

Fast Facts: Platforms From Personalization to Centralization

Andrew Iliadis¹  and Heather Ford²

Social Media + Society
July-September 2023: 1–8
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20563051231195546
journals.sagepub.com/home/sms



Abstract

In recent years, there has been a noticeable shift in the economies of search and information retrieval concerning how the products of giant platform companies consolidate and represent facts directly in their search results. Concomitantly, media, communication, and information scholars have recently refocused on how media technology companies variably create, collect, connect, and commercialize data related to facts about the world and how such processes have implications for how we know the world. Such approaches often counter popular narratives that seek to frame the problems of platforms in terms of personalization and personalized content. While research on the personalization afforded by media is widespread, platforms also engage in the centralization of facts by merging web data representing factual claims and offering answers directly in search engines and virtual assistants' results and responses (what we refer to as "fast facts"). These processes considerably affect how knowledge is constructed and shared in a networked society. This special issue collects empirical investigative research on the platformization, exploitation, and centralization of facts while offering a variety of perspectives from which to study these developments, including semantic and infrastructural techniques. This article provides an overview of this field and contextualizes recent media studies on search and information retrieval in broader debates around facts and truth claims.

Keywords

infrastructure, platforms, centralization, semantics, knowledge graphs

Introduction

There is a long and significant history of research on personalization and personalized content afforded by digital media (Beniger, 1987; Bennett & Segerberg, 2011). Personalization has been a fixture of social media research since its beginnings and can be traced back to some of the earliest studies of platforms, including how profiles, news feeds, and search results are customized or customizable for individuals (Baym, 2010; Kennedy, 2008). Such personal affordances were a key selling point for manufacturers and consumers who sought to engage the supposedly liberatory potential of platforms for self-expression, the curated self, and personal content (Gehl, 2011; Livingstone, 2008; Scolere et al., 2018; Thorson & Wells, 2016). Personalization and personalized content are critical aspects of platform marketing and platform-focused research, especially in politics, news, and journalism (Creech & Maddox, 2022; McGregor, 2018; McGregor et al., 2017; Molyneux, 2019). Digital media, too, are a popular source of expression for young people, even though there are tensions between personal and private social spheres (boyd, 2014; Jenkins et al., 2016). Yet not all forms of platform

personalization were seen as positive, and indeed several classic studies have examined the negative implications of personalization in the context of search engine results (Bozdag, 2013; Introna & Nissenbaum, 2000; Khopkar et al., 2003) and, more recently, in research on misinformation, radicalization, and manipulation (Culloty & Suiter, 2021; Krafft & Donovan, 2020). But it is safe to say that, regardless of frames or research questions, personalization practices have been essential to the study of platforms in the past and will continue.

Like all technologies, perspectives on platforms change over time regarding their functions and how they are theorized and discursively constructed (Gillespie, 2010). Platforms function as technical products or sources of

¹Temple University, USA

²University of Technology Sydney, Australia

Corresponding Author:

Andrew Iliadis, Department of Media Studies and Production, Klein College of Media and Communication, Temple University, 2020 N. 13th St., Philadelphia, PA 19122, USA.
Email: andrew.iliadis@temple.edu
Twitter: @andrewiliadis



creative expression for their users (or sometimes both). While platforms have been the subject of personalization studies, often through the lenses of the algorithms and recommender systems that curate and sort personalized third-party content for individuals (Gillespie, 2014; Seaver, 2017), significant developments in search and information retrieval over the last decade have indicated a shift toward the centralization and consolidation of third-party content among giant platform companies (Ford, 2022; Iliadis, 2022). This shift is particularly noticeable in platforms' question-answering functions and the looking up of facts. Rather than acting as a gateway to third-party content (links to websites, documents, and so on), search engines and virtual assistants (as well as social media platforms) are increasingly involved in directly producing answers and facts in response to individuals' questions. For example, currently, Google searches for cities such as "Philadelphia" or "Sydney" will return menus that Google calls "knowledge panels" that contain facts and descriptions of each place. The information in these knowledge panels is authoritatively marked "—Google" at the end and appears not to come from a third-party source like Wikipedia (there is no Wikipedia link). The facts are contained in a centralized source which, in this case, is Google. Ford and Graham (2016a, 2016b) have critically addressed the semantic power of companies like Google in naming, managing, and establishing popular understandings of places in knowledge panels. Google describes these knowledge panels as "information boxes that appear [. . .] when you search for entities (people, places, organizations, things) [. . .] to help you get a quick snapshot of information on a topic based on Google's understanding" (Google, 2023). The company has released white papers describing how it is giving "high-level facts about a person or issue" to provide "users with contextual information [. . .] to help them be more informed consumers of content on the platform" (Google, 2019). But where does Google source the information for this "quick snapshot," and what does it mean by "understanding"?

