

## The Quest for Digital Preservation — MathFest 2015 Bibliography

### Understanding checksums and hash functions

- “Checksum.” *Wikipedia, the Free Encyclopedia*, June 19, 2015.  
<https://en.wikipedia.org/w/index.php?title=Checksum&oldid=667621054>.
- “Comparison of SHA Functions.” *Wikipedia, the Free Encyclopedia*, July 20, 2015.  
[https://en.wikipedia.org/w/index.php?title=SHA-3&oldid=672206476#Comparison\\_of\\_SHA\\_functions](https://en.wikipedia.org/w/index.php?title=SHA-3&oldid=672206476#Comparison_of_SHA_functions).
- Crowley, Paul. “LShift - Why I Support the US Government Making a Cryptography Standard Weaker,” October 1, 2013.  
<http://www.lshift.net/blog/2013/10/01/why-i-support-the-us-government-making-a-cryptography-standard-weaker/>.
- “Cryptographic Hash Function.” *Wikipedia, the Free Encyclopedia*, July 28, 2015.  
[https://en.wikipedia.org/w/index.php?title=Cryptographic\\_hash\\_function&oldid=673479707](https://en.wikipedia.org/w/index.php?title=Cryptographic_hash_function&oldid=673479707).
- “Cyclic Redundancy Check.” *Wikipedia, the Free Encyclopedia*, June 1, 2015.  
[https://en.wikipedia.org/w/index.php?title=Cyclic\\_redundancy\\_check&oldid=66503428](https://en.wikipedia.org/w/index.php?title=Cyclic_redundancy_check&oldid=66503428).
- LeFurgy, Bill. “Hashing Out Digital Trust | The Signal: Digital Preservation.” Webpage, November 15, 2011. <http://blogs.loc.gov/digitalpreservation/2011/11/hashing-out-digital-trust/>.
- “List of Hash Functions.” *Wikipedia, the Free Encyclopedia*, July 7, 2015.  
[https://en.wikipedia.org/w/index.php?title=List\\_of\\_hash\\_functions&oldid=670380038](https://en.wikipedia.org/w/index.php?title=List_of_hash_functions&oldid=670380038).
- “MD5.” *Wikipedia, the Free Encyclopedia*, July 28, 2015.  
<https://en.wikipedia.org/w/index.php?title=MD5&oldid=673517851>.
- “SHA-1.” *Wikipedia, the Free Encyclopedia*, July 30, 2015.  
<https://en.wikipedia.org/w/index.php?title=SHA-1&oldid=673838295>.
- “SHA-2.” *Wikipedia, the Free Encyclopedia*, July 30, 2015.  
<https://en.wikipedia.org/w/index.php?title=SHA-2&oldid=673712593>.
- “SHA-3.” *Wikipedia, the Free Encyclopedia*, July 20, 2015.  
<https://en.wikipedia.org/w/index.php?title=SHA-3&oldid=672206476>.
- “The Keccak Sponge Function Family.” Accessed August 2, 2015.  
[http://keccak.noekeon.org/yes\\_this\\_is\\_keccak.html](http://keccak.noekeon.org/yes_this_is_keccak.html).

### Checksum collisions

- “Peter Selinger: MD5 Collision Demo.” Accessed August 2, 2015.  
<http://www.mathstat.dal.ca/~selinger/md5collision/>.
- Stevens, Marc. “Cryptanalysis of MD5 & SHA-1.” Special Purpose Hardware for Attacking Cryptographic Systems (workshop), March 17, 2012. <http://2012.sharcs.org/slides/stevens.pdf>.

### Checksum and Hash Function performance

- “Crypto++ 5.6.0 Benchmarks.” Accessed August 2, 2015.  
<http://www.cryptopp.com/benchmarks-amd64.html>.
- Duryee, Alex. “What Is the Real Impact of SHA-256? A Comparison of Checksum Algorithms.” AVPreserve.com, October 2014.  
[http://www.avpreserve.com/wp-content/uploads/2014/10/ChecksumComparisons\\_102014.pdf](http://www.avpreserve.com/wp-content/uploads/2014/10/ChecksumComparisons_102014.pdf).

