Co-constitutive complexity

Unpacking Google's privacy policy and terms of service post-GDPR

Bjarki Valtysson,[I] Rikke Frank Jørgensen,[II] & Johan Lau Munkholm[III]

[I] Department of Arts and Cultural Studies, University of Copenhagen, Denmark
[II] Danish Institute for Human Rights, Copenhagen, Denmark
[III] Department of Communication, University of Copenhagen, Denmark

Abstract
Google is the gateway to the Internet for billions of people. However, to use Google's multiple platforms and services, users must accept Google's terms. With the advent of the EU's GDPR (General Data Protection Regulation), Google made significant changes to these terms. In this article, we scrutinise the intertextual relations between Google's privacy policies and terms of service (ToS) and the GDPR – and the discursive co-constitutive complexity within and between these frameworks. We argue that the material and communicative articulation of Google's privacy policies and ToS should be understood as deliberative data politics delimiting users' agency, consent, and privacy. Furthermore, we emphasise complexity and the demands of reducing complexity as two opposing dynamics. While the GDPR required Google to make its terms and policies clearer and more understandable, ironically, in the process of accommodating GDPR's demand of increased transparency, the discursive complexity of Google's policies has in fact increased.

Keywords: data politics, user rights, privacy policy, Google, GDPR

Introduction

In January 2019, the French National Data Protection Commission (CNIL) imposed a financial penalty of EUR 50 million on Google for breaches of the EU's General Data Protection Regulation (GDPR) (CNIL, 2019). CNIL found that Google's contractual agreements were structured in a way that undermines transparency, provides inadequate information about the purposes of data processing, and lacks valid user consent for advertisement personalisation. CNIL emphasises three problematic aspects of Google's contractual agreements: the complicated, interrelated structuring of various documents; the vagueness and general nature of the language used; and the convoluted ways of acquiring user consent. In this article, we focus specifically on the first aspect, the complex nature of Google's interrelated document structure and CNIL's (and the GDPR's) re-

Valtysson, B., Jørgensen, R. F., & Munkholm, J. L. (2021). Co-constitutive complexity: Unpacking Google's privacy policy and terms of service post-GDPR. *Nordicom Review*, *42*(1), 124–140. https://doi.org/10.2478/nor-2021-0033

quirement that this complexity is relieved. Our ambition is to both scrutinise the structure of Google's policy landscape and to problematise their means of reducing complexity, thus accommodating CNIL. We argue that the material (inter-related) and communicative (vague and generic) articulation of Google's policies are best understood as deliberative data politics. These data politics are embedded within a specific socioeconomic system with its own interest in keeping policies generic and complex. It further proposes that the GDPR – itself a complex legal document – may be used to produce more, rather than less, complexity, when used in regulatory negotiations by powerful actors. While the complexity of a legal instrument such as the GDPR is unsurprising, this still creates a paradox in terms of the issues which CNIL points to, such as the need for increased transparency and the demand that users – and laypeople – should be able to understand the terms and policies put forward by platforms.

The ambition is therefore to demonstrate how complexity plays out in the language and interrelated structure of Google's privacy policies and terms of service (ToS), and how the implementation of the GDPR has enhanced the discursive complexity for users. Even though we engage with socio-legal and socioeconomic perspectives, our primary focus is on the discursive complexity of these frameworks and how the generation of such complexity creates less transparency for users, rather than providing socio-legal solutions to this problem.

Empirically, we focus on Google's privacy policy, ToS, and the EU's GDPR. The former is not limited to Google's privacy policy, but further branches out into multiple Google services partly covered under the privacy policy. By critically scrutinising the intertextual relations of these documents, we lay the foundation for emphasising complexity, as well as the demand to reduce such complexity, as two opposing dynamics at stake both in relation to Google and the GDPR. To account for the political economy of Google's policies, we consider terms such as platforms, data politics, and user rights, and how these notions clarify and illuminate the document analysis. We do this to further account for how privacy policies and ToS, produced by Google, are intertwined with an economic rationale, as well as with contestant actors such as EU regulators, whose objective of transparent practices may be interlocked in a paradoxical production of ever more discourse.

Methodological framework

Our analysis of the intertextual relations of Google's privacy policies and ToS and the GDPR is inspired by Norman Fairclough's discourse theory, and the relationships between text, discourse, and social practice. The analysis is attentive to the dialectical relationships between discourse and social practice, "with how discourse figures within processes of change, and with shifts in the relationship between discourse and more broadly semiosis and other social elements within networks of practices" (Fairclough, 2003: 205). Our main focus is on how key terms change and stabilise over time, and how these compare between Google's privacy policies and the GDPR. We are therefore particularly interested in Fairclough's (1992: 232) notion of intertextual chains, as it is useful to "specify the distribution of a (type of) discourse sample by describing the intertextual chains it enters into, that is, the series of text types it is transformed into or out of". This concept is useful to demonstrate the association between different documents,

how discourses are constructed over time, and how some of these gain dominance and form clusters. We further evolve the notion of intertextuality to *vertical* intertextuality and *horizontal* intertextuality.

Vertical intertextuality should be understood as changes in contractual agreements and policies *over time*. Vertically, Google's privacy policies date back to 9 June 1999, and there are 35 different versions to date; the current version was effective from 4 February 2021. Similarly, there are 19 versions of Google's ToS, with the first dating back to 20 September 1999 and the most recent to 31 March 2020 (Google Privacy & Terms, 2020). Horizontal intertextuality should be understood as the intertextual relations between different kinds of documents, be they *internal* – for instance, Google's privacy policy and ToS – or *external*, marking the intertextual relations between Google's contractual agreements and the GDPR.

