NIPFP Working paper series

Disclosures in privacy policies:
Does "notice and consent" work?

No. 246
11-December-2018
Rishab Bailey, Smriti Parsheera, Faiza Rahman, Renuka Sane

National Institute of Public Finance and Policy
New Delhi

Disclosures in privacy policies: *Does "notice and consent" work?*

December 2018

Rishab Bailey
Smriti Parsheera
Faiza Rahman
Renuka Sane

The authors are with the National Institute of Public Finance and Policy (NIPFP), New Delhi.

Abstract

This paper evaluates the quality of privacy policies of five popular online services in India from the perspective of access and readability. We ask – do the policies have specific, unambiguous and clear provisions that lend themselves to easy comprehension? We also conduct a survey among college students to evaluate how much do users typically understand of what they are signing up for. We find that the policies studied are poorly drafted, and often seem to serve as check-the-box compliance of expected privacy disclosures. Survey respondents do not score very highly on the privacy policy quiz. The respondents fared the worst on policies that had the most unspecified terms, and on policies that were long. Respondents were also unable to understand terms such as "third-party", "affiliate" and "business-partner". The results suggest that for consent to work, the information offered to individuals has to be better drafted and designed.

We thank Omidyar Network for funding support. We thank Vrinda Bhandari for vetting our analysis of the privacy policies and the participants at the NIPFP Technology Policy Conference in Bangalore, and the NIPFP Data Protection Conference in Delhi, for their comments. We also thank the following persons for their help in organising the surveys: Aparna Chandra, Anant Sangal, Arushi Vats, John Sebastian, Tanvee Nandan, Tridip Ray and research scholars in the School of Social Sciences, Jawaharlal Nehru University. All views expressed are personal and all errors are our own.

Contents

1 Introduction

The "notice and consent" framework has been the basis for much of the thinking in modern data protection and privacy laws. It relies on the ability of providers to collect and process personal data conditional on providing adequate information to, and obtaining the consent of, the data subject. Its intuitive appeal lies in the normative value of individual autonomy that is the cornerstone of modern liberal democracies. Seeking consent ensures an individual's autonomy and control over her personal information, enabling "privacy self-management" (Solove, 2013). There is, however, a growing concern around the inability of this model to provide individuals with meaningful control over their data in light of evolving technologies and data practices (Matthan, 2017).

First, research shows that most people do not read privacy policies.[1] Those that do read privacy policies do not opt out or change the default privacy settings.[2] Cognitive issues such as bounded rationality, the availability heuristics, and framing effects also limit an individual's ability to make rational choices about the costs and benefits of consenting to the collection, use, and disclosure of their personal data (Solove, 2013).

Second, structural issues such as the presence of too many entities processing personal data impose a burden on user time and lead to consent fatigue.[3] Third, many privacy harms flow from an aggregation of pieces of data over a period of time through interconnected databases of different entities, or from the use of complex machine learning algorithms that process the data. It is, therefore, unrealistic to expect people to assess the impact of permitting the downstream use

[1] See Consumer Policy Research Centre (2018), where a survey of the digital behaviour of 1004 Australians found that only 6% of the participants had read the privacy policies or terms and conditions for all the products they signed up to in the preceding 12 months and 33% admitted to not reading all privacy policies in the said period. Also see Internet Society (2012), where the survey results demonstrated that 80% of users do not usually read privacy policies even when aware that they are sharing personal data with a website or service, and that 12% of users admitted to never having read privacy policies.

[2] In addition, see Steinfeld (2016) where the experiment showed that nearly 79.7% of the participants who were asked to agree to the terms and conditions of the study without being presented with the policy by default (where policy was only given via a link and not present entirely while clicking) agreed to the terms without clicking the link to read the policy. On the other hand, the eye-tracking device showed that participants who were given a complete copy of the terms and conditions as a default setting spent 59,196.11 ms on an average (nearly a minute) on reading the terms.

[3] In a study conducted in the context of American Internet users, McDonald and Cranor (2008) found that reading privacy policies would take each user approximately 201 hours a year, worth about USD 3,534 annually per user.

and transfer of their data (Solove, 2013). Fourth, it is infeasible for companies to comprehensively detail possible uses of personal information at all times.[4] Finally, privacy policies are often binary in nature where they allow people to either fully opt-in or completely opt-out of using the services (Cate & Mayer-Schönberger, 2013).

These points about the evolving nature of consent have also been acknowledged in policy and legal debates. In August 2017, the Supreme Court of India recognised the fundamental right to privacy.[5] Around the same time the Government of India constituted a committee under the chairpersonship of Justice B. N. Srikrishna (Srikrishna Committee) to draft a data protection law. The Srikrishna Committee submitted its report and a draft bill on data protection to the Government in July, 2018. The committee has affirmed the central role of notice and consent in its draft law, making consent one of the grounds for processing of data. In Europe also, the recently enforced European General Data Protection Regulations (GDPR) has continued (and attempted to strengthen) the consent model implemented under the Data Protection Directive of 1996, while setting out several duties of data controllers.

As per the Srikrishna Committee's recommendations, for consent to be valid it should be *"free, informed, specific, clear and capable of being withdrawn"* (The Personal Data Protection Bill, 2018). In case of "sensitive personal data", the draft law proposes a higher standard of "explicit consent" with additional requirements on what would amount to informed, clear and specific consent with respect to such data.[6] Given the critical role of consent in the draft law, it becomes important to question how can consent be made to work better?

In this paper, we lay the groundwork for answering this question by first asking – *is the present notice and consent regime broken because of the way in which privacy policies are designed?* We evaluate the quality of privacy policies of five

[4]Barocas and Nissenbaum (2013) argue that complexity is a huge challenge for achieving meaningful consent by giving the example of online behavioural advertising where it is virtually impossible to either know or predict the manner in which tracking, analysis, and use (present and future) of data will pan out. They note that, *"There is potentially an unending chain of actors who receive and may make use of behavioral and other data. New companies bloom, novel analytical tools emerge, business relationships begin and end"*.

[5]Justice K.S. Puttaswamy (Retd.) v Union of India and Ors (2017).

[6]Section 18(2) of the draft Personal Data Protection Bill, 2018 defines "explicit consent" as consent which is *"(a)informed, having regard to whether the attention of the data principal has been drawn to purposes ofor operations in processing that may have significant consequences for the data principal; (b)clear, having regard to whether it is meaningful without recourse to inference from conduct in a context; and (c)specific, having regard to whether the data principal is given the choice of separately consenting to the purposes of, operations in, and the use of different categories of sensitive personal data relevant to processing"*.

popular online services being offered in India. We analyse the identified privacy policies from the perspective of access – how easy they are to find, how easy they are to read, and on issues of substantive content – how well do they conform to well recognised principles of a model data protection law. In doing so, we evaluate whether the policies have specific, unambiguous and clear provisions that lend themselves to easy comprehension.[7] This sort of analysis also finds support in the recent recommendations of the Justice Srikrishna Committee, which emphasised the need for improvements in privacy policies on grounds like approachability, comprehensibility, helpfulness, legibility and readability of the policies.

We then evaluate how much do users typically understand of what they are signing up for, and if this can inform us on whether consent is an effective tool to enable individuals to control their personal data in the online environment. We conduct surveys in five universities in and around New Delhi, and randomly assign one of the five privacy policies to students in the classroom. We then quiz the students on the policy, and evaluate how many questions the students were able to answer. Our questions fall into one of three categories – 'easy', 'intermediate', and 'difficult'. The easy questions have a simple and direct answer in the policy. The intermediate questions require a closer reading of the policy making it slightly harder to figure the correct response. The difficult questions require careful readings and some inference.

We find that the policies studied are poorly drafted, and often contain language that appears as though it is meant to insulate the company from liability claims rather than genuinely informing the user. In some cases, there are rights that are considered to be essential in modern privacy frameworks but are not included in the policies (for instance clauses covering data breach notification, or data retention periods). Sometimes, the policies also seem to assume that the user has knowledge of legal terms and is up-to-date with statutory and other regulatory requirements in their jurisdiction.

Survey respondents do not score very high on the privacy policy quiz. The average score of the sample is about 5.3 on 10, i.e., on an average respondents were able to answer 5 out of the 10 questions. The respondents fared the worst on policies that had the most unspecified terms, and on policies that were long. Respondents were also unable to understand terms such as "third-party", "affiliate" and "business-partner". The results suggest that for consent to work, the information offered to individuals has to better drafted and designed.

While surveys of this nature have been conducted in other jurisdictions, we are not

[7] Our analysis is limited to the text of the policy policies, without attempting to evaluate how the terms are being implemented or adhered to in actual practice.

aware of any similar study (to understand how users interact with privacy policies) involving Indian participants. The peculiarities of the Indian context throw up new challenges of diversity in language, literacy, modes of Internet access and other variations among the over 494 million Internet users in India. All of these factors will have to play a role in determining the appropriate design of disclosures and consent frameworks for Indian users. This study makes a modest start in that direction by questioning how well do educated, English-speaking users fare in terms of understanding privacy policies. Making the same privacy policies accessible to the larger set of Indian users, many of whom are first time adopters of technology, is a larger challenge that this paper does not attempt to address.

The paper is organised as follows. Section 2 describes the research design. This includes a discussion on the choice of online services, the parameters for the study of the privacy policies, and the design of the survey. Section 3 provides an analysis of the privacy policies on parameters such as length, readability, visual presentation as well as substantive content. Section 4 presents the survey results. Section 5 sets out our concluding remarks.

2 Research design

The use of surveys to understand how individuals engage with the subject of privacy is not novel. Broadly, there are two kinds of surveys in this field – those that evaluate user perceptions and concerns regarding their privacy, and those that evaluate their ability to understand the terms of privacy policies.

The most influential surveys of the former kind were conducted by Alan Westin. These surveys sought to measure the American public's attitudes and concerns towards privacy (Kumaraguru & Cranor, 2005). Consumers were asked to choose between options such as (i) very concerned; (ii) not too concerned; (iii) somewhat concerned; (iv) not concerned at all, about different aspects of their privacy.[8] Their responses were used to classify Americans into various categories with regard to their concern of privacy leading to the conclusion that most individuals are "privacy rationalists" i.e. they weigh costs and benefits before making rational choices that guide the market towards a direction that balances concerns pertaining to privacy with the widespread adoption of technology (Hoofnagle & Urban, 2014).

