



## UvA-DARE (Digital Academic Repository)

### Platform and app histories: Assessing source availability in web archives and app repositories

Helmond, A.; van der Vlist, F.

**DOI**

[10.1007/978-3-030-63291-5\\_16](https://doi.org/10.1007/978-3-030-63291-5_16)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

The Past Web

**License**

Article 25fa Dutch Copyright Act (<https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment>)

[Link to publication](#)

**Citation for published version (APA):**

Helmond, A., & van der Vlist, F. (2021). Platform and app histories: Assessing source availability in web archives and app repositories. In D. Gomes, E. Demidova, J. Winters, & T. Risse (Eds.), *The Past Web: Exploring Web Archives* (pp. 203–214). Springer. [https://doi.org/10.1007/978-3-030-63291-5\\_16](https://doi.org/10.1007/978-3-030-63291-5_16)

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)*

# Platform and App Histories: Assessing Source Availability in Web Archives and App Repositories



Anne Helmond and Fernando van der Vlist

**Abstract** In this chapter, we discuss the research opportunities for historical studies of apps and platforms by focusing on their distinctive characteristics and material traces. We demonstrate the value and explore the utility and breadth of web archives and software repositories for building corpora of archived platform and app sources. Platforms and apps notoriously resist archiving due to their ephemerality and continuous updates. As a consequence, their histories are being overwritten with each update rather than written and preserved. We present a method to assess the availability of archived web sources for social media platforms and apps across the leading web archives and app repositories. Additionally, we conduct a comparative source set availability analysis to establish how, and how well, various source sets are represented across web archives. Our preliminary results indicate that despite the challenges of social media and app archiving, many material traces of platforms and apps are in fact well preserved. We understand these contextual materials as important primary sources through which digital objects such as platforms and apps co-author their own “biographies” with web archives and software repositories.

## 1 Introduction

Contemporary digital objects, such as digital platforms and mobile apps, pose significant challenges to archiving and research practice. With millions or even billions of monthly active users, some of those platforms and apps are among the most popular products and services around the world (Statista 2017; Statista 2019a, b). Yet, despite their social, economic, and cultural significance, many of their histories are at risk of getting lost. As a result of rapid release cycles that enable developers to

---

A. Helmond (✉)  
University of Amsterdam, Amsterdam, The Netherlands  
e-mail: [a.helmond@uva.nl](mailto:a.helmond@uva.nl)

F. van der Vlist  
Utrecht University, Utrecht, The Netherlands  
e-mail: [f.n.vandervlist@uu.nl](mailto:f.n.vandervlist@uu.nl)

develop and deploy their code very quickly, large web platforms such as Facebook and YouTube change continuously, overwriting their material presence with each new deployment. Similarly, the pace of mobile app development and deployment is only growing, with each new software update overwriting the previous version.

In this chapter, we consider how one might write the histories of these new digital objects, despite such challenges. We reflect on the materiality of platforms and apps as specific types of digital objects and outline a method to take inventory of their archived materials for historical studies. As we argue, these archived sources offer various opportunities for historical studies of platforms and apps. That is, the *routine overwriting* of digital objects and their data through continuous incremental software updates constitutes both a core problem and a source of research opportunities for historians—at least, as long as those changes are documented by these digital objects themselves or preserved by web archives. We, therefore, look into the source availability of preserved material traces of platforms and apps.

In the first section, we consider how, from a material perspective, platforms and apps are different from other digital objects such as websites. As a consequence, there are challenges with regard to their archiving and study as well as new opportunities. In the second section, we describe a method of taking inventory of the available materials for writing platform and app histories. The method is not just useful for building corpora of historical platform or app sources but also potentially valuable for determining significant omissions in web archives and for guiding future archiving practices. In the third section, we describe the outcomes of an exploratory case study of the availability of leading platforms and apps today. We conclude with a reflection on the future of platform and app historiography.

## 2 The Archived Materiality of Platforms and Apps

The early Web mainly consisted of websites and interlinked webpages. As a consequence, the website has become the main unit of archiving as well as the main unit of historical analysis (Brügger 2018). However, in the past decade, we have witnessed the emergence of new types of digital objects, in particular, digital platforms and apps for social media and beyond. But what characterizes these specific digital objects as archived objects, as compared to the website or webpage?

