



The CLOCKSS Archive: A Primer

August 2021

A “dark archive” currently housing over 46 million journal articles, 260,000 books, and a wide array of supplementary materials and metadata information, CLOCKSS preserves a growing corpus of scholarly content. This unique service assures publishers and libraries that the content they steward will withstand potential technological, economic, environmental, and political disruptions and failures and will always be available to those who want to access it.

Technology

CLOCKSS runs on the LOCKSS technology that was developed by Stanford University libraries in 1999. Digital content is stored for the long term in the CLOCKSS Archive with no user access unless a “trigger” event (defined below) occurs.

CLOCKSS (Controlled Lots of Copies Keep Stuff Safe) operates [12 archive nodes](#) at leading academic institutions worldwide preserving the authoritative versions of journal articles, books, supplementary materials, and metadata information. The core of CLOCKSS's preservation capabilities is a polling-and-repair mechanism, by which peer systems (the nodes) continuously validate the integrity of the common data they each hold. In this way, though the contents of the archive are “dark”, CLOCKSS maintains a high degree of confidence that the data are healthy over time, and that the original content can be made available intact in the case of a trigger event. CLOCKSS was [certified](#) as a Trusted Repository by the Center for Research Libraries in 2014, and recertified in 2018. CLOCKSS has the highest score for Technologies, Technical Infrastructure, and Security, and for Organizational Infrastructure. CLOCKSS's total score of 14 out of a possible 15 is the highest of any of the certified repositories.

Publishers submit their content by giving CLOCKSS permission to [crawl](#) their platform, or by providing the content by [file-transfer](#).

Governance by the Community

An independent 501(c)(3) non-profit charitable organization, CLOCKSS is run by a [Board of Directors](#) representing 12 libraries and 12 publishers. Together, they discuss issues and make policy. CLOCKSS began operations in 2010 and is financially strong.

Triggers

If content that is held in the Archive disappears (or is about to disappear) from the Web, CLOCKSS will “trigger” it for Open Access. Since CLOCKSS launched in 2009, 66 journals comprising ~15,000 articles have been [triggered](#) – a very small fraction of the ~25,000 journal titles that the Archive preserves, and no books have been triggered.

Publisher Participation and Library Support

Publishers sign an agreement with CLOCKSS, which defines rights, responsibilities, and fees. Libraries make contributions to CLOCKSS, in recognition of their responsibility as stewards of scholarly literature. The [fees](#) are posted on the CLOCKSS website.