[8]For instance, in Westin's survey on health information privacy conducted in 1993, he asked, *"How concerned are you that many health care providers you use today employ computers in some of their operations, such as patient billing and accounting, laboratory work, and keeping some medical records?"* See (Kumaraguru & Cranor, 2005).

This assumes that consumers have understood what is at stake and made rational choices about the trade-offs between their privacy and services.

However, research has demonstrated that privacy concerns (or lack thereof) may be misplaced, for consumers often do not understand the contracts that they have signed into. For example, a study to understand the way Indians perceive privacy, *inter alia* concluded that some participants incorrectly believed that India had a privacy law, participants believed that passwords were more protected personal information as compared to religion, financial information, mobile numbers, and health related information. Participants also admitted to storing passwords, credit card numbers, permanent account numbers, and personal identification numbers on their mobile phones (Kumaraguru & Sachdeva, 2012).

Other surveys that have specifically evaluated what consumers understood of privacy policies also show a significant divergence between the consumer understanding of online rules and common business practices. For instance, a study found that participants who shop online believed that privacy policies inherently proscribe third-party information sharing. Further, a majority of participants appeared to believe (not always correctly) that privacy policies created rights such as access and correct information, data breach notification, assistance in case of identity thefts (Hoofnagle & King, 2008). Often these discrepancies arise because the policies are not drafted in a manner that aids understanding (Contissa, 2018).

The recognition of these limitations in users' understanding of privacy policies has been accompanied by research to show that there are several tools that can help to improve these outcomes. This could include simplified notices that adopt good design techniques such as white spaces, bold words, subheadings, bullet lists, a larger font size, and standard formats to improve readability (Kleimann Communication Group Inc., 2006); layered notices and privacy finder formats.[9] McDonald et al. (2009) found that layered policies and the privacy finder report format assisted faster decision making in comparison to natural language standardised policies. However, with regard to accuracy, they were not necessarily better than natural language policies and were in fact worse in many instances – respondents often did not seem to go past the first layer which led them to make incorrect inferences of the company's privacy practices.

This paper is closer to the literature on the analysis of consumers' understanding

[9]In a layered notice, the first layer gives a brief overview of the key terms with requisite standardised headings, usually covered in only one screen of text, It includes links to the second layer, which contains the entire natural language policy. Privacy finder format refers to a tool developed by AT&T and the Cylab Usable Privacy and Security laboratory to standardise the text descriptions of privacy practices in a brief bullet format. See McDonald, Reeder, Kelley and Faith (2009).

of policies. We distinguish ourselves from previous research on the basis of the three key decisions taken in terms of research design – (i) basis for selection of online services, (ii) criteria of assessment, and (iii) design of survey. We turn next to explaining the decisions on these parameters.

2.1 Selection of online services

This paper focuses on the privacy policies of five commonly used online services in India across different sectors. In making this selection we relied on two broad parameters. First, the popularity of the selected services, which determines the number of users bound by their privacy policies. The study therefore includes the policies of India's top five applications in terms of reach, namely WhatsApp and four Google services (Google Play, YouTube, Gmail and Google search) (Comscore, 2018).[10] Second, we looked at the relevance and diversity of different sectors. We selected services that span a spectrum of sectors, such as online search, messaging, e-commerce, payments and transportation, each of which constitutes a critical element of the digital ecosystem in India. It is pertinent to highlight that the versions of the privacy policies that were accessed for the study were dated as of March, 2018 i.e. before the GDPR was enforced. Therefore, any changes that may have been effected to the privacy policies of these services in light of the GDPR are not captured in this study.

The online services identified are as follows:

1. *WhatsApp (Messaging)*: WhatsApp is a popular messaging app that holds 98% of the instant messaging market in India (Comscore, 2018). It has over 200 million active monthly users in the country, making India one of the largest markets for Whatsapp.[11] Two privacy-related issues for which the company has attracted attention relate to the terms of the data sharing arrangements between WhatsApp and Facebook and the adoption of end-to-end encryption by the company. Facebook's acquisition of WhatsApp in February, 2015 was followed by changes to their privacy policies to allow for sharing of data between the two companies. The implications of this arrangement on the right to privacy of users led to a challenge, which is currently

[10]The services are ranked using ComScore's Mobile Metrix, which captures total mobile audience behaviour on browsers and apps across smartphones and tablets. Google follows a common privacy policy for all its services, which means that the policy studies here applies to a range of services besides the four more popular ones identified above.

[11]Data as of February, 2017. See Statista (2017) and LiveMint (2017).

pending before the Supreme Court of India.[12] In April, 2016 WhatsApp adopted full end-to-end encryption on its service, offering stronger privacy and security features to its users. However, this measure has also been linked to the spread of misinformation and fake news on the platform due to the challenges in detecting and controlling the flow of encrypted messages.

2. *Google (Online search and other services)*: Google is the market leader in online search services in India, currently holding over ninety seven percent of the market share in that segment (Statcounter, 2018).[13] Some of the other popular services in its bouquet include Gmail, its email service; YouTube, its video hosting platform; and Google Maps, its online mapping service. Google follows a common privacy policy across its services.

3. *Uber (Transportation)*: Uber is a technology-based transportation platform that launched its services in India in 2013. By July, 2017 it reported a presence in 29 Indian cities counting about 5 million active riders per week (Uber, 2017). Uber and competing Indian cab aggregator Ola have faced allegations of abuse of dominance under competition law, although, thus far, none of the cases have resulted in an adverse finding against them. In other jurisdictions, such as the United States (US), Uber has also faced regulatory scrutiny for its privacy related practices. In August, 2017, following a data breach incident, Uber entered into a consent agreement with the US Federal Trade Commission (FTC) regarding misrepresentation of information about how Uber's employees access users' personal data and the steps taken to secure that data.[14] Subsequently, on April 12, 2018, Uber agreed to an expanded settlement after the FTC found that the company had failed to inform it of a data breach which took place while the earlier investigation was in progress.

4. *Flipkart (E-commerce)*: Launched in 2007, Flipkart is India's leading e-commerce marketplace, and claims to host over 80 million products. The platform has 100 million registered users, sees 10 million daily page visits and does around 8 million shipments per month.[15] Even though it was recently purchased by another US company Walmart, in popular imagination it is seen as the "Indian" rival to Amazon.

[12]Special Leave to Appeal (C) No.804/2017. The case has arisen as an appeal against the decision of the Delhi High Court in Karmanya Sareen v Union of India (2016).

[13]In February, 2018, the Competition Commission of India (CCI) found Google to be a dominant player in the market for online general web search services as well as online search advertising services in India (Matrimony.com Limited and Google LLC and others, 2018). It imposed a penalty of Rs. 1.36 billion on Google along with orders to desist from indulging in certain practices.

[14]See Federal Trade Commission (2017).

[15]https://www.flipkart.com

5. *Paytm (Payments)*: Paytm is one of India's leading payment gateways. It offers payment solutions to about 7 million merchants. In 2017, Paytm became the first payment app in India to cross over 100 million downloads,[16] and also became of the first payment services to get a "Payment Bank License."[17] Paytm has been in the news on account of privacy concerns for two reasons. Frist, there have been allegations about Paytm being asked by the government to share confidential user data (The Wire, 2018). Second, Paytm has been fined by the Reserve Bank of India for illegaly opening payment bank accounts of users using the KYC that was done for getting a mobile phone (Mehta, 2018).

The inclusion of Flipkart and Paytm along with multi-national corporations (MNCs) like Google, WhatsApp amd Uber ensures a fair mix of domestic and foreign players in the study. This makes for an interesting comparison as the privacy policies of companies may vary depending on the governing laws of the jurisdictions in which they are incorporated or established. Further, companies that are part of larger conglomerates, such as WhatsApp and the Google services, may also have data sharing agreements with other group companies, that can have implications on user privacy.

Notably, to the best of our knowledge, this approach to selection of services has not been adopted so far as similar survey-based studies are concerned.[18] Other studies often limited their approach to privacy notices belonging to a single sector, such as e-commerce[19] or banking,[20] choosing specific policies based on varying degrees of readability or distinct design choices.

[16]https://bit.ly/2Q0lNXI

[17]https://bit.ly/2www3yG

[18]In their study, Contissa (2018) use artificial intelligence to automate legal evaluation of privacy policies, under the GDPR. While the content of privacy policies of distinct companies such as Google, Facebook, Amazon, Apple, Uber etc. were examined in this analysis, a survey-based study was not undertaken.

[19]See McDonald et al. (2009), where the survey restricted its analysis to the study of six popular e-commerce websites. These websites were selected since they represented all three formats, were likely to collect a variety of personal information and exhibited varying degrees of readability

[20]See Kleimann Communication Group Inc. (2006), where the study restricted itself to testing privacy notices of three banks that were identified on the basis of varying amount of personal information they collected and different notice design choices they adopted.

2.2 Criteria for assessing policies

As a precursor to understanding a privacy policy, the policy itself must be easy to locate. In the online space, user attention is generally short[21] making it all the more important for privacy policies to be easily accessible.[22] Accessibility also implies that policies must be easy to understand.

While gauging the "readability" of a policy is not an exact process, certain uniform tests may be applied to each policy, providing a workable comparative analysis. In addition, the availability of privacy policies in local languages, visual representations and clarity in explaining legal terms can influence a better understanding of the policy. In section 3, we evaluate the policies from these points of view before moving on to analysing the substantive content of the policies. The focus of this section is to explain the rationale behind the selection of the specific themes around which the policies were assessed.

The Srikrishna Committee's recommendations and draft bill on data protection as well as earlier work by a Group of Experts headed by Justice AP Shah that submitted its report to the Government in October, 2012 drew to a large extent from what have come to be recognised as well as accepted principles of data protection globally. Some of these basic principles of data protection relate to: collection and purpose limitation; obligations of data quality and storage limitation; requirement of personal data breach notification; data security obligations; individual's right to access and correction of their information; and the right to seek redress. We draw from these principles to identify the types of information that service providers would typically be expected to share in their privacy policies in order to enable users to make informed choices.