When thinking of how platforms and apps are archived today, we contend that we need to consider their specific materiality. With the term materiality, we refer to the material form of those digital objects themselves as well as the material circumstances of those objects that leave material traces behind, including developer resources and reference documentation, business tools and product pages, and help and support pages (Ankerson 2012; Fuller 2008; Gillespie 2003; Kirschenbaum 2003). Furthermore, developers commonly keep changelogs, release notes, and do versioning. Importantly, rather than secondary sources, which are commonly used for web histories of platforms and apps (Brügger 2015; Poulsen 2018), these materials are primary sources that offer particular research opportunities or that

may be supplemented and triangulated for accuracy. These material traces may “tell stories” about the evolving production, preferred usage, and embedded politics of software objects (Gillespie 2003).

We understand these contextual materials as important primary sources through which digital objects such as platforms and apps write, or indeed *overwrite*, their own “biographies”, thus building on the emerging genre of media biography, including “software biography”, “website biography”, and “platform biography” (Burgess and Baym 2020; Natale 2016; Rogers 2017; Pollock and Williams 2008). The dual materiality of platforms and apps, as software objects and as sets of material contextual traces, opens up a productive avenue for historical analysis. Even when a platform or app as such is not archived, we may turn to web archives to look for their contextual material traces instead. These traces “provide a potential entryway to the web cultures, production practices, and symbolic systems informing lost cultural artifacts” (Ankerson 2012: 392). Furthermore, these “textual supplements are perhaps even more potent because they seem to be part of the tool itself” as they document a “self-interpretation” of the software object that we may employ for its history writing (Gillespie 2003).

## 2.1 Web Archives

The materiality of a web platform manifests as a collection of interrelated web pages that are meaningfully arranged to address different groups of users on different “sides”. That is, platforms are programmable infrastructures as well as digital intermediaries that bring together different groups of users (Gillespie 2010; Helmond 2015; de Reuver et al. 2018). For each user group, there are different sets of resources and documentation that describe the operational logics, stakeholder relations, and preferred uses of a platform. For example, social media platforms provide such materials for their various user groups, which include end-users, developers, businesses, advertisers, partners, creators, media and publishers, politicians, investors, and researchers. As we have outlined previously, these different sets of materials are well archived and afford and privilege different types of social media and platform history (Helmond and van der Vlist 2019). To locate historical platform resources and documentation, we may turn toward web archives.

The materiality of apps is different from platforms. While many digital platforms exist principally on the Web and operate tools, products, and services on multiple “sides” to different groups of users, apps are software *bundles* (or packages) that are downloaded directly onto mobile devices from app stores. In contrast to websites and web platforms, mobile apps are not web “native” and instead reside on mobile devices and in app stores, which makes them even more difficult to archive and study. Yet they are entangled with a variety of other web services (Dieter et al. 2019). App stores, arguably, are a “native” environment for apps. For end-users, apps present themselves as contained digital objects that are purchased and downloaded from platform-specific app stores, such as Google Play for Android or the App Store

for the iOS operating system. Yet by their design, app stores only provide access to the latest version of an app bundle and not to former versions. With each new software update, a former app version is overwritten—both inside the app store and on the user’s mobile device. As a result, neither app stores nor mobile devices keep former versions of apps, which poses challenges for historical app studies.

## 2.2 *App Repositories*

To locate former app bundle versions, we may turn to several third-party software repositories, such as Cydia for iOS apps or APKMirror for Android apps.<sup>1</sup> Contrary to traditional institutional archives, these repositories are noninstitutional storage locations for the retrieval of software that were never designed for permanent preservation (Allix et al. 2016). While they may share commonalities with archives, software repositories do not curate collections of “records” for permanent historical preservation and do not necessarily consider their value as evidence or as a source for historical research (Brügger 2018). Additionally, the use of software repositories as app archives raises issues with regard to archive incompleteness and software insecurity. They are incomplete because they rely on users manually uploading app versions; they pose security risks because not all repositories scan package uploads for malicious code injections. When app code is tampered with, this may directly limit or influence historical code-based analyses. And even if we find former app versions in repositories, we still face software emulation challenges with apps as they typically require a complex set of dependencies and will only “run” or operate on specific devices and operating systems of the past (Boss and Broussard 2017; Helmond and van der Vlist 2019; Stevenson and Gehl 2018).