Studies on search engines describe present-day searches of this nature as "zero-click" searches; in these searches, individuals supposedly find the information they need with just one query without further exploring other sources, although Google has contested this characterization (Ferguson, 2021; Fishkin, 2019, 2021; Sullivan, 2021). Still, the convenience of conducting zero-click searches is perceived as beneficial and time-saving for many individuals. Most notably, according to one study, "62 percent of mobile searches in June 2019 were no-click," and "people ages 13 to 21" are "twice as likely" as people over 50 "to consider their search complete" once they have seen a "knowledge panel" (Kelley, 2019). A recent search engine survey showed 51% of respondents "indicated that they 'very frequently' or 'often' make important life decisions based on Google information" and that "95% of respondents across all age groups find the Knowledge Panel results to be at least 'trustworthy'" (Ray, 2020). Companies seek to take advantage of these

consumer behavior changes by increasing their content, including centralizing facts in knowledge panel results. As shown in a recent investigative report by *The Markup*, currently, "41% of the first page of Google search results is taken up by Google products" (Jeffries & Yin, 2020). These practices led the US Department of Justice to file a federal lawsuit, accusing Google of "illegally monopolizing the market for search through anticompetitive behavior" (Jeffries, 2020). The filing states, "Google has taken steps to close the ecosystem from competition and insert itself as the middleman between app developers and consumers." Among Google's properties in its top search results, the reporting found that some included knowledge panels that "show summaries and facts drawn from the 'knowledge graph,' Google's database of facts and entities curated from various sources." According to a study cited in the report, customers are often mistaken about where the content that populates these items comes from, confusing Google and Wikipedia information (McMahon et al., 2017).

Every day, media technologies are changing the way people access information. Instead of only guiding people to various sources, search products now directly provide facts and answers to questions. Search engines, applications, platforms, and virtual assistants have taken on the role of generating consumer information. Consequently, searches no longer lead to other sources, such as ranked search results or Wikipedia pages. Instead, consumers are presented with answers that seem to come directly from the companies that make search products. These developments indicate a gradual shift to a new era in media, where companies attempt to centralize and consolidate factual content for individuals engaged in question-answering searches. But the rationalization of such techniques is tied to long-standing business practices seeking to grow companies' market share. Throughout history, large media companies have often monopolized markets, consolidating power and ownership (Bagdikian, 1983/2004). This trend extends to contemporary internet companies mediating facts (Mosco, 1996/2009; McChesney, 2013). Studies have demonstrated how capitalism influences and shapes products like search engines by privatizing search (Mager, 2012). Researchers have also highlighted the emergence of new forms of "semantic capitalism" (Feuz et al., 2011; Floridi, 2018), in which market logic governs the production and dissemination of meaning and facts. According to Thornton (2017, 2018), Google search and advertising strategies involve monetizing language within regular search results. The outcome often leads to a semantic mismatch between the words users input into the search engine and the descriptions appearing in search results. For instance, currently, searching for the term "cloud" on Google will primarily yield a knowledge panel for cloud technology, even if the intention is to find information about clouds in the sky, while searching for "sky" will present a knowledge panel for the British telecommunications company with the same name (this is true at the time of this

writing, but search results change frequently). Although Google is gradually introducing disambiguation options for knowledge panels to address these issues, many sociocultural biases persist.

Platform companies strive to deliver facts and information to users, aiming to keep them engaged with their own media products and associated offerings instead of directing them to the web products of their competitors. While companies like Google have traditionally positioned themselves as the primary gateway to information, their approaches have evolved to provide users with answers to their queries directly. Various strategies are employed to achieve this, including platform companies taking advantage of large language models and generative AI. For example, Microsoft's Bing uses ChatGPT (Mehdi, 2023), a generative pretrained transformer, while Google Search uses Bard (Pichai, 2023), a language model for dialogue applications. Each product provides answers as facts in search results that are often not linked to any sources (currently, Google does not provide them, while Bing does, though they are difficult to parse). Thus, the content these companies' language models produce becomes a source. There are also perhaps lesser-known but long-standing tools such companies use to populate facts in knowledge panels, such as semantic technologies (standardized metadata vocabularies and ontologies) for making web data exchangeable and interoperable. These semantic technologies include things such as open knowledge bases like Wikidata, which packages facts using metadata for retrieval by platforms from its database, or universal web schemas such as Schema.org, whose metadata vocabulary is used by web developers to mark up their pages with facts so that platforms can represent them directly in search results (Ford, 2022; Iliadis et al., 2023). Text-based search engines like Baidu, Google, Bing, Yahoo, and Yandex, along with virtual assistants and voice search platforms such as Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google Assistant, rely on (and in some cases, exploit) these various semantic technologies designed to retrieve and present facts to users, which has the concomitant effect of sometimes stripping them of verifiability and provenance (Orlowski, 2014). The use of these techniques in labeling data is a double-edged sword for web developers, who want their information to appear directly in search results but who are negatively impacted by platforms that extract facts and thus dissuade people from visiting the sources from where those facts originated (leading to loss of traffic, revenue, verifiability, etc.). Still, companies are increasingly also using proprietary databases of facts they have created or purchased, as in the case of Google's acquisition of the Metaweb company and its Freebase knowledge base (Iliadis, 2022). All major platform companies, including Amazon, Microsoft, and Apple, use such techniques, leading to the centralization and consolidation of facts and information.