Table 1 *Two kinds of intertextuality*

Concept	Explanation	Examples
Vertical	Intertextuality over time (within one kind of document)	Google's privacy policies (1999-present); Google's terms of service (1999-present)
Horizontal	Intertextuality between different kinds of documents or texts	How privacy policies are related to terms of services; how privacy policies and terms of service are related to GDPR

The analysis starts by accounting for the vertical intertextuality that shapes Google's privacy policy and the GDPR. Next, it accounts for horizontal intertextuality, focusing specifically on how Google's privacy policy and the GDPR treat the notions of third party and personal information.[1]

The CNIL decision will not be used directly as an object of analysis, but rather serve as an indicator of the various problems which can be identified in the discursive construction of Google's privacy policies and ToS. Even though this article primarily focuses on the discursive complexity of policies and regulation pertaining to privacy and data protection, it will also briefly relate to the recent EU proposals for a Digital Services Act and a Digital Markets Act. First, in order to interrogate the theoretical context of our analysis, we discuss platforms, data politics, and user rights.

Platforms, data politics, and user rights

In Benjamin Bratton's (2015: 41) account, the most novel challenge introduced by digital platforms is that they "are simultaneously organizational forms that are highly technical, and technical forms that provide extraordinary organizational complexity to emerge", and "as organizations, they can also take on a powerful institutional role, solidifying economies and cultures in their image over time". As a technical entity, a platform is characterised as a "standards-based technical-economic system that simultaneously distributes interfaces through their remote coordination and centralizes their integrated control through the same coordination" (Bratton, 2015: 42). The technical system operationalises a particular institutional program that determines the coordina-

tion of graphical user interfaces. These, in turn, set the terms for user agency within the platforms' policies: "Platforms are generative mechanisms – engines that set the terms of participation according to fixed protocols" (Bratton, 2015: 44). It is because platforms can determine the technical conditions for user participation that they can be considered institutional models that govern the possible actions and participation of users. According to Van Dijck (2013), platforms should therefore be perceived as techno-cultural constructs and socioeconomic structures. From a techno-cultural perspective, this entails attentiveness towards how technology shapes usage and users, and what kind of content platforms allow:

> Technologically speaking, platforms are the providers of software, (sometimes) hardware, and service that help code social activities into a computational architecture; they process (meta) data through algorithms and formatted protocols before presenting their interpreted logic in the form of user-friendly interfaces with default settings that reflect the platform owner's strategic choices. (Van Dijck, 2013: 29)

A platform like Google, for example, is structured by a design logic that generates an architectural, computational, and political program that includes strategies for organising the publics that generate information translatable to surplus value. As such, the business model is integrated into both software and hardware in ways that make Google capable of integrating additional services. The functions of the code are made available to developers through the application programming interface (Evans & Schmalensee, 2011).

This ties further into what Van Dijck refers to as socioeconomic structures, thereby directing attention towards platforms business models, ownership, and governance. Even though the governing logics of major platforms are partly accessible through ToS and privacy policies, the real effect of their status as regulatory devices, and how they chime into current regulatory frameworks at national and supranational levels, is ambiguous (Klass, 2019; Solove & Hartzog, 2014; Van Dijck, 2013). This ambiguity is closely linked to the interconnections of multinational platforms within a socioeconomic system that increasingly relies on the collection, processing, and monetisation of information. Understanding the data-based modulations within existing modes of production has led to several conceptual proposals for understanding this recent form of capitalism: platform capitalism (Srnicek, 2016), surveillance capitalism (Zuboff, 2015, 2019), and Big Data capitalism (Fuchs, 2019), to name a few. Couldry and Mejias (2019), on the other hand, argue that the value extraction from human relations into data is inherent in the process of capitalist accumulation and expansion. For this reason, they reject new qualifications of capitalism and contend that recent social developments occurring around information-based economies reflect capitalism as it has always been: "The systematic organization of life so as to maximize value, resulting in the concentration of power and wealth in very few hands" (Couldry & Mejias, 2019: 32).

The few hands of major platform providers such as Apple, Google, Microsoft, Amazon, and Facebook offer an interface which works as an access point for the consumption and production of information connecting users to their proprietary and transnational services, where granular bits of user data flow. The GDPR was developed partly in response to such data processing to update and strengthen existing rules on data processing (Hoofnagle et al., 2019), to address an increasing imbalance between platform giants and

their users (Vanberg, 2021), and to ensure user rights such as right of access (Ausloos & Dewitte, 2018). In response to the GDPR, Google's privacy policy proposes that users may control their privacy via the platform's privacy settings, and further explain the volume and "aftermath" of the collection of user data. As we will see in the analysis of Google's privacy policy and ToS, this "aftermath" is discursively far too multifaceted to be contained within an easily understood privacy policy and ToS. Moreover, contrary to platform-neutral settings that would allow for do-no-track, Google's privacy settings require the user to create an account and identify themselves, which ironically allows for the collection of more personal data. In fact, even though Google claims its policies are transparent, unfolding the manifold and connected layers entangled within these policies reveals comprehensive volumes of data. Transparency is therefore unlikely to equate with intelligibility, even if Google were committed to revealing their mode of operations. As Marilyn Strathern (2000) demonstrated, transparency is not the same as knowledge, and amplified visibility in a certain domain simultaneously works to conceal other processes. Similarly, Mike Annany and Kate Crawford assert that a multifaceted assemblage such as Google cannot be made transparent simply by peering inside it, but must be understood as a system interrelated with other socio-technical systems "that do not *contain* complexity but *enact* complexity by connecting to and intertwining with assemblages of humans and non-humans [emphasis original]" (Ananny & Crawford, 2018: 974).