Again, prior studies seeking to understand the manner in which individuals interact with privacy notices did not formulate questions on the basis of the above mentioned fundamental privacy principles. For instance, Kleimann Communication Group Inc. (2006) sought to assess identified privacy policies on the basis of the following three criteria –(i) comprehensibility of their content; (ii) the design of the policies and (iii) the perception of participants in relation to identified policies.

Other surveys, such as McDonald et al. (2009), have focused on finding statistically significant differences in the mean scores of participants on the basis of accuracy in answering questions that require comprehension of terms of the policies and the

[21] For instance it is estimated that users spend an average of 15 seconds dwelling time per webpage (Haile, 2014).

[22] Studies also indicate that most users who are not presented with a policy by default never click to read it (Steinfeld, 2016).

time taken to complete. Participants in that study were asked questions pertaining to typical privacy concerns such as use of cookies by websites, sharing of details with third parties etc. While the nature of these questions asked is similar to the ones asked in our study, as stated earlier, this survey restricted its study to e-commerce websites and specifically sought to conduct a comparative analysis of how well different formats of privacy policies worked. To do this, the study compared standardised policies, layered notices and privacy finder reports.

2.3 Survey design

The choice of sample was influenced by three criteria. First the respondents had to be aware of the online services with a very high probability of being users, and second, they had to have the ability to read and understand English. We then chose five higher education (undergraduate as well as postgraduate) institutions in and around New Delhi. Two of these institutions gave us access to students with a law background, while two gave us access to students studying economics, and one was a management programme, with a mix of students from various disciplines.

The sample considered in this paper is a skewed sample, and certainly not representative of either the general population, or that of students in India. However, the sample is useful in that if this even this group does not fare well with regard to understanding the privacy policies, then there may be little hope for everyone else. This hypothesis ties in with the readability scores of the selected policies using the Flesch-Kincaid test. As explained further in section 3, the contents of each of the policies were found to be of a reading level that either college students or university graduates would be best placed to understand.

A survey instrument was prepared consisting of 10 questions on the privacy policy of each company based on the principles discussed above.[23] The questions were divided into three categories –

- *Easy*: These questions were relatively easy to answer based on a basic reading of the policy;
- *Intermediate*: These questions had an obvious answer, but required careful reading of the policy; and
- *Difficult*: These questions required the respondent to infer if a particular provision would apply to the situation presented before her.

[23]The instrument is available online at https://macrofinance.nipfp.org.in/releases/BPRR2018_Disclosures-in-privacy-policies.html. These are also discussed in more detail when discussing the results.

Table 1 The questions and difficulty levels

Q1: Collection Q5: Sharing with government Q9: Right to seek clarification	Easy
Q2: Permitted use Q4: Use by affiliated entities Q8: Data retention	Intermediate
Q3: Sharing with third party Q6: Data breach notification Q7: Access to own data Q10: Exporting of data	Difficult

Table 1 summarises the categories of questions and their difficulty level. The questions were finalised following an iterative process of conducting smaller pilot studies which helped in modifying the language and scope of the questions to address any comprehension issues that were being faced by the respondents.

Each question could have four possible answers – (i) Yes; (ii) No; (iii) Policy does not specify; or (iv) Can't say. That is, if the question was about whether the online service collected a specific type of data, the answer could be a yes, the company does, or a no, the company does not collect this data, or the policy does not specify anything about collection of the data at all. The respondents could choose the final option ("cant say") if they were unable to evaluate what the policy said on the question. The correct answers were agreed upon after an in-depth analysis by the authors, and also vetted by an external lawyer. The explanation of the difficulty level of each question and the relation with the privacy principle on which it was based is as follows:

1. **Collection** – *Category - Easy* – The first question sought to understand whether users were able to understand from the policy the types of information that was being collected about them. The question asked users whether the service provider could automatically collect details about the model/ make of the computer/ smartphone being used by the individual. This was classified as an easy question since most of the policies did make an upfront disclosure in their policies about the types of data that could be collected by them.[24]

2. **Permitted use** – *Category - Intermediate* – The second question sought to find out whether users were able to understand the purposes for which the information collected on them could be deployed by the service provider, with a particular focus on the ability of services to profile users. The question was

[24]This particular information about access to the user's device information was not explicitly listed in Flipkart's policy.

customised in each case depending on the nature of the service being rendered by the provider. For instance, the persons answering the questionnaire for Flipkart were asked if the provider could use their prior product browsing history to advertise products to them while Uber's respondents were asked if the policy allowed the battery status of their phone to be used for determining the applicable fares.

A review of the policies showed that each of them clearly provided this sort of information in their policies. Yet, this was classified as an intermediate question as it required some application of mind by the respondent about the possible uses for which the collected data may be deployed by the service provider.[25]

3. **Sharing with third parties** – *Category - Difficult* – The third question dealt with the issue of sharing of information by the service provider with third parties. We introduced four types of variations in this question. The first was in case of WhatApp's questionnaire where we explicitly mentioned that the data was proposed to be shared with an *unaffiliated* third party. In two other questionnaires (Uber and Google), we gave examples of data sharing arrangements while specifying that the entity with which the data was being shared was a third party. In case of Flipkart, the question gave an example of a data sharing arrangement with a hospital that was interested in marketing its services to a user who regularly orders medicines using Flipkart, without using the term "third party". Finally, for Paytm the question specifically mentioned that the entity with whom the data was being shared was a "business partner" of Paytm.

   We classified this as a difficult question in view of the uncertainty around the meaning of terms like "third party", "business partners", etc., which were being used in the privacy policies without clarifying the meaning of those terms.[26] This imposed an obligation on the respondent to infer the meaning of those terms and then apply the provisions contained in the policy to the given fact situation.

4. **Use by affiliated entities** – *Category - Intermediate* – This question dealt with the sharing of information by the service providers with affiliated entities, such as a parent company or a wholly owned subsidiary, and the purposes for which that affiliated entity could use the information. For instance, the

[25] Only in case of the WhatsApp questionnaire, the second question also related to the collection of information. It asked whether Whatapp may collect **and retain** the location information that we may share with friends using the functionality offered on the app.

[26] One caveat is that it is possible that some terms are defined by the companies elsewhere and not within the privacy policy.

Google questionnaire asked if the contents of a person's Gmail communications could be used by a group company, YouTube, for making relevant video recommendations. For the sake of clarity, each of the questionnaires, except in case of WhatsApp, specified the relationship between the service provider and the affiliated entity in the question itself. In WhatsApp's case, this relationship was not mentioned as the question referred to the use of WhatsApp metadata by Facebook and its privacy policy clearly specifies that WhatsApp constitutes a *"part of the Facebook family of companies"*. Paytm is the only provider that explicitly states that they will not share the data with any "third party". We have interpreted this term to include affiliates.

This was classified as an intermediate question because, as with the previous question on third parties, it also required the respondent to interpret the meaning of the term 'affiliates' and apply the provisions contained in the policy to the given fact situation, which could be a source of confusion for respondents.

5. **Sharing with the government** – *Category - Easy* – It is a well established principle that an individual's right to data protection is not an absolute right and in appropriate circumstances this right may be overridden by other requirements, in accordance with a procedure laid down by law. Legitimate access by law enforcement agencies generally falls under this exception. Against this background, this question sought to check the extent to which respondents were able to understand and apply the exceptions contained in the privacy policies relating to sharing of data with government bodies, including the police and other law enforcement and regulatory agencies.

   This was classified as an easy question as we found that all the policies had clearly spelt out the circumstances under which they would be sharing the data with government authorities.

6. **Data breach notification** – *Category - Difficult* – This question sought to test whether the privacy policies provided users with the right to receive a notification in case of any unauthorised access to their data. It asked that if the service provider's servers were hacked and the personal data of users was accessed by a third party, would the provider be required to notify the user of the data breach? This question was included in all the questionnaires except in case of Whatsapp.[27] However, a review of the policies showed that none of them provided users with this information, becoming the only question that had "not specified" as the correct response for all the service providers.

   When a policy is not specified, it requries greater effort for the respondent

[27] This exclusion was a result of oversight and not part of the research design.

to ensure that it is actually not specified, and that she has not missed interpreting some clause that may actually have an answer. Therefore, this question was classified as a difficult one.

7. **Access to own data** – *Category - Difficult* – This question dealt with the right of the user to access the information held on them by the service provider. It asked that if a user were to write to the service provider requesting for all the available information about her, would the provider be required, as per the terms of the privacy policy, to make this information available to her. We found that only Uber and Google had provisions regarding sharing of the user's data with her while the other three policies were silent on this issue. The unspecified nature of most of the policies classified this question as a difficult question.

8. **Data retention** – *Category - Intermediate* – This question asked respondents about the provisions contained in the privacy policies relating to the scope and duration for which information relating to the user may continue to be held by the service provider.

   The questionnaires for WhatsApp, Google, Flipkart and Paytm asked whether the provider would delete all the information that it holds on the user upon the deletion or discontinuation of her account by the user. In Uber's case the question was drafted slightly differently to ask whether Uber would delete all the information it has about the user after deletion of the app even though there is an outstanding dispute between the user and Uber. This framing was done taking into account a specific provision in Uber's privacy policy which provides that Uber may continue to hold the information in certain circumstances, including when there is an outstanding dispute.

   While assessing the correct answers to the questions we found that the user data is usually retained after deleting the account or in some cases this information is not specified (Paytm and Flipkart) in the policies. Moreover, the text in many of the policies is not direct – they do not clearly state that the entire data will be unconditionally deleted. This requires the respondent to judge whether the situation described in the question would qualify for deletion and therefore this question is classified as an intermediate question.

9. **Right to seek clarification** – *Category - Easy* – This question was based on the principle of requiring "openness" in engaging with users about their privacy policies. It questioned whether the policy explicitly provided for a mechanism through which a user who did not understand some of the terms of the policy could seek clarifications from the service provider. While assessing the correct answer to this question we found that each of the policies

did provide for a mechanism to seek such additional information. Notably, the policies of Flipkart and WhatsApp contained a specific section setting out the email address and contact information of the concerned person, which made it easier to locate these details while perusing the policies. Hence this qualifies as an easy question.

