The 2016 AAAI Fall Symposium Series:
Privacy and Language Technologies
Technical Report FS-16-04

Privacy Policies and Their Lack of Clear Disclosure Regarding the Life Cycle of User Information

Priya Kumar
University of Maryland College of Information Studies
pkumar12@umd.edu

Abstract

Companies, particularly those in the information and communications technology sector, collect, aggregate, and store immense amounts of information about billions of people around the world. Privacy policies represent the primary means through which companies articulate to the public how they manage this user information. Extensive research has documented the problems with such policies, including that they are difficult to understand. This paper presents an analysis of 23 policies from 16 of the world's largest internet and telecommunications companies and shows the specific ways that vague or unclear language hinders comprehension of company practice. It argues that the lack of clarity in such policies presents a significant barrier toward empowering people to make informed choices about which products or services to use. The incoherent language in privacy policies can also hinder the widespread adoption of machine learning or other techniques to analyze such policies. Clearer disclosure from companies about how they use, share, and retain all types of information they collect will shed light on what the life cycle of user information looks like.

Introduction

Companies in the information and communications technology (ICT) sector are at the vanguard of surveillance capitalism, a political, economic, and social principle that, "aims to predict and modify human behavior as a means to produce revenue and market control" (Zuboff 2015). Their ability to do so rests on the accumulation of immense amounts of information about people, including their communication, whereabouts, online browsing habits, purchases, and more. As the minutiae of daily life become increasingly digitized, companies that collect, aggregate, and store this information gain significant influence over individuals (MacKinnon 2012).

Data protection and privacy frameworks, including laws and regulations, are meant to check the unrestrained collection and use of information about people. At minimum, such frameworks should require notice, among other principles (Greenleaf 2014). In other words, people should be aware of what an entity does with their information.

The primary mechanism for providing this notice has been through a so-called "privacy policy," though these documents contain well-known problems (Schwartz and Solove 2009). They tend to be long and unintelligible, written to satisfy regulators rather than provide clarity to people about what happens to their information (Cate 2010). Indeed, 52 percent of Americans incorrectly believe that simply having a privacy policy means a company keeps user information confidential (Smith 2014). Campaigns such as "That's Not Privacy"[1] encourage organizations to call their policy a "data use policy" rather than a "privacy policy."

The notion that people will read a privacy policy before deciding whether to transact with an organization remains unrealistic; one study suggests that doing so would take each American internet user about 200 hours per year and cost more than $3,500 per person in lost time (McDonald and Cranor 2008). Several efforts have sought to facilitate analysis of these policies. These include efforts to develop machine-readable versions of policies (Cranor 2012) as well as techniques to automatically parse text of existing policies (Breaux, Hibshi, and Rao 2014) and compare the computer judgments with human interpretations (Zimmeck and Bellovin 2014; Liu et al, 2014;) or with human expectations (Rao et al, 2016). The lack of standardization across existing policies and the lack of clarity within policies about how companies handle user information complicate such efforts. Based on an analysis of privacy policies from 16 of the world's largest ICT companies, this paper highlights specific ways that vague or unclear language hinders understanding of company practices. It argues that the lack of clarity in such policies presents a significant barrier toward empowering people to make informed choices about

[1] https://www.thatsnotprivacy.com/

which products or services to use. This lack of clarity can also hinder the widespread adoption of machine learning or other techniques to analyze such policies.

The Development of the Ranking Digital Rights Corporate Accountability Index

This analysis of company privacy policies was part of the first comprehensive, globally applicable framework to measure the extent to which ICT companies respect their users' rights to freedom of expression and privacy.

The past two decades have seen growing awareness of the need to hold ICT companies accountable to human rights standards, given their critical role in connecting individuals and organizations. Existing efforts, such as Freedom House's Freedom on the Net Index (Freedom House 2015), evaluate government influence on internet freedom. The Ranking Digital Rights[2] (RDR) Corporate Accountability Index complements such work by highlighting the role that companies play in respecting users' rights to freedom of expression and privacy. This in turn provides responsible investment firms and civil society organizations with knowledge they can use to push companies to improve (Maréchal 2015).