In further accounting for the consequences of platform practices, Ruppert and colleagues carve out the notion of data politics. They maintain that "data politics is concerned with not only political struggles around data collection and its deployments, but how data is generative of new forms of power relations and politics at different and interconnected scales" (Ruppert et al., 2017: 2). Data politics emerge from the interrelated conditions of worlds, subjects, and rights – that is, how the creation and governing of material infrastructures (worlds) is generative of politics and struggles directly affecting citizens and data subjects, and thereby also their rights. Similarly, Van Dijck and colleagues (2018) refer to data politics as complex systems of interdependencies around and between dominant platforms, and Couldry and Mejias (2019) refer to data relations and the social quantification sector as a conscious development of infrastructures focused on profiting from human life through data. Indeed, in further scrutinising Google as a dominant provider of infrastructural platforms, Van Dijck and colleagues (2018) point to its influence as a search engine, mobile operating system, web browser, social network service, app store, pay services and advertising service program, video-sharing site, geospatial information system, cloud platform provider, provider of hardware with the pre-installed Google software package, as well as its huge investments in artificial intelligence and machine learning. As a part of Google's deliberative data politics, it attempts to reduce this "platformised" complexity by combining its services within one privacy policy, but as the forthcoming analysis will show, the policy remains a glossy surface hiding the complex data ecology influencing user experience and agency. As creators of the conditions that situate subjects with certain rights within a particular platformised infrastructure (Plantin et al., 2018), or world, platforms become key stakeholders in the politics emerging around data, and each platform represents its own politics (Gillespie, 2010). These politics are only partially found in the platform's ToS and privacy policies, as, for instance, the performance of proprietary algorithms escapes the intention of transparency suggested in such policies.

Here, the central question pertaining to politics is how Google's contractual agreements reflect the socioeconomic logic of the platform while responding to the standards for transparency and user rights set by the GDPR. In terms of regulation, Google's ToS and privacy policy constitute a contractual relationship between the platform and its users, which defines and delimits their privacy rights. While Google has the power to set the terms of this relationship, they must ensure their policies and practices meet the standards set by the GDPR. As such, their ToS and privacy policy are always relative to another governing body. A key question is therefore how these policies have changed vertically from 1999 to 2021 and how they horizontally tie into increasing services and partners absorbed into the platform, and externally to the GDPR.

Contested space is often defined by struggle, in this case, a struggle over the right to define privacy rights and enforce them in specific contexts (Nissenbaum, 2010). One could argue that while GDPR, as the regulatory framework, has the authority to define and enforce privacy norms, Google attempts to deliver their own interpretation of these norms. This would, however, be a simplification of a much more complex interrelation between the two actors, as it tendentiously highlights a binary conflict while glossing over interdependencies and necessary collaborations between state, supra state, and corporate actors in constructing governmental and corporate services (Easterling, 2014). Google, for instance, has invested heavily in cloud computing operations within the European continent. It provides services to a wide variety of European businesses that the EU has a vested interest in serving with its unwavering commitment to economic growth and a business-friendly environment, as well as collaboration in other policy areas such as cybersecurity, the fight against terrorism, and the distribution of illegal content. While the main focus of this article is centred around a misalignment between Google's platform practices and the EUs legal norms, this should not be interpreted to mean that unmitigated conflict is the only game in town. This would offer a one-sided depiction of the multifaceted interrelation between the two actors. From the viewpoint of data politics, a crucial question to be asked, then, is how citizens can claim rights when they engage with data-driven services entwined with complex power structures. What are the mechanisms in place, and how are these tailored as tools for citizens to strike back, to subvert, to take control of their data, and to affect the data streams that actors such as Google take advantage of? This are not easy questions to answer from a socio-legal and socioeconomic perspective, as the delineation between personal data and non-personal data in the GDPR is fraught with difficulty (Finck & Pallas, 2019), as well as how the GDPR affects and challenges personal data usage worldwide (Hoofnagle et al., 2019), how concretely American tech companies can fulfil the requirements of the GDPR, and how these further impact their business models (Houser & Voss, 2018).

As previously noted, our focus is less on these socio-legal and socioeconomic implications, and more on the discursive complexity of Google's privacy policy and ToS and the GDPR, as they are likely to be seen through the eyes of regular users. As our analysis will demonstrate, the documents examined in this article represent two powerful actors that both claim to provide the individual with privacy rights – in both cases, from positions that are discursively removed from the individuals they claim to serve. Google's framing of its users' privacy rights as the ability to adjust privacy settings is thus carefully tailored to allow for the platform's continuous collection and processing of data about its users (Jørgensen, 2017).

In her recent work, legal scholar Julie E. Cohen (2019) turns to the functioning of regulatory frameworks and their role in protecting and promoting human rights, such as the right to privacy. A key challenge, according to Cohen's account, is the unaccountability of private economic power, including the tendency to rely on corporate social responsibility and defer to opaque and often privatised arrangements for the expert supervision of algorithmic processes over more stringent legal obligations. While policy-makers and international organisations can appeal to companies such as Google to adhere to privacy norms, these companies are not subject to binding human rights obligations. With opaque algorithmic arrangements, as, for instance, discussed by Pasquale (2015), Kitchin (2017), and Amoore (2020), Cohen also identifies a challenge in terms of the impenetrability of propriety algorithms and that of complex regulatory frameworks. In both cases, citizens are lost in terms of understanding their rights when using specific platforms: "Effective control of highly informationalized processes requires governance institutions capable of responding in kind, but the very process of optimizing regulatory controls to highly informationalized processes makes governance processes more opaque and less accountable to broader global publics" (Cohen, 2019: 234).

