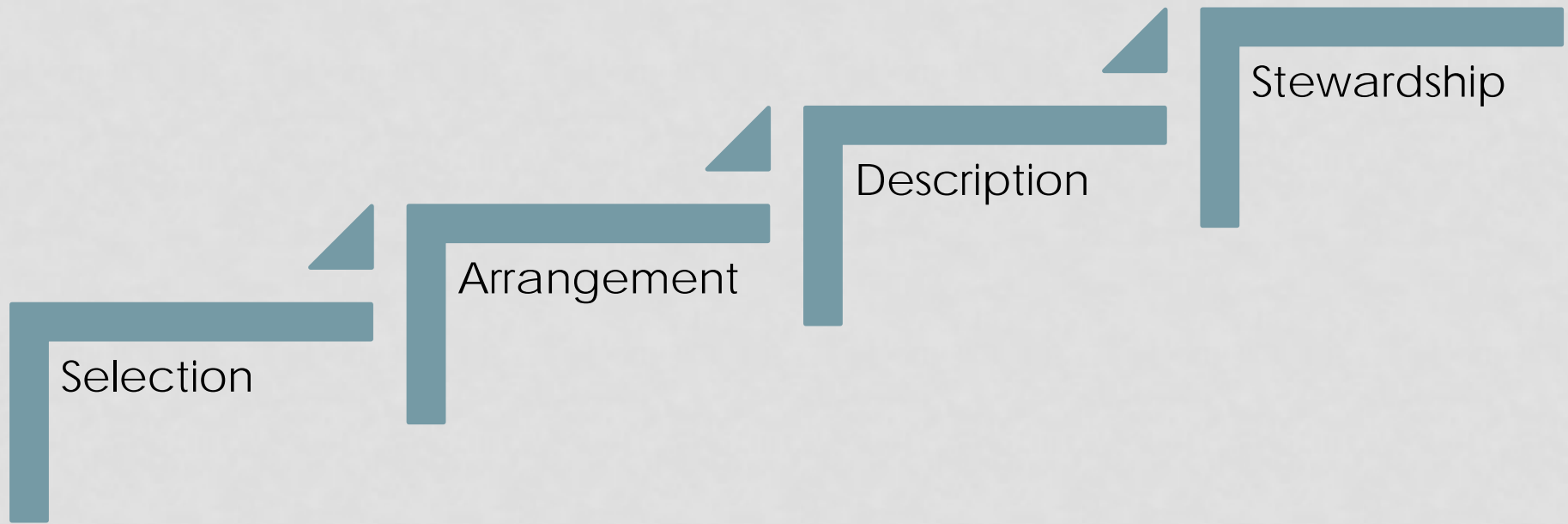
Procedures

Organizational
Commitments

Documentation

Consistent
Actions

Shared Purpose



Basic Archival Actions necessary for Digital File and Electronic Records Management



Digital File
Management

Documentation

Definitions, policies, roles and
responsibilities, and
procedures and workflows.

Organizational Commitment

Short/long term commitment to what is
developed (projects/processes), commitment
to the funding the humans, technologies,
space, time, training, etc.