10. **Exporting of data** – *Category - Difficult* – The availability of a user's transaction and browsing history and other types of data with a service provider, which enhances the user's experience of that service, can pose as a barrier for switching to another provider. This problem can be addressed to some extent if the user has the option to easily migrate her data from one service provider to another provider, in a convenient and cost effective manner. Accordingly, this question sought to understand whether the privacy policies of the selected providers had provisions to facilitate the download/export of their data. While assessing the correct answers we found that none of the service providers, except Google, had included a provision along these lines in their privacy policies, though the provision was not clearly drafted. Accordingly, we chose to classify this as a difficult question.

3 Analysis of the privacy policies

We turn next to analysing the accessibility of the selected policies, based on parameters such as readability and the languages that they are available in; the visual design of the policies; and the terminology used in them. This is followed by an assessment of the substantive content of the policies using the ten issues identified by us in Section 2 as the basis for comparison of the policies. Several of the tools adopted by us in conducting this analysis also find support in the recommendations made by the Justice Srikrishna Committee on data protection.

The Committee also suggests a broader set of methods to aid user comprehension of privacy policies, which include: (i) *simplifying text* – use crisp sentences and easier words, replace statements with relevant questions;[28] (ii) *designing for easier comprehension* – provide summaries of sections, use non-textual elements (videos, sound, etc), proper spacing and font size; (iii) *emphasising important points* – avoid the use of all-uppercase, use visual markers; and (iv) *ensuring broader accessibility* – offer text in multiple languages, optimise for access across devices and for offline use.

[28]For example, replace 'Data sharing policy' with 'who has access to my information?'.

3.1 Access to privacy policies

The number of clicks taken to locate a privacy policy is a simple but effective test to determine the accessibility of a policy.[29] If a policy is deeply embedded within a website or platform, it will require more time and patience on the part of the user to access it. It is therefore logical to infer that the more clicks it takes to lead a user to a privacy policy, the less likely that the user will be to bother to actually read or locate it.[30] Table 2 presents the results from our analysis of the policy on metrics such as number of clicks, length, and readability.

Table 2 Access to the policies

Online service	(1) Number of clicks	(2) Length Pages (A4)	(3) Length Words	(4) Language	(5) Readability Reading ease
Uber	2	11	3,355	Eng.	16.44
WhatsApp	2	10	3,352	Eng.	36.56
Google	1	9	2,890	Eng., Ind.	18.30
Flipkart	1	5	1,767	Eng.	41.03
PayTM	3	3	819	Eng.	20.55

As noted in Table 2, privacy policies can generally be accessed through 1-3 clicks (from the main webpage). However, the links to the privacy policies are usually positioned at the bottom of the main webpage, and in relatively small font size.[31] This does not lend itself to easy discoverability, particularly as links to the privacy policies are usually not highlighted.

The length of a policy impacts the time it takes for a user to read the policy. It is not unreasonable to expect more users to give up reading a longer policy than one that takes a comparatively shorter time to read (or to merely skim longer policies).[32] At the same time, certain policies may be longer than others if they explain concepts in a more detailed or comprehensive manner, or for instance, if

[29] While the privacy terms would normally be displayed to (and have to be accepted by) the user while signing up for the service, this analysis is more relevant in situations where a user wants to refer to the terms at a later point of time.

[30] As noted previously, the Committee of Experts under the Chairmanship of Justice BN Srikrishna (2018) suggests providing a separate page for all privacy related information, as a distinct tab.

[31] In the case of Paytm, the privacy policy could be accessed both through a link on the main webpage directing users to the terms and conditions (following which another click takes you to the privacy policy), as well as links in the 'About Us' section of the website (users then need to click on the section 'Our Policies' followed by another click on the 'Privacy' button).

[32] Longer policies are more likely to lead to consent fatigue. Interestingly, research indicates that it would take an average user 76 work days to read all the privacy policies encountered in the course of one year - an amount of time that is clearly impractical for a normal user to spend

examples are provided to aid the user's understanding of the application of privacy policies. There is, therefore, no absolute measurement possible of what an ideal length of a policy is or how a policy should be designed to aid easy understanding (particularly as it is possible that users may read only specific portions of a policy at any given time, based on need).[33]

We see from Columns (2) and (3) of Table 2 that the privacy policies of the Indian companies (Paytm and Flipkart) are significantly shorter than the multinational companies ("MNCs") studied. This is largely due to the greater number of issues touched upon as well as more detailed explanations of rights and obligations by the MNCs. Some of this may be due to the fact that the MNCs' policies may be following some of the obligations under data protection laws of foreign countries that contain more onerous requirements than India's Information Technology Act, 2000 (and rules thereunder).

Given that all the services studied are available across India, it is essential that privacy policies are also made available in multiple Indian languages. While roughly 125 million Indians speak English, less than a quarter of this number considers English as their first language (Kroulek, 2018). Even Hindi (including variants and dialects) is recognised as a first language by less than half the population (GOI, 2001). This metric may take on added significance when one considers reports that indicate that online content in local languages tends to be trusted more and circulated more widely (Bhattacharya, 2017; Pal & Bozarth, 2018).

As seen in Column (4) of Table 2, Google is the only company amongst those studied that provides a copy of its privacy policy in a language other than English. Despite some of the other websites being made available in non-English languages (for instance, Uber's website can be accessed in Hindi), the privacy policy continues to be accessible only in English. The Committee of Experts under the Chairmanship of Justice BN Srikrishna (2018) underscores the importance of providing options to view privacy policies in the 'common languages of the regions where the service is available'.

While measuring readability is not an exact science, tools such as the Flesch-

reading legal documents (Chyi, 2018). This is recognised in the Committee of Experts under the Chairmanship of Justice BN Srikrishna (2018) which notes that policies should aim to minimise the intimidating nature of policies to encourage engagement, and simplify content to make it easier to understand.

[33] In this context, the recommendation in the Committee of Experts under the Chairmanship of Justice BN Srikrishna (2018) to categorise the content of the policy into sections, as a collapsible set of links, appears to be sensible as this would provide an overview of the various issues covered in the policy at a glance. Users could then choose to click on specific links pertaining to the part of the policy they are interested in.

Kincaid reading ease and grade level tests have been used for decades to analyse certain metrics such as word and sentence length and their impact on readability. Such readability scores provide a useful comparative matrix within which to evaluate readability. However, it should be noted that the model does not actually analyse the meaning of words used, whether they could have multiple or ambiguous meanings, whether words used in the text are commonly used, etc. It is therefore possible for a completely un-understandable text (consisting of short but rarely used, complex or ambiguous words) comprising short sentences with short words to have a high readability score.[34] Column (5) of Table 2 shows the reading ease level for the 5 policies under study. The readability scores are graded in the following Table 3 (Wikipedia, 2018):

Table 3 Readability

Name of company	Readability	Level of reading
Uber	Very difficult	University
WhatsApp	Difficult	College
Google	Very difficult	University
Flipkart	Difficult	College
Paytm	Very difficult	University

To provide some context, one may note that Reader's Digest has a readability score of about 65, David Copperfield by Charles Dickens has a readability score of 62, Harry Potter books have a readability score of 80.6, while the Harvard Law Review has a readability score in the low 30s (Lively, 2015).

Interestingly, many states in the US require insurance policy forms and endorsements to have minimum Flesch scores between 40 and 50.[35] In our sample, only WhatsApp and Flipkart came even close with readability scores of 36.56 and 41.03, respectively.

[34]The Flesch Reading Ease uses the total words, sentences, and syllables to produce a readability score on a scale of 0-100. The formula to calculate the score is based on two factors: sentence length as judged by the average number of words per sentence, and word length as judged by the average number of syllables in a word. The rationale for this is that long sentences, and sentences with long words (with multiple syllables) are tougher to read. A high Flesch reading ease score indicates that the material is easy to read, while a low score indicates difficulty of reading the content. As a rule of thumb, readability scores of 90-100 can be understood by an average 5th grader or a 10/11 year old student; 8th and 9th grade students can understand documents with a score of 60-70 or a 13-15 year old student; college students are best placed to understand material with scores of 30-50; and university graduates can understand documents with a score of 0-30. The results in Table 3 have been compiled by inputting the privacy policies of the five companies under study into an open source readability application - Flesch (Biagini & Frink, 2003)

[35]For instance, Arkansas, Hawaii, North Dakota, Ohio, Oklahoma, South Dakota and Tennessee require minimum scores of 40; Florida requires a score of 45; Maine, Massachusetts, Minnesota, Montana, Nebraska, Nevada, New Jersey, New Mexico and North Carolina require minimum scores of 50 (Hawkins, 2011).

The results above indicate that all the privacy policies under study are complicated documents and require a firm grasp of English and reasonably advanced comprehension abilities to be understood. Given that the target audience for many of these online services ranges from adolescents upwards, it appears that the privacy policies will *prima facie*, be too complicated for many users to comprehend.

3.2 Visual presentation

Another way in which reading a privacy policy can be made easier, both in terms of readability and comprehension, is through the use of highlights, marginal notes and by properly segregating and identifying overarching topics.[36]

Uber's privacy policy for instance, is divided into multiple sections with each subheading in bold font. The policy also contains marginal notes that summarise each section, thereby making the policy easier to understand at a glance.[37] Notably, Uber also provides an easy-to-read summary of their privacy policies in a seperate 'overview' page. Certain specific services also enable click-throughs for more information (for instance, privacy policies concerning Uber's freight carriage service, cookie statement, etc.).

Google's privacy policy also contains segregated sections, and a table of contents which permits easy access to different portions of the policy. Interestingly, the policy also frequently uses layered information or pop-ups – additional information is presented pertaining to certain terms and activities when a user moves the cursor over highlighted words. To illustrate, certain crucial words or phrases (for example the words 'unique identifiers') are highlighted or underlined and moving a cursor over them opens a pop-up or sidebar with a simpler explanation.