The first edition of RDR's index was published in November 2015 after three years of consultation and research. It evaluated 16 of the world's largest internet and telecommunications companies on a set of criteria grounded in international human rights norms and standards (Ranking Digital Rights 2015a.). Work on the project began in late 2012, when the team convened with individuals from civil society, academia, the investment community, and companies to discuss what the index should examine. Case study research focused on five countries and three multinational corporations informed the creation of draft indicators. The team sought public feedback on three versions of the draft indicators, and it tested the third iteration in a pilot study on 12 companies (Ranking Digital Rights 2015b.). The final indicators for the 2015 index were published in June 2015 (Ranking Digital Rights 2015c.).

The index focused on publicly traded companies under the assumption that they would be more receptive to public advocacy than privately held or fully state-owned companies. The companies were selected based on a variety of factors, including:

- Headquarters location
- Countries of operation
- Size of user/subscriber base
- Market capitalization
- Market share

[2] Until August 2016, the author was a research analyst with Ranking Digital Rights.

While the team sought a balance of companies that would provide an informative, global snapshot of the ICT industry, this sample is not representative in the statistical sense. RDR did not aim to produce generalizable results, but rather to conduct an in-depth analysis on a cross-section of the ICT industry to uncover trends and highlight specific areas for company improvement.

For each internet company, the team selected two to three services to examine, based in part on the size of their user base and comparability across companies. For telecommunications companies, the team evaluated the mobile (and fixed broadband, if offered) service of the operating company in the home market. Table 1 provides information on the companies and services evaluated in the index.

Data collection, review, and analysis for the index were conducted from June-November 2015. The research was based on publicly available information; researchers examined company websites for relevant disclosure in reports, policies, and other web pages. The full research process included seven steps of data collection and review, and any discrepanies in responses to the indicator questions were discussed and resolved among the team.

The index's privacy analysis reviewed disclosure related to corporate handling of user information, responses to third-party requests for user information, and security practices. This paper focuses on the findings and recommendations related to corporate handling of user information. On this topic, the index sought disclosure about what user information companies collect, how they collect it and why, with whom they share user information and under what circumstances, how long they retain user information, the extent to which users can control such collection and sharing, and whether users can access the information companies hold on them.

Defining User Information

The terms "personal information," "personal data" and "personally identifiable information" (PII) are commonplace in privacy policies because they demarcate what is protected by privacy regulations. Under the law, the existence of a privacy harm turns on whether the information in question is personal or personally identifiable; "[I]n the absence of PII, there is no privacy harm" (Schwartz and Solove 2011 p. 1816). Problematically, U.S. law lacks a clear definition of PII, and information that often is not considered PII, such as an IP address, can be linked to an identifiable person (Schwartz and Solove 2011). The European Union's General Data Protection Regulation takes a broader approach to personal data, defining it as "any information relating to an identified or identifiable natural person"; this can include "an identifier such as a name, an identification number, location data, an online identifier

or...one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" that can be used to directly or indirectly link a piece of information to a person (European Union 2016 p. 33).

Parent Company / Operating Company (where applicable)	Headquarters Location	Services Evaluated
Internet Companies		
Facebook	United States	Facebook Instagram WhatsApp
Google	United States	Google Search Gmail YouTube
Kakao	South Korea	Daum Search Daum Mail KakaoTalk
Mail.ru	Russia	VKontakte Mail.ru Mail.ru Agent
Microsoft	United States	Bing Outlook Skype
Tencent	China	Qzone QQ WeChat
Twitter	United States	Twitter Vine
Yahoo	United States	Yahoo Mail Flickr Tumblr
Telecommunications Companies		
América Móvil / Telcel	Mexico	Mobile
AT&T	United States	Mobile Fixed broadband
Axiata / Celcom	Malaysia	Mobile
Bharti Airtel	India	Mobile Fixed broadband
Etisalat	United Arab Emirates	Mobile Fixed broadband
MTN	South Africa	Mobile
Orange	France	Mobile Fixed broadband
Vodafone	United Kingdom	Mobile

Table 1: Companies and Services Evaluated in the 2015 Corporate Accountability Index

The determination that a particular piece of information is PII holds significant implications for companies, because it triggers the need to comply with privacy regulations that govern the flow of that information. Yet company practices can still threaten people's privacy regardless of whether the law recognizes those actions as privacy harms. Indeed, "the business model supporting much of the Internet industry is predicated on users relinquishing individual privacy in exchange for free information and software" (DeNardis 2014 p. 231).

