WEB ARCHIVING AS CULTURE: TUMBLR AND THE CULTURAL CONSTRUCTION OF THE ARCHIVED WEB

Jessica Ogden
University of Bristol

Introduction

Recent trends in communication and Internet research show a growing interest in the significance and impact of content moderation practices on the circulation and accessibility of particular forms of online media. This research has often focused on the platform policies and technical apparatuses of content delivery, the invisible labour of content moderators (both human and algorithmic), and the effects of this labour on the actors themselves and digital life online (Gillespie, 2018; Roberts, 2019). Currently missing from this area of research, is an examination of the ways in which various web archiving initiatives, organisations and community groups are intervening and shaping the availability of information/media in the face of platform moderation policies. In this paper I bring together these fields of enquiry to discuss the cultural significance of web archiving through the example of Tumblr's 2018 efforts to remove so-called 'adult content' and 'Not Safe for Work' (NSFW) posts from the platform.

This paper examines efforts on the part of Archive Team - a self-described 'loose collective of rogue archivists, programmers, writers and loudmouths dedicated to saving our digital heritage' - to archive Tumblr NSFW posts.[1] The study forms part of a wider ethnographic investigation of web archiving across three field sites, with the aim of examining how web archival practices shape the archived Web. Borrowing from 'facet methodology' (Mason, 2011), the wider research findings are framed elsewhere through the lens of three facets, where each facet provides one 'methodological-substantive' plane or surface view of web archiving (Ogden, 2020). Here, I use the case of Archive Team to reflect on *web archiving as culture* and emphasise the ways that practices 'as an observable object for the study of culture' (Swidler, 2001) are filtered through and structured by 'strategies of action' that shape how the Web is archived. In doing so, the paper contributes to broader discussions of platform politics and the role of power in the

---

[1] http://archiveteam.org (visited 1 Mar. 2020)

production of online memory, and raises further questions about the ethics of web archives and their positioning as historical representations of online culture.

Background and Methods

Web archives (WAs) - broadly conceived as any attempt to capture and preserve the Web for future use - are evermore central to discussions of digital access in the public sphere, as they provide tools for accessing parts of the Web that have been subject to neglect, removal or censorship. This research is therefore motivated by the observation that despite their positioning as critical resources for a range of scholarly Internet research agendas (Rogers, 2013) and their widespread use as tools for evidence-based accountability online, WAs remain relatively understudied. As such, recent scholarship has framed the need for further research into the practices of web archiving, arguing the inherent connections between the ways the Web is archived and our future understanding of the Web's past (Ogden et al., 2017; Summers and Punzalan, 2017).

Although featured in numerous news media articles since their formation in 2009, the specifics regarding the processes that underpin Archive Team web archiving activities have yet to receive critical attention within the field. Situated within a broader, longitudinal study of Archive Team's WA activities (using interviews, non/participant observation and documentary research), this research focuses on data collected during their attempts to archive Tumblr NSFW blogs between December 2018 and January 2019. During this period, I observed activities as they occurred on public EFNet Internet Relay Channels (IRC), Archive Team's GitHub repositories, social media (including Tumblr, Twitter, Reddit) and in popular news media. Interviews and IRC chat with Archive Team participants supplemented observations and provided a mechanism for clarifying findings and focusing further observations. The findings also draw on publicly available historical IRC logs, the Archive Team wiki and media interviews to further contextualise observations.

Findings and Discussion

Web archiving is produced within cultural worlds that produce meaning through practice.

The first observation is rooted in the view that web archiving is situated within particular cultural worlds that are observable through local meanings, practices and strategies of action. Archive Team's origin story is first used to contextualise the early visions for a volunteer 'emergency response team' dedicated to archiving online user-generated content. Their origins, accompanying analogies and imagery then frame both 'the opposition' (corporations, bureaucracy, censorship) and an 'A-Team' imaginary with a moral cause to 'save the Web'. Next, web archiving as culture is observable through the practices used to both action and organise the collective work of web archiving. Here, I discuss some of the ways that community identity is conveyed through socio-technical protocols and tools that: facilitate a form of 'institutional memory-making' through self-archiving; shape remote communication and project organisation; create a sense of community belonging and ultimately reflect a broader attentiveness to the core tenets of liberalism and 'hacker culture' famously described by Levy (2010).

Web archiving transforms the Web through practice.

The second observation is that web archiving is transformative; where a focus on cultural practices works to frame the ways the Web is altered through archival practices, as well as highlights the contribution of culture in both sustaining and constructing strategies of action. Practice dilemmas are revealed that highlight the impact of 'folk theories' and claims regarding which particular components of the site were deemed worth saving (and conversely, not worth saving) and why. Archive Team's persistence in the collection of NSFW posts despite Tumblr's attempts to thwart their efforts revealed an observable 'fetishisation of abundance' that trumped the perceived risks of breaching the platform's access restrictions. Findings demonstrate the prioritisation of the scale of collection over the 'integrity' or completeness of WAs, where the juxtaposition of these two principles reveals the shifting ethical boundaries of practice in response to the challenges of archiving. In turn, the study reflects the performative ways that web archiving is shaped by culture, and critically positions practices as a transformative force that shapes the nature of our interaction with the past Web.

References

Gillespie, T., 2018. *Custodians of the Internet: platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press, New Haven and London.

Levy, S., 2010. *Hackers: Heroes of the Computer Revolution*. O'Reilly, Sebastopol, CA.

Mason, J., 2011. Facet Methodology: The Case for an Inventive Research Orientation. Methodological Innovations Online 6, 75–92. https://doi.org/10.4256/mio.2011.008

Ogden, J., 2020. Saving the Web: Facets of Web Archiving in Everyday Practice (PhD Thesis). University of Southampton, Southampton, UK.

Ogden, J., Halford, S., Carr, L., 2017. Observing Web Archives: The Case for an Ethnographic Study of Web Archiving, in: *Proceedings of WebSci '17*. Presented at the Web Science, ACM, Troy, NY USA. https://doi.org/10.1145/3091478.3091506

Roberts, S.T., 2019. *Behind the Screen: Content Moderation in the Shadows of Social Media*. Yale University Press, New Haven and London.

Rogers, R., 2013. *Digital Methods*. MIT Press, Cambridge, MA.

Summers, E., Punzalan, R., 2017. Bots, Seeds and People: Web Archives As Infrastructure, in: *Proceedings of CSCW '17*. ACM, New York, NY, USA, pp. 821–834. https://doi.org/10.1145/2998181.2998345

Swidler, A., 2001. What anchors cultural practices, in: Schatzki, T.R., Knorr Cetina, K., von Savigny, E. (Eds.), The Practice Turn in Contemporary Theory. Routledge, London and New York, pp. 83–102.