Platform and app histories NEW

Assessing source availability in web archives and app repositories

Anne Helmond (University of Amsterdam)
@silvertje

Fernando van der Vlist (University of Siegen / Utrecht University)
@fvandervlist

This presentation is based on the following work:

Helmond A and van der Vlist FN (2021) Platform and app histories: Assessing source availability in web archives and app repositories. In: Gomes D, Winters J, Demidova E, et al. (eds) *The Past Web*. Springer. [pre-print PDF]

Helmond A and van der Vlist FN (2019) Social Media and Platform Historiography: Challenges and Opportunities. *TMG – Journal for Media History* 22(1): 6–34. DOI: 10.18146/tmg.434/.

Overview

The archived materiality of platforms and apps

- From websites to platforms and apps
    - The website as main unit of historical analysis (Brügger, 2012; 2018)
- The material traces of platforms and apps
- A dual materiality (Helmond and Van der Vlist, 2019; 2021)
    - 1) The material form of digital objects
    - 2) The material circumstances of their production and use (Ankerson, 2012; Fuller, 2008; Gillespie, 2003; Kirschenbaum, 2003)

The material traces of software

"Software is the product of white papers, engineering specs, marketing reports, conversations and collaborations, intuitive insights and professionalized expertise, venture capital (in other words, money), late nights (in other words, labor), Mountain Dew, and espresso. **These are material circumstances that leave material traces** - in corporate archives, in email folders, on whiteboards and legal pads, in countless iterations of alpha versions and beta versions and patches and upgrades, in focus groups and user communities, in expense accounts, in licensing agreements, in stock options and IPOs, in carpal tunnel surgeries, and in the [former] Bay Area real estate market (to name just a few)" (Kirschenbaum, 2003).

Material traces for storytelling

- The material traces may 'tell stories' about the evolving production, preferred usage and embedded politics of software objects `(Gillespie, 2003)`
- Web/platform/app historiography through the lens of software studies `(Ankerson, 2009)`.

Platforms and apps

- Ephemeral digital objects;
- Continuous updates (cf. Acker and Beaton, 2016);
- Where and how are they archived?

Platforms in web archives

https://twitter.com/

http://business.twitter.com/

Types of social media resources and histories

Table A-2. User groups of social media platforms.

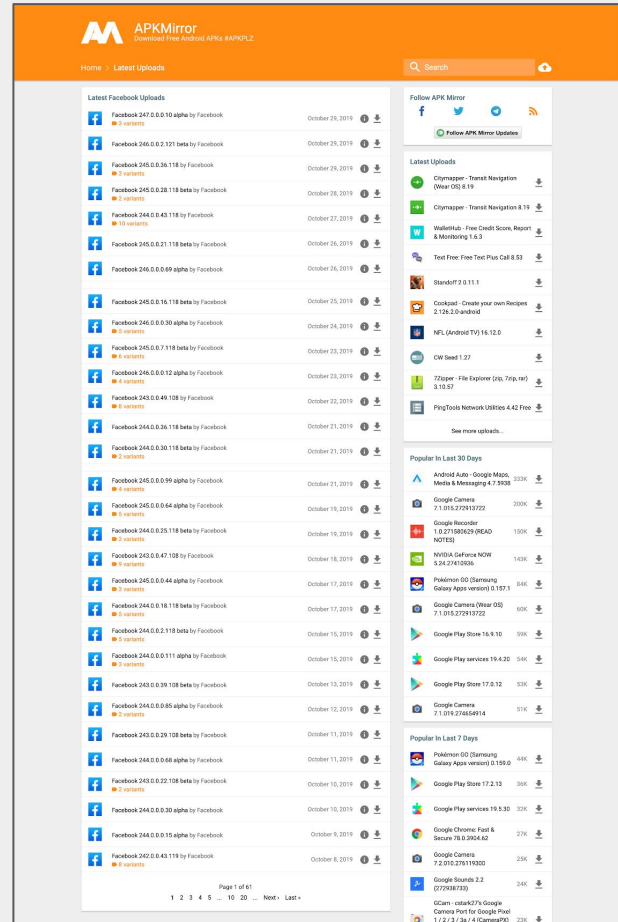| user group	archived resources (sample)	afforded histories (sample)
end users	- graphical user interfaces and 'empty' frames - sign up and registration procedures - about and feature pages - data and privacy policies - terms of service and use - end-user license agreements (EULA) - account security pages - help pages - language support	- self-description histories[17] - feature and practice histories[18] 'stakeholder politics' histories[19] - revenue model histories[20] - data and privacy policy histories - terms of service and use histories[21]
developers	- tools and product pages - application programming interfaces (APIs) and endpoints - software development kits (SDKs) - integrated development environments (IDEs) - software and developer tools and frameworks - guides for app development, best practices, app review, and privacy and consent - online training courses for developers - developer support, help pages, and frequently asked questions (FAQs) - API reference documentation, version histories, and changelogs - developer news, blog posts, and blog archives - open source projects and code repositories - programming, query, and markup languages - bug reports - platform status - annual developer conferences - developers community groups, meetups and local developer communities - startup accelerator programmes - platform policies - careers	- API-based data sharing histories[22] - data strategy and 'intraoperability' histories[23] 'programmability' and app development histories[24] - app ecosystem histories - tracking technology histories - standards and protocol histories 'datastructuring' histories[25] - platform architecture design and governance and control histories[26] - platform growth and embedding histories[27] - platform status, maintenance, repair, and 'issue' histories - platform and privacy policy histories[28]
business	- tools and product pages - business news, blog pages, and archives	- ad creation and targeting fields histories - platform growth and embedding histories[29]

