
Web Archiving Comes of Age: Averting the Digital Dark Age of September 11th, 2001

Ian Milligan*†¹

¹University of Waterloo [Waterloo] – 200 University Avenue West, Waterloo, ON, Canada N2L 3G1, Canada

Abstract

The terrorist attacks of 11 September 2001 against the United States were among the first major disasters of the web age. History seemed to be made in the minutes, weeks, and months after the first plane hit the World Trade Center's North Tower that morning. As phone networks failed, many turned to the web and e-mail to ask about the safety of security of friends, family, and colleagues. Office workers, working in cubicles and offices away from television, tried to make sense of the day on soon overloaded online news sites. Within hours of the attacks, archivists and librarians began to capture webpages, television feeds, voice mail messages, and email messages. Much of this started the next day, when Library of Congress staff returned to their campus in Washington DC and collaborated with the Internet Archive and academics on the task ahead. Within weeks, new and emerging forms of memory websites were created, allowing anybody to record reactions, thoughts, and details about how their lives had been transformed. A web portal was launched a month later, providing immediate and indexed access to the archived web.

All of this happened quickly. Just six years before the attacks, the spectre of a digital dark age had haunted those worried about the future of collective cultural heritage. By 2001, an infrastructure was ready. The Internet Archive was then over five years old. The Library of Congress' web archive had just celebrated its second anniversary. That infrastructure, put in place to either record long-foreseen elections or the everyday evolution of web culture, suddenly needed to preserve the very specific and ephemeral online experiences of how people understood and responded to the attacks.

This presentation opens with a brief context of the attacks and their specific impact on web infrastructure. It then focuses on the work, beginning on the afternoon of the day of the attacks, of preserving events at scale. Drawing on oral histories with researchers and librarians (at both Internet Archive and Library of Congress), my paper presents a novel history of these efforts. Crucially, an in-progress project that had been about to look at K-12 School Board websites was pivoted overnight to focus on the attacks and their aftermath. Between September and December 2001, 30,000 seeds were suggested by an interdisciplinary team. On 11 October 2001, september11.archive.org launched with over 500,000 pages, representing daily news reports to personal memorials to beyond. The Internet Archive compelled their colleagues to go "live," to showcase its collection, in order to find holes as quickly as possible.

*Speaker

†Corresponding author: i2millig@uwaterloo.ca

Ultimately, I argue that the 11 September 2001 attacks were a watershed moment in digital preservation and collecting. The digital dark age, which had preoccupied so many over the previous decade, would be measured for much of the web on that day not in weeks and months but in hours and days. People turned to the web for information, to memorialize, to console each other, to post flags and ribbons, and in so doing, contributed to a rich and evolving historic record. When historians write the social and cultural history of the attacks and their aftermath, these sources will be critical.

Keywords: web history, event collecting, web archives