“Hash Benchmark.” Accessed August 2, 2015.

<http://www.febooti.com/products/filetweak/members/hash-and-crc/hash-benchmark/>.

“Hash - When Is CRC More Appropriate to Use than MD5/SHA1? - Stack Overflow.” Accessed August 2, 2015.

<http://stackoverflow.com/questions/996843/when-is-crc-more-appropriate-to-use-than-md5-sha1>.

(See comment from Martin with speed test results.)

“Speed Comparison of Popular Crypto Algorithms.” Accessed August 2, 2015.

<http://www.cryptopp.com/benchmarks.html>.

“Speed Hashing.” Accessed August 2, 2015. <http://blog.codinghorror.com/speed-hashing/>.

### Data Loss

Steinhart, Gail. “A picture is worth 1000 words... back it up!” August 20, 2012. Accessed August 6, 2015.

<https://twitter.com/gailst/status/237525591547580416>

Vines, T. H., Albert, A. Y. K., Andrew, R. L., Débarre, F., Bock, D. G., Franklin, M. T., ... Rennison, D. J. (2013). *The Availability of Research Data Declines Rapidly with Article Age*. *Current Biology*, 24(1), 94–97. [doi:10.1016/j.cub.2013.11.014](https://doi.org/10.1016/j.cub.2013.11.014)

### How Many Copies Keep Stuff Safe

Greenan, Kevin M., James S. Plank, and Jay J. Wylie. “Mean Time to Meaningless: MTDDL, Markov Models, and Storage System Reliability.” In *Proceedings of the 2nd USENIX Conference on Hot Topics in Storage and File Systems*. Boston, Mass.: USENIX Association, 2010.

[http://www.usenix.org/legacy/events/hotstorage10/tech/full\\_papers/Greenan.pdf](http://www.usenix.org/legacy/events/hotstorage10/tech/full_papers/Greenan.pdf).

Rosenthal, David S. H. “Bit Preservation: A Solved Problem?” *International Journal of Digital Curation* 5, no. 1 (June 22, 2010): 134–48. [doi:10.2218/ijdc.v5i1.148](https://doi.org/10.2218/ijdc.v5i1.148).

<http://www.ijdc.net/index.php/ijdc/article/view/151>.

———. “DSHR’s Blog: ‘Petabyte for a Century’ Goes Main-Stream.” *DSHR’s Blog*, October 6, 2010.

<http://blog.dshr.org/2010/09/petabyte-for-century-goes-main-stream.html>.

———. “DSHR’s Blog: The Half-Empty Archive.” *DSHR’s Blog*, March 31, 2014.

<http://blog.dshr.org/2014/03/the-half-empty-archive.html>.

———. “Keeping Bits Safe - ACM Queue.” *ACM Queue* 8, no. 10 (October 2010) (October 1, 2010).

<http://queue.acm.org/detail.cfm?id=1866298>.

### Digital Preservation Networks

“Academic Preservation Trust.” Accessed August 2, 2015. <http://academicpreservationtrust.org/>.

“DuraCloud Introduction | DuraCloud.” 2015. Accessed August 6. <http://www.duracloud.org/>.

German, Chip, Andrew Diamond, Nathan Tallman, Linda Newman, and Jamie Little. “The Academic Preservation Trust: Report on First Months of Production.” In *CNI: Coalition for Networked Information*. Seattle, Washington, 2015.

<http://www.cni.org/topics/digital-preservation/the-academic-preservation-trust-report-on-first-months-of-production/>.

“Global & Private LOCKSS Networks | LOCKSS.” Accessed August 2, 2015.

<http://www.lockss.org/community/networks/>.

MetaArchive. “Methodology | MetaArchive.” Accessed August 2, 2015.

<http://www.metaarchive.org/methodology>.

“Portico - A Digital Preservation and Electronic Archiving Service.” 2015. Accessed August 6.  
<http://www.portico.org/digital-preservation/>.