Log data, location information, and telephone metadata are not typically considered PII, yet research continues to document how this information can be used to identify people and sensitive information about those people. Schwartz and Solove (2011) describe how IP address records from Wikipedia were used to track down a user who wrote false statements on the Wikipedia page of former Justice Department official John Seigenthaler. Hoang, Asano, and Yoshikawa (2016) demonstrate how an adversary could use location information from popular LGBT-focused dating apps to identify users, despite the fact that some apps have implemented measures to help users obscure this information. Liccardi, Abdul-Rahman, and Chen (2016) show that a Twitter user's home and workplace can be identified through as few as eight geo-location tagged tweets. Mayer, Mutchler, and Mitchell (2016) analyzed smartphone logs and explain how easily such data can be used to infer sensitive personal information, including medical conditions, drug use, and gun ownership.

If the public is to clearly understand the privacy implications of using ICTs, they must know the full scope of what information companies collect, use, share, and retain. Knowing what companies do with personal information is insufficient. Consequently, Ranking Digital Rights sought company disclosure related to "user information," a term that includes information that people actively provide (e.g., name, text of messages), as well as information that companies automatically collect when people use a service (e.g., IP address, GPS coordinates). RDR defines user information as "any data [that] is connected to an identifiable person, or may be connected to such a person by combining datasets or utilizing data-mining techniques" and may or may not be connected to a user account. Examples include "personal correspondence, user-generated content, account preferences and settings, log and access data, data about a user's activities or preferences collected from third parties either through behavioral tracking or purchasing of data, and all forms of metadata." (Ranking Digital Rights 2015c).

The term "user information" has a broader definition than many of the terms that companies use in their privacy policies. This makes it extremely difficult to obtain a complete picture of how companies handle user information.

Company Definitions of Personal Information

Of the 35 services evaluated in the index, privacy policies were available for all but four: Orange mobile and fixed

broadband and Mail.ru's Mail and Agent (messaging) services. Orange and Mail.ru provided limited disclosure related to user information practices in their terms of service, and those policies were evaluated in this research. Several companies published privacy policies that covered many or all of their services, and in total, the analysis included 23 policies. For this paper, the author used versions of the policies that were available on May 4, 2016.

Nine policies defined "personal information" or a similar term, such as PII or personal data. Seven companies provided definitions that were focused on information about an identified person. While definitions from WhatsApp and Etisalat were tautological or overly vague, others contained examples of types of information that fell into the PII category. Notably, definitions from the remaining two companies, Bharti Airtel and Google, included identifiable information, or information that, when linked with existing information, could be used to identify a specific person. Table 2 includes the nine company definitions. Axiata/Celcom and Etisalat provided policies in English as well as Malay and Arabic, respectively; the English versions of their definitions are in the table. The remaining 14 policies did not define personal information, despite the fact that most policies use the term.

These definitions, and subsequent disclosure about what information companies collect, show that companies from around the world hew fairly closely to the legal and regulatory conceptions of personal information outlined by Schwartz and Solove (2011). These companies do not appear to consider information such as log data or cookie data as personal information. For example, the policies of Bharti Airtel, Etisalat, and Yahoo contain paragraphs or lists of the personal information they collect, but paragraphs on information collected through cookies or similar technologies are separate. As explained in the previous section, the focus of this research was not to compare company practices related to personal information, but to user information. However, as the subsequent analysis highlights, use of the term "personal information" obscures, rather than clarifies, how companies handle the information they have on their users.

Collection of User Information

Positively, all policies provided at least some disclosure about what user information companies collect. Eighteen policies listed the types of user information companies collect; these explanations generally included examples of pieces of information that fall under that category. For example, Instagram's policy stated the company collects the following types of information: "information you provide us directly," "finding your friends on Instagram" (which explains what information it collects if users enable this feature), analytics information, information from cookies and similar technologies, log file information, information related to device identifiers, and metadata. For each of these types, the policy included a bulleted list that describes what information that type includes and, in some cases, how the company obtains such information.