As an alternative or additional strategy, app historians may turn to archived app metadata sources as preserved in web archives that hold “snapshots” of app details pages in app stores or repositories. While apps and app stores both exist primarily on mobile devices, the leading app stores—Google Play and Apple’s App Store—also provide web-based graphical user interfaces to their stores. These stores contain a wealth of information about specific apps as well as their relations to other, “similar” apps and the store categories or app collections to which they belong (Dieter et al. 2019). For each app, there is a details page with the app’s title, developer, bundle version, screenshots, description, requested app permissions, download statistics, reviews, ratings, and more. Fortunately, these app store details pages are preserved in web archives, which generates opportunities for historical app studies. In short, to locate historical app materials, we may thus either turn to app repositories to retrieve former app versions or to web archives to retrieve contextual information.

---

<sup>1</sup>Cydia, <https://cydia-app.com/>; APKMirror, <https://www.apkmirror.com>.

### 3 Assessing the Availability of Platform and App Sources

To determine whether these materials have been preserved and where they are located, we conducted an exploratory study of the availability of archived sources for platform and app history. Building on previous work (Helmond and van der Vlist 2019), we first detail a method for assessing the availability of archived web sources for platforms and apps in web archives and app repositories.

Making use of market data portals Statista and App Annie, we selected the current top-20 most popular social media platforms and top-10 mobile apps for Android and iOS combined, both based on the current number of active users worldwide (App Annie 2019; Statista 2019a, b). For the first source set of social media platforms, we made an inventory of their most prominent “sides” and created a list of URLs pointing to the location of their principal materials (e.g., [twitter.com](https://twitter.com), [developer.twitter.com](https://developer.twitter.com), [business.twitter.com](https://business.twitter.com), [marketing.twitter.com](https://marketing.twitter.com), [investor.twitterinc.com](https://investor.twitterinc.com)). For the second source set of mobile apps, we created a list of URLs pointing to the app store details pages for each app.<sup>2</sup> These URLs contain the unique bundle identifier of each app, which remains stable even when apps are continuously updated and overwritten. App store links are constructed with these bundle identifiers and thus also remain stable over time.<sup>3</sup> So, although apps are updated continuously, they have a stable bundle identifier and a stable web URL that points to a details page that we may track in archives over time. In addition, we used these unique bundle identifiers to locate these apps in ten prominent third-party software repositories for Android apps.<sup>4</sup>

To assess which web archives actually hold archival records of a particular resource, we employed Memento’s Time Travel Service (Van de Sompel et al. 2009).<sup>5</sup> The service functions as a search engine “on top of” the 25 leading

---

<sup>2</sup>Over the past decade, app store URLs changed only once or twice: Google Play (since 2012) was formerly called Android Market (2008–2012), and the domain changed from [android.com/market](https://android.com/market) to [play.google.com/store](https://play.google.com/store); Apple’s App Store (since 2008) was formerly called App Store (iTunes Preview) (2012–2019) and before that, Web Apps for iPhone (2008–2012), and its domains changed from [apple.com/webapps](https://apple.com/webapps) to [itunes.apple.com](https://itunes.apple.com) to [apps.apple.com](https://apps.apple.com). For our exploratory study, we focused only on the current URLs at the time of writing.

<sup>3</sup>App store URLs are constructed as follows: for Google Play, [https://play.google.com/store/apps/details?id=bundle\\_id](https://play.google.com/store/apps/details?id=bundle_id); for the App Store, there are three URL formats, [https://itunes.apple.com/app/bundle\\_id](https://itunes.apple.com/app/bundle_id), [https://itunes.apple.com/us/app/bundle\\_id](https://itunes.apple.com/us/app/bundle_id), [https://itunes.apple.com/us/app/appname/bundle\\_id](https://itunes.apple.com/us/app/appname/bundle_id).

<sup>4</sup>We included the following app repositories: [AndroidAPKsBox.com](https://androidapksex.com), [AndroidAPKsFree.com](https://androidapksex.com), AndroidDrawer, APKMirror, APKMonk, APKPure, APKPure.ai, APKPure.co, Aptoide, and Uptodown.

<sup>5</sup>Time Travel, <http://timetravel.mementoweb.org/>.

international web archives and may be queried for specific URLs (Memento 2016).<sup>6</sup> For end-users, it offers a graphical user interface (GUI) that may be deployed to manually query and locate a URL across multiple web archives. Additionally, it offers an application programming interface (API) to programmatically request that data. Both methods return a list of web archives that hold one or more Mementos (i.e., time-stamped archived copies of a specific URL). For each Memento, the service returns the first and last Memento available as well as links to all available captures across archives. Time Travel thus provides a simple method to assess the availability of specific archived sources across web archives. To determine the total number of Mementos held or the number of archives holding them, users may follow the “All captures from” link for each web archive and manually count the number of Mementos held.