Knowledge Graphs

Until approximately a decade ago, search engines primarily generated a list of web pages based on keyword matching in response to a search query. However, around 2012, Google shifted significantly in its search approach (as did other companies). Instead of solely focusing on keyword matching, it prioritized the interpretation of search queries to uncover their intended meaning. This transformation led to the introduction of various products and updates, the most notable being the Google Knowledge Graph. According to a blog post by Amit Singhal, the former Senior Vice President of Search at Google, this change marked a transition in Google's focus from merely identifying algorithmic "strings" of content through keyword matching and ranking to comprehending the significance of conceptual "things" (Singhal, 2012). Whether such comprehension has been realized is debatable. Still, it did result in Google shifting its attention toward recognizing and describing objects, relationships, and processes in the physical world that could be provided as factual responses to user queries. "The perfect search engine should understand exactly what you mean," Singhal wrote. This release marked Google's official entry into the domain of computational ontology building (Iliadis, 2018, 2019) as it endeavored to define entities and their relationships to gain a deeper understanding of their meaning in the real world. Once this understanding was established, Google could produce, connect, and link relevant facts and information accordingly.

The announcement released by Singhal describes how the Knowledge Graph "enables you to search for things, people or places that Google knows about [. . .] and instantly get information that's relevant to your query." The project is described as "a critical first step towards building the next generation of search, which taps into the collective intelligence of the web and understands the world a bit more like people do." Once, when reporters asked Singhal about the purpose of the Knowledge Graph, he said, "Sometimes you only need an answer" (Schwartz, 2014). Further information about Google's Knowledge Graph is provided in a co-authored research paper by engineers at competing platform companies (such publications are a relatively rare occurrence in the corporate world), where Google researchers describe the Knowledge Graph as "a long-term, stable source of class and entity identity that many Google products and features use behind the scenes" and that it "helps Google products interpret user requests as references to concepts in the world of the user" (Noy et al., 2019). The researchers at these companies (Allhutter [2019] describes such people as "working ontologists") further describe how knowledge graphs help with actions; they recognize "that certain kinds of interactions can take place with different entities"—for example, a search for "'Russian Tea Room' provides a button to make a reservation, while a query for 'Rita Ora' provides links to her music on various music services" (Noy et al., 2019). The

Google Knowledge Graph currently contains roughly 1 billion entities and 70 billion assertions related to facts about the world—in comparison, Microsoft’s contains 2 billion entities and 55 billion facts, and Facebook’s has 50 million entities and 500 million assertions (Noy et al., 2019). Amazon, Facebook, Microsoft, Apple, and other companies have each developed proprietary iterations of knowledge graphs, which are actively used in each company’s respective products.

Fast Facts

Media, communication, and information researchers have noticed subtle shifts in the centralization and consolidation of facts on platforms that have evolved over the last decade, and several books have recently been released that address these issues. Jutta Haider and Olof Sundin’s (2019) *Invisible Search and Online Search Engines: The Ubiquity of Search in Everyday Life* describes Google (and other search companies) in the context of searching for documents versus searching for content and how search results now provide both while also blurring the lines between them (e.g., search results will give the references to facts in links as well as bare facts represented in knowledge panels). The book examines how “Google changes its role as an index and provider of links, framed as unbiased, to very openly becoming an arbiter and even a producer of knowledge presented as factual answers to questions” (Haider & Sundin, 2019, p. 15) and how “Google is changing from a pure reference database to a fact provider” (p. 95).

One of our recent books, *Writing the Revolution: Wikipedia and the Survival of Facts in the Digital Age* (Ford, 2022), takes up these issues, discussing how knowledge graphs “represent the future of knowledge discovery as search engines and voice assistants have been redesigned to encourage us to ask questions of our *devices* rather than *search for information on websites*” (Ford, 2022, p. 2), and how such products take advantage of how facts are wrapped in metadata (semantic technologies) which is “rocket fuel for facts because it enables machines to recognize and extract facts that can later be represented as answers to user queries” (p. 8). Facts are encoded through semantic technologies (such as those offered by Wikidata and Schema.org), and the knowledge graphs of giant platform companies extract these data (see Ford & Iliadis, in this special issue). Ford (2022) describes how the facts in knowledge graphs often lack information about their sources, leaving users unaware of their origin. Furthermore, there is a lack of transparency regarding the selection of one factual statement over others and the process for correcting inaccurate or misleading information. The development of knowledge graphs has primarily focused on computational rules rather than rules that account for the inherent ambiguity of knowledge, particularly about rapidly evolving political events. Automating facts from Wikipedia to knowledge graphs in

Wikidata and Google can result in significant information loss, altering the meaning of events and making it challenging, and sometimes impossible, for users to rectify false or misleading information.