From: Helmond and Van der Vlist (2019), "Social Media and Platform Historiography: Challenges and Opportunities," *TMG – Journal for Media History*.

App archiving challenges

- Mobile phones and app stores overwrite older app versions with each update
- Where are apps archived?
  - Apps ≠ web (are they the focus of web archives?)

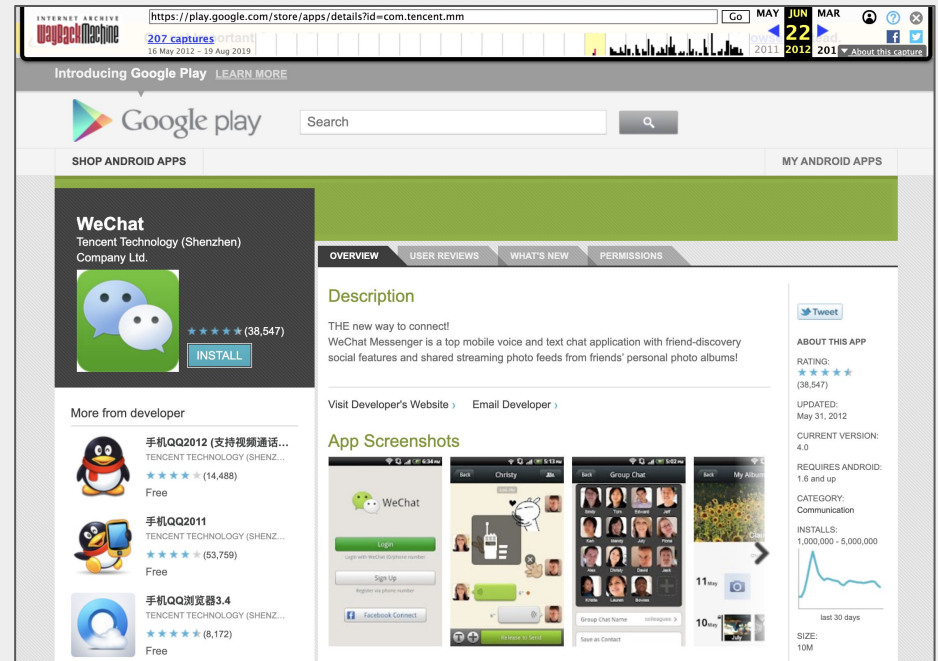From: Helmond and Van der Vlist (2021), "Platform and app histories: Assessing source availability in web archives and app repositories," *The Past Web*.