To scale and automate this process for a large source set of URLs, researchers may use MemGator, an open-source command-line interface utility that is built “on top of” the Memento API and aggregates Mementos.<sup>7</sup> MemGator programmatically requests Memento TimeMaps from a list of web archives that support the Memento protocol (Alam and Nelson 2016). Each TimeMap provides a time-stamped list of all Mementos held in that archive for a given URL (Memento 2015). It also lets researchers customize the list of web archives from which to request TimeMaps. For present purposes, we extended MemGator’s list of web archives that natively support the Memento protocol, as specified in “archives.json”, with a number of web archives listed in the Time Travel Archive Registry that run Memento proxies (Memento 2015), so as to be as inclusive as possible in our exploratory study. Our custom list included 20 web archives from which to programmatically retrieve data. More specifically, we used MemGator to programmatically retrieve the available platform and app materials from across these 20 web archives and then analyzed the results to assess the availability of sources.<sup>8</sup> In what follows, we describe the results of our exploratory study.

---

<sup>6</sup>As of June 2019, “Time Travel Find” supported the following web archives: archive.today, Archive-It, Arquivo.pt.: the Portuguese web-archive, Bayerische Staatsbibliothek, Bibliotheca Alexandrina Web Archive, DBpedia archive, DBpedia Triple Pattern Fragments archive, Canadian Government Web Archive, Croatian Web Archive, Estonian Web Archive, Icelandic web archive, Internet Archive, Library of Congress Web Archive, NARA Web Archive, National Library of Ireland Web Archive, National Records of Scotland, perma.cc, PRONI Web Archive, Slovenian Web Archive, Stanford Web Archive, UK Government Web Archive, UK Parliament’s Web Archive, UK Web Archive, Web Archive Singapore, and WebCite.

<sup>7</sup>MemGator, <https://github.com/oduwsdl/MemGator>.

<sup>8</sup>All data were collected between May and June 2019.

## 4 The Availability of Platform and App Sources

We analyzed the source availability of platform and app materials according to three criteria: first, the *volume of availability* or the total number of Mementos held; second, the *depth of availability*, specified as the number of days, months, or years between the first and last Mementos; and third, the *breadth of availability*, referring to the number of web archives holding those Mementos (Helmond and van der Vlist 2019). The first two criteria determine the amount of available material and the possible levels of granularity for historical analysis, while the third criterion enables researchers to triangulate and verify historical sources, such as when certain elements are corrupted or missing.

Table 1 provides a summary of our exploratory study results. We counted the total number of Mementos held across web archives (i.e., volume), determined the time span between the first and last Mementos held (i.e., depth, expressed in number of days), and counted the number of web archives holding those Mementos (i.e., breadth, expressed as a single number up to 20 web archives). Taken together, these three dimensions provide a useful account of source availability and allow researchers to determine the feasibility of certain historical projects or allow archiving practitioners to reconsider their archiving strategy. Based on these counts, we then calculated an availability rank for each platform and app by calculating the number of captures per day (volume divided by depth) and then multiplying that number by breadth.

**Table 1** Availability of archived app store page details (Mementos) for each one of the top 10 Android and iOS apps across web archives (accumulated)

App title	Android (Google Play)				iOS (App Store)			
	Volume	Depth	Breadth	Rank	Volume	Depth	Breadth	Rank
Facebook	8198	2637	8	3	390	3389	6	4
WhatsApp messenger	4092	2600	7	4	548	3395	6	2
Facebook messenger	85,222	2638	10	1	99,581	2708	9	1
WeChat	442	2557	5	5	120	3019	4	5
Instagram	13,215	2611	11	2	447	3153	6	3
QQ	38	2551	1	8	16	2547	1	8
Alipay	26	2188	1	9	26	800	1	6
Taobao	31	2147	2	7	50	3168	1	7
WiFi master key	31	1890	3	6	0	0	0	n/a
Baidu	24	2196	1	10	0	0	0	n/a

## ***4.1 Social Media Platforms in Web Archives***

As we have analyzed elsewhere, social media platforms have been relatively well archived on all of their “sides” (Helmond and van der Vlist 2019). The five best-archived social media platforms represent an average of 913,440 Mementos, followed by an average of 130,036 for the next 15 platforms (Max = 1,783,855; Min = 3007; Median = 166,412).

