Archives Unleashed

Fall 2020 Newsletter

Welcome to the Fall edition of the Archives Unleashed newsletter, where you can catch up on all the latest news.

- The Next Phase
- Article Highlights
- AU Reading List
- Recent Workshops + Presentations
- Special Congratulations

The Next Phase

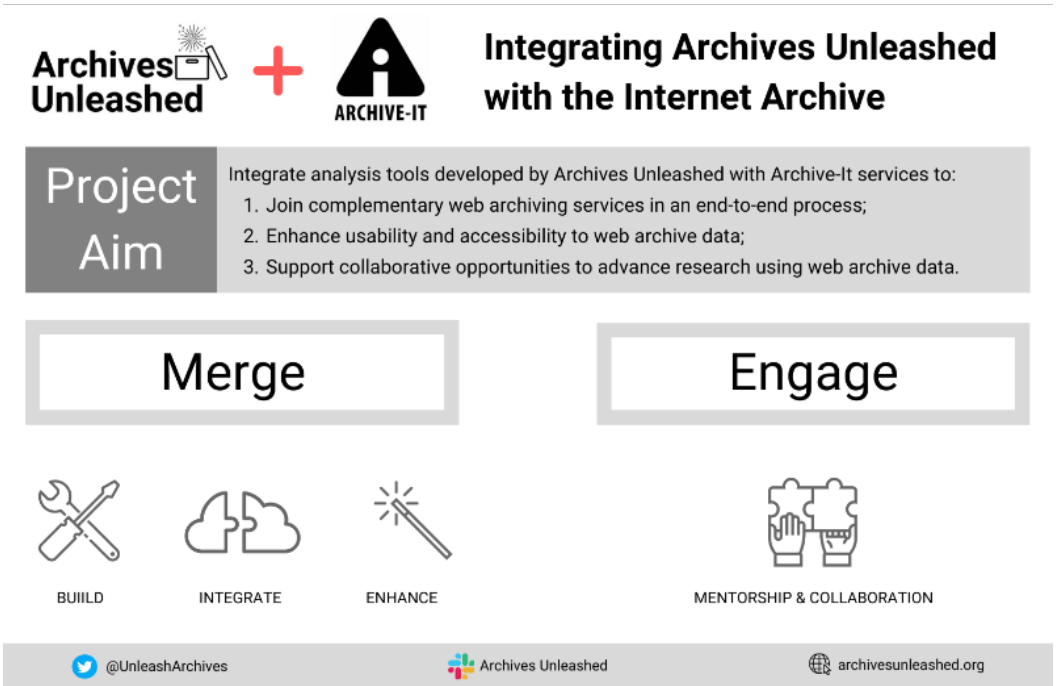
Archives Unleashed + ARCHIVE-IT

This past summer, the Archives Unleashed Project announced we would be continuing our work in a second Andrew W. Mellon Foundation grant. Our original team of investigators will be joined by colleagues from the Internet Archive to integrate and blend complementary services and tools to broaden

access to web archives.

As our team begins work, we wanted to share what our priorities will look like for this second phase!

You can read more in our recent Medium Post: [Ding, Ding, Round Two](#).

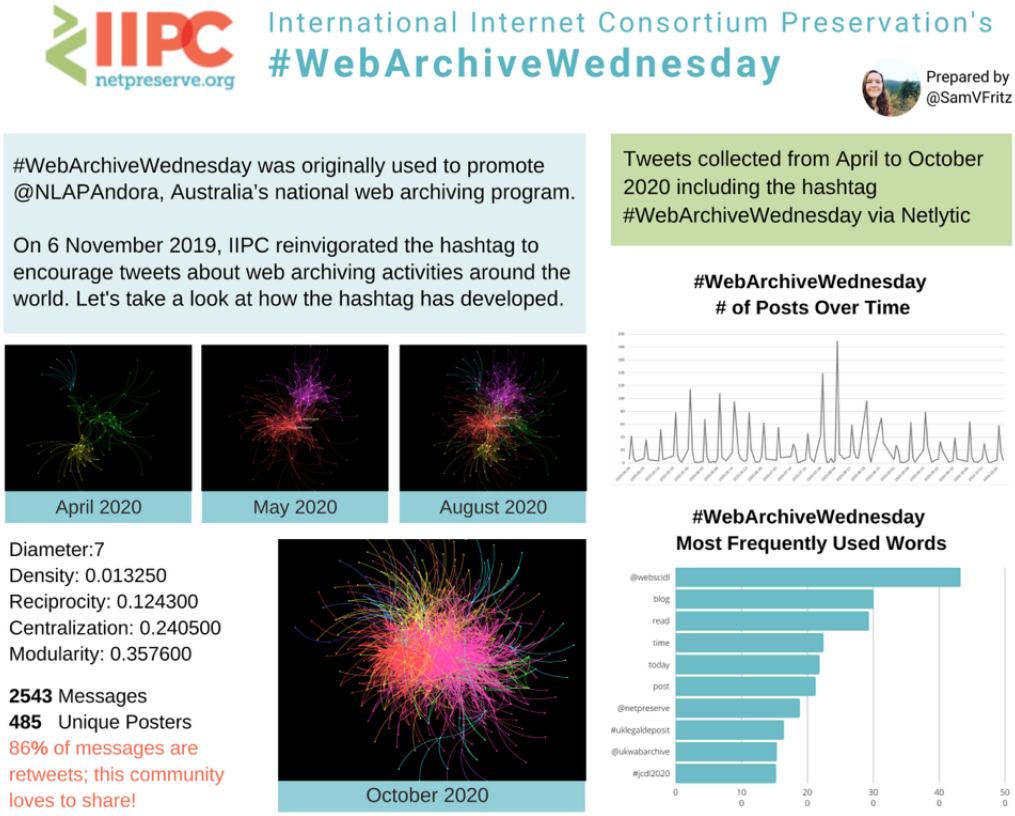


Article Highlights

#WebArchiveWednesday: A Community Conversation

Thanks to the folks at the International Internet Preservation Consortium (IIPC), the community has an opportunity to contribute to regular Wednesday discussions using the #WebArchiveWednesday Twitter hashtag.

