

Finding Community in the Ruins of GeoCities: Distantly Reading a Web Archive

Ian Milligan

Department of History

University of Waterloo

200 University Ave. W, Waterloo ON

N2L 3G1 Canada

i2millig@uwaterloo.ca

ABSTRACT

This paper provides a brief overview of my work with the GeoCities web archive. Asking the question of “can we find community,” I use it as a case study to explain various methods for distantly reading web archives.

1. INTRODUCTION

In the early 1990s, as the Web became part of everyday life, GeoCities provided a sense of community to facilitate users' first, often hesitant, steps into the world of Web publishing. At that time, GeoCities was one of the largest collections of non-commercialized, user-generated, public speech ever assembled. Along the digital “GeoAvenues” and in the provocatively-named “homesteads,” users developed a relationship with the Web that built a foundation for the blogging and social networking explosion of the 2000s. Rather than simply being a platform upon which users created individualized webpages, GeoCities users helped each other develop the site as a whole.

I explore how we can find community in the ruins of GeoCities. Using a combination of distant, computational reading (studying websites as a collective whole, rather than as individual documents) and more focused, targeted reading, I argue that we can understand the GeoCities web archive as ruins of a vibrant, interconnected community.[1]

Overcoming these technical challenges is of critical importance to the historiography on the early web, however. In the digital ruins of GeoCities, new users figured out their relationship to the Web. They were not alone; they were part of a larger community. As so much of our relationship with web services is now mediated by user interfaces and glossy tutorials, it is worthwhile to look back to the early 1990s and how -- spread out across time and space -- users figured out what the Web would mean to them. GeoCities, a massive assemblage of non-commercialized public speech, presents an interesting way to explore the history of the early World Wide Web.

2. IMPORTANCE

Big Data is coming to history. The advent of web-archived material from 1996 onwards presents a challenge. In my work, I explore what tools, methods, and approaches historians need to adopt to study web archives.

GeoCities lets us test this. It will be one of the largest records of the lives of non-elite people *ever*. The Old Bailey Online can rightfully describe their 197,000 trials as the “largest body of texts detailing the lives of non-elite people ever published” between 1674 and 1913.[2] But GeoCities, drawing on the material we have between 1996 and 2009, has over thirty-eight million pages.

GEOCITIES USERS:

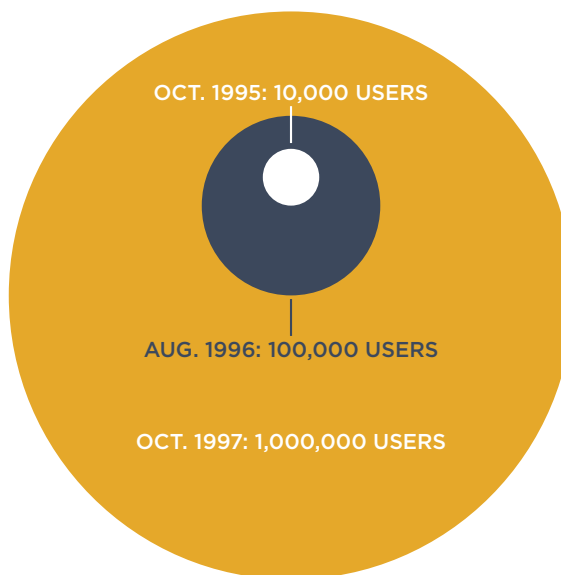


Figure 1. Rapid growth of an online community.

These are the records of everyday people who published on the Web, reaching audiences far bigger than previously imaginable.

GeoCities was a unique place. It rapidly grew, reaching millions of users in a few years. It concept helped to bridge the locally-based networks of the early 1990s and BBSes with the wide open Web - it clustered users together in neighbourhoods, from the children-focused EnchantedForest, the family-focused Heartland, the education-focused Athens. Users relied on each other to find content: from living next to each other in neighbourhoods, to linking to each other using Web rings. In an era before the widespread adoption of search engines, which often required explicit SEO techniques and effort to be part of, these were important.

This unique early experiment in the history of the Web came to halt in 2009 when Yahoo! shuttered it (after acquiring it in a major 1999 deal).[3] If it hadn't been for the timely intervention of Archive Team and others such as the Internet Archive, GeoCities would have been lost forever.[4]

It would be as if the Old Bailey had been thrown on the fire pit of history.

Ian Milligan. “Finding Community in the Ruins of GeoCities: Distantly Reading a Web Archive.” Bulletin of IEEE Technical Committee on Digital Libraries, Vol. 11, Issue. 2 (October 2015). *This article is distributed under the terms of the Creative Commons Attribution 3.0 Unported License (<https://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.*