Another of our recent books, *Semantic Media: Mapping Meaning on the Internet* (Iliadis, 2022), outlines how the history, theories, and technologies related to the early “Semantic Web” project initiated by Sir Tim Berners-Lee and the World Wide Web Consortium (Berners-Lee et al., 2001; Berners-Lee & Fischetti, 1999) are associated with the development of semantic media products such as knowledge graphs. Knowledge graphs are described as the latest iteration of semantic technologies, preceded by terms such as “linked data” and the “semantic web.” “Knowledge graphs” as a term is described as having become the main buzzword of such technologies due to Google’s use of the term, and the book argues that Google’s use of such tools represents somewhat of a corporate capture of early Semantic Web technologies, which were once envisioned as helping to grow an open and free internet. The book explains how semantic technologies came to be used by large media technology companies to provide facts to people and consumers in products like search engine results and virtual assistants. It offers three case studies focusing on Google’s Knowledge Graph, Schema.org, and Wikidata while explaining that, contrary to popular accounts which view the Semantic Web as a failed project, the Semantic Web never really “died” but instead went to work for companies like Google, Amazon, and Microsoft. The book ends by articulating some of the problems of semantic media technologies like knowledge graphs and virtual assistants, including loss of verifiability, consolidation of knowledge, exploitation of social data, creation of gatekeepers to knowledge, and typification of logic, socio-cultural biases, and misinformation.

Recent studies have further examined these technologies and the significance of their political implications (McDowell & Vetter, 2022; Nielsen & Ganter, 2022; Tripodi, 2022). Multiple studies have examined the semantic functionalities of platforms like knowledge graphs (Kejriwal et al., 2021). Vang (2013) analyzes Google’s role as a semantic entity modeler, focusing on its Knowledge Graph product, arguing that Google’s knowledge panels serve to consolidate its control over information, keeping users within the Google ecosystem. Monea (2016) critically analyzes contemporary knowledge databases, including Google’s Knowledge Graph, discussing limitations in differentiating and representing diverse meanings. Uyar and Aliyu (2015) assess Google’s Knowledge Graph and Bing’s Satori through a series of searches, highlighting how these semantic search systems impose constraints on the conceptual complexity allowed within their respective platforms. Furthermore, others have observed that Google’s search results marginalize smaller companies and organizations that rely on direct interaction with their products, indicating Google’s dominant position and its “structural tendencies towards monopoly” (Rieder &

Sire, 2014). Like its knowledge panels, Google's web-based products frequently give users direct answers to their inquiries and provide immediate opportunities to take specific actions rather than functioning solely as a search engine that directs users to external sources.

Future Directions

The articles in this special issue represent just a few future research directions in fast facts, platforms, and the centralization of information. The first article, "Abortion Near Me? The Implications of Semantic Media on Accessing Health Information," by Francesca Bolla Tripodi and Aashka Dave, explores the impact of semantic media on finding health information through search engines. The researchers investigated how 42 individuals from different counties in North Carolina with varying political affiliations searched for information about abortion. The study revealed that participants' ability to find accurate details was primarily influenced by their stance on abortion (whether they supported or opposed it). In addition, the study found that search engine optimization (SEO) and advertising posed challenges in accessing reliable abortion information.

The second article, "Valuating Words: Semantic Practices in Web Search Advertising," by Anna Jobin, focuses on the role of advertising in digital media industries, specifically web search advertising, which relies on words for targeting purposes. The lack of empirical research on how advertising with words is implemented in practice is addressed. The article presents findings from in-depth interviews with web search advertising professionals, exploring their understanding of Google's "linguistic capitalism." The study identifies seven themes, including three contextual factors (locality, semantic footprints, and governance) that influence the valuation of words and four semantic practices (attaching meaning, ascribing intention, associating algorithmically, and measuring relevance) employed by advertising professionals to interpret the meaning of words. The analysis reveals that the value of words is not fixed but constantly re-evaluated by advertising professionals. It also highlights the significance of semantic practices in commodifying words, reinforcing Google's semantic power. The article contributes to the critical literature on web search and sheds light on the role of meaning-making in algorithmic media.