It is this discursive paradox that will be scrutinised further in our analysis, casting light on the vertical and horizontal intertextual relations which Google and the EU establish, using Google's different versions of its privacy policies and ToS, and the GDPR, as a point of departure.

Vertical intertextuality: Google and the GDPR

As previously mentioned, Google's privacy policies have a relatively long history. The focus here will be on the most recent versions, as these are updated to respond to the GDPR, which came into effect on 25 May 2018. It is, however, still important to bear in mind that the latest version is an amendment which has intertextual relations to prior versions that should be regarded as agenda setters, to which newer versions respond. The first version of the privacy policy (1999) is not extensive in scope, but it entails some of the same privacy concerns to which the most recent version responds:

> Google is sensitive to the privacy concerns of its users. The Internet allows individuals to explore and communicate with unprecedented ease, but it also allows websites to collect and distribute personal information with equal ease. We at Google know that many users are, understandably, concerned about such practices, and we wish to make clear our policy for collecting and using personal information. (Google Privacy & Terms, 1999: para. 1)

The policy states that Google may share information about users with advertisers, business partners, sponsors, and other third parties, but only in aggregate, and not as individual users. This is not clear cut, however, as the following excerpt demonstrates:

> From time to time, there may be situations where Google asks you for personal information. When we intend to use your personal information, we tell you up front. This way you can decide whether you want to give us the information or not. In case you change your mind or some personal information changes, we will

> endeavour to provide a way to correct, update or remove the personal data you give us. (Google Privacy & Terms, 1999: para. 4)

As the CNIL ruling implies, the lack of clarity as to how data is collected and used is a recurring challenge in the history of Google's privacy policies. Formulations such as "we will endeavour" are examples of the kinds of vagueness that has legal significance as a low level of commitment, compared to, for example, formulations such as "we are obliged to" or "we will endeavour by all means". This ambiguity as to what actually happens with user data also relates to the clause on Google and cookies: "Most browsers are initially set up to accept cookies. You can reset your browser to refuse all cookies or indicate when a cookie is being sent. However, note that some parts of Google may not function properly if you refuse cookies" (Google Privacy & Terms, 1999: para. 7). The generic language about what data is collected, and what is shared, is a recurrent element in Google's long list of privacy policies. Discursively, these do, of course, change over time and become more capacious, particularly the post-GDPR versions. Even though Google maintains that it has responded adequately in the latest update, the need to share data across the ever-growing ecosystem of Google services is fundamental to their business model and therefore remains stable throughout the different privacy policies. It is important to note that in 2012, Google made significant changes to its privacy policies, as one common policy was adopted across all of its services. This was a controversial move seen through the lens of EU regulators, as it allowed Google to use personal data from one service for another service. Google claimed this would provide users with greater transparency, but as our analysis on horizontal intertextuality demonstrates, different services and external partners are still widely integrated into the structure and function of Google's policies, making it anything but simple. This adds not only to the discursive complexity, but also to the structural complexity of, particularly, post-GDPR Google policies.

The most recent privacy policy is more detailed in accounting for the different services and platforms to which the policy applies. Still, the text is similar in its indeterminate formulations. The services the policy mentions are only a handful of the services and platforms Google actually operates. Phrases such as "include" and "platforms like" are indicative of uncertainties which leave users on unsure ground in terms of the scope of the data collection across different platforms, and thus the reach of the policy. The current policy also retains vague rhetoric when explaining why Google collects information on users:

> We collect information to provide better services to all our users – from figuring out basic stuff such as which language you speak, to more complex things like which ads you'll find most useful, the people who matter most to you online, or which YouTube videos you might like. The information Google collects, and how that information is used, depends on how you use our services and how you manage your privacy controls. (Google Privacy & Terms, 2021: 2 of 30)

The policy furthermore provides information on how to manage privacy controls. While providing such information, the policy at the same time admonishes that if users limit Google's access to this information, the services and platforms will not function properly. This is a very similar wording – from the first policy to the most recent one – and demonstrates discrepancies in the transition from policy to practice, as, for instance,

users' acceptance of cookies becomes a repetitive, habitual, automated user practice in order for services to actually work.

While our focus in the remainder of this article is the latest versions of Google's privacy policy, these excerpts demonstrate certain vertical intertextual stability through different versions. This is also true regarding Google's ToS, for instance, in terms of rights management. The latest version of Google's ToS (effective from 31 March 2020; Google Privacy & Terms, 2020) responds most directly to the GDPR. It is more detailed and explains key concepts more thoroughly than former versions. For example, in relation to content and licences, concepts such as "worldwide", "non-exclusive", and "royalty free" are specifically explained. The structure is different, but essentially, the content remains the same.

By demonstrating the textual affinities in this manner, we wish to account for the vertical relationships between the different versions of Google's privacy policy and ToS. What these reveal are convoluted patterns of generic formulations and scope, as well as iterations of the importance of protecting users' privacy rights. At the same time, the formulations determining who owns the data and how users can expect their data to be utilised by Google have been consistent for 20 years. This is aligned with Peslak and colleagues' (2020) longitudinal study of Google privacy policies, which indicates that the complexity of Google's policy has increased over time and that it has discursively evolved to a more personalised and friendlier document by using more positive and enjoyment-laden words. They also note that Google has further obfuscated their privacy policy by adding links to other pages.

