The Demographics of Readability

Grace Aronsohn*

* Correspondence: garonsohn@middlebury.edu; Tel..: +1-203-952-2457

Submitted: 26 May 2023

Abstract: Consumers are clearly aware of the dangers posed by blindly accepting the terms of a company's privacy policy, pointing to a major contradiction: despite this awareness and near-universal feelings of wariness and suspicion, United States adults continue to agree to terms and conditions when asked without actually reading the terms in question. One likely explanation for this behavior is that privacy policies are not readable in the first place. By using Python to calculate Flesch Reading Ease (FRE) and Flesch-Kincaid (F-K) scores for terms of service and privacy policy agreements published by the "Big Five": Amazon, Apple, Google, Meta, and Microsoft, I determine that users of the products and services offered by these tech giants are being held responsible for the written terms of unreadable contracts. Further, I analyze how responses to Wave 49 of Pew Research Center's American Trends Panel vary across demographic groupings by sex, race and ethnicity, age, education, and income to identify subpopulations that are at greater risk of exploitation.

Keywords: age, Amazon, Apple, Big Tech, contract, education, end-user license agreement, Flesch-Kincaid, Flesch Reading Ease, Google, income, Microsoft, Meta, privacy policy, race, readability, sex, terms and conditions, terms of service,

1 Introduction

According to a 2019 survey of the American Trends Panel conducted by Pew Research Center, 82 percent of United States adults are asked to agree to the terms and conditions of a company's privacy policy at least monthly, with 25 percent of survey respondents reporting that they are asked to do so almost every day. Despite the frequency at which Americans are signing online contracts, survey results also revealed that these contract signers are rarely reading privacy policies in full before agreeing to their terms of service. In fact, 36 percent of survey respondents report that they never read a company's privacy policy before accepting. Another 38 percent report that they only read before accepting "sometimes", meaning only 1 in 4 United States adults regularly read a company's privacy policy when asked to agree to its terms of service (Pew Research Center 2019).

The writers of these contracts, which often take the form of end-user license agreements (EULA), are under no misconception that their privacy policies and community guidelines are being read in full before their users agree to their terms. In 2010, British video game retailer GameStation took advantage of this behavior for an April Fools joke. The company added the following "immortal soul clause" to its website's terms and conditions:

> By placing an order via this Web site...you agree to grant Us a non transferable option to claim, for now and forever more, your immortal soul. Should We wish to exercise this option, you agree to surrender your immortal soul, and any claim you may have on it, within 5 (five) working days of receiving written notification from gamestation.co.uk or one of its duly authorized minions.

Despite the addition of the immortal soul clause, GameStation claimed 88 percent of their customers still accepted the website's terms and conditions (BBC 2013).

Unfortunately, the knowledge that most people are not reading privacy policies before agreeing to them is not limited to companies playing harmless pranks or academic research–there have been countless instances of powerful corporations taking advantage of this fact to exploit their consumers. In 2019, social media giant Facebook Inc. (now Meta) was blasted for paying hundreds of contract workers to transcribe audio clips of its users. The whistleblowers were the contract workers themselves, who questioned the purpose of the task after hearing audio clips of users' conversations, despite having no information from Facebook on the source of the clips or the purpose for their transcription. In response to public backlash, Facebook admitted to collecting audio clips of its users–but stated that a user's acceptance of the terms of the Messenger app granted the company permission to do so (Frier 2019).

In recent years, an increase in controversy surrounding the major corporations that make up Big Tech has led the media to draw public attention to the odd, confusing, and even dangerous clauses hidden deep within their privacy policies and terms of service. In 2016, Business Insider published an article exposing a clause in the service terms of Amazon Web Services that outlined the restrictions around using their products in the event of a zombie apocalypse "that will likely result in the fall of organized civilization", as determined by the Center for Disease Control (Volokh 2016). Apple's iTunes Licensed Application EULA agreement includes a similar clause, prohibiting users from using their products for "the development, design, manufacture, or production of nuclear, missile, or chemical or biological weapons" (Apple 2022).

These are just a few examples of the ease with which abusive clauses are able to remain inconspicuous, hidden deep in the terms of service agreements of powerful corporations.

2 Literature

In 2018, professors Jonathan A. Obar of York University and Anne Oeldorf-Hirsch of the University of Connecticut published "The Biggest Lie on the Internet: Ignoring the Privacy