

DIGITAL BEDROCK

Managing Off-Cloud Digital Preservation

Linda Tadic, Founder/CEO
Digital Bedrock
ltadic@digitalbedrock.com

Designing Storage Architectures Meeting
Library of Congress, September 9, 2019

Digital Bedrock

- Managed OAIS-compliant digital preservation service, built specifically for active preservation actions to support any type organization or individual
- No license subscription, no hardware – we do the work as your extended staff and infrastructure.
- Any format and content type, although specialize in audiovisual formats
- No required metadata. Flexible database supports all types of data and metadata

Deep archive with object storage benefits and cloud connectivity

Deep archive:

Preserved files stored offline in 3 geographically dispersed secure locations

Object storage benefits:

All technical and embedded metadata extracted from files and indexed for unstructured data searching in clients' portals

Cloud connectivity (optional):

Location in direct connect facility with 10GB connections to major cloud provider servers

What we do

- Verify clients' hashes (checksums) upon receipt, but also create and verify SHA-512 for ongoing scheduled SHA-512 fixity checks
- Extract extensive technical and embedded metadata (becomes indexed)
- Retain original directory structure for context; track file relationships
- Monitor obsolescence vulnerabilities (Digital Object Obsolescence Database; aka the "DOOD")
- Retain audit trail of all actions
- Write three copies on LTO7 (LTFS), storing offline in geographically dispersed secure locations. Operations in an ISO 27001 compliant data center
- Migrate to future storage media
- Implement an open architecture (**no vendor lock-in and easy exit**)

Who we help

Clients are diversified, so our system must support all types of organizations and individuals, as well as all file format types.

While we can help any type organization, our current clients come from these categories:

- Media and Entertainment
- Creators (artists, filmmakers, composers, photographers)
- Public Broadcasters
- Libraries, Archives, Museums
- Law Firms

Approved FEDLINK vendor.

Preserved to date (since January 2017)

- 11.5 million files (x3 = 34+ million)
- 618 TB (x3 = 1.8+ PB written to tape)
- 100% accuracy

Digital Bedrock's cloud storage connectivity

Our operations are in a high security data center in Los Angeles. This enables us to offer clients hybrid off-cloud managed digital preservation services with cloud connectivity/storage.

- The data center is a direct connect facility for US-West regions for:

AWS, Microsoft Azure, Google Cloud, IBM Cloud

Benefit: secure, faster (10GB) connection than going over public internet. Data can be pushed up from DB's servers to the cloud provider, or pulled down from cloud to DB's servers for digital preservation or migration actions.

Add New SIP

Search SIPs

Filter by All Job status

SIP ID	Client	Active	Destination Name	Destination Type	Region	Files	Size	Created
S000000016 (B) Documents	Studio 6							
S000000015 (B*) Music	Studio 6	Completed	C, A, V, O, P, W, T			34	797.18 MB	2019-07-10
S000000014 (B*) Images_Documents	Studio 6	Completed	C, A, V, O, P, W, T			178	1.25 GB	2019-07-16
S000000013 (A*) Camera_Tests_RED	Studio 6	Completed	C, A, V, O, P, W, T			10	3.05 GB	2019-07-10
S000000012 (B*) Camera_Tests_ARRI_other	Studio 6	Completed	C, A, V, O, P, W, T			1,099	75.12 GB	2019-07-10
S000000011 (A*) Newsreels	Studio 6	Completed	C, A, V, O, P, W, T			23	175.28 MB	2019-07-10
S000000010 (A*) Movies_Cineraama_Others	Studio 6	Completed	C, A, V, O, P, W, T			16	220.08 GB	2019-07-10
S000000009 (A*) Axle_Video	Moving pictures, Inc. 09	Completed	C, A, V, O, P, W, T			16	1.85 GB	2017-06-19

Edit SIP S000000016

DETAILS MANIFESTS DESTINATIONS

Active	Destination Name	Destination Type	Region
+ Add destination			
There are no destinations to display.			

Target Destination

Destination Type

Amazon S3

Name

name of the destination

Notes

Region

region

Service URL

amazon s3 service url

Access Key

amazon s3 access key

Secret Key

••••••••

Bucket

amazon s3 bucket

Cancel

Save

302.3 GB STORAGE USED

Clear All Filters

Total Results 8 < 1 >

Deeper dive into Digital Bedrock

Digital Bedrock's off-cloud storage and fixity checks

- Client's files received on HDDs (we can loan), LTO, FileCatalyst, or via cloud.
- SIPs built using our package creator tool (de-dupes, verifies client checksums, creates SHA-512, retains directory structure and modification dates, creates manifest, copies files to delivery media)
- SHA-512 checked upon ingest into the digital preservation system.
- After digital preservation processing, simultaneously written to 3 copies LTO7 using LTFS. Only the client's data is written to their tapes.
- SHA-512 verified after files are written on all copies, before storage. Once verified, files deleted from servers (security and small carbon footprint)

Digital Bedrock's off-cloud storage and fixity checks

Scheduled fixity checks:

- For our first two years, ran fixity every 6 months on our local LTO7.
- After 8M files and no errors, we changed our policy to annual checks.
- No need to run a tape through the “read” cycle so frequently, so long as the tapes were verified throughout the writing process, and are stored in a stable environment with no use. We find LTO7 (barium ferrite magnetic particles) robust.

Our software

- **Package creator tool** (client selects files; tool de-dupes & packages files with file-level SHA-512 checksums, retaining original context)
- **Digital Preservation Application (DPA)**
- **Client portal**
- **LTFS software** to manage the tape libraries (formerly used 3rd party system but was too slow. Our software manages 8 drives and is 3x faster.)
- **Digital Object Obsolescence Database (DOOD)**

Patent pending for DIGITAL OBSOLESCENCE AVOIDANCE SYSTEMS AND METHODS, the key algorithm behind the DOOD. The DOOD tracks and monitors format vulnerabilities through research and proprietary algorithms. Just filed response to USPTO examiner's comments.

Digital Preservation Application (DPA)

Technical metadata extraction

We apply these tools:

Tika

Sox

JHOVE

ffprobe

ExifTool

Arri MetaExtractor

MediaInfo

Questions?

Linda Tadic

Founder/CEO

ltadic@digitalbedrock.com

www.digitalbedrock.com