Judging by this, CNIL's financial penalty to Google for inadequate information and lack of transparency about their data processing practices comes as no surprise. To the contrary, it testifies to the importance the GDPR places on the individual's ability to understand and consent to the processing of their personal data. Ironically, the process which paved the way for the GDPR, and the actual text which constitutes the regulation, is itself tremendously complicated. While complex processes and language are to be expected when deliberating legal documents, they still constitute intricacies difficult for citizens, users, and laypeople to grasp. When these interrelated convolutions between the GDPR and the Google policy language meet, citizens have no way of deciphering how they relate to each other, nor to their right to privacy and protection of personal information.

The vertical intertextual process leading up to the GDPR can be analysed using similar methods as for Google's privacy policy and ToS. In its vertical intertextuality, the GDPR reflects the former regulatory framework, the Data Protection Directive of 1995 (Hoofnagle et al., 2019). The directive established rules and conditions for a citizen's right to data protection, while supporting the free flow of personal data between EU member states. The socio-legal context of the GDPR is the rapid technological developments which have brought new challenges for the protection of personal data both in terms of the digital economy and social life, as argued by the Commission in the preamble to the GDPR (EUR-Lex, 2016: 2). So, while there are vertical intertextual relations with former regulatory frameworks dealing with personal data, there are also vertical intertextual relations within the EU concerning the making of a new regulatory framework. The Commission acts as the organ which proposes regulations, taking on the role of agenda-setter, whereas the other EU bodies react to the proposal submitted

by the Commission. However, it is not just the Commission's voice which characterises the proposal, as it follows detailed prescriptions in terms of public consultations, impact assessments, and consultations with stakeholders, companies, industry actors, civil society actors, powerful lobbyism, and so on. This part of the process alone lasted more than two years, involving several preparatory papers and complex deliberations between stakeholders representing different interests (Laurer & Seidl, 2020).

Once a proposal is put forward, the EU's formal path to legislation is further activated and, in this case, involved a number of opinions, for example, from the European Data Protection Supervisor and the Economic and Social Committee. In all, there were over 20 deliberations within the Council of the European Union or its preparatory bodies, two readings by the European Parliament, and a Commission position on the European Parliament's amendments. It would exceed the scope of this article to go further into these processes,[2] but they do indicate the extent of vertical intertextuality centred around the GDPR. While these documents can be accessed and scrutinised, they inevitably reflect Cohen's (2019) concern regarding increased specialisation and what we conceive to be the incremental complexity of regulatory frameworks.

Horizontal intertextuality

Wider contexts of complexity

In terms of horizontal internal intertextuality at Google, the different versions of the privacy policy and ToS are far from the only documents that explain the relations between Google and its users. As the formulation in the ToS indicates, there are many "additional services" which further contribute to the structural complexity of the policies. Interestingly, even though the GDPR's aim was, amongst other things, to make information about personal data processing more explicit and understandable, the post-GDPR version of Google's privacy policy is actually more extensive than prior versions. Arguably, this is because Google has been forced to be more specific about the type of information collected and the manner in which data is used.

The increased complexity is also manifested in the modality of Google's latest privacy policy, which is not only based on text, but also on YouTube videos and numerous illustrations providing additional ways to understand the policy. This is in line with GDPR's requirement of transparency that stresses, for example, visual tools as a means to inform users. In addition to illustrations and videos, the policy is accompanied by over 120 hyperlinks, many of which provide a further explanation of terms, such as personal information, sensitive personal information, unique identifiers, payment information, IP address, cookies, pixel tags, application data catches, server logs, personalised ads, sensitive categories, algorithms, third parties, and affiliates. While this is a common way to present information on the web, it adds to the complexity of the policy – both in terms of structure and scope of provided information – as, for instance, discussed by Peslak and colleagues (2020). The most recent ToS follows the same hyperlinked structure, which again indicates that the GDPR has prompted Google to add more explanations to their policies, which actually increases the textual and structural complexity.

The GDPR itself is an extensive legal document which, as previously stated, attempts to react to emerging gaps between technological advances and regulation. It therefore (re)defines key terms such as personal data, processing, profiling, pseudonymisation,

controller, processor, recipient, consent, third party, biometric data, genetic data, and data concerning health. In line with its predecessor, the Data Protection Directive, the GDPR provides the data subject with several rights (Vanberg, 2020). These include the right to access, the right to rectification, the right to object, the right to data portability, the right to erasure ("right to be forgotten"), and rights regarding automated individual decision-making. The latter is stipulated in Article 22 of the GDPR: "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her" (EUR-Lex, 2016: 46). In some cases, the rights are conditioned by specific situations, but this enumeration is illustrative of how the GDPR defines key terms and relates to citizens as data subjects. As case law under GDPR develops – as well as the monitoring and guidance of the European Data Protection Board and member states' data protection authorities – the scope of this and other rights will be substantiated, for example, when someone is "significantly" affected by automated decision-making. Another example is the provision on data protection by design and default stipulated in Article 25 of the GDPR: "The controller shall implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed" (EUR-Lex, 2016: 48). Again, as case law develops, the scope and delimits of these norms will be clarified.

There is no doubt that the GDPR is tailored towards data protection, including different forms of rights given to data subjects; however, there are certainly other prominent discourses inherent in the framework aimed at protecting the EU's digital single market and furthering its flow of data. One of the recitals of the GDPR formulates this quite clearly: "Those developments require a strong and more coherent data protection framework in the Union, backed by strong enforcement, given the importance of creating the trust that will allow the digital economy to develop across the internal market" (EUR-Lex, 2016: 2). As such, the GDPR is yet another piece in the EU's puzzle, as put forward in the i2010 and i2020 strategies and the Digital Agenda for Europe. In intertextual terms, the GDPR mirrors the EU's overall strategy within the information society, which is attentive to economic activity and innovation, economic growth, and creation of jobs, among other things. In terms of complexity, it would therefore be too simplistic to see the GDPR as primarily a reaction to technological and infrastructural changes caused by global actors such as Amazon, Microsoft, Apple, Facebook, and Google – that is, to protect the personal information of EU citizens from the "Big Five". One must also consider the GDPR as a regulation to further promote the robustness of the EU's digital single market.