While WhatsApp also provides segregated sections, it does not generally provide additional information in a layered manner or highlight particularly important information (though, certain highlighted terms do allow click-throughs – for instance 'Facebook family of companies' and 'cookies'). A table of contents allows for easy navigation to different portions of the privacy policy.

[36] These methods are also recommended in the Committee of Experts under the Chairmanship of Justice BN Srikrishna (2018), which further encourages the strategic use of non textual design elements (icons, colour codes, etc.) and multi-media solutions to aid user engagement and comprehension. The Report also points to the importance of ensuring proper pagination and text layout – including by ensuring appropriate use of different fonts and line spacing.

[37] The Committee of Experts under the Chairmanship of Justice BN Srikrishna (2018) highlights the importance of such marginal notes or summaries, while recognising that these would be merely indicative in nature and not binding.

The two Indian companies studied – Flipkart and Paytm – do not provide layered information or any further click-throughs in their privacy policies. Flipkart demarcates sections using a bold font (in the same font size as the rest of the document), while Paytm utilises a larger font size, in bold, for section headings.

3.3 Terminology in the policy

We focus next on the kind of terminology used in the privacy policies. Our focus here remains on the text of the policies, without getting into the manner in which the policies may be implemented in actual practice. We note that the use of legal and technical terminology in a privacy policy can lead to a decrease in comprehensibility for the user. Unless specifically defined, a user may not be aware of the true import of a particular word, particularly if technical in nature.

For instance, WhatsApp's privacy policy mentions that

> *"we do not retain your messages in the ordinary course of providing our Services to you"*

This does not define what the phrase "ordinary course" implies or explain what the exceptions are. A user may, on a thorough reading of the policy come to understand that it may apply to situations where for instance, law enforcement is involved. However, there is no clarity on this.

Similarly, the use of words and phrases such as 'third party', 'affiliate', 'profiling', etc., may also lead to confusion in the minds of users particularly in the absence of any specific definitions. An average user may not be aware of what information a "profile" would consist of, the possible uses that it may be put to or the risks associated with profiling. Similarly, a user not familiar with legal terms may not understand the import of words such as "third party" or "affiliate" which are used frequently in privacy policies to refer to the entity with whom the information is being shared.

Equally important to note is that policies sometimes fail to disclose sufficient or clear information to users. For instance, WhatsApp's privacy policy specifies that

> *"we offer end to end encryption for our Services, which is on by default, ...End to end encryption means that your messages are encrypted to protect against us and third parties from reading them."*

The privacy policy appears therefore to give a user the impression that the use of end-to-end encryption, guarantees privacy of the user's data. However, this need not be the case - for instance, the policy does not mention if and what kinds of

metadata with respect to the encrypted message are produced and/or collected and retained by the company.[38]

Separately, it is important to note that none of the privacy policies, except Google, contain their own definitions section. This implies that either the terms used are left undefined or they require the user to locate the more generic terms and conditions of service to confirm if the terms are defined there.[39]

The absence of definitions is exacerbated by the use of open-ended statements (or inclusive lists), that often require a user to guess what the company intends. For instance, Flipkart's privacy policy states that

> *"we collect and store information which is provided by you from time to time..this allows us to provide services and features that most likely meet your needs and to customise our website to make your experience safer and easier.* ***More importantly, while doing so we collect personal information from you that we consider necessary for achieving this purpose****."* (emphasis added).

The last line quoted above does not *prima facie* clarify what specific kinds of information are collected by the company, leaving it for the user to try and understand what information the company may consider necessary to achieve the stated purpose. Given the expansive and varying nature of business interests of most companies, such a necessity could be difficult to gauge for a normal user.

As users are often not provided complete or verifiable information about the services they are using, Chyi (2018) has arrgued that *"is impossible for people to grasp the wider potential harms of sharing their data with these platforms"*. To illustrate the issue of inclusive lists, WhatsApp's privacy policy states that

> *"We collect device specific information when you install, access, or use our Services. This* ***includes*** *information such as hardware model, operating system information, browser information, IP address, mobile network information* ***including*** *phone number, and device identifiers."* (emphasis added).

[38]Further, the policy implies that the use of end-to-end encryption renders any messages unreadable by third parties. This ignores the fact that encryption technology too can be breached. Notably, reports indicate that hackers can morph the identity of a sender of a message on WhatsApp and even alter the text of another person's message (Barda, Zaikin & Vanunu, 2018). The WhatsApp policy therefore arguably conveys a greater sense of security to users than is strictly warranted.

[39]Certain basic terms are defined in most policies. For instance, the use of pronouns (such as 'we' and 'our'). However most policies do not define specific legal and technological terminology, with the exception of the term 'cookies' which most policies studied explain in fair detail – often linking to a separate cookie related policy.

While the use of examples should normally aid in understanding the relevant clauses, providing an indicative list arguably does not fulfill the company's obligation to ensure that the user is fully aware of her rights and obligations (in this case, the user is not made aware of all the device specific information collected by the company). It would be preferable, from a privacy protection perspective, for companies to provide exhaustive lists, wherever possible. That said, excessive quantities of information can also prove pointless or confusing to a user, particularly given the limited granularity of consent forms and the lack of real choice that a user may face.[40]

Connected to the problem of lack of adequate information is the issue of whether the information being provided is trustworthy and reliable. While it is outside the scope of this paper to examine the issue of trust in online services, it must be kept in mind that online businesses frequently appear to treat user privacy rights with less than due respect (not least due to the lack of bargaining power and information asymmetry between the parties). Recent news reports, for instance, indicate that despite users on Android and iPhones opting not to share location data with certain Google services, the company continued to record their location and movements (Associated Press, 2018). Trust could also affect how users read privacy policies - particularly in the absence of clarity in specific cases. Verifiability therefore becomes important and is possibly a space for legislative intervention.[41]

3.4 Substantive content

Table 4 provides a snapshot of whether policies have specific provisions on the ten issues identified as a basis for analysis of policies, i.e. collection; permitted use of data; information sharing with a third party, use by affiliated entities, and law enforcement; uses of personal information/purposes of processing; data breach

[40]Recently, PayPal released a full list of third parties it shares user information with. The list has over 600 entities on it - a user will have to scrutinise a 98 page list, and is thereafter given the binary choice of either stopping use of the service or accepting its practices in totality (Chyi, 2018). The Committee of Experts under the Chairmanship of Justice BN Srikrishna (2018) attempts to overcome such issues by recommending the use of ordered or unordered lists wherever possible, simplifying phrasing and using pop-ups or links to provide addition or tangential information.

[41]There are currently no external independent verification or auditing systems in place to ensure that a company adheres to its policy statements or that the policy statements represent a truthful version of what practices the company undertakes. The Personal Data Protection Bill (2018) attempts to enhance accountability and verifiability of data practices through the implementation of external audit mechanisms. Interestingly, the draft law proposes the assignment of 'data trust scores' to data controllers (referred to as data fiduciaries in the draft law) thereby enabling users to easily judge reliability of a website's privacy policies and practices).

Table 4 Analysing the privacy policies

Q1: Collection
Q2: Permitted use
Q3: Sharing with third party
Q4: Use by affiliated entities
Q5: Sharing with government
Q6: Data breach notification
Q7: Access to own data
Q8: Data retention
Q9: Seek clarification
Q10: Exporting of data

Policy	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
WhatsApp	Y	Y	**NS**	Y	Y		**NS**	N	Y	**NS**
Uber	Y	Y	N	Y	Y	**NS**	Y	N	Y	**NS**
Paytm	Y	Y	N	N	Y	**NS**	**NS**	**NS**	Y	**NS**
Google	Y	Y	N	Y	Y	**NS**	Y	N	Y	Y
Flipkart	**NS**	Y	N	Y	Y	**NS**	**NS**	**NS**	Y	**NS**

notification; access rights; data retention/deletion; right to seek clarifications; and exporting of data.[42]

3.4.1 Collection of personal information

The five companies studied collect as much information on the user as is possible. The onus is on the user to limit her interactions with the platforms and decide what information to make available as any information provided is generally retained by the companies. Personal information collected generally includes demographic information, transaction history, usage and log information, financial information, location information, device and connection information, etc. The policies usually use indicative lists or examples to illustrate the kinds of information they collect - meaning that a user may often have to guess if certain types of information are collected or not.

While the three MNCs studied – Google, Uber and WhatsApp – provide indicative lists of the information collected under broad headings,[43] the two Indian companies tend to use more generic language implying that any information that is provided by the user can be collected (or to put it another way, the categories of data collected are not limited in any way). To illustrate, Paytm states that personal information includes

> *"... any details that may have been voluntarily provided by the user in connection with availing any of the services on Paytm."*

[42]The detailed reasoning for the answers to the questions contained in the survey (including references to the relevant provisions of the privacy policies) are available online at https://macrofinance.nipfp.org.in/releases/BPRR2018_Disclosures-in-privacy-policies.html.

[43]For instance, WhatsApp differentiates between information provided by the user, automatically collected information and third party information; similarly Uber differentiates between information provided by the user and that collected from other sources; Google differentiates between (a) information given by the user, (b) that collected from the use of Google's services by the user, and (c) information associated with the user's Google account.

All the policies studied refer to the use of cookies and provide a brief explanation on what cookies do and are used for or provide a link to a separate cookie policy. However, these statements are not necessarily thorough – for instance, there is little information on whether a user continues to be tracked after leaving the webpage of the relevant service provider.

3.4.2 Permissible uses of personal information

The five policies tend to allow fairly expansive uses of the data by the data controllers. In general however, one may identify the following uses that personal information is put to: (a) to provide services and improve service quality (b) to resolve disputes (c) to promote safe services, ensure safety of the system, prevent breach of terms and conditions, prevent fraud, administer the service/website, etc. (d) for marketing purposes (e) to customise user experiences. There are rarely any details on what constitutes use under any of these categories.

Use of personal information is often determined by whether the data is classified as "sensitive personal data".[44] Interestingly, Google appears to be the only one that segregates information depending on whether it constitutes sensitive personal information or not. This is a notable difference given that all the services studied do actually collect certain types of sensitive personal information - for instance, Paytm has access to financial data of users. It would appear logical for more precautions to be taken with respect to such data than say, demographic data.