Company	Definition
Definitions Focused on "Identified" Information	
AT&T	"**Personal Information:** Information that directly identifies or reasonably can be used to figure out the identity of a customer or user, such as your name, address, phone number and e-mail address. Personal Information does not include published listing information."
Axiata/Celcom	"'Personal Data' means any personal information relating to CELCOM's customer that the customer has provided to CELCOM or made available to CELCOM due to his/her contract with CELCOM, e.g. name, Identity Card / Passport No., address, information about his/her transactions with CELCOM such as contact number, account number, account balances, payment history, and account activity."
Etisalat	"Personal information is any information that is specific to a particular customer."
Facebook	"[P]ersonally identifiable information is information like name or email address that can by itself be used to contact you or identifies who you are"
MTN	"[P]ersonal information is information which identifies you as an individual, such as your first and last name, your ID number, your phone number, credit vetting and payment information, your preferences and opinions."
WhatsApp	"Personally identifying information": "information that can be used to identify you". The policy also uses the term "personally identifiable information."
Yahoo	"Personal information is information about you that is personally identifiable like your name, address, email address, or phone number, and that is not otherwise publicly available."
Definitions Focused on "Identifiable" Information	
Bharti Airtel	"'**Personal information**' is any information that can be used by itself to uniquely identify, contact, or locate a person, or can be used with information available from other sources to uniquely identify an individual."
Google	Personal information: "This is information which you provide to us which personally identifies you, such as your name, email address or billing information, or other data which can be reasonably linked to such information by Google, such as information we associate with your Google account."

Table 2: Company Definitions of Personal Information

However, company disclosure varied in clarity. For example, MTN listed several types of information that fall under its definition of personal information, and it refer-

enced personal information throughout its policy. But it did not explicitly state what personal information it collects. This is particularly problematic because its list included:

- "Information relating to the race, gender, sex, pregnancy, marital status, national, ethnic or social origin, colour, sexual orientation, age, physical or mental health, well-being, disability, religion, conscience, belief, culture, language and birth of that person,"
- "Information relating to the education or the medial [sic], financial, criminal or employment history of the person;"

"The personal opinions, views of [sic] preferences of the person;"

- "Correspondence sent by the person that is implicitly or explicitly of a private or confidential nature or further correspondence that would reveal the contents of the original correspondence;"
- "[T]he views of [sic] opinions of another individual about the person."

Thus, the policy appears to leave open the option for the telecommunications company to collect a wide swath of extremely sensitive information about its users.

In addition, policies included language that clearly indicated that the disclosure was not comprehensive. Vodafone's policy prefaced the list of user information it collects by saying the list "includes (but isn't limited to) the following." Etisalat's list included "any other information we need to provide a particular service." Companies also included vague descriptions of information. Yahoo's policy stated that the company "may combine information about you that we have with information we obtain from business partners or other companies," but provided no further explanation about what this information may include.

The remaining five policies mentioned specific pieces of information that companies collect, but the disclosure was incomplete or scattered throughout the policy in a way that made it difficult to clearly understand what the company collected. The user agreement for Mail.ru stated that the company collects usernames and passwords, and referenced "user data" but did not specify what else it collects. Axiata/Celcom's policy included examples of information in its definitions of "personal data" and "sensitive personal data," and then referenced these terms throughout the policy. The policy did not state whether these examples constitute a list of all the user information the company collects. Several sections of the policy also referenced other types of information the company collects, including IP addresses, cookie data, and web browsing activity. It organized its policy into sections based on the Fair Information Practices, rather than sections that relate to the life cycle of user information (e.g., collection, sharing, retention, etc.), which makes it difficult to understand how the company handles a particular type of user information.

Sharing of User Information

All policies included at least some disclosure about what user information companies share with third parties, though this disclosure provided even less clarity than disclosure related to collection of information. Disclosure fell into three general categories: it provided some specific examples of types of information that the company shares, it described the sharing of personal information but not other types of information, or it suggested that all user information could be shared. Companies used general terms to describe the types of third parties with which they shared data; Yahoo was the only company to provide a list of the names of third parties with which it shares information.

Instagram's policy stated that it shares "User Content and your information (including but not limited to, information from cookies, log files, device identifiers, location data, and usage data)" with other businesses within the Facebook, Inc. group and with service providers. It also stated that it shares cookie data with third-party advertising partners.