App repositories

- Software repositories are non-institutional storage locations.
- Not designed for permanent preservation (Allix, Bissyandé, Klein, & Le Traon, 2016).
- Repositories ≠ archives (cf. Brügger, 2018)

From: Helmond and Van der Vlist (2021), "Platform and app histories: Assessing source availability in web archives and app repositories," *The Past Web*.

App metadata in app stores

- App stores archived by web archives
- Snapshots of individual app details pages
- Details: app's title, developer, bundle version, screenshots, description, requested app permissions, download statistics, reviews, ratings, and more (see Dieter et al., 2019).
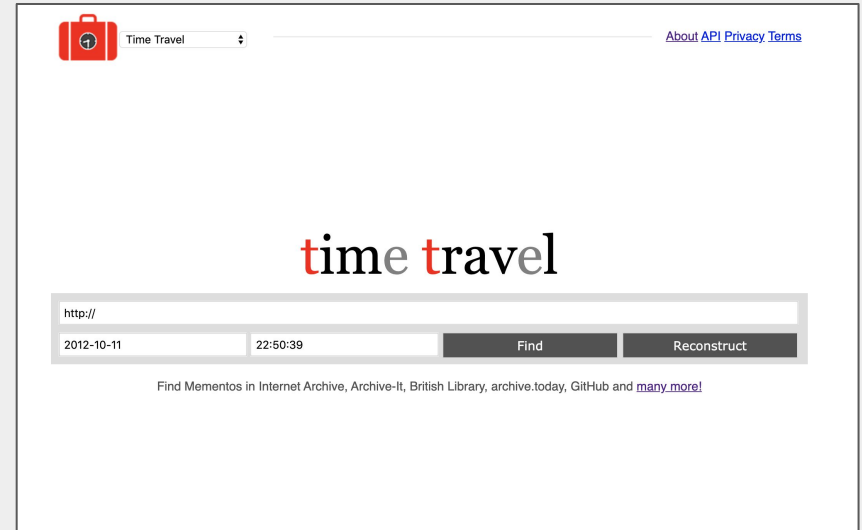

From: Helmond and Van der Vlist (2021), "Platform and app histories: Assessing source availability in web archives and app repositories," *The Past Web*.

Assessing the availability of platform and app sources

- Top-20 most popular social media platform (Statista, 2019)
- Create inventory of their most prominent 'sides' and URL lists of principal materials:
  - e.g. twitter.com, developer.twitter.com, business.twitter.com, marketing.twitter.com, investor.twitterinc.com

- Top-10 mobile apps for Android and iOS combined (App Annie, 2019)
- Create list of URLs pointing to the app store details pages/app repositories for each app.
  - Google Play: https://play.google.com/store/apps/details?id={bundle_id}
  - App Store: https://itunes.apple.com/app/{bundle_id}
    https://itunes.apple.com/us/app/{bundle_id}
    https://itunes.apple.com/us/app/appname/{bundle_id}.

From: Helmond and Van der Vlist (2021), "Platform and app histories: Assessing source availability in web archives and app repositories," *The Past Web*.

Assessing availability: Memento & Memgator

- Retrieve availability of social media platform URLs and app detail pages URLs in 20+ web archives.
- Tool: Memento Time Travel API via Memgator (Alam & Nelson, 2016).