The third article, "Search Fluency Mistaken for Understanding: Ease of Information Retrieval from the Internet Inflates Internal Knowledge Confidence," by Kristy A. Hamilton and Li Qi, investigates the impact of digital search fluency, specifically the use of featured snippets in search engines, on internal knowledge confidence. The study finds that participants with immediate access to semantic information report higher confidence levels in their inner knowledge than those with delayed or no access to such information. This effect is observed for topics directly related to the retrieved information and unrelated topics. The findings

suggest the enormous impact of features found in semantic search on users' knowledge of the world.

The fourth article, "Wikidata as Semantic Infrastructure: Knowledge Representation, Data Labor, and Truth in a More-than-Technical Project," by Heather Ford and Andrew Iliadis, focuses on the social and political implications of Wikidata, a knowledge base project by the Wikimedia Foundation that contains editable facts and serves as a data source for platform companies and researchers. While previous analyses have praised Wikipedia for its collaborative nature, less attention has been given to the political and economic implications of Wikidata. The article introduces the concept of semantic infrastructure and explores how Wikidata acts as the primary vehicle for Wikipedia to become infrastructural for digital platforms. Two key themes, knowledge representation and data labor, are developed to address power dynamics in infrastructure studies and their relevance to Wikidata. Examining these issues helps situate infrastructural technologies like Wikidata within media and communication studies, emphasizing the contingencies that shape their outcomes.

The fifth article, "Reproductive Health and Semantics: Representations of Abortion in Semantic Models and Search Applications," by Brian Dobreski, Laura Ridenour, and Melissa Resnick, examines the representation of abortion in Wikidata, a knowledge base often used by search applications and its implications for user searches in the context of reproductive health. Following the Supreme Court's decision to overturn *Roe v. Wade*, online information on reproductive health care in the United States has become particularly relevant. The study compares Wikidata's treatment of abortion with three medical domain models and assesses its impact on web search results. The findings reveal that Wikidata attempts to represent abortion from multiple perspectives while simplifying the topic, resulting in logical inconsistencies compared to domain models. Determining Wikidata's influence on semantically supported web searches is challenging due to search engines' exceptional treatment of abortion, although a strong influence from Wikipedia was observed. The study emphasizes how semantic models address the medical domain and advocates for greater transparency in how health care information is handled within web search applications.

The sixth article, "The Hidden History of the Like Button: From Decentralized Data to Semantic Enclosure," by Harry Halpin, highlights the role of semantic technologies in artificial intelligence and how the story of their use is exemplified by the Facebook "Like" button. The "Like" button utilized the decentralized and open Semantic Web standards to collect personal data for advertising across the entire web. Shortly after this, Google introduced the Google Knowledge Graph, a private corporate version of the Semantic Web, and other major companies in Silicon Valley followed suit by creating their proprietary knowledge graphs. This transformation shifted the Semantic Web from a democratic project aiming for standardized open knowledge to a project focused

on control. The shift blurred the line between the actual object and its representation in a knowledge graph, collapsing semantics in the process.

The seventh article, “Comic Vine: Participatory and Idiosyncratic Documentation of a Semantic Platform,” by Hervé Saint-Louis, discusses Comic Vine (CV), a semantic platform created in 2006 by developers Dave Snider, Ethan Lance, and Tony Guerrero. Initially, CV served as a news and review website and discussion forum focusing on the comic industry. The article examines how CV evolved to provide descriptive features tailored to the comics industry using a proprietary architecture. The author uses the walk-through method approach to analyze the platform’s development and finds that although CV does not adhere to open-web standards, it contributes to architectural design diversity (ADD). The concept of ADD emphasizes the existence of various technical schemes in computing sciences that do not rely on a few standardized approaches. The article draws insights from information studies, communication studies, and human-computer interaction to provide contextual understanding.

The eighth and final article, “Semantic Search Engine Optimization in the News Media Industry: Challenges and Impact on Media Outlets and Journalism Practice in Greece,” by Dimitrios Giomelakis, focuses on the role of search engines as gateways to news and the importance of top rankings in search results for online news outlets. It explores how SEO has become essential in newsrooms for disseminating content, presenting new practices and challenges for media professionals. With the transition to the Semantic Web and advancements in search engines, Semantic SEO has emerged as a new approach, bringing additional challenges to the news media industry. The study investigates the application of Semantic SEO in newsrooms and its impact on journalism and news organizations. Through interviews with Greek SEO experts and a systematic review of semantic search and related technologies, the study analyzes how Semantic SEO influences news content and identifies technological practices that can improve the discoverability of news content in the evolving landscape of online search.