Twitter's privacy policy included fairly detailed information about the types of user information it collects. However, the company summarized its sharing of information by stating, "We do not disclose your private personal information except in the limited circumstances described here." The term "private personal information" did not appear in any other section of the policy. Twitter did not define the term but provided examples at the end of the section, stating that this includes a user's name or contact information. The policy included explanations of six types of sharing. Its explanation of sharing for commerce transactions identified the types of data that can be shared, and its explanation of sharing with user consent and sharing of non-private or non-personal information provided examples that indicate what information can be shared in those circumstances. However, its explanation of sharing with service providers stated that "private personal information" may be shared. Its explanation of sharing related to law and harm as well as business transfers and affiliates stated that "your information" may be shared, suggesting that this includes all user information the company holds.

Ten policies focused their disclosure on how the companies shared personal information, which is a subset of user information. The "Information we collect" section of Google's privacy policy stated that it collects seven types of information: personal information, device information, log information, location information, unique application numbers, information from local storage, and information from cookies and similar technologies. However, the "Information we share" section of the policy only references "personal information," sweeping all other types of information into the statement that "We may share non-

personally identifiable information publicly and with our partners – like publishers, advertisers or connected sites." Its definition of "non-personally identifiable information" as information "that is recorded about users so that it no longer reflects or references an individually identifiable user" did not clearly state whether this category includes the other six types of information Google said it collects.

AT&T's disclosure similarly failed to provide clarity on what user information it shares. In a section of the policy focused on information collection, the company stated that it collects several types of user information and provided examples of each: account information (includes contact information and billing information), technical and usage information (includes equipment information, network performance and usage information, web browsing and mobile application information), location information, and information about U-verse services. However, the policy's section on sharing of information only referenced sharing of personal information, which it defined as information relating to an identified person. The policy did not specify which types of collected information it considers to be personal information, and as such, a reader cannot know how each of those types of information are shared.

Finally, Microsoft, Vodafone, and Mail.ru's policies appeared to suggest that all data they collect may be shared. Microsoft's policy called the section on collection, "Personal Data We Collect," and the section on sharing, "Reasons We Share Personal Data." It states that the company collects name and contact data, login credentials, demographic data, information about interests and favorites, payment data, usage data, contacts and relationships, location data, and content. The policy's section on sharing did provide examples of different types of sharing, but the explanations did not map directly to the list information in the collection section. That is, a reader is unable to know, for example, how the company shares the information it collects about a user's interests and favorites. The company stated that it shares personal data "with vendors or agents working on our behalf for the purposes described in this statement," but a reader is unable to discern whether these vendors can access all types of information the company collects, or only specific types.

Retention and Deletion of User Information

With regard to retention of user information, people would expect a company to keep information they actively submit to the service (e.g., posts, messages, photos, videos, etc.), until they delete it themselves. But companies collect several other types of user information, and they typically fail to disclose how long they retain those types of information.

Ten policies contained statements that companies would retain information for as long as necessary. The lack of explanation of what types of information this includes and the lack of examples of what the company would consider necessary for different types of information render such disclosure unhelpful. For example, AT&T's policy stated, "We keep your Personal Information as long as we need it for business, tax or legal purposes. After that, we destroy it by making it unreadable or undecipherable." Similar to the company's disclosure on sharing of information, this provided no clarity on how long the company retains information it does not consider personal information. Other policies referenced legal requirements with regard to retention, but did not provide further information on what those requirements entail. This reinforces the notion that privacy policies are written for a regulatory, rather than consumer, audience. Facebook's disclosure was less detailed, stating, "We store data for as long as it is necessary to provide products and services to you and others, including those described [in the policy.]"

Seven policies included a time frame for the retention of specific types of user information. KakaoTalk's policy explained what information the company retains for three and six months and one, three, and five years. A Vodafone document about how users can request their information from the company stated that the company retains "detailed records of incoming calls, texts and other messages for a 12-month period only." WhatsApp's policy stated that it deletes undelivered messages from its servers after 30 days. Microsoft's policy stated, "For interest-based advertising, we retain data for no more than 13 months, unless we obtain your consent to retain the data longer." It did not provide additional information about this consent. Twitter stated, "After a maximum of 10 days, we start the process of deleting, de-identifying, or aggregating Widget Data…"and that it "will either delete Log Data or remove any common account identifiers, such as your username, full IP address, or email address, after a maximum of 18 months." Yahoo's privacy policy linked to a page on data storage and anonymization that stated, "IP addresses within search user log data will be anonymized or deleted within 6 months from the time of collection."