Furthermore, the GDPR is not alone in the EU's regulatory jungle, as topics and concerns related to data subjects, privacy, and personal information arise in other legal sources as well (Vanberg, 2021). The GDPR, in fact, has intertextual horizontal relations to the ePrivacy directive, directives on collective rights management, the eCommerce directive, the directive on consumer rights, the telecoms framework, and the framework within audiovisual media services, just to name a few. The recently proposed EU's Digital Services Act and Digital Markets Act are also examples of regulation specifically designed to respond to such complex horizontal intertextuality. But as these are currently being negotiated, their wider effect and implications remain to be seen. The

internal horizontal intertextuality within the EU legal regime is therefore rather comprehensive, and its complexity is increased when some of the key definitions in the GDPR are compared to similar notions in Google's privacy policy.

Personal information and third party

In order to demonstrate the intricate relationship between Google's privacy policy and the GDPR, we focus on the notions of personal information and third party. In Google's privacy policy, users must follow the personal information hyperlink to reach the following definition:

> This is information that you provide to us which personally identifies you, such as your name, email address, or billing information, or other data that can be reasonably linked to such information by Google, such as information we associate with your Google Account. (Google Privacy & Terms, 2021: 20 of 30).

In the GDPR's Article 4, personal data is defined as follows:

> Any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. (EUR-Lex, 2016: 33)

Compared with Google's definition of personal information, the EU version is much broader in its account of what may count as personal data. As such, Google and the EU are at odds over a fundamental definition; in fact, they do not even provide comparable characteristics of said information.

CNIL specifically criticises Google's vague and unclear language regarding third parties and external partners, and when these terms are scrutinised, it is obvious why. Consider this sentence from the privacy policy: "We process your information for our legitimate interests and those of third parties while applying appropriate safeguards that protect your privacy [underlining indicates links in the original]" (Google Privacy & Terms, 2021: 16 of 30). When users click on the link for third parties, they are met with the following explanation:

> For example, we process your information to report usage statistics to rights holders about how their content was used in our services. We may also process your information if people search for your name and we display search results for sites containing publicly available information about you. (Google Privacy & Terms, 2021: 29 of 30)

Moreover, it is unclear how Google's notion of third parties relates to its framing of the word "partner", which is fuzzy to say the least:

> We may share non-personally identifiable information publicly and with our partners – such as publishers, advertisers, developers, or rights holders. For example, we share information publicly to show trends about the general use of our services. We also allow specific partners to collect information from your browser

> or device for advertising and measurement purposes using their own cookies or similar technologies [underlining indicates links in original]. (Google Privacy & Terms, 2021: 12 of 30)

When users click on the link that directs them to "specific partners", they are met with this description, and yet another link:

> For example, we allow YouTube creators and advertisers to work with measurement companies to learn about the audience of their YouTube videos or ads, using cookies or similar technologies. Another example is merchants on our shopping pages, who use cookies to understand how many different people see their product listings. Learn more about these partners and how they use your information [underlining indicates links in original]. (Google Privacy & Terms, 2021: 29 of 30)

In this, several matters remain unclear: What precisely constitutes a partner? what is a merchant? And what is a measurement company? When users activate the "learn more" link, they are redirected to a page which explains this in the following terms:

> Google works with businesses and organizations in a variety of ways. We refer to these businesses and organizations as "partners". For example, over 2 million non-Google websites and apps partner with Google to show ads. Millions of developer partners publish their apps on Google Play. Other partners help Google with securing our services [underlining indicates links in original]. (Google Privacy & Terms, n.d.: para. 1)

Google's partners are measured in millions of developers and apps partners, but the extent of Google's collaborative network goes further still. Google also works with data processors, which are different from partners:

> Note that we also work with trusted businesses as "data processors" rather than partners, meaning that they process information on our behalf, to support our services, based on our instructions and in compliance with our Privacy Policy and other appropriate confidentiality and security measures. (Google Privacy & Terms, n.d.: para. 2)

In addition to the above, Google refers to "specific partners", who collect information for advertising and ad measurement purposes, using their own cookies or similar technologies. These specific partners are Nielsen, comScore, Integral Ad Science, DoubleVerify, Oracle Data Cloud, Kantar, and RN SSI Group. Google provides links to these partners, and, when following the first link, users are redirected to Nielsen's web page and its digital measurement privacy statement, which starts with the following line: "Nielsen uses its proprietary software and products to provide digital measurement services, which measure and analyze how consumers engage with media across online, mobile and emerging technologies, and enable Nielsen to offer research insights into consumer behavior" (Nielsen, n.d.: para. 1). Nielsen's privacy information branches out into numerous documents, such as the Website Privacy Statement, Digital Measurement Privacy Statement, Nielsen Marketing Cloud Privacy Statement, Careers Privacy Statement, Marketing Privacy Statement, and Market Segmentation Privacy Statement. This demonstrates the intricate network a user will encounter when seeking to uncover how Google interacts with partners and data processors.