As these results suggest, there are many opportunities for historical platform studies about different “sides” and user groups, albeit at different levels of granularity, depending on source availability. In particular, developer and business materials have been well archived and enable researchers to write histories beyond the “front-end” interface for end-users. They may look at platforms’ influential roles as development platforms, advertising platforms, content creation platforms, media publishers, and platform companies (Helmond and van der Vlist 2019; Helmond et al. 2019; Nieborg and Helmond 2019). These materials also enable researchers to examine how the technological architectures and economic business models of platforms evolve side by side. In short, platform histories would benefit from considering more than just their end-users and contents and include their multiple user groups to examine how they coevolved with respect to other “sides”.

## ***4.2 App Details in Web Archives***

Contrary to most popular social media platforms, apps have been less well archived in general, at least when we look at the preservation of their app store details pages in web archives (Table 1). For Android apps, Facebook Messenger is the best-archived app by far, leaving all other apps behind. In fact, other apps have hardly been archived at all. While the four best-archived top Android apps—Facebook Messenger, Instagram, Facebook, and WhatsApp Messenger—represent an average of 27,681 Mementos each, the next six top apps have an average of just 98.6 Mementos (Max = 85,222; Min = 24; Median = 240). For top iOS apps, Facebook Messenger accounts for nearly 99,581 Mementos, while the next nine top apps have an average of just 177.4 Mementos (Max = 99,581; Min = 0; Median = 85). In particular, pages of non-Western apps have been poorly archived, in line with a previously identified imbalance of source availability in archived websites between the United States and other countries (Thelwall and Vaughan 2004).

The archived app materials enable researchers to examine the evolution of individual apps, or app collections and genres. In a previous project, we examined the emergence of secure or encrypted messaging and chat apps on Android and used their descriptions to determine how those apps offered new and different ways of “doing privacy” (e.g., the emergence of new encryption protocols and tradeoffs between security, privacy, and usability). Tracking app descriptions over time thus enabled us to understand how apps or app developers responded to Edward

Snowden’s surveillance revelations in June 2013, when digital surveillance became a “matter of concern” on the web and mobile ecosystem (Dieter et al. 2019; van der Vlist 2017). App details pages enable app historians to tell stories about an app’s rhetorical positioning (e.g., using taglines, descriptions), production (e.g., using developer names, app versions, changelogs), distribution (e.g., using app collections, relations, pricing models), and reception (e.g., using app downloads, reviews, ratings).

### 4.3 App Bundles in App Repositories

With regard to the preservation of Android app bundles in third-party software repositories, we found more promising results (Table 2). All of the 10 top apps in our set are relatively well archived based on all three criteria. In terms of volume, the four Facebook-owned top apps—WhatsApp Messenger, Facebook, Instagram, Facebook Messenger—have been stored an average of 3722 times, while the next six, all non-Western, top apps have been stored 297 times on average (Max = 4585; Min = 166; Median = 469). The oldest versions of the apps in our dataset date back to May 2012.

These results suggest that app repositories are promising sources for historical app studies, both to study app bundles themselves and to triangulate app details between app repositories and official app stores. Most importantly, these primary app materials enable researchers to devise historical methods based on “static” app analysis (Dieter et al. 2019). That is, app bundles may be decompiled and analyzed as source code to study requested app permissions, embedded code, and external relationships to other infrastructural web services such as advertising and content delivery networks, for example (Gerlitz et al. 2019). Or, researchers may emulate those app bundles to conduct “dynamic” app analysis and study evolving interface design patterns and the network connections that mobile devices establish on behalf of apps.

**Table 2** Availability of top-10 Android apps across app repositories (accumulated)

App title	Android			
	Volume	Depth	Breadth	Rank
Facebook	4585	2584	9	2
WhatsApp messenger	4268	2585	10	1
Facebook messenger	2765	2609	10	4
WeChat	315	2364	10	6
Instagram	3271	2600	10	3
QQ	229	2187	9	9
Alipay	193	1362	8	7
Taobao	258	1844	7	8
WiFi master key	623	1401	8	5
Baidu	166	2242	5	10

## 5 Conclusions: Platform and App Historiography

In this chapter, we have demonstrated how researchers may use web archives and app repositories to write histories of new digital objects such as platforms and apps, despite their archiving challenges. We have reflected on the materiality of platforms and apps as specific types of digital objects and have outlined a method to make an inventory of their archived materials. Existing archived sources offer many opportunities for historical platform and app studies, and it is our hope that their affordances for research are further explored.