Google, Microsoft, and Yahoo provided information on the time frame in which they de-identify certain types of information. Google's privacy policy linked to a page on advertising that stated the company removes part of the IP address and cookie information from server log data after 18 months and 9 months, respectively. The Bing-specific section of Microsoft's policy stated that the company removes "the entirety of the IP address [from stored search queries] after 6 months, and cookie IDs and other cross-session identifiers after 18 months." Yahoo's data storage and anonymization page stated the company de-identifies "search user log data within 18 months of collection, with limited exceptions to meet legal obligations."

Eleven policies contained some disclosure related to users deleting their accounts or user information, while

four policies only included general statements about how the companies delete or destroy user information. Eight of the policies mentioned that deleting an account may not remove all user information from company servers. Facebook, Microsoft, Twitter, Yahoo, and Tumblr provided specific examples, typically relating to information that has been copied elsewhere or is not associated with a user's account (e.g., text of messages sent to another user, search engine logs, content indexed by search engines, information in backups or archives, content that other users have shared publicly). The remaining three policies, from Instagram, Mail.ru, and MTN, provided more general statements that all user information may not be deleted. MTN went so far as to say, "You may request that we delete and destroy any of your personal information provided that you have finalised your relationship with MTN. MTN will however retain and use your personal information for as long as is necessary to comply with our legal and business obligations, resolve disputes and enforce this policy."

Google's policy described several tools that enable users to "make meaningful choices about how [the information Google collects] is used." These tools allow users to control whether information is linked to the user's account or stored on the user's device; they do not control the company's collection or retention of information. For example, users can "decide what types of data, such as videos you've watched on YouTube or past searches, you would like saved with your account when you use Google services" as well as "manage whether certain activity is stored in a cookie or similar technology on your device when you use our services while signed-out of your account."

Google did not clearly state whether it deletes user information when a user terminates her account. A company help page related to deleting a Google account stated, "If you delete your account, you won't have access to your data," but the page did not clearly disclose whether the company itself can access the information.

How Privacy Policies Can Provide More Clarity to Users

The overall result of attempting to understand how these companies handle user information by reading their privacy policies is incoherence. While all policies list at least some types of user information that companies collect, disclosure of other aspects of information handling, such as sharing or retention of user information, are not presented using the same list. If user information flows through a life cycle, or "distinct periods and conditions from its first collection to its disposal or destruction," then individuals who use an ICT product or service should be able to discern how a company manages their information through this cycle (Schwartz and Solove 2014 p. 892). However, the disclosure in these policies does not make this possible.

Breaux, Hibshi and Rao (2014) examined three privacy policies in their work to develop a specification language to support developers in mapping data flows based on disclosures in privacy policies. The found the format of the disclosure (i.e. listing the types of information the company collects separately from the explanation of how companies use the information) to be problematic:

> This separation yields a logically inferred many-to-many mapping between information types and purposes, because the analyst must reasonably assume that any data type maps to any purpose within the entailment of subsumption...Many-to-many tracing is likely an indicator of a less privacy protective policy, because it affords companies more opportunities to use data in difficult to comprehend ways or ways that are unforeseeable by examining any one statement in the policy (p. 299).

Nissenbaum (2010) states "that a right to privacy is neither a right to secrecy nor a right to control but a right to *appropriate* flow of personal information" (p. 127, emphasis in original). This paper's analysis of 23 policies from 16 companies shows that these companies' disclosure does not enable a reader to create a one-to-one mapping of the flow of different types of user information that companies handle. In particular, the language that such policies use to name or define the information they handle contributes to the lack of clarity. Company policies generally provide a list of the different types of information they collect, but when they discuss how they share such information, several companies focus their disclosure on personal information. This makes it difficult for a reader to understand whether their disclosure is complete. This can also complicate efforts to develop machine-learning techniques to analyze privacy policies.

Companies, particularly those in the ICT sector, aggregate massive amounts of data that are generated as people use networked products and services. Zuboff (2015) calls these data "surveillance assets" and explains that companies use them to generate revenue and attract investment, or "surveillance capital." The result is not simply an erosion of privacy, she states, but a redistribution of the rights to choose what information remains with whom. This power is increasingly becoming concentrated among a few corporate actors.

Obtaining more clarity from companies about how they handle such assets is one way to check such power. As Schwartz and Solove (2011) write, "increased transparency will go far toward correcting the asymmetry of knowledge between consumers and the companies that track their online behavior" (p. 1890). One way companies can do so is by framing their privacy policies as documents that

clearly depict the life cycle of each type of information they collect rather than highlight their compliance with law or regulation.