Conclusion

As fast facts become more commonly embedded in search engines and virtual assistants, media, communication, and information researchers must examine these technologies’ social and political stakes. This special issue is one attempt at starting this work, and other special issues have likewise sought to turn a critical eye back on search engine studies and Google. A recent *Big Data & Society* issue on “The State of Google Critique and Intervention,” edited by Astrid Mager, Ov Cristian Norocel, and Richard Rogers, contains papers that touch on some of the themes represented here. Two special issues in the *Journal of the Association for Information Science and Technology*, one on “Healthier

Information Ecosystems” and one on “Re-orientating Search Engine Research in Information Science,” also focus on such themes. Studies on fast facts will also become important as large language models and knowledge graphs are increasingly used in search engines and virtual assistants, which will likewise require proper data labeling and annotation for verifiability and provenance (Bender et al., 2021; Gebru et al., 2021). Like fast foods, where long struggles over adding nutritional labels have been waged over several years, fast facts also need proper labeling to maintain a healthy informational diet. But more than this, giant platform companies must be held accountable for extracting, centralizing, and consolidating facts produced by internet sources and users, which effectively strips information of its historicity.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Temple University Grant-in-Aid.

ORCID iD

Andrew Iliadis  <https://orcid.org/0000-0002-8345-6251>

References

- Allhutter, D. (2019). Of “working ontologists” and “high-quality human components”: The politics of semantic infrastructures. In D. Ribes & J. Vertesi (Eds.), *DigitalSTS: A field guide for science & technology studies* (pp. 326–348). Princeton University Press.
- Bagdikian, B. H. (1983/2004). *The new media monopoly: A completely revised and updated edition with seven new chapters*. Beacon Press.
- Baym, N. K. (2010). *Personal connections in the digital age*. Polity.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610–623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>
- Beniger, J. R. (1987). Personalization of mass media and the growth of pseudo-community. *Communication Research*, 14(3), 352–371. <https://doi.org/10.1177/009365087014003005>
- Bennett, W. L., & Segerberg, A. (2011). Digital media and the personalization of collective action: Social technology and the organization of protests against the global economic crisis. *Information, Communication & Society*, 14(6), 770–799. <https://doi.org/10.1080/1369118X.2011.579141>
- Berners-Lee, T., & Fischetti, M. (1999). *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*. Harper.

- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), 34–43. <https://www.scientificamerican.com/article/the-semantic-web/>
- boyd, D. (2014). *It's complicated: The social lives of networked teens*. Yale University Press.
- Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15, 209–227. <https://doi.org/10.1007/s10676-013-9321-6>
- Creech, B., & Maddox, J. (2022). Thus spoke Zuckerberg: Journalistic discourse, executive personae, and the personalization of tech industry power. *New Media & Society*. Advance online publication. <https://doi.org/10.1177/14614448221116344>
- Culloty, E., & Suiter, J. (2021). *Disinformation and manipulation in digital media: Information pathologies*. Routledge.
- Ferguson, J. (2021, April 29). SEO professionals: Stop sharing debunked zero-click search statistics. *Search Engine Journal*. <https://www.searchenginejournal.com/stop-quoting-zero-click-search-studies/403763/>
- Feuz, M., Fuller, M., & Stalder, F. (2011). Personal Web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation. *First Monday*, 16(2). <https://doi.org/10.5210/FM.V16I2.3344>
- Fishkin, R. (2019, August 13). Less than half of Google searches now result in a click. *SparkToro*. <https://sparktoro.com/blog/less-than-half-of-google-searches-now-result-in-a-click/>
- Fishkin, R. (2021, March 22). In 2020, two thirds of Google searches ended without a click. *SparkToro*. <https://sparktoro.com/blog/less-than-half-of-google-searches-now-result-in-a-click/>
- Floridi, L. (2018). Semantic capital: Its nature, value, and curation. *Philosophy & Technology*, 31(4), 481–497. <https://doi.org/10.1007/s13347-018-0335-1>
- Ford, H. (2022). *Writing the revolution: Wikipedia and the survival of facts in the digital age*. MIT Press.
- Ford, H., & Graham, M. (2016a). Provenance, power and place: Linked data and opaque digital geographies. *Environment and Planning D: Society and Space*, 34(6), 957–970. <https://doi.org/10.1177/0263775816668857>
- Ford, H., & Graham, M. (2016b). Semantic cities: Coded geopolitics and the rise of the semantic web. In R. Kitchin & S.-Y. Perng (Eds.), *Code and the city* (pp. 200–214). Routledge.
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Iii, H. D., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92. <https://doi.org/10.1145/3458723>
- Gehl, R. W. (2011). Ladders, samurai, and blue collars: Personal branding in Web 2.0. *First Monday*, 16(9). <https://firstmonday.org/ojs/index.php/fm/article/view/3579>
- Gillespie, T. (2010). The politics of “platforms.” *New Media & Society*, 12(3), 347–364. <https://doi.org/10.1177/1461444809342738>
- Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167–194). MIT Press.
- Google. (2019). *How Google fights disinformation* [Google blog]. https://www.blog.google/documents/37/How_Google_Fights_Disinformation.pdf
- Google. (2023). *About knowledge panels* [Google knowledge panel help]. <https://support.google.com/knowledgepanel/answer/9163198?hl=en>
- Haider, J., & Sundin, O. (2019). *Invisible search and online search engines: The ubiquity of search in everyday life*. Routledge.
- Iliadis, A. (2018). Algorithms, ontology, and social progress. *Global Media and Communication*, 14(2), 219–230. <https://doi.org/10.1177/1742766518776688>
- Iliadis, A. (2019). The Tower of Babel problem: Making data make sense with Basic Formal Ontology. *Online Information Review*, 43(6), 1021–1045. <https://doi.org/10.1108/OIR-07-2018-0210>
- Iliadis, A. (2022). *Semantic media: Mapping meaning on the Internet*. Polity.
- Iliadis, A., Acker, A., Stevens, W., & Kavakli, B. (2023). One schema to rule them all: How Schema.org models the world of search. *Journal of the Association for Information Science and Technology*. Advance online publication. <https://doi.org/10.1002/asi.24744>
- Introna, L. D., & Nissenbaum, H. (2000). Shaping the web: Why the politics of search engines matters. *The Information Society*, 16(3), 169–185. <https://doi.org/10.1080/01972240050133634>
- Jeffries, A. (2020, October 20). In historic antitrust lawsuit, DOJ cites banned words and Google self-preferencing as supporting evidence. *The Markup*. <https://themarkup.org/google-the-giant/2020/10/20/google-antitrust-lawsuit-markup-investigations>
- Jeffries, A., & Yin, L. (2020, July 28). Google’s top search result? Surprise! It’s Google. *The Markup*. <https://themarkup.org/google-the-giant/2020/07/28/google-search-results-prioritize-google-products-over-competitors>
- Jenkins, H., Shresthova, S., Gamber-Thompson, L., Kligler-Vilenchik, N., & Zimmerman, A. (2016). *By any media necessary: The new youth activism*. New York University Press.
- Kejriwal, M., Knoblock, C., & Szekely, P. (2021). *Knowledge graphs: Fundamentals, techniques, and applications*. MIT Press.
- Kelley, L. (2019, September 23). The Google feature magnifying disinformation. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2019/09/googles-knowledge-panels-are-magnifying-disinformation/598474/>
- Kennedy, H. (2008). New media’s potential for personalization. *Information, Community & Society*, 11(3), 307–325. <https://doi.org/10.1080/13691180802025293>
- Khopkar, Y., Spink, A., Giles, C. L., Shah, P., & Debnath, S. (2003). Search engine personalization: An exploratory study. *First Monday*, 8(7). <https://firstmonday.org/ojs/index.php/fm/article/view/1063>
- Krafft, P. M., & Donovan, J. (2020). Disinformation by design: The use of evidence collages and platform filtering in a media manipulation campaign. *Political Communication*, 37(2), 194–214. <https://doi.org/10.1080/10584609.2019.1686094>
- Livingstone, S. (2008). Taking risky opportunities in youthful content creation: Teenagers’ use of social networking sites for intimacy, privacy and self-expression. *New Media & Society*, 10(3), 393–411. <https://doi.org/10.1177/1461444808089415>
- Mager, A. (2012). Algorithmic ideology: How capitalist society shapes search engines. *Information, Communication & Society*, 15(5), 769–787. <https://doi.org/10.1080/1369118X.2012.676056>
- McChesney, R. W. (2013). *Digital disconnect: How capitalism is turning the internet against democracy*. The New Press.
- McDowell, Z. J., & Vetter, M. A. (2022). Fast truths and slow knowledge: Oracular answers and Wikipedia’s epistemology. *Fast*