Our exploratory study of source availability for the most popular social media platforms and mobile apps provides important insights into the current state of platform and app archiving, which should be of interest to researchers and historians of web platforms and mobile apps. Furthermore, our assessment of source availability provides relevant starting points and example case studies for different types of platform and app history and may guide future historians in the process of corpus building. Our exploratory study should also be of interest to web and app archiving practitioners. In particular, our source availability assessment method and the preliminary results of our exploratory study may guide or inspire a reconsideration of archiving efforts going forward. Current web archiving strategies or protocols may not capture all of the relevant materials, as in the case of app store details pages which are located deep within app stores. We particularly recommend a more comprehensive archiving strategy that captures the multiple “sides” of popular social media platforms and the app details pages of popular app stores beyond the top apps.

Although we only looked at a small selection of top platforms and apps, we already observed large discrepancies in source availability between both types of digital objects, which inevitably determines and limits the future histories that may be written about and with those apps. Our selection of popular apps is expected to be far better archived than the millions of apps in the “long tail” of app stores. We should note, however, that even with a hundred or fewer Mementos it is, of course, possible to write the histories of platforms and apps. Depending on the historical project, differences in source availability may have implications with regard to volume (e.g., limiting the afforded level of granularity or resolution), depth (e.g., constraining the historical period), and breadth of availability (e.g., limiting the possibilities of triangulation or source verification). Existing services and utilities such as Memento and MemGator offer the opportunity to move beyond the Internet Archive as the primary or even only source of web history. They also enable researchers to triangulate and verify sources and thereby address common issues of archive incompleteness and software insecurity (including corrupt app files).

The ephemerality of digital platforms and mobile apps may be understood as the result of a continuous stream of incremental software updates that overwrite the material presence of a platform or app every time. We may conceive of this process of overwriting as a challenge of material erasure, or as a “native” mode

of software history-writing. That is, even though these ephemeral digital objects change continuously, web archives and software repositories, fortunately, capture many of those changes, thereby arresting the ongoing material transformation of platforms and apps at certain time intervals (e.g., with hourly, daily, or monthly captures or “snapshots”). Consequently, we argue that the biographies of platforms and apps are co-written by these digital objects themselves and by web archives and, in the case of apps, also by software repositories. We can employ their different types of primary and contextual sources to “reconstruct” these processes of overwriting at different levels of granularity—from the minute, incremental changes to the longer-term evolution of a platform or app. We can use web archives and repositories to reconstruct what was written on top of other writing and narrate the drama of changes, updates, and versions.

## References

- Alam S, Nelson ML (2016) MemGator: a portable concurrent memento aggregator: cross-platform CLI and server binaries in Go. In: Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries, pp 243–244. ACM. <https://doi.org/10.1145/2910896.2925452>
- Allix K, Bissyandé TF, Klein J, Le Traon Y (2016) AndroZoo: collecting millions of Android apps for the research community. In: Proceedings of the 13th International Conference on Mining Software Repositories, 468–471. ACM. <https://doi.org/10.1145/2901739.2903508>
- Ankerson MS (2012) Writing web histories with an eye on the analog past. *New Media Soc* 14(3):384–400. <https://doi.org/10.1177/1461444811414834>
- App Annie (2019, January 16) The state of mobile in 2019. App Annie. Retrieved from: <https://www.appannie.com/en/insights/market-data/the-state-of-mobile-2019/>
- Boss K, Broussard M (2017) Challenges of archiving and preserving born-digital news applications. *IFLA J* 43(2):150–157. <https://doi.org/10.1177/0340035216686355>
- Brügger N (2015) A brief history of Facebook as a media text: the development of an empty structure. *First Monday* 20(5). <https://doi.org/10.5210/fm.v20i5.5423>
- Brügger N (2018) The archived web: doing history in the digital age. The MIT Press
- Burgess J, Baym NK (2020) Twitter: a biography. NYU Press
- de Reuver M, Sørensen C, Basole RC (2018) The digital platform: a research agenda. *J Inf Technol* 33(2):124–135. <https://doi.org/10.1057/s41265-016-0033-3>
- Dieter M, Gerlitz C, Helmond A, Tkacz N, van der Vlist FN, Weltevrede E (2019) Multi-situated app studies: methods and propositions. *Social Media Society* 5(2). <https://doi.org/10.1177/2056305119846486>
- Fuller M (2008) Software studies: a lexicon. The MIT Press
- Gerlitz G, Helmond A, Nieborg DB, van der Vlist FN (2019) Apps and infrastructures—a research agenda. *Computat Cult* 7. Retrieved from: <http://computationalculture.net/apps-and-infrastructures-a-research-agenda>
- Gillespie T (2003) The stories digital tools tell. In: Everett A, Caldwell J (eds) *New media: theses on convergence media and digital reproduction*, pp. 107–126. Routledge
- Gillespie T (2010) The politics of ‘platforms’. *New Media Soc* 12(3):347–364. <https://doi.org/10.1177/1461444809342738>
- Helmond A (2015) The platformization of the web: making web data platform ready. *Social Media Society* 1(2). <https://doi.org/10.1177/2056305115603080>