Reich, Victoria, and D. Rosenthal. “Distributed Digital Preservation: Lots of Copies Keep Stuff Safe.” In *Proceedings Indo-US Workshop on International Trends in Digital Preservation March, 24–25, 2009*. <http://humgroup.lockss.org/locksswiki/files/ReichIndiaFinal.pdf>.

“The Digital Preservation Network.” *The Digital Preservation Network*. Accessed August 2, 2015.  
<http://www.dpn.org/>.

### Digital Repository Open Source Software

“DSpace | DSpace is a turnkey institutional repository application.” Accessed August 6, 2015.  
<http://www.dspace.org/> .

“Fedora Repository | Fedora is a general-purpose, open-source digital object repository system.” Accessed August 6, 2015. <http://fedora-commons.org/> .

“Hydra Project.” Accessed August 6, 2015. <http://projecthydra.org/> .

“Islandora Website.” Accessed August 6, 2015. <http://islandora.ca/> .

### Format Obsolescence

Arms, Caroline R., Carl Fleischhauer, and Kate Murray. “Sustainability of Digital Formats: Planning for Library of Congress Collections.” Webpage, 2013. <http://www.digitalpreservation.gov/formats/>.

Holden, Matt. “Preserving the Web Archive for Future Generations : Practical Experiments with Emulation and Migration Technologies.” In *Conference Proceedings, 783–95*. Vancouver, Canada: United Nations Educational, Scientific and Cultural Organization, 2012.  
[http://cisra.org/docs/UNESCO\\_MOW2012\\_Proceedings\\_FINAL\\_ENG\\_Compressed.pdf](http://cisra.org/docs/UNESCO_MOW2012_Proceedings_FINAL_ENG_Compressed.pdf).

Jackson, Andrew N. “Formats over Time: Exploring UK Web History.” In *arXiv:1210.1714 [cs]*. Toronto, Canada, 2012. <http://arxiv.org/abs/1210.1714>.

“Recommended Format Specifications - Resources (Preservation, Library of Congress).” Web page, 2015. <http://www.loc.gov/preservation/resources/rfs/TOC.html>.

Rosenthal, David S. H. “DSHR’s Blog: Formats through Time.” *DSHR’s Blog*, October 9, 2012.  
<http://blog.dshr.org/2012/10/formats-through-time.html>.

### Format Analysis Tools

Apache Software Foundation. “Apache Tika - Apache Tika.” Accessed August 2, 2015.  
<https://tika.apache.org/>.

“Archivematica: Open-Source Digital Preservation System.” 2015. Accessed August 6.  
<https://www.archivematica.org/en/>.

“JHOVE2: The Next-Generation Architecture for Format-Aware Characterization.” Accessed August 2, 2015. <https://bitbucket.org/jhove2/main/wiki/Home>.

The National Archives. “Download DROID: File Format Identification Tool.” *The National Archives*. Accessed August 2, 2015.

<http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>.

## Trusted Digital Repositories and Certification

Center for Research Libraries. "ISO 16363 / TDR | CRL." Accessed August 2, 2015.

<https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/iso16363>.

Consultative Committee for Space Data Systems. "AUDIT AND CERTIFICATION OF TRUSTWORTHY DIGITAL REPOSITORIES." Consultative Committee for Space Data Systems, 2011. <http://www.ccsds.org/publications/archive/652x0m1.pdf>.

Dale, Robin L., and Bruce Ambacher. "Trustworthy Repositories Audit and Certification: Criteria and Checklist. Report of the RLG-NARA Task Force on Digital Repository Certification, February 2007." Online Computer Library Center and the Center for Research Libraries, February 2007. [https://www.crl.edu/sites/default/files/d6/attachments/pages/trac\\_0.pdf](https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf).

International Organization for Standardization, (first). "ISO 16363:2012(en) Audit and Certification of Trustworthy Digital Repositories," September 2011. <https://www.iso.org/obp/ui/#iso:std:iso:16363:ed-1:v1:en>.