

How Do Levels of Description Affect Discoverability of the Web Archives at the Library of Congress?

CARLYN OSBORN

Abstract: Since its creation in 2000, the Web Archiving program at the Library of Congress has created varying levels of description for different categories of their collections. These various levels of description have created an opportunity to observe the effect that description level may have on discoverability and access.

This poster will explore how various levels of metadata influence user interactions with web archive collections. Using data drawn from the publicly available JSON API and web traffic statistics made available by the Library of Congress, we ask: Do more subject headings make web archives more discoverable? Is there an effect from how many subject headings there are? Do web archives with detailed records receive more traffic from search engines than those with minimal records? Does being a part of a larger collection have more of an impact on findability than the number of subject headings?

As more and more web archives are described and made available (17,425 web archives to date), the Library of Congress seeks to be as strategic as possible in making its materials discoverable and accessible. This work is a direct result of an intra-institutional partnership at the Library of Congress, where different divisions within the Library have collaborated on improving discovery for web archives. How do we take the results of this research and apply it to existing practices at the Library of Congress? Do more exhaustive records really increase the use of the collections?

About the Author:

Carlyn Osborn is a Librarian (Digital Collections Specialist) in the Digital Content Management Section at the Library of Congress. She graduated from the Johns Hopkins University with a B.A. in the History of Science, Medicine, and Technology (2014) and from the University of Maryland College of Information Studies with a M.L.I.S., specializing in Archives and Digital Curation (2017). At the Library of Congress, Carlyn manages born-digital collections, works with metadata for digital objects, and provides digital content management support to staff across the Library.