- Helmond A, Nieborg DB, van der Vlist FN (2019) Facebook's evolution: development of a platform-as-infrastructure. *Internet Hist* 3(2):123–146. <https://doi.org/10.1080/24701475.2019.1593667>
- Helmond A, van der Vlist FN (2019) Social media and platform historiography: challenges and opportunities. *TMG – J Media Hist* 22(1):6–34. <https://www.tmgonline.nl/articles/434/>
- Kirschenbaum MG (2003) Virtuality and VRML: software studies after Manovich. *Electronic book review*. Retrieved from: <http://www.electronicbookreview.com/thread/technocapitalism/morememory>
- Memento (2015, April 23) Time Travel APIs. Memento. Retrieved from: <http://timetravel.mementoweb.org/guide/api/>
- Memento (2016, September 22) About the time travel service. Memento. Retrieved from: <http://timetravel.mementoweb.org/about/>
- Natale S (2016) Unveiling the biographies of media: on the role of narratives, anecdotes, and storytelling in the construction of new Media's histories. *Commun Theory* 26(4):431–449. <https://doi.org/10.1111/comt.12099>
- Nieborg DB, Helmond A (2019) The political economy of Facebook's platformization in the mobile ecosystem: Facebook messenger as a platform instance. *Media Cult Soc* 41(2):196–218. <https://doi.org/10.1177/0163443718818384>
- Pollock N, Williams R (2008) *Software and organisations: the biography of the enterprise-wide system or how SAP conquered the world*. Routledge
- Poulsen SV (2018) Becoming a semiotic technology—a historical study of Instagram's tools for making and sharing photos and videos. *Internet Hist* 2(1–2):121–139. <https://doi.org/10.1080/24701475.2018.1459350>
- Rogers R (2017) Doing web history with the internet archive: screencast documentaries. *Internet Hist* 1(2):160–172. <https://doi.org/10.1080/24701475.2017.1307542>
- Statista (2017, October 10) Mobile app usage. Statista. Retrieved from: <https://www.statista.com/topics/1002/mobile-app-usage/>
- Statista (2019a, March 28) Social media statistics. Statista. Retrieved from: <https://www.statista.com/topics/1164/social-networks/>
- Statista (2019b, May 27) Global social media ranking 2019. Statista. Retrieved from: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
- Stevenson M, Gehl RW (2018) The afterlife of software. In: Papacharissi Z (ed) *A networked self and birth, life, death*. Routledge, pp. 190–208
- Thelwall M, Vaughan L (2004) A fair history of the web? Examining country balance in the internet archive. *Libr Inf Sci Res* 26(2):162–176. <https://doi.org/10.1016/j.lisr.2003.12.009>
- Van de Sompel H, Nelson ML, Sanderson R, Balakireva LL, Ainsworth S, Shankar H (2009) Memento: time travel for the web. *ArXiv Preprint*. ArXiv:0911.1112. Retrieved from: <https://arxiv.org/abs/0911.1112>
- van der Vlist FN (2017) Counter-mapping surveillance: a critical cartography of mass surveillance technology after Snowden. *Surveill Soc* 15(1):137–157. <https://doi.org/10.24908/ss.v15i1.5307>