(Van de Sompel et al., 2009)

From: Helmond and Van der Vlist (2021), "Platform and app histories: Assessing source availability in web archives and app repositories," *The Past Web*. Helmond and Van der Vlist (2019), "Social Media and Platform Historiography: Challenges and Opportunities," *TMG – Journal for Media History*.
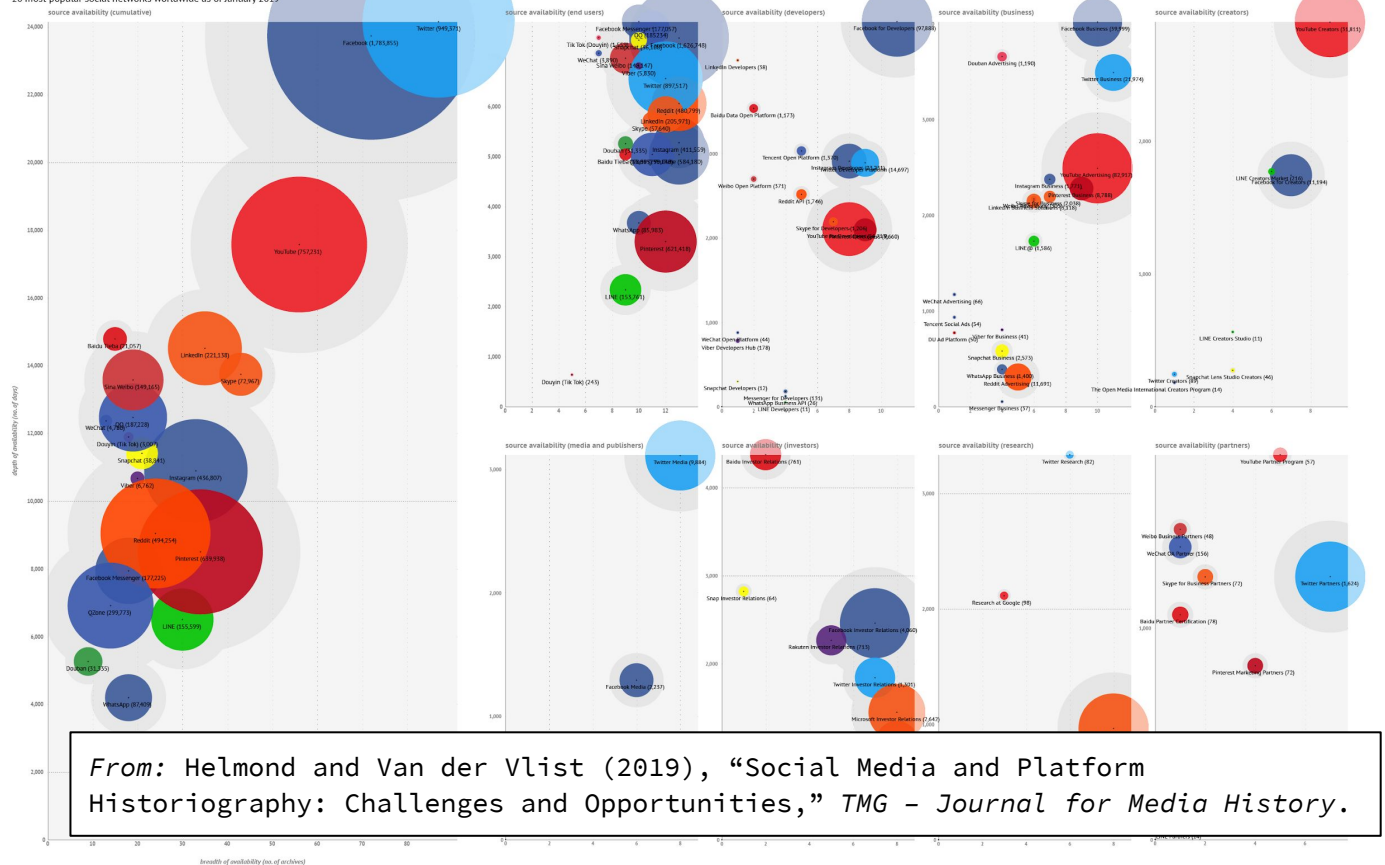
Assessing source availability in web archives

Three key dimensions:

1. the *volume* of availability – the total number of Mementos held
2. the *depth* of availability – the number of days, months, or years between the first and last Mementos
3. the *breadth* of availability – the number of archives holding Mementos

From: Helmond and Van der Vlist (2021), "Platform and app histories: Assessing source availability in web archives and app repositories," *The Past Web*.
Helmond and Van der Vlist (2019), "Social Media and Platform Historiography: Challenges and Opportunities," *TMG – Journal for Media History*.

