Archives Unleashed

Summer 2023 Newsletter

Welcome to the final issue of the Archives Unleashed Newsletter!

Over the past six years, our team has thoroughly enjoyed connecting and working with individuals who are curious about web archives.
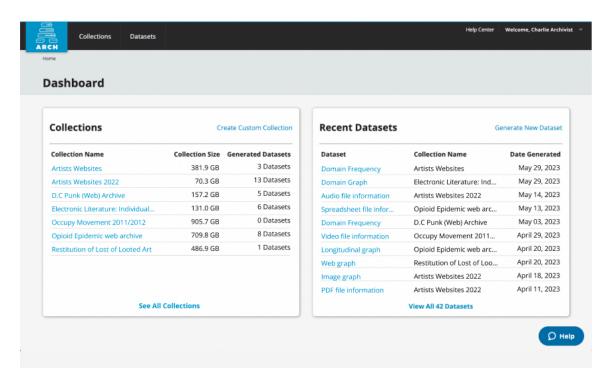
Our work has been motivated by the vision of reducing the barriers facing scholars when working with born-digital data at scale.

It's been a privilege to witness the changing shape of this community and the many opportunities that still exist to support research, skills-building, and knowledge-sharing within and

beyond the web archiving landscape.

Thank you for joining us on this journey!

Just In!

Launching the Archives Research Compute Hub

Build, Access, Analyze: Introducing ARCH (Archives Research Compute Hub)

Alongside our friends at the Internet Archive, we [were very excited to announce the launch of ARCH](#) (Archives Research Compute Hub), a new research and education service that lets users easily build, access, and analyze digital collections computationally at scale. ARCH brings together the Internet Archive's experience supporting computational research for more than a decade by

providing large-scale data to researchers and dataset-oriented service integrations like ARS (Archive-it Research Services) with the Archives Unleashed project. The development of ARCH was supported in part by generous funding from the Mellon

supported in part by generous funding from the Mellon Foundation.

Read the Full Announcement from the Internet Archive: https://blog.archive.org/2023/06/26/build-access-analyze-introducing-arch-archives-research-compute-hub/

To learn more about ARCH, please reach out via the following form.

ARCH Website: https://webservices.archive.org/pages/arch
ARCH Help Center: https://arch-webservices.zendesk.com/hc/en-us

Celebrating Research Opportunities with Web Archive through Archives Unleashed Cohort Projects

Following the successes of the Archives Unleashed datathon series (2017–2020), the Archives Unleashed Cohort Program launched in 2021 to foster research engagement with web archives.

The program has seen two cycles [2021/2022 and 2022/2023] of

year-long intensive collaborations to support research with web archives as a primary data source and scholarly objects. The Archives Unleashed Cohort roster includes ten incredible projects by fifty researchers from twenty institutions.

At the end of June, we hosted our second group of Cohort participants at the beautiful Internet Archive Canada facility in Vancouver. This was our final event and marked the culmination of our six years as the Archives Unleashed project.

Teams showcased their remarkable research accomplishments. We celebrate the achievements of these innovative researchers through an exclusive blog post: Celebrating Research Opportunities with Web Archive through Archives Unleashed Cohort Projects. The post provides a comprehensive overview of the ten projects that emerged from our Cohort Program, demonstrating the potential and opportunities of working with web archives.

Cohorts have published their research through a number of scholarly avenues — from journal publications to conferences to community engagement and advocacy work. As these projects offer high-quality use cases for working with web archive data, we invite you to explore our Cohort Projects Summaries Report, accessible at: https://bit.ly/AUCohortProjects. We extend our sincere gratitude to the cohort researchers for generously sharing their content for this report, enriching our understanding and inspiring future research endeavours.

Archives Unleashed Toolkit (AUT) 1.2.0 Release

Archives Unleashed Toolkit 1.2.0: A Scalable Way to Explore Web Archives

Archives Unleashed
36 subscribers

Subscribe

👍 1   👎   ➦ Share   ☰+ Save   ...

5 views Jul 10, 2023
The Archives Unleashed Toolkit is an open-source platform for exploring and analyzing web archives. The Toolkit was developed by the Archies Unleashed Project, with funding from the Mellon Foundation.

In November 2022, the 1.2.0 version of the Archives Unleashed Toolkit was released, making it a stable product available to the community under the broader Mellon-funded Archives Unleashed Project.

Learn about all of the features available in the 1.2.0 release!

Archives Unleashed Toolkit via GitHub (open source):
https://github.com/archivesunleashed/aut

User Documentation: https://aut.docs.archivesunleashed.org/

Computational Notebooks:
https://github.com/archivesunleashed/notebooks

Ready-to-use datasets are available through the Web Archives

for Historical Research Group via Zenodo and Borealis.

Project Updates ⚙️

Launching the Archives Research Compute Hub (ARCH). After three years of development, testing, and research application, we

are happy to announce the ARCH platform. Bringing together engineering, operations, and engagement activities, ARCH is the culminating piece for the Archives Unleashed Project and our integrated collaborations with the Internet Archive. ARCH adds value to the web archiving and research community by allowing users to easily build, access, and analyze digital collections computationally at scale.