3.4.3 Data sharing with third parties

Each of the policies studied has a fairly detailed section explaining the circumstances under which personal information may be disclosed to third parties. However, as discussed in the previous section, the terminology used is often ambiguous and policies are replete with illustrative lists. Negative covenants are few and far between.

Generally speaking, rights to share personal information with third parties are reserved due to: (a) a legal request or reason (for instance to protect property rights of a third party) (b) as the third party provides a data processing service or is a subcontractor/vendor to the data controller (c) for marketing purposes.

[44]Notably, the Information Technology Act, 2000 and Information Technology (Reasonable security practices and procedures and sensitive personal data or information) Rules, 2011, recognise a difference between "personal information" and "sensitive personal information", with processing of the latter category requiring compliance with a greater number of obligations.

WhatsApp, Google and Uber provide detailed disclosures with regard to the sharing of information. For instance, WhatsApp[45] and Google[46] classify personal information into different categories based either on the nature of information or the method/parties with whom such information is shared. The policies then provide a short explanation on the sharing of each of these categories of personal information.

Paytm is a positive outlier in so far as sharing of information with third parties is concerned - its policy states that it does not *"sell, share or rent"* a users' personal information to any third party (though aggregated statistical information may be released from time to time). The only possibility of sharing information with third parties is if Paytm is served with "a legally compliant request for disclosure".[47]

3.4.4 Data sharing with affiliated entities

Most of the privacy policies studied contain some information on sharing of personal data with related companies, though the use of legalese does not necessarily help in comprehension. For instance, Uber's privacy policy notes that *"it may share data with subsidiaries and affiliates who either help Uber provide services, or conduct data processing on Uber's behalf."* The term "affiliate" is not defined, which could possibly lead to confusion in the minds of users about the exact nature of the parties with whom information could be shared. Similarly, Paytm's privacy policy states that *"it does not share (or sell or rent) personal information to any third party"*. It is unclear, from a reading of the policy, whether the term "third party" includes related or group companies (though it could be argued that it should as these are third parties to the contract between the service provider and the user).

From the five policies studied, it appears that information sharing within a group of companies is fairly common (and is not something that the user can object to or withdraw consent from). For example, WhatsApp uses a stand-alone/ delineated section to state that personal information is exchanged within the Facebook family

[45] Personal information is categorised into account information, contact and other information, information provided to third party providers, and information provided to third party service providers. The policy provides details on when each of these are triggered and what information is shared.

[46] Google segregates personal information based on (i) explicit consent being taken for sharing, (ii) sharing with domain administrators, (iii) sharing for external processing, and (iv) sharing for legal reasons.

[47] We assume that this refers only to compliance under statute / regulation and not instruments that may have the force of law such as contracts. A contract for sale/sharing of information would go against the negative covenant mentioned previously.

of companies (of which WhatsApp is a part). Its policy (a) clarifies that user messages will not be shared publicly on Facebook and will not be used for any purpose other than to provide or improve services; (b) attempts to explain the purposes behind information sharing between the companies (largely to provide customised services, improve service provision, secure systems, for marketing etc.). It also encourages users to read Facebook's privacy policy. It is possible that the detail in this section is due to the differences between the companies on how to handle user data and the public concerns raised in this regard (Grind & Seetharaman, 2018).

3.4.5 Data sharing with law enforcement

Each of the policies studied informs the user that their personal information may be shared with law enforcement authorities. This is generally done on request (by the enforcement agency/government), and in all five cases, without prior or post facto notification to the user concerned.

To illustrate, Paytm reserves the right to *"communicate your personal information to any third party that makes a legally compliant request for its disclosure."* This implies that Paytm would disclose personal information to law enforcement authorities pursuant to a legal request, even though the language does not specifically refer to government agencies/ law enforcement authorities.

Interestingly, Paytm and Google are the only ones that refer to a request for disclosure having to be *'legally compliant'* or *'enforceable'* respectively in order for the company to respond to them. This implies that requests for disclosure by law enforcement agencies that do not follow relevant procedural norms/safeguards set out by law will not be acquiesced to.

3.4.6 Data breach notification

None of the policies studied contain a data breach notification clause. Users will therefore be unaware if any of their personal information is compromised or accessed/used without authorisation.

3.4.7 Access to personal information

The three MNCs studied provide access and rectification rights to users – though the exact contours of this right are not always clearly defined. Uber, for instance, permits users to correct personal information associate with the account at

any time. WhatsApp recognises that users have the right to manage information through the use of service settings (used to manage contacts, groups and broadcast lists), change their phone number, picture, etc. or delete the account. However, there does not appear to be any way for the user to access the information the company holds on the person, for example, by way of profiles.

Google has arguably the most detailed section pertaining to access rights. Broadly, a user has access to the personal information about her collected by Google. Users are also able to check what information is collected and retained by the company through the use of various embedded links.

Both Paytm and Flipkart do not provide for any right to access for users (or do not specify any such right in their privacy policies). This is despite Rule 5(6) of the Information Technology (Reasonable security practices and procedures and sensitive personal data or information) Rules, 2011, requiring users to be given the opportunity to access or correct their personal information (whether sensitive in nature or not).

3.4.8 Data retention and deletion

The policies studied are fairly ambiguous about their data retention practices, with little information provided to the user about when (or if at all) user data will be deleted. Paytm and Flipkart, both do not provide any details as to how long personal information will be retained by them, and also if personal information will be deleted upon deactivation or deletion of a user account.

The information provided by Uber, WhatsApp and Google also leaves much to be desired. Uber for instance, does not specify any data retention period but implies that user data will be deleted or anonymised, subject to certain exceptions, on de-activation of the account or a user request.[48]

WhatsApp specifies that deletion of an account leads to deletion of all information (that is no longer required to provide the relevant services), including undelivered content. Information retained includes that which is necessary for providing services and that already in the custody of 'third parties'.[49]

[48] Uber's policy provides that subject to certain exceptions (and applicable law), personal data will be deleted or anonymised at the user's request / deletion of account. The scope of the exceptions includes, (a) if there is an unresolved issue relating to the account say an unresolved claim or dispute, (b) if required to retain any data by any applicable law, and / or in aggregated or anonymised form, (c) if necessary for legitimate business interests of Uber such as to aid in fraud prevention and enhance user safety and security.

[49] Note that it is not entirely clear who the term "third parties" extends to. For instance,

Google permits the deletion of portions of information that the user does not want linked with her Google account. However, the privacy policy specifies that deletion of a user's information may not result in the immediate deletion of all residual copies of that information from Google's active servers. At the same time, Google *may not* delete such information from backup systems at all. Users are therefore informed that certain parts of their personal information may be retained in perpetuity. It is the only company that explicitly informs users that their data may be retained in perpetuity.

3.4.9 Right to seek clarifications

Each of the five companies studied provides some contact details or mechanism to enable users to connect with the company in case of queries, clarifications or grievances.[50] However, the quality of information provided – and consequently the ability for users to interact with the company – differs.

WhatsApp provides a US based address for users to write to regarding any queries or grievances. It also provides a clickable link allowing users to either send the company a question pertaining to the privacy policy, or redirect to a set of frequently-asked-questions (FAQs). Similarly, Google provides FAQs and online methods of filing complaints. Upon receiving a complaint, Google will follow up directly with the individual concerned.

Uber on the other hand, only provides a physical Netherlands based address (identified as the data controller) whom users can presumably contact in the case of grievances. While this system does provide an avenue for outreach, the utility of the system is questionable particularly for users based other than in the European Union. The absence of an email ID or online complaints mechanism in particular could act to limit the opportunity for non-European Union based users to easily file complaints or make queries.

Paytm provides contact details of a designated grievance redress officer as required under the Information Technology (Reasonable security practices and procedures and sensitive personal data or information) Rules, 2011. Interestingly, Paytm has established a grievance redress process to deal with wallet / payment related complaints and queries, though it is unclear if privacy related issues can be addressed

would it cover sub-contractors of WhatsApp i.e., those who process information on WhatsApp's behalf?

[50] Notably, Rule 11 of the Information Technology (Intermediaries guidelines) Rules, 2011 requires intermediaries to publish on their website the name and contact details of the Grievance Officer as well as mechanism by which users or any victim who suffers as a result of access or usage of computer resource by any person in violation of Rule 3 can notify their complaints.

through this mechanism (this grievance mechanism is also specified in a separate document). Similarly, Flipkart provides a contact (in India) – complete with address, phone number and email address. The entry is relatively easy to locate in the policy.

3.4.10 Exporting of data

Out of the five privacy policies studied, only that of Google provides the user with information and an opportunity to extract personal information or port data to another service.[51] This implies that in all the other cases users have little to no ability to transfer their data to a competing service provider. This issue directly affects not just the ability of individuals to control their information, but also competition in the sector. It is far more likely that a user will consider switching to a competing service provider, if she could take all historical data to the new service. The absence of a data portability provision therefore acts to entrench the position of incumbent service providers.

The legal analysis of the policies shows several problems with how far the policies protect users from privacy harms and how well the information is communicated to users. We turn next, to an evaluation of how respondents fare on understanding the terms of the policies.

4 Analysis of survey responses

The sample consists of 155 respondents across five colleges around New Delhi. The group is between 19 and 25 years of age, and has at least 12-15 years of education. Of the sample, 33% (N=51) are from a law background, 67% (N=104) from a non-law (mostly economics) background. A little over half, 59% (N=92) are post-graduate students while 41% (N=63) are under-graduate students. These students would also be conversant in English as all of these institutes conduct studies in English, and in all likelihood most students at these places would also have completed their schooling in English.

The responses across policies are as follows: Flipkart: 21% (N=32), Google: 21% (N=33), Paytm: 24% (N=37), Uber: 10% (N=16), WhatsApp: 24% (N=37). Each respondent took about 15-20 minutes to fill up the questionnaires.

[51] The policy indicates that certain Google services give the user the option of data portability, though the relevant services are not specified.