Assessing source availability in web archives

Source availability in web archives (no. of Mementos held)
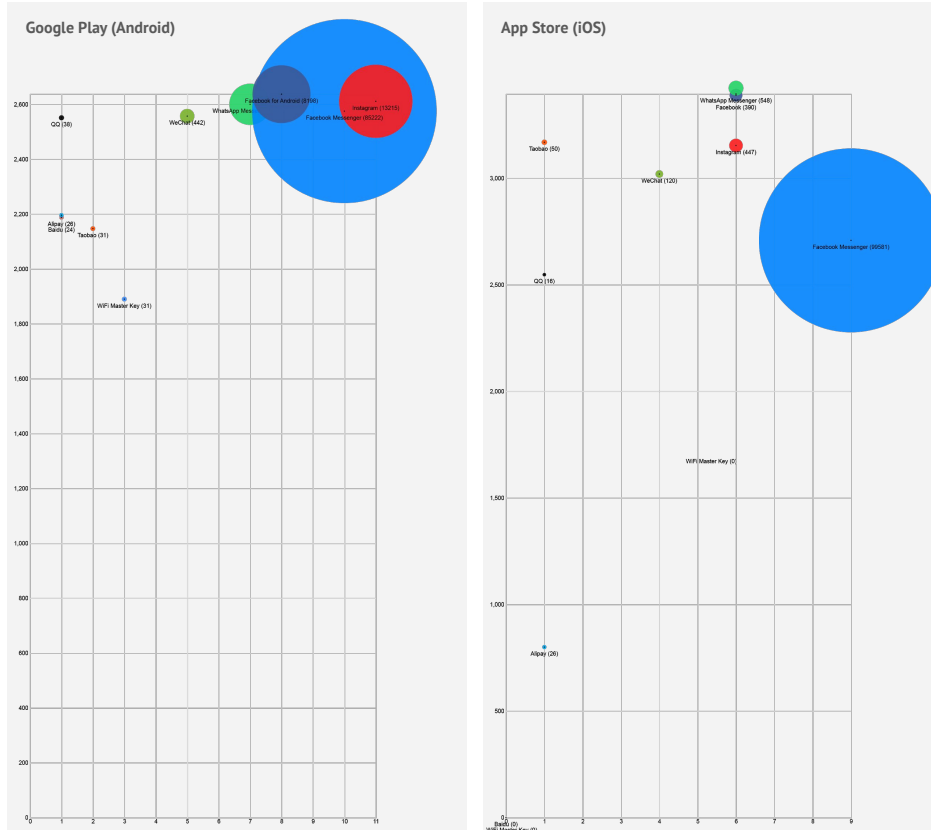20 most popular social networks worldwide as of January 2019

From: Helmond and Van der Vlist (2019), "Social Media and Platform Historiography: Challenges and Opportunities," *TMG – Journal for Media History*.

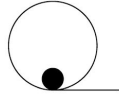# Assessing source availability in web archives

Source availability in web archives (no. of Mementos held)
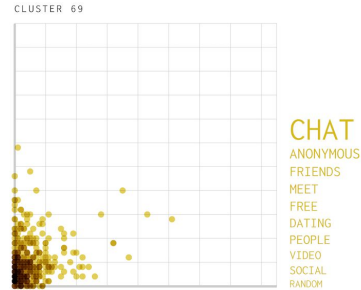
10 most popular apps worldwide as of 2018

Google Play (Android)

QQ (36)
WeChat (442)
WhatsApp Mes...
Facebook for Android (8196)
Instagram (13215)
Facebook Messenger (85222)
Alipay (26)
Baidu (24)
Taobao (31)
WiFi Master Key (31)

App Store (iOS)