- Capitalism*, 19(1), 104–112. <https://doi.org/10.32855/fcapital.202201.009>
- McGregor, S. C. (2018). Personalization, social media, and voting: Effects of candidate self-personalization on vote intention. *New Media & Society*, 20(3), 1139–1160. <https://doi.org/10.1177/1461444816686103>
- McGregor, S. C., Lawrence, R. G., & Cardona, A. (2017). Personalization, gender, and social media: Gubernatorial candidates' social media strategies. *Information, Communication & Society*, 20(2), 264–283. <https://doi.org/10.1080/1369118X2016.1167228>
- McMahon, C., Johnson, I., & Hecht, B. (2017, May). The substantial interdependence of Wikipedia and Google: A case study on the relationship between peer production communities and information technologies. *Proceedings of the 11th International Conference on Web and Social Media*, 11(1), 142–151. <https://doi.org/10.1609/icwsm.v11i1.14883>
- Mehdi, Y. (2023). Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web. *Microsoft Blog*. <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>
- Molyneux, L. (2019). A personalized self-image: Gender and branding practices among journalists. *Social Media + Society*, 5(3). <https://doi.org/10.1177/2056305119872950>
- Monea, A. (2016). The graphing of difference: Numerical mediation and the case of Google's Knowledge Graph. *Cultural Studies ↔ Critical Methodologies*, 16(5), 452–461. <https://doi.org/10.1177/1532708616655763>
- Mosco, V. (1996/2009). *The political economy of communication* (2nd ed.). Sage.
- Nielsen, R. K., & Ganter, S. A. (2022). *The power of platforms: Shaping media and society*. Oxford University Press.
- Noy, N., Gao, Y., Jain, A., Narayanan, A., Patterson, A., & Taylor, J. (2019). Industry-scale Knowledge Graphs: Lessons and challenges. *Communications of the ACM*, 62(8), 36–43. <https://doi.org/10.1145/3331166>
- Orlowski, A. (2014, January 13). Google stabs Wikipedia in the front: Is Knowledge Graph killing its readership? *The Register*. https://www.theregister.com/2014/01/13/google_stabs_wikipedia_in_the_front
- Pichai, S. (2023). *An important next step on our AI journey* [Google blog]. <https://blog.google/technology/ai/bard-google-ai-search-updates/>
- Ray, L. (2020, March 2). 2020 Google search survey: How much do users trust their search results? Moz. <https://moz.com/blog/2020-google-search-survey>
- Rieder, B., & Sire, G. (2014). Conflicts of interest and incentives to bias: A microeconomic critique of Google's tangled position on the Web. *New Media & Society*, 16(2), 195–211. <https://doi.org/10.1177/1461444813481195>
- Schwartz, B. (2014, March 12). Google to publishers concerned over the Knowledge Graph; Searchers still need your content. *Search Engine Land*. <https://searchengineland.com/google-publishers-concerned-knowledge-graph-searchers-still-need-content-186325>
- Scolere, L., Pruchniewska, U., & Duffy, B. E. (2018). Constructing the platform-specific self-brand: The labor of social media promotion. *Social Media + Society*, 4(3), 1–11. <https://doi.org/10.1177/2056305118784768>
- Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4(2), 1–12. <https://doi.org/10.1177/2053951717738104>
- Singhal, A. (2012, May 16). *Introducing the Knowledge Graph: Things, not strings* [Google blog]. <https://blog.google/products/search/introducing-knowledge-graph-things-not/>
- Sullivan, D. (2021, March 24). *Google search sends more traffic to the open web every year* [Google blog]. <https://blog.google/products/search/google-search-sends-more-traffic-open-web-every-year/>
- Thornton, P. (2017, November 28). Geographies of (con)text: Language and structure in a digital age. *Computational Culture*. <http://computationalculture.net/geographies-of-context-language-and-structure-in-a-digital-age/>
- Thornton, P. (2018). A critique of linguistic capitalism: Provocation/intervention. *GeoHumanities*, 4(2), 417–37. <https://doi.org/10.1080/2373566X.2018.1486724>
- Thorson, K., & Wells, C. (2016). Curated flows: A framework for mapping media exposure in the digital age. *Communication Theory*, 26(3), 309–328. <https://doi.org/10.1111/comt.12087>
- Tripodi, F. B. (2022). *The propagandists' playbook: How conservative elites manipulate search and threaten democracy*. Yale University Press.
- Uyar, A., & Aliyu, F. (2015). Evaluating search features of Google knowledge graph and bing satori. *Online Information Review*, 39, 197–213. <https://doi.org/10.1108/OIR-10-2014-0257>
- Vang, K. J. (2013). Ethics of Google's knowledge graph: Some considerations. *Journal of Information, Communication and Ethics in Society*, 11(4), 245–60. <https://www.emerald.com/insight/content/doi/10.1108/JICES-08-2013-0028/full/html>

Author Biographies

Andrew Iliadis is an Assistant Professor at Temple University in the Department of Media Studies and Production within the Klein College of Media and Communication and serves on the faculties of the Media and Communication Doctoral Program, Cultural Analytics Graduate Certificate Program, and Science, Technology, and Society Network. He is the author of *Semantic Media: Mapping Meaning on the Internet* (Polity, 2022) and co-editor of *Embodied Computing: Wearables, Implantables, Embeddables, Ingestibles* (MIT Press, 2020).

Heather Ford is an Associate Professor at the University of Technology Sydney (UTS) in the School of Communications, Coordinator of the UTS Data and AI Ethics Cluster, Affiliate of the UTS Data Science Institute, and Associate of the UTS Center for Media Transition. She was formerly Google Policy Fellow at the Electronic Frontier Foundation and Executive Director of iCommons. She is the author of *Writing the Revolution: Wikipedia and the Survival of Facts in the Digital Age* (MIT Press, 2022).