



2022 Canadian Web Archiving Symposium

Summary and Next Steps

14 March 2022

Corey Davis, Co-chair of the CWAC Coordinating Committee

The Canadian Web Archiving Coalition (CWAC), with generous support from the Canadian Association of Research Libraries (CARL), held its second national symposium—this time via Zoom—on January 26th and 27th, 2022, with over 100 people in attendance across both days. We were privileged to welcome Internet Archive founder Brewster Kahle as the keynote speaker on Day 1, and spent the following day talking about the challenges facing the community, and how we might move forward together. Below is a summary of the presentations and discussions, and plans for next steps to be undertaken by the CWAC Coordinating Committee.

(Note: The Internet Archive did a great summary too! Check it out [on their website](#).)

Brewster Kahle addresses the future of web archiving

This was a fascinating session which focused on where we are today, and some areas of emerging importance.

Kahle started the Internet Archive more than 25 years ago and reflected on how in the early days, it was very much a technical endeavor. Back then, the web was still seen as discrete. Today, it's everywhere! This means that web archiving has moved from the technical challenges of grabbing everything and putting it on a server, to cultivating communities of practice that can help identify which content is most important, and why.

The Internet Archive sits on a vast trove of data, one that Kahle thought in the early days would be transformative for society if used to its full extent. There is a growing realization that between this massive amount of data, and the Archive-It interface, there is a gap that can be filled with specific middleware tools that will allow non-technical experts to interrogate the data in new and innovative ways.

Some areas of growth in the future include: data mining services and middleware; using web archives to fight misinformation by connecting users to high-quality data on platforms like Wikipedia; and, personal digital archives.

The community discusses the biggest challenges that face them

The community gathered on the second day to discuss a number of questions:



- What are the top three challenges facing your institution?
- How do we increase community and researcher engagement?

Most challenges relate to a lack of resources, both in terms of the ability to advocate for more people to do the work, and also the fact that very few institutions have dedicated staff for web archiving activities. We're all juggling too many priorities.

When we are able to do the work, performing quality assurance and assigning metadata can be challenging, as well as dealing with new and emerging technologies that existing platforms struggle with. There are also challenges around how to scope collections, which became obvious during the pandemic, when many of us started a COVID-19 collection, only to see the pandemic become an event of such proportions, our efforts were overwhelmed.

For this and many other reasons, participants felt that we needed to coordinate our efforts better, perhaps along to lines of web archiving efforts at organization like the Ivy+, where curatorial scoping and seed identification happen in a distributed manner, but the collecting happens on a centralized platform with support for QA and metadata application. Efforts to move in this direction, however, will require resources.

Where we go from here

The CWAC Coordinating Committee is taking a sustainable approach to how it might support the community in the next year or so, with a realistic expectation as to how much can be accomplished by such volunteer efforts. Our strategy is two-fold.

First, we will work with the CARL Digital Preservation Working Group (CARL DPWG) to explore funding opportunities that could resource a position at the national level to support the web archiving community, and help build capacity and scale our infrastructure.

Secondly, we want to encourage more communication and conversation within the CWAC community, especially focused on informal events and helping beginner practitioners. We are planning to:

- Have bi-monthly Zoom drop-in sessions for practitioners to connect on specific issues, or just get to know each other better!
- Help organize small volunteer peer-groups within the CWAC that can help each other as they develop and refine their web archiving programs.
- Coordinating Committee members will take turns monitoring and contributing to the CWAC listserv, in order to help reinvigorate community connects.

We want to hear from you!

The Coordinating Committee is here to answer any questions you have. Please feel free to reach out to either co-chair if you have any questions or concerns:



Dana McFarland, Vancouver Island University <dana.mcfarland@viu.ca>

Corey Davis, University of Victoria, <coreyd@uvic.ca>

About the CWAC

The Canadian Web Archiving Coalition (CWAC) is an inclusive community of practice within Canadian libraries, archives, and other memory institutions engaged or otherwise interested in web archiving. The Coalition's mission is to identify gaps and opportunities that could be addressed by nationally coordinated strategies, actions, and services, including collaborative collection development, training, infrastructure development, and support for practitioners and researchers.

Join the CWAC

The community of practice communicates via a listserv, which can be joined here:

<https://listserv.uottawa.ca/scripts/wa.exe?SUBED1=CWACCAW-L&A=1>