WhatsApp Messenger (548)
Facebook (380)
Taobao (50)
Instagram (447)
WeChat (120)
Facebook Messenger (99581)
QQ (16)
WiFi Master Key (0)
Alipay (26)
Baidu (0)
WiFi Master Key (0)

app title	Android (Google Play)				iOS (App Store)			
	volume	depth	breadth	rank	volume	depth	breadth	rank
Facebook	8,198	2,637	8	3	390	3,389	6	4
WhatsApp Messenger	4,092	2,600	7	4	548	3,395	6	2
Facebook Messenger	85,222	2,638	10	1	99,581	2,708	9	1
WeChat	442	2,557	5	5	120	3,019	4	5
Instagram	13,215	2,611	11	2	447	3,153	6	3
QQ	38	2,551	1	8	16	2,547	1	8
Alipay	26	2,188	1	9	26	800	1	6
Taobao	31	2,147	2	7	50	3,168	1	7
WiFi Master Key	31	1,890	3	6	0	0	0	n/a
Baidu	24	2,196	1	10	0	0	0	n/a

Table 1. Availability of archived web sources for top 10 Android and iOS apps across web archives (accumulated). In: Helmond and Van der Vlist (2021), "Platform and app histories: Assessing source availability in web archives and app repositories," *The Past Web*.

Mapping (Secure) Messaging App Ecologies (2016) - Helmond, Van der Vlist, Weltevrede, Ossevoort, Randy, de Gaetano.
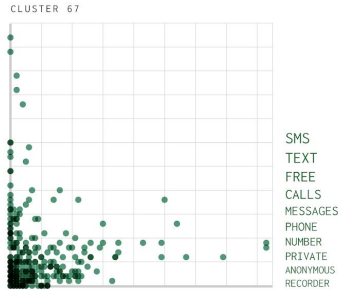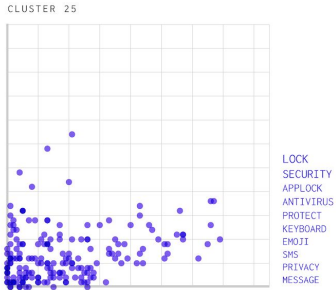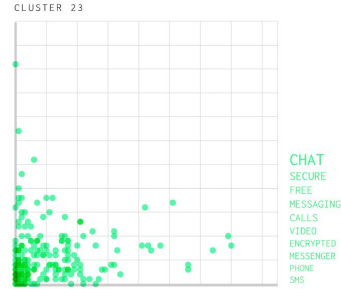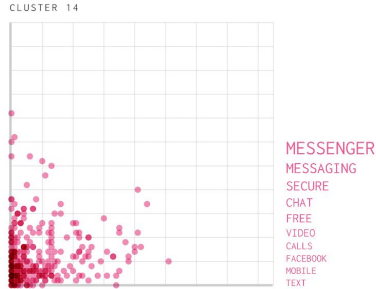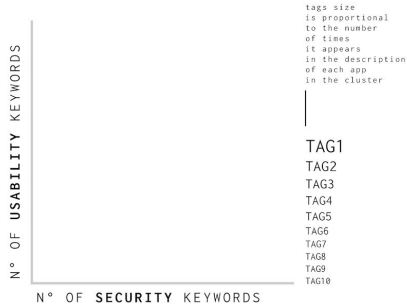
app title	Android			
	volume	depth	breadth	rank
Facebook	4,585	2,584	9	2
WhatsApp Messenger	4,268	2,585	10	1
Facebook Messenger	2,765	2,609	10	4
WeChat	315	2,364	10	6
Instagram	3,271	2,600	10	3
QQ	229	2,187	9	9
Alipay	193	1,362	8	7
Taobao	258	1,844	7	8
WiFi Master Key	623	1,401	8	5
Baidu	166	2,242	5	10

Table 2. Availability of top-10 Android apps across app repositories (accumulated). In: Helmond and Van der Vlist (2021), "Platform and app histories: Assessing source availability in web archives and app repositories," *The Past Web*.