Table 5 Average scores by sample characteristics

	Average Score
Overall average	5.3
By study area	
Non-law	5.3
Law	5.2
By degree	
Under graduate	5.1
Post graduate	5.3

Table 6 Average scores by policy

	Average score	
	Overall	Can't say
By policy		
Flipkart	5.3	0.7
Google	5.4	1.2
Paytm	5.5	0.8
Uber	5.9	0.9
WhatsApp	4.6	0.8

4.1 Overall performance

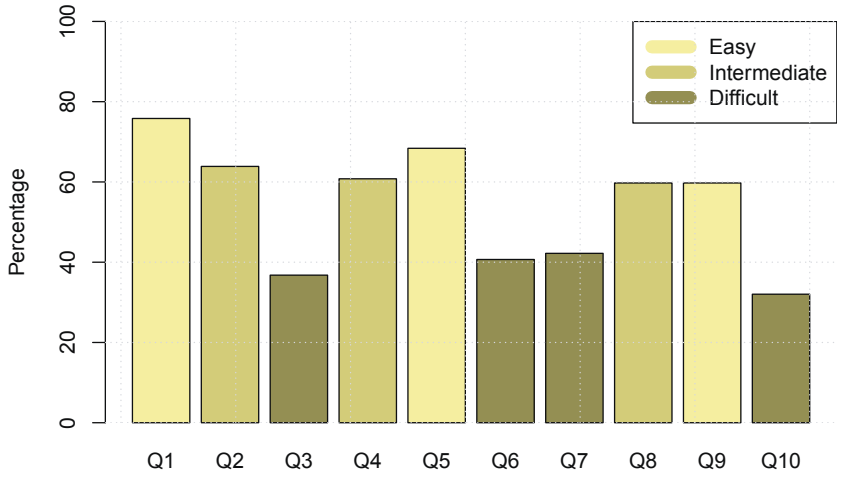
We calculated the total score for each respondent. The highest possible score is 10, if the respondent gets all the answers correct, while the lowest is zero, if none of the answers are correct. Table 5 shows the average scores of the sample.

We find that the overall average is about 5.30. This suggests that on an average, respondents got five questions right. The score is not very surprising, as we expected that respondents should get the obvious questions correct, and this is perhaps driving the average score. There is not much of a difference between the law and non-law students, and between under-graduate and post-graduate students.

What is of greater interest is the average score by policy. Table 6 shows the average scores by policy. Column (1) shows the overall score. The highest average was obtained by those who read the Uber policy, while the lowest was that of WhatsApp. Flipkart and WhatsApp had the most number of "Not Specified" provisions in their policies. It was our contention that when features are not specified, respondents would make mistakes in evaluating the content of the policy.

Another metric of evaluating understanding is the score on the "Can't say" option. If a policy generates a large number of "Can't say" responses, this suggests that the policy has not been understood properly by the respondent. Towards this, we calculate a "Can't say" score for each person for each policy. Here also, the

Figure 1 Number correct by question



maximum score is 10, if the person answers "Can't say" for every question. Column (2) shows the average of this score by policy. The overall low average suggests that very few respondents actually felt that they could not find the answer to the question in the policy. This, however, does not necessarily mean that they got the correct answer, which is reflected in the overall low score. The high score of Google is surprising – one explanation could be that because the Google policy is detailed, respondents felt more confused.

Moving to the question-wise analysis of the responses, Figure 1 shows the percentage of respondents that got each question correct. The difficulty level of the question increases with the shade – the darker the colour, the more difficult the question.

Not surprisingly, we find that a greater percentage of respondents got the easier questions (as classified by us) correct. For example, almost 76% of the respondents got the correct answer to Q1, about 68% got the correct answer to Q5, and almost 60% got the correct answer to Q9. The more difficult the question, the fewer people got the correct answers. The difficulty level of the question is determined based on factors such as the use of complex legal terms in the policy, or ambiguity about specific provisions. We find that when these conditions are present, respondents make mistakes. This offers hope for improving the notice and consent framework – when the policy is straightforward, respondents are able to better understand

what is on offer.

4.2 Analysing responses to the difficult questions

We turn next to an analysis of the difficult questions. Table 7 shows the percentage of correct answers for the difficult questions. Column (1) reflects the answers to the question on sharing data with third-parties (Q3), column (2) the answers to data breach (Q6), column (3) answers to access to own data (Q7), and column (4) answers to exporting data (Q10).[52]

The question on sharing data with a third party begins by painting a hypothetical scenario of the online services business partner asking for information about the respondent's habits or preferences. The question ends by asking, can the online service share this information? The question is as follows:

> *"The [online service] has prepared a profile of you based on your usage. Can it share/sell this information to [business partner/affiliate]?"*

We find that overall 37% of the respondents got the correct answer to this question. When we analyse by policy, we find that about 50% of Google, Paytm and Uber respondents got the correct answers. In the case of WhatsApp, however, only 13% got the correct answer, while in the case of Flipkart only 25% got the correct answer. We, therefore, analyse the content of the WhatsApp and Flipkart clauses relating to sharing of data with third parties.

In the case of WhatsApp, while the policy specifies sharing of data with third parties for processing on behalf of WhatsApp, it is silent on "selling" the data. The question on WhatsApp was worded as follows:

> *"Would WhatsApp be able to sell the data to an unaffiliated third party for direct marketing?"*

It is possible that respondents did not differentiate between sharing the data with a third party for processing on behalf of WhatsApp and selling the data to a third party for the third-parties' own use, and therefore answered incorrectly.

In the case of Flipkart, the question was worded differently. In all the other questions, we specifically asked, if the online service can share the information with either a third party or a business partner. In the case of Flipkart, we asked, if Flipkart can share the information "with a hospital", and did not mention the word

[52]The correct answers to these questions are specified in Table 4 in Section 3. The proportion of respondents that got the correct answer for each question is shown in Table A.1 in the Appendix.

Table 7 Percentage correct answers

Policy	% correct responses			
	Sharing third-party (1)	Data breach (2)	Access own data (3)	Export data (4)
WhatsApp	13		24	35
Uber	50	50	69	38
Paytm	51	38	54	43
Google	52	33	37	19
Flipkart	25	47	41	26

third-party. It may be that most respondents answered this incorrectly because they did not interpret the hospital to be a third party.

While not a difficult question, the question on processing data by affiliated entities also reveals that perhaps people do not understand the term "affiliate or third party" and make mistakes. For example, in the case of Paytm the question reads:

> *"Gemtm, a subsidiary of Paytm, offers a customised jewelery manufacturing service. Gemtm is interested in getting the names and contact information of Paytm users who normally carry out transactions worth more than INR 50,000 on consumables. Can Paytm provide this information to Gemtm?"*

We find that only 35% of Paytm respondents got the correct answer. Interestingly, Paytm is the only company which says that it will not share the data with "third party". While the policy uses the term "third party", we have interpreted the term to include affiliates of the company. Therefore, the correct answer for the Paytm policy would have been "No", that is, Paytm cannot provide the information to Gemtm. But this may be the source of confusion – respondents do not think affiliates are third parties – 54% have therefore responded with a "Yes", which is the incorrect answer.

The question on data breach was phrased as follows:

> *"[Online service's] servers are hacked and your information is accessed by a third party. Is [online service] required to inform you of the data breach?"*

Overall, only 41% of respondents got the correct answer.[53] No policy has specified provisions related to data breach. Yet a large proportion of respondents thought that the company was required to send a notification (41% for Flipkart, 51% for Google, 49% for Paytm and 19% for Uber), suggesting that often respondents just assume the existence of some features by virtue of the fact that a privacy policy

[53] We did not ask this question on the WhatsApp survey.

exists.

The access to data question was worded as follows:

> *"You are curious about the information that [online service] has on you. You write requesting them to share this information. Is [online service] required to provide it?"*

Overall, 42% respondents got the correct answer. WhatsApp had the lowest percentage of correct responses followed by Google. Only Uber and Google have provisions on sharing of the user's data with them. The others do not specify this provision. Yet, we see very different results for Google and Uber – 37% of Google respondents answered a Yes, as opposed to 69% for Uber. The Google policy says that it *"aims to"* provide access to data – there does not seem to be an explicit obligation to do so. This convoluted language could possibly explain why only 37% got it right for Google.

The question on exporting data asks if the online service facilitates the download or export of the respondent's data. The question is worded as follows:

> *"You are considering switching from [online service] to a competing service provider. Will [online service] facilitate the download/export of your data from your account?"*

We find that overall only 32% of the respondents got the correct answer. Here again, with the exception of Google, no other policy specifies anything on data export or data portability. Even though Google specifies something, the wording of the policy is very vague. It says, *"Google allows you to take information associated with your Google Account out of many of the services."* As a result, almost 50% of the respondents answered that the policy did not specify a provision on data export. For most of the other policies, a significant proportion of respondents just answered a "No" although the correct response was "Not specified".

The analysis of the difficult questions suggests that understanding of privacy is a function of several complex factors such as length of the policy, clarity of legal terms, clarity of explanations within the policy. These may, in turn, interact with ex-ante perceptions of the respondents on the particular online service they are using.

5 Conclusion

A central concern in privacy debates around the world has been the inefficacy of the notice and consent model. A large body of literature has evolved that demonstrates that consent is broken, and yet, accepts the necessity of finding ways to make the notice and consent regime work better. This paper asks if the notice and consent framework is broken *because of the way in which it is currently designed*?

Towards this end, it evaluates the quality of privacy policies of five popular online services in India from the perspective of access, and whether the policies have specific, unambiguous and clear provisions that lend themselves to easy comprehension. It also conducts a survey to evaluate if respondents understand the privacy policy.

The legal analysis of the policies suggests that it is quite likely that the policies under review are primarily written with a view to protect the service providers from liability claims rather than provide the user with useable information. This is indicated not just by the excessive use of legalese and ambiguous terminology, but the absence of many rights considered essential in the current thinking on data protection. Some of the policies assume that the user has a knowledge of not just complex legal or technical terms, but is also up-to-date with statutory and other regulatory requirements in their jurisdiction. This is illustrated by the frequent usage of terms such as "to the extent permitted by law", "as permitted by law", etc.