As a special way of celebrating a successful year of engagement, our project manager has put together a short analysis of community tweets.



AU Reading List

Check out recent articles written by the Archives Unleashed Team and collaborators!

The Archives Unleashed Project: Technology, Process, and Community to Improve Scholarly Access to Web Archives

The main contribution articulated in this paper is a process model that decomposes scholarly inquiries into four main activities: filter, extract, aggregate, and visualize. Based on the insight that these activities can be disaggregated across time, space, and tools, it is possible to generate "derivative products", using our Archives Unleashed Toolkit, which serves as useful starting points for scholarly inquiry.

Scholars can download these products from the Archives Unleashed Cloud and manipulate them just like any other dataset, thus providing access to web archives without requiring any specialized knowledge. Over the past few years, our platform has processed over a thousand different collections from over two hundred users, totalling around 300 terabytes of web archives.

This was prepared and presented for the Joint Conference on Digital Libraries in 2020.

Article Link
Preprint
Presentation

Building Community at Distance: A Datathon During COVID-19

by Samantha Fritz, Ian Milligan, Nick Ruest, and Jimmy Lin.

This article reflects on the challenges and opportunities that underlie the transition of in-person events to online ones in the wake of COVID-19. The authors use their experience of running the first virtual Archives Unleashed Datathon to discuss the broader implications for transitioning events online. It gives both practical recommendations to event organizers as well as broader reflections on the role of digital libraries during the COVID-19 pandemic and beyond.

This article is part of a special issue of Digital Library Perspectives on "Digital libraries and COVID-19: responding to a global emergency.

Open Access Article

Content-Based Exploration of Archival Images Using Neural Networks

by Tobi Adewoye, Xiao Han, Nick Ruest, Ian Milligan, Samantha Fritz, and Jimmy Lin

We present DAIRE (Deep Archival Image Retrieval Engine), an image exploration tool based on latent representations derived from neural networks, which allows scholars to "query" using an image of interest to rapidly find related images within a web archive. This work represents one part of our broader effort to move away from text-centric analyses of web archives and scholarly tools that are direct reflections of

methods for accessing the live web. This short piece describes the implementation of our system and a case study on a subset of the GeoCities web archive.

This piece was prepared and presented for the Joint Conference on Digital Libraries in 2020.

Article Link
Preprint
Poster Demo
Video Summary

We Could, but Should We? Ethical Considerations for Providing Access to GeoCities and Other Historical Digital Collections

by Jimmy Lin, Ian Milligan, Douglas Oard, Nick Ruest, and Katie Shilton.

This article provides an introspective and cross-disciplinary approach to the questions raised around technology ethics. The authors consider the implications and impact of analyzing digital collections, using a 4-Terabyte GeoCities web archive collection as a case study.

This paper was originally intended to be presented at CHIIR 2020 in Vancouver (Conference on Human Information Interaction and Retrieval), however, the event was cancelled due to COVID-19.

This paper can be accessed through the ACM Digital Library and is also available as a preprint.

Workshops + Presentations

Archive-It Partner Meeting

The 2020 Archive-It Virtual Partner Meeting was held on October 7, 2020, with over two dozen presentations, sessions and discussion groups.

Ian Milligan presented on the project's second phase which will see the integration

of the Archives Unleashed Cloud with Archive-It.

- Presentation Slides
- Presentation Recording

Special Congratulations

The Archives Unleashed Project would like to extend congratulations to Dr. Sarah McTavish who recently and successfully defended her dissertation "The Queer Eternal September: LGBTQ Identity on the Early Internet and Web." We are so grateful to have had Sarah as a collaborator on our project, she created several of our Cloud Learning guides and was an avid tester of the Toolkit and Cloud. We look forward to seeing where research opportunities will take Dr.McTavish next!

Get Involved

Interested in getting involved with the Archives Unleashed Project? Connect with our team and help grow our community

Follow us on Twitter
Join our Slack group
Participate on Github
Subscribe to our newsletter
Provide feedback on your experiences with our tools and platforms
Share our news with colleagues and friends

Twitter GitHub Website Email

The Archives Unleashed Project aims to make petabytes of historical internet content accessible to scholars and others interested in researching the recent past. Supported by a grant from the Andrew W. Mellon Foundation, the project will develop web archive search and data analysis tools to enable scholars and librarians to access, share, and investigate recent history since the early days of the World Wide Web.

The Archives Unleashed Project
200 University Avenue West | Waterloo, ON N2L 3G1

This email was sent to <<Email Address>>
why did I get this? unsubscribe from this list update subscription preferences
Archives Unleashed · 200 University Ave W · Waterloo, On N2L 3G1 · Canada