Simultaneous erasure and history writing

"That is, the routine overwriting of digital objects and their data through continuous incremental software updates constitutes both a core problem as well as 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."
(Helmond and Van der Vlist, 2021)

Writing software biographies

- Material erasure as a 'native' mode of software history-writing (Helmond and Van der Vlist, 2021).
- Writing platform and app 'biographies' with archived web sources: 'software biography', 'website biography', and 'platform biography' (cf. Brügger, 2015; Bory & Natale, 2017; Burgess & Baym, 2018/2020; Helmond, Nieborg, and Van der Vlist, 2019; Helmond and Van der Vlist, 2021, Rogers, 2017; Williams & Pollock, 2008).
- Platform and app biographies are co-written by these digital objects and by web archives (Helmond and Van der Vlist, 2021).

Thank you

Further reading:

1. Helmond A and van der Vlist FN (2021) Platform and app histories: Assessing source availability in web archives and app repositories. In: Gomes D, Winters J, Demidova E, et al. (eds) *The Past Web*. Springer. [pre-print PDF]

2. Helmond A and van der Vlist FN (2019) Social Media and Platform Historiography: Challenges and Opportunities. *TMG – Journal for Media History* 22(1): 6–34. DOI: 10.18146/tmg.434/.

Anne Helmond (University of Amsterdam)
a.helmond@uva.nl / **@silvertje**

Fernando van der Vlist (University of Siegen / Utrecht University)
@fvandervlist

References

- Acker A and B Beaton (2016) Software Update Unrest: The Recent Happenings Around Tinder and Tesla. *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pp. 1891–900. doi:10.1109/HICSS.2016.240.
- Ankerson MS (2009) Historicizing Web Design: Software. Style, and the Look of the Web. In: Staiger J and Hake S (eds) *Convergence, Media, History*. New York, NY: Routledge, pp. 192–203.
- Brügger N (2015) A brief history of Facebook as a media text: The development of an empty structure. *First Monday* 20(5). Available at: http://firstmonday.org/ojs/index.php/fm/article/view/5423 (accessed 8 December 2015).
- Bucher T (2013) Objects of Intense Feeling: The Case of the Twitter API. Computational Culture 3. Available at: http://computationalculture.net/article/objects-of-intense-feeling-the-case-of-the-twitter-api (accessed 6 May 2014).
- Burgess J and Baym N (2020). Twitter: A Biography. NYU Press.
- Burgess J and Green J (2018) YouTube: Online Video and Participatory Culture. 2 edition. Cambridge, UK ; Medford, MA: Polity.
- Gillespie, T. (2003). The stories digital tools tell. In A. Everett & J. Caldwell (Eds.), New Media: Theses on Convergence Media and Digital Reproduction (pp. 107–126). Routledge.
- Helmond A and van der Vlist FN (2021) Platform and app histories: Assessing source availability in web archives and app repositories. In: Gomes D, Winters J, Demidova E, et al. (eds) *The Past Web*. Springer. [pre-print PDF]
- Helmond A and van der Vlist FN (2019) Social Media and Platform Historiography: Challenges and Opportunities. *TMG – Journal for Media History* 22(1): 6–34. DOI: 10.18146/tmg.434/.
- Kirschenbaum MG (2003) Virtuality and vrml: Software Studies after Manovich. *Electronic Book Review* (The Politics of Information). Available at: http://www.electronicbookreview.com/thread/technocapitalism/morememory (accessed 21 July 2014).
- Nieborg DB and Helmond A (2019) The political economy of Facebook's platformization in the mobile ecosystem: Facebook Messenger as a platform instance. *Media, Culture & Society* 41(2): 196–218. DOI: 10.1177/0163443718818384.
- Poulsen SV (2018) Becoming a semiotic technology – a historical study of Instagram's tools for making and sharing photos and videos. *Internet Histories* 2(1–2): 121–139. DOI: 10.1080/24701475.2018.1459350.