Feature Implementations with ARCH. In preparation for sharing ARCH with the community, three additional features were implemented to enhance and complement the research process. Visit ARCH for a comprehensive look at all features and services

- **Introducing direct links to Notebooks**. By seamlessly connecting Google Colaboratory with ARCH-generated datasets, users can conduct initial analysis on web archives through pre-set computational notebooks.
- **Building Collections + Custom Filtering**. One of the biggest challenges of working with web archive collections is their size! Users can refine datasets by applying filters which empower users to focus analysis on relevant content and create manageable datasets.
- **Publishing**. Users can now openly publish datasets in line with best practices in reproducible research. All published datasets will be preserved in perpetuity, which promotes transparency, accessibility, and collaboration.

AUT 1.2.0 Release. In November 2022, the 1.2.0 version of the Archives Unleashed Toolkit was released, making it a stable product available to the community under the broader Mellon-funded Archives Unleashed Project. Check out the AUT 1.2.0 feature highlights.

Cohort Program. Over the past two years, we have worked with 50 interdisciplinary scholars through year-long intensive collaborations to support research with web archives as primary data sources and scholarly objects. During the course of the program, research teams were provided technical and academic mentorship to steward projects.

Community Engagement. Our team has engaged with supporters, collaborators, and the community by attending and presenting at recent conferences to share project progress.

- **Supporting Computational Research on Large Digital Collections**, Jefferson Bailey and Nick Ruest, CNI Fall 2022, Washington, DC. [Slides]
- **Archives Unleashed Cohort Program: Opportunities to Access, Explore, and Engage with Web Archives**, Samantha Fritz, Archive-It Partner Meeting, November 2022, Online [Slides]
- **Community Engagement and Research Support as Solutions for Expanding Data Access and Use of Web Archives**, Samantha Fritz, RDAP Summit 2023, Online [Slides]
- **Through the ARCHway: Opportunities to Support Access, Exploration, and Engagement with Web Archives**, Samantha Fritz, IIPC WAC 2023, Hilversum, Netherlands [Slides]
- **Building the Next Generation of Web Archive Analysis Service**, Ian Milligan, Helge Holzmann, Nick Ruest, Samantha Fritz, Thomas Padilla, Panel, RESAW Conference 2023, Marseille, France [Slides]

Farewell from our Team

```
status = input("what is the project
status")
if status = complete:
print ("Archives Unleashed Project
Completed!")
exit()
```

As we bring the Archives Unleashed Project to a close, we would like to express our heartfelt gratitude to our project collaborators, supporters and our vibrant community! THANK YOU! Your engagement, active participation, and unwavering support throughout this remarkable journey have enriched and fueled our progress to make this project a true success in lowering barriers and unleashing web archives.

In our final blog post, "**Unleashing Web Archives: A Final Letter to Our Community**," we reflect on the past six years of collaboration and celebrate the tools that we proudly share (and leave) with the broader community.

Reading + Resources

Our team strongly believes in and is dedicated to open access, a reminder that the following resources developed by the Archives Unleashed Project are freely available.

- **Archives Unleashed GitHub repositories**:
  https://github.com/archivesunleashed
- **Archives Unleashed Toolkit 1.2.0**:
  https://github.com/archivesunleashed/aut/
- **Archives Unleashed Toolkit User Documentation**:
  https://aut.docs.archivesunleashed.org/
- **Archives Unleashed Tutorial**: Creating JUXTA Collages with Web Archive Images
- **Project Website**: https://archivesunleashed.org/
- **Publications List**:
  https://archivesunleashed.org/publications/
- **Zotero**:

  https://www.zotero.org/groups/2500681/archives_unleashed_project, please feel free to add to this space
- **Datasets** via Zenodo + Borealis through the Web Archives for Historical Research Group
- **Slack**: @archivesunleashed Space, sign up to continue conversations. While the project comes to an official end on 1 July 2023, our Slack space will continue to remain open for discussion and sharing, and we do encourage you to use this space and make it your own!
- **YouTube Channel**:
  https://www.youtube.com/@archivesunleashed
- **Medium Blog**: https://news.archivesunleashed.org/

Acknowledgements

The Archives Unleashed Project would not have been possible without the generous support and funding from the Mellon Foundation from 2017-2023.

Financial and in-kind contributions from the Internet Archive, the University of Waterloo, York University, Compute Canada, Smart Start Labs, the Social Sciences and Humanities Research Council of Canada, and the Ministry of Research, Innovation, and Science also supported the work of this project.

We also express gratitude to members of our advisory boards, cohort participants (2021 and 2022), past Archives Unleashed alumni, collaborators, and project supporters. These individuals have been generous with their time in providing guidance, feedback, and suggestions, and we are so proud to have been

part of such a welcoming and supportive community.

Twitter   GitHub   Website   Email

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