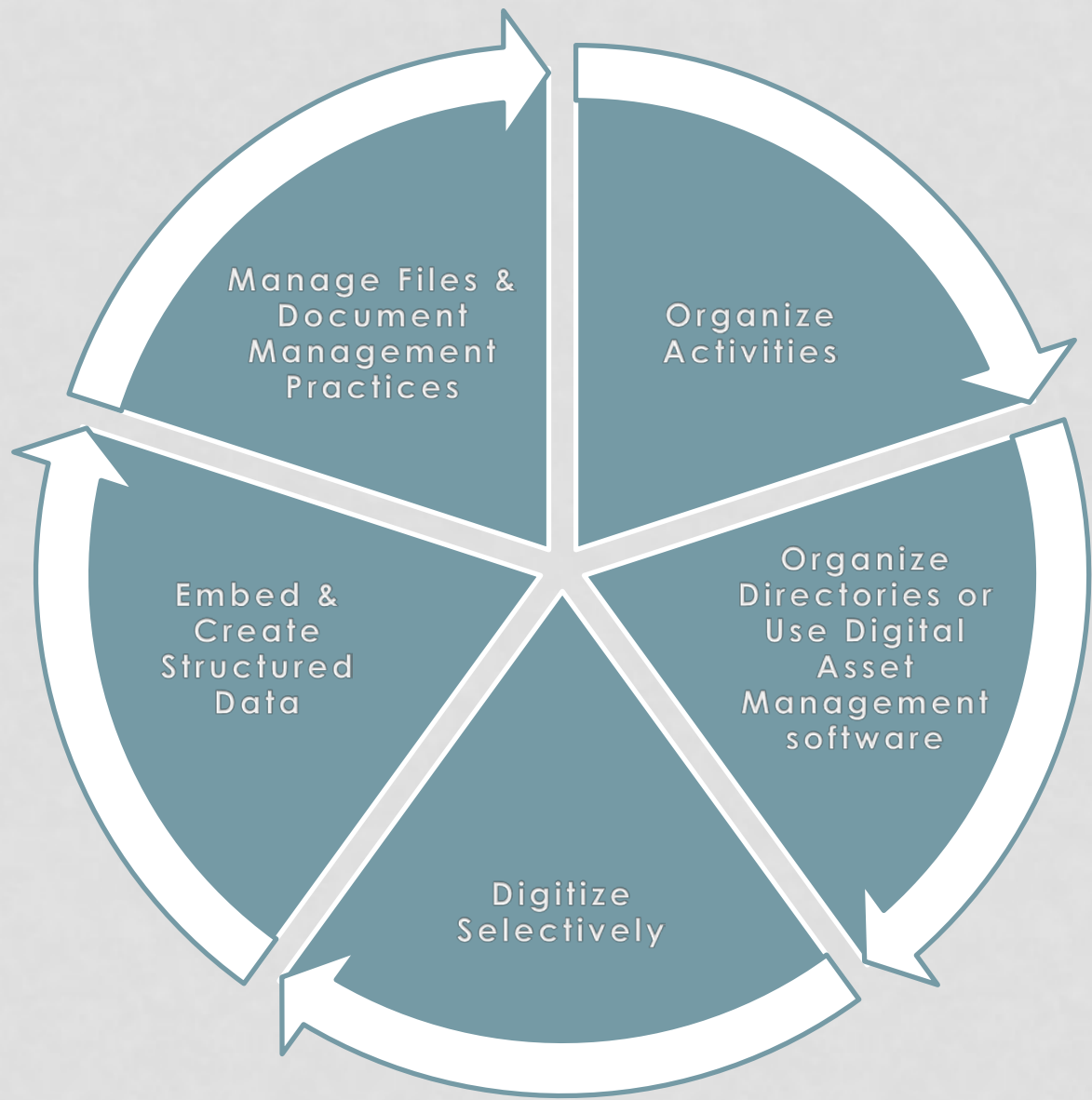
Foundation for Sustainability

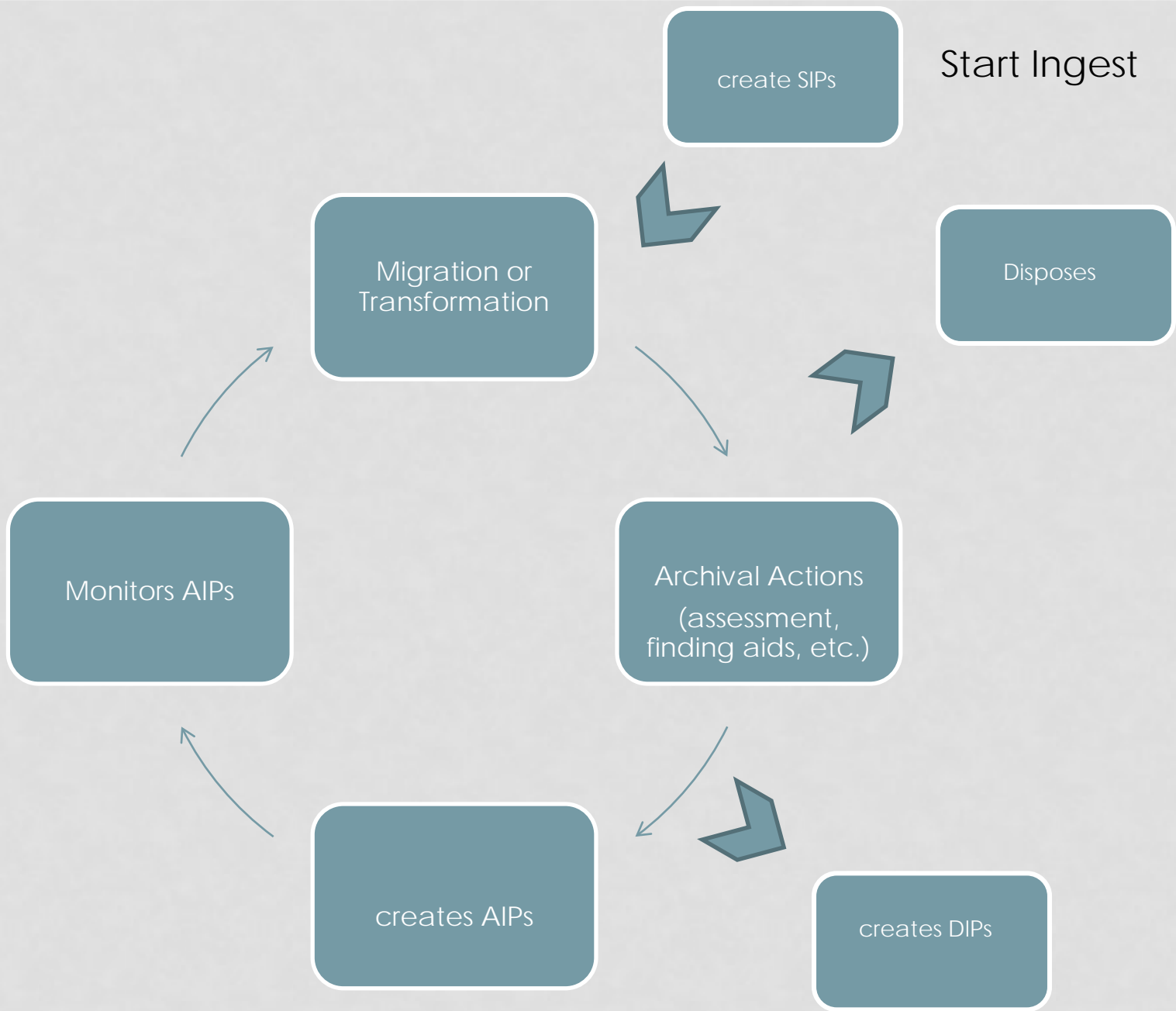



PRACTICAL ACTIONS & VOCABULARY



PLAN FOR PRESERVATION







	Level 1
Storage & Geographic Location	<ul style="list-style-type: none">-Two complete copies that are not collocated-For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system
File Fixity & Data Integrity	<ul style="list-style-type: none">-Check file fixity on ingest if it has been provided with the content-Create fixity info if it wasn't provided with the content
Information Security	<ul style="list-style-type: none">-Identify who has read, write, move and delete authorization to individual files-Restrict who has those authorizations to individual files
Metadata	<ul style="list-style-type: none">-Inventory of content and its storage location-Ensure backup and non-collocation of inventory
File Formats	<ul style="list-style-type: none">-When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs



2 disparate storage locations
(non-heterogeneous)

1 off-site (preferably)

STORAGE AND GEOGRAPHIC LOCATION





69CB82153BD4CC62394832C1B1954368

FILE FIXITY AND DATA INTEGRITY



5102BA3537092B6AD89D803A17CD7EFB

FILE FIXITY AND DATA INTEGRITY



Alan Holyoak. Speakeasy door with sliding panel & two-way mirror in Rookwood cloakroom. Butte, Montana (c) 2014 Alan Holyoak. Retrieved 16 September 2016
<http://idaholyoaks4.blogspot.com/2014/05/the-rookwood-speakeasy-butte-montana.html>

INFORMATION SECURITY



FILE FORMATS



Discussion