Overall, the paper finds that privacy policies are fairly widely drafted to permit service providers the broad power to collect and process information in pursuance of their business interests. Users currently have little to no leeway in amending the contracts entered into by them and must usually sign up for the entire contract if they wish to access the service.[54]

The complexity of the language and inadequacy of specific details are reflected in the low understanding of respondents. What is interesting about the responses of the survey is that when provisions are clearly drafted, or when users can be expected to find the answers in the policy, they are more likely to evaluate the questions correctly. However, when terms whose meaning is not precisely defined are used (such as "third-party" and "affiliate", for example), then respondents make mistakes. This suggests that in an environment where respondents actually do read the policy, and when the policy is unambiguously drafted, respondents are able to make better sense of what is being offered to them.

[54]Certain services such as Google and WhatsApp do permit users to change certain limited settings that affect their privacy.

The study raises further questions on what drives understanding of privacy policies – whether factors such as age, education, intelligence quotient, comfort with English, urbanisation, familiarity with Internet-based services, all play a role in how an individual evaluates what is on offer? It also raises questions on how privacy policies should be designed so that users are able to understand them.

The proposed data protection law in India provides that in order for consent to be valid it should be *free, informed, specific, clear and capable of being withdrawn.* Businesses that collect and use personal data will bear the burden of establishing that these requirements have been met. The methodology deployed in this study can become one of the tools for meeting this end. Conducting a survey designed around their privacy policies among different user groups (for instance, based on the categories identified above) can help businesses in testing whether the terms are in fact comprehensible to a broad range of their users.

Ultimately, the goal of privacy policies should be to make it possible for individuals to evaluate trade-offs between privacy and service and make choices that suit their preferences, which might themselves change over time. Finding ways to make the notice and consent process more meaningful is an essential part of this process.

References

Associated Press. (2018). Google tracks your location, even when you tell it not to do so. Indian Express, 14 August, 2018. Retrieved from https://bit.ly/2P5WMcj

Barda, D., Zaikin, R. & Vanunu, O. (2018). Fakesapp: A vulnerability in whatsapp. Check Point Research, 7 August 2018. Retrieved from https://bit.ly/2ANe6Ak

Barocas, S. & Nissenbaum, H. (2013). On notice: The trouble with Notice and Consent. Retrieved from https://bit.ly/2wzyUGg

Bhattacharya, A. (2017). India's internet users have more faith in content that's not in English. Quartz India, 2 May 2017. Retrieved from https://bit.ly/2pB24B3

Biagini, C. & Frink, J. (2003). Flesh. Sourceforge, 3 April 2013. Retrieved from https://bit.ly/2L4oHuH

Cate, F. & Mayer-Schönberger, V. (2013). Notice and consent in a world of big data. *International Data Privacy Law*, *3, No.2*, 67–73.

Chyi, N. (2018). Going beyond transparency. Privacy International, 12 July 2018. Retrieved from https://bit.ly/2wqmJLN

Committee of Experts under the Chairmanship of Justice BN Srikrishna. (2018). A free and fair digital economy: Protecting privacy, empowering indians. Ministry of Electronics and Information Technology, Government of India. Retrieved from http://meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf

Comscore. (2018). Global digital future in focus, 2018. Retrieved from https://bit.ly/2FmPksN

Consumer Policy Research Centre. (2018). Australian consumers 'soft targets' in big data economy. CPRC. Retrieved from https://bit.ly/2wvd87q

Contissa, G. (2018). CLAUDETTE meets GDPR: Automating the Evaluation of Privacy Policies using Artifical Intelligence. Study Report, Funded by the European Consumer Organisation (BEUC). Retrieved from https://bit.ly/2NtloyC

Federal Trade Commission. (2017). Uber settles FTC allegations that it made deceptive privacy and data security claims. https://bit.ly/2uY81Jj.

GOI. (2001). Statement 4, Scheduled languages in descending order of Speaker's strength - 2001. Census of India, Government of India. Retrieved from https://bit.ly/2NoZUjb

Grind, K. & Seetharaman, D. (2018). Behind the messy, expensive split between facebook and whatsapps founders. The Wall Street Journal, 5 June 2018. Retrieved from https://www.wsj.com/articles/behind-the-messy-expensive-

split-between-facebook-and-whatsapps-founders-1528208641?mod=trending_now_3

Haile, T. (2014). What you think you know about the web is wrong. Time, 9 March 2014. Retrieved from https://ti.me/1ei3ti9

Hawkins, K. (2011). Many consumers confused by language in insurance policies. Insurance Quotes. Retrieved from https://bit.ly/2wqYRaV

Hoofnagle, C. & King, J. (2008). Research report: What Californians understand about privacy online. Retrieved from https://papers.ssrn.com/abstract=1262130

Hoofnagle, C. & Urban, J. (2014). Alan Westin's privacy Homo Economicus. *Wake Forest L. Rev. 49*, 261.

Internet Society. (2012). Global internet user survey, 2012. ISOC. Retrieved from https://bit.ly/2whTrQx

Justice K.S. Puttaswamy (Retd.) v Union of India and Ors. (2017). W.P. (Civil) No. 494 of 2012, Supreme Court of India.

Karmanya Sareen v Union of India. (2016). W.P. (Civil) No. 7663 of 2016, High Court of Delhi.

Kleimann Communication Group Inc. (2006). Evolution of a prototype financial privacy notice: A report on the form development project. Retrieved from https://bit.ly/2P8cbsm

Kroulek, A. (2018). Which countries have the most English speakers? K International. Retrieved from https://bit.ly/2mi9t94

Kumaraguru, P. & Cranor, L. (2005). Privacy Indexes: A survey of Westin's studies. Retrieved from https://www.cs.cmu.edu/~ponguru/CMU-ISRI-05-138.pdf

Kumaraguru, P. & Sachdeva, N. (2012). Privacy in India: Attitudes and Awareness v 2.0. Indraprastha Institute of Information Technology, Delhi. Retrieved from https://bit.ly/2wC42F9

Lively, G. (2015). Readability. Retrieved from https://bit.ly/2BWU5YC

LiveMint. (2017). WhatsApp has 1 billion active users daily, 28 July 2017. Retrieved from https://bit.ly/2IwXJa3

Matrimony.com Limited and Google LLC and others. (2018). Case Nos.07 and 30 of 2012, Competition Commission of India.

Matthan, R. (2017). Beyond consent: A new paradigm for data protection. Takshashila Discussion document 2017-03. Retrieved from https://bit.ly/2NxUW6P

McDonald, A. & 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. Retrieved from https://bit.ly/1qbLQJ9

McDonald, A., Reeder, R., Kelley, P. & Faith, L. (2009). A comparative study of online privacy policies and formats. Retrieved from https://bit.ly/2vQhh4N

Mehta, T. (2018). Airtel payments bank fined Rs.5 crore for violating Aadhaar biometric KYC rules, 10 March 2018. Retrieved from https://beebom.com/airtel-payments-bank-fined-5-crore-aadhaar/q

Pal, J. & Bozarth, L. (2018). Is tweeting in Indian languages helping politicians widen their reach? *Economic and Political Weekly*, *53*(25). Accessed 6 August 2018.

Solove, D. (2013). Privacy self-management and the consent dilemma. *126 Harv. L. Rev. 1880*.

Statcounter. (2018). Statcounter Global stats, search engine market share in India, June 2018. Retrieved from https://bit.ly/2MyeDvr

Statista. (2017). WhatsApp: Number of monthly active users in India as of February 2017. Retrieved from https://bit.ly/2hF9JLg

Steinfeld, N. (2016). "I agree to the terms and conditions": (How) do users read privacy policies online? An eye-tracking experiment. *Computers in Human Behaviour*, *55*, 992–1000.

The Personal Data Protection Bill. (2018). Ministry of Electronics & Information Technology, Government of India.

The Wire. (2018). Cobrapost Expose: Senior VP at Paytm claims firm was asked to share user data with PMO. The Wire, 24 May 2018. Retrieved from https://bit.ly/2wyLoPa

Uber. (2017). Uber hits 500 million rides milestone in India, 17 July 2017. Retrieved from https://ubr.to/2whRT9b

Wikipedia. (2018). Flesch Kincaid readability tests. Retrieved from https://bit.ly/2NNo2MF

Appendix

Table A.1 Percentage of respondents who got the correct answer

Q1: Collection
Q2: Permitted use
Q3: Sharing with third party
Q4: Use by affiliated entities
Q5: Sharing with government
Q6: Data breach notification
Q7: Access to own data
Q8: Data retention
Q9: Seek clarification
Q10: Exporting of data

Policy	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
WhatsApp	89	73	13	44	54		24	57	78	35
Uber	94	6	50	88	81	50	69	69	44	37
Paytm	81	51	51	35	81	38	54	72	51	43
Google	97	64	52	70	58	33	37	67	42	19
Flipkart	25	97	25	87	75	46	41	37	75	26

MORE IN THE SERIES

- Mohanty, R. K., and Bhanumurthy, N.R. (2018). Analyzing the Dynamic Relationship between Physical Infrastructure, Financial Development and Economic Growth in India, WP No. 245 (November).
- Kaur, A., and Chakraborty, L. (2018). UDAY Power Debt in Retrospect and Prospects: Analyzing the Efficiency Parameters, WP No. 244 (November).
- Gupta, A., Patnaik, Ila, and Shah, A. (2018). Exporting and firm performance: Evidence from India, WP No. 243 (November).

Rishab Bailey, is Technology Policy Researcher, NIPFP
Email: rishab.bailey@nipfp.org.in

Smriti Parsheera, is Technology Policy Researcher, NIPFP
Email: smriti.parsheera@nipfp.org.in

Faiza Rehman, is Consultant, NIPFP
Email: faiza.rehman@nipfp.org.in

Renuka Sane, is Associate Professor, NIPFP
Email: renuka.sane@nipfp.org.in.in

National Institute of Public Finance and Policy,
18/2, Satsang Vihar Marg,
Special Institutional Area (Near JNU),
New Delhi 110067
Tel. No. 26569303, 26569780, 26569784
Fax: 91-11-26852548
www.nipfp.org.in