References

Cate, F.H. 2010. The Limits of Notice and Choice. *IEEE Security and Privacy* 8(2): 59-62.

Cranor, L.F. 2012. Necessary But Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice. *Journal on Telecommunications and High Technology Law* 10(2): 273-307.

Breaux, T.D.; Hibshi, H.; and Rao, A. 2014. Eddy, a Formal Language for Specifying and Analyzing Data Flow Specifications for Conflicting Privacy Requirements. *Requirements Engineering* 19(3): 281-307.

DeNardis, L. 2014. *The Global War for Internet Governance*: Yale University Press.

European Union. 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). *Official Journal of the European Union* 119: 1-88.

Freedom House. 2015. Freedom on the Net 2015, Freedom House, New York City, New York.

Greenleaf, G. 2014. Sheherezade and the 101 Data Privacy Laws: Origins, Significance and Global Trajectories. *Journal of Law, Information & Science* 23(1): 4-49.

Hoang, N.P.; Asano, Y.; and Yoshikawa, M. 2016. Your Neighbors Are My Spies: Location and Other Privacy Concerns in GLBT-focused Location-based Dating Applications. *ICACT Transactions on Advanced Communications Technology* 5(3): 851-860.

Liccardi, I.; Abdul-Rahman, A.; and Chen, M. 2016. I Know Where You Live: Inferring Details of People's Lives by Visualizing Publicly Shared Location Data. In Proceedings of the ACM Conference on Human Factors in Computing Systems, 1-12. San Jose, Calif: CHI'16.

Liu, F.; Ramanath, R.; Sadeh, N.; and Smith, N.A. 2014. A Step Towards Usable Privacy Policy: Automatic Alignment of Privacy Statements. In Proceedings of the International Conference on Computational Linguistics, 1-11. Dublin, Ireland; COLING '14.

MacKinnon, R. 2012. *Consent of the Networked: The Worldwide Struggle for Internet Freedom*: Basic Books.

Maréchal, N. 2015. Ranking Digital Rights: Human Rights, the Internet and the Fifth Estate. *International Journal of Communication* 9(10): 3440-3449.

Mayer, J.; Mutchler, P.; and Mitchell, J.C. 2016. Evaluating the Privacy Properties of Telephone Metadata. *Proceedings of the National Academy of Sciences* 113(20): 5536-5541.

McDonald, A.; and Cranor, L.F. 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society* 4(3): 543-568.

Nissenbaum, H. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*: Stanford University Press.

Ranking Digital Rights. 2015a. 2015 Corporate Accountability Index, Ranking Digital Rights, Washington, D.C.

Ranking Digital Rights. 2015b. Methodology Development, Ranking Digital Rights. Last modified November 2, 2015. https://rankingdigitalrights.org/methodology-development/.

Ranking Digital Rights.2015c. Corporate Accountability Index 2015 Research Indicators, Ranking Digital Rights, Washington, D.C.

Rao, A.; Schaub, F.; Sadeh, N.; Acquisti, A.; and Kang, R. 2016. Expecting the Unexpected: Understanding Mismatched Privacy Expectations Online. In Proceedings of the 12th Symposium on Usable Privacy and Security, 1-21. Denver, Colo.; SOUPS'16.

Schwartz, P.M.; and Solove, D. 2009. Notice and Choice: Implications for Digital Marketing to Youth. Memo prepared for The Second NPLAN/BMSG Meeting on Digital Media and Marketing to Children, Berkeley, California, 29-30 June.

Schwartz P.M.; and Solove, D.J. 2011. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review* 86: 1814-1894.

Schwartz P.M.; and Solove, D.J. 2014. Reconciling Personal Information in the United States and European Union. *California Law Review* 102: 877-916.

Smith, A. 2014. What Internet Users Know About Technology and the Web: The Pew Research Center's "Web IQ" Quiz, Pew Research Center, Washington, D.C.

Zimmeck, S.; and Bellovin, S.M. 2014. Privee: An Architecture for Automatically Analyzing Web Privacy Policies. In Proceedings of the 23rd USENIX Security Symposium, 1-17. San Diego, Calif.

Zuboff, S. 2015. Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology* 30: 75-89.