To add to the density of the privacy policy, Google lists specific services subject to additional information: Chrome and the Chrome operating system, Payments, Fiber, Google Fi, G Suite for Education, YouTube Kids, Read Along, Google Accounts Managed With Family Link For Children Under 13, and Voice and the Audio Collection from Children's Features on the Google Assistant. Google refers to these as related privacy practices for specific Google services, and some are quite detailed. The horizontal intertextuality these are enmeshed with in terms of Google's privacy policy demonstrates the level of complexity users must navigate in order to understand their privacy rights. Moreover, it gives a sense of the underlying business model. We refer to this as internal horizontal complexity. Consider, in turn, what we refer to as external horizontal complexity, the intertextual chains of which these documents are part, with links to partners such as Nielsen, and the complexity soars.

We have already accounted for Google's way of defining third parties. Comparing this to GDPR's wording further carves out the external horizontal intertextuality, as "third party" here refers to "a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data" (EUR-Lex, 2016: 34). This is in stark contrast to Google's much looser formulation, particularly when Google's definition of partners and data processors is included. Google's response to the GDPR is structurally more complex than earlier versions of its privacy policy. Textually, however, explanations of key terms such as "personal information" and "third party" are over-simplified, making it more difficult for users to understand exactly how and when their personal data is processed, and by whom. The GDPR, on the other hand, is much more precise in its definitions, yet textually it does not present comprehensible language accessible to laypeople.

Concluding remarks

In this article, we have examined the co-constitutive complexity between Google's privacy policies and ToS and the GDPR. Motivated by CNIL's financial penalty against Google, we were curious to examine how Google's privacy policies address their users' right to transparency stipulated by the GDPR, and how they reconcile the GDPR's demand for transparency with the data collection imperative of their business model. Our analysis demonstrates that the vertical and horizontal, and external and internal intertextuality of Google's privacy policies and ToS undermine users' ability to understand and foresee the types of data processing to which they consent. This complexity has increased over time, and the post-GDPR versions are particularly dense, despite the GDPR's efforts to the contrary.

This article also demonstrates that the GDPR is itself a complex document with its own vertical and horizontal, and external and internal intertextuality, making it challenging to understand for general users and laypeople. This complexity hampers users' understanding of how their rights are secured, not only as data subjects, but as citizens. The EU objectives of clear and understandable policies and transparent data processing practices, therefore, seem far away – if they are even achievable. While this may be an unwanted, but unsurprising, side-effect of a comprehensive piece of EU regulation, it counters the very purpose of the regulation. The CNIL decision points to the conflict between specific

Google practices and the GDPR; however, it remains an open question to which extent Google is capable of complying with the GDPR while retaining its data-driven business model and the complex and interrelated format of its privacy policy. From a socioeconomic perspective, Google's power resides within its dominant position in the supply chain of information, which is central to the very functioning of platforms. Google's privacy policy and ToS represent only a fraction of what constitutes Google's operations as a whole, and even though Google has been required to respond to the GDPR, our analysis demonstrates that whereas the vertical intertextuality has remained relatively stable, the horizontal intertextuality has increased in complexity. This indicates the limitations of the GDPR as a normative corrective to the kind of complexity Google enacts. While the GDPR has required Google to reconsider its terms and policies, to make them clearer and more understandable to users, the discursive complexity of Google's privacy policies have in fact increased. In the process of accommodating GDPR's demand of increased transparency, Google has thus made the policies even more obscure to general users of their services. This is what we refer to as Google's deliberative data politics, thus stressing contractual complexities as an integral part of the platform's business model, and as such, serving a specific (and strong) socioeconomic purpose for Google.

Our analysis illustrates the enormous tasks that lie ahead in providing EU citizens with transparency and privacy rights in the current regime of data politics and data capitalism. Arguably, the GDPR – despite good intentions – cannot truly address the fundamental power imbalance between giant platforms such as Google and individual users. As previously mentioned, new regulatory proposals such as the EU's Digital Services Act and the Digital Markets Act seem better suited for a fight, as they address the challenges that complex interrelated services such as Google pose to competition and a fair market (Digital Markets Act), as well as to users' fundamental rights (Digital Services Act). The latter focuses particularly on the duties of platforms in relation to illegal content, but also demands increased transparency and human rights risk assessment according to company size and potential impact on user rights. As such, the policy discourse around Digital Services Act and Digital Markets Act takes the infrastructural role of giants like Google as its starting point, yet the proposals have only now started their long journey through the EU's negotiation process, and it is too early to anticipate their effect on users' ability to understand and claim privacy rights.

Funding

The manuscript is written by members of the collective research project "'Don't take it personal': Privacy and information in the algorithmic age", funded by the Independent Research Fund Denmark (grant nr. 8018-00041B).

Notes

1. The notions of personal data and personal information are often used interchangeably, although there is a preference for personal *information* in the American privacy literature (Cohen, 2019; Nissenbaum, 2010; Schwartz & Solove, 2014), whereas European data protection is anchored in the notion of personal *data*. The GDPR uses the notion of personal data, but defines it (in Art. 4) as any *information* relating to an identified or identifiable natural person. In this article, we use the notion of personal data when we refer to the GDPR and its legal provisions, but otherwise rely on personal information, which is also the concept used in Google's privacy policy.
2. See Valtysson (2018) for a study that demonstrates the discursive discrepancies that unfold in the institutional negotiation of specific EU programmes.

References

Amoore, L. (2020). *Cloud ethics: Algorithms and the attributes of ourselves and others*. London: Duke University Press.

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, *20*(3), 973–989. https://doi.org/10.1177/1461444816676645

Ausloos, J., & Dewitte, P. (2018). Shattering one-way mirros: Data subject access rights in practice. *International Data Privacy Law*, *8*(1), 4–28. https://doi.org/10.1093/idpl/ipy001

Bratton, B. H. (2015). *The stack: On software and sovereignty*. London: MIT Press.

CNIL. (2019, January 21). The CNIL's restricted committee imposes a financial penalty of 50 Million euros against GOOGLE LLC. https://www.cnil.fr/en/cnils-restricted-committee-imposes-financial-penalty-50-million-euros-against-google-llc

Cohen, J. E. (2019). *Between truth and power: The legal constructions of informational capitalism.* Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780190246693.001.0001

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford, California: Stanford University Press.

Easterling, K. (2014). *Extrastatecraft: The power of infrastructure space*. London: Verso Books.

EUR-Lex. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Evans, D. S., & Schmalensee, R. (2011). The industrial organization of markets with two-sided platforms. In D. S. Evans (Ed.), *Platform economics: Essays on multi-sided businesses* (pp. 2–29). Competition Policy International.

Fairclough, N. (1992). *Discourse and social change*. Cambridge: Polity Press.

Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. New York: Routledge.

Finck, M., & Pallas, F. (2019). *The who must not be identified: Distinguishing personal from non-personal data under the GDPR* [Max Planck Institute for Innovation and Competition Research Paper no 19-14].

Fuchs, C. (2019). Karl Marx in the age of Big Data capitalism. In D. Chander, & C. Fuchs (Eds.), *Digital objects, digital subjects: Interdisciplinary perspectives on capitalism, labour and politics in the age of Big Data* (pp. 53–71). London: University of Westminster Press.

Gillespie, T. (2010). The politics of 'platforms'. *New Media & Society*, *12*(3), 347–364. https://doi.org/10.1177/1461444809342738

Google Privacy & Terms. (1999). *Google and privacy*. Acessed April 27, 2020, from https://policies.google.com/privacy/archive/19990609?hl=en&gl=dk

Google Privacy & Terms. (2020, 31 March). *Google terms of service*. https://www.gstatic.com/policies/terms/pdf/20200331/ba461e2f/google_terms_of_service_en_eu.pdf

Google Privacy & Terms. (2021, 4 February). *Google privacy* policy. https://www.gstatic.com/policies/privacy/pdf/20210204/3jla0xz1/google_privacy_policy_en-GB_eu.pdf

Google Privacy & Terms. (n.d.). *Who are Google's partners?* Accessed April 27, 2020, from https://policies.google.com/privacy/google-partners?hl=en&gl=dk

Hoofnagle, C. J., van der Sloot, B., & Borgesius, F. Z. (2019). The European Union general data protection regulation: What it is and what it means. *Information & Communication Technology Law*, *28*(1), 65–98. https://doi.org/10.1080/13600834.2019.1573501

Houser, K. A., & Voss, G. (2018). GDPR: The end of Google and Facebook or a new paradigm in data privacy? *Richmond Journal of Law & Technology*, *25*(1).

Jørgensen, R. F. (2017). Framing human rights: Exploring storytelling within Internet companies. *Information, Communication & Society*, *21*(3), 340–355. https://doi.org/10.1080/1369118X.2017.1289233

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, *20*(1), 14–29. https://doi.org/10.1080/1369118X.2016.1154087

Klass, G. (2019). Empiricism and privacy policies in the restatement of consumer contract law. *Yale Journal on Regulation*, *36*(1), 45–116. https://heinonline.org/HOL/P?h=hein.journals/yjor36&i=51

Laurer, M., & Seidl, T. (2020, June 25). Regulating the European data-driven economy: A case study on the general data protection regulation. *Policy & Internet*. Early view. https://doi.org/10.1002/poi3.246

Nielsen. (n.d.). *Digital measurement privacy statement* (last updated August 2020).

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life.* Stanford: Stanford Law Books.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. London: Harvard University Press.

Peslak, A., Kovalchick, L., & Conforti, M. (2020). A longitudinal study of Google privacy policy. *Journal of Information Systems Applied Research*, *13*(2), 54–64.

Plantin, J. C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, *20*(1), 293–310. https://doi.org/10.1177/1461444816661553

Ruppert, E., Isin, E., & Bigo, D. (2017, July 3). Data politics. *Big Data & Society*, 1–7. https://doi.org/10.1177/2053951717717749

Schwartz, P., & Solove, D. J. (2014). Reconciling personal information in the United States and European Union. *California Law Review*, *102*(4), 877–916. http://www.jstor.org/stable/23784355

Solove, D. J., & Hartzog, W. (2014). The FTC and the new common law on privacy. *Columbia Law Review*, *114*(3), 583–676.

Srnicek, N. (2016). *Platform capitalism*. Cambridge: Polity Press.

Strathern, M. (2000). The tyranny of transparency. *British Educational Journal*, *20*(3), 309–321. https://doi.org/10.1080/713651562

Valtysson, B. (2018). Camouflaged culture: The 'discursive journey' of the EU's cultural programmes. *Croatian International Relations Review*, *24*(82), 14–37. https://doi.org/10.2478/cirr-2018-0008

Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford: Oxford University Press.

Van Dijck, J., Poell, T., de Waal, M. (2018). *The platform society: Public values in a connective world.* Oxford: Oxford University Press.

Vanberg, A. D. (2021). Informational privacy post GDPR – end of the road or the start of a long journey? *The International Journal of Human Rights*, *25*(1), 52–78. https://doi.org/10.1080/13642987.2020.1789109

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*, 75–89. https://doi.org/10.1057/jit.2015.5

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power.* London: Profile Books.