

HTRC Data Capsule

Create an HTRC Data Capsule from the HTRC website.

About the HTRC Data Capsule

The HTRC Data Capsule service gives a researcher or educator a secure, virtual computer (called a "Capsule") for a limited time for analytical access to the digitized works of the HathiTrust. Currently the Data Capsule service can only access a snapshot of public domain content from the HathiTrust Digital Library, but soon will provide analytic access to the full HathiTrust Digital Library. The HTRC Data Capsule gives researchers and educators the ability to carry out analytics that is non-consumptive** only. A researcher's capsule therefore has restrictions on its use, particularly on how and when the products created by analysis tools leave the Capsule. Data products leaving a Data Capsule must undergo results review prior to release to ensure they meet the HTRC's policy for [non-consumptive data exports](#).

The HTRC Data Capsule currently enables access to the works in HT that are in the public domain. Preliminary access through the capsules to digitized works in-copyright is anticipated winter 2018.

The HTRC Data Capsule system was prototyped through funding from the Alfred P. Sloan Foundation (2011-2015); the final report is available here: [Final report](#). Extension of the HTRC Data Capsule project to larger compute resources and better integration with the HTRC worksets was recently funded by a grant from the Andrew T. Mellon Foundation (2016-2018).

** From the [2010 Authors Guild vs Google amended settlement agreement](#): "Non-Consumptive Research" means research in which computational analysis is performed on one or more Books, but not research in which a researcher reads or displays substantial portions of a Book to understand the intellectual content presented within the Book." Non-consumptive analytics includes image analysis, text extraction, textual analysis and information extraction, linguistic analysis, automated translation, and indexing and search.

Get More Information

- [Frequently Asked Questions \(FAQ\)](#)
- [Getting started with the HTRC Data Capsule](#)
- [HTRC Data Capsule Step-by-Step Guides](#)
- [Quick Video Demos](#)
- [For Developers: Developer API](#)
- [Non-consumptive Use Research Policy](#)

Availability

The HTRC Data Capsule system is available via the HTRC Portal as part of HTRC v3.0 released in beta version 16 January 2015 at <https://analytics.hathitrust.org/>

References

Kevin Borders, Eric Vander Weele, Billy Lau, and Atul Prakash, Protecting Confidential Data on Personal Computers with Storage Capsules. Proceedings of the 18th USENIX Security Symposium, Aug. 2009.

Zeng, J., Ruan, G., Crowell, A., Prakash, A., & Plale, B. (2014, June). Cloud Computing Data Capsules for Non-Consumptive Use of Texts. In Proceedings of the 5th ACM workshop on Scientific cloud computing (pp. 9-16). ACM.

Plale, Beth; Prakash, Atul; McDonald, Robert (2015). The Data Capsule for Non-Consumptive Research: Final Report. Available from <http://hdl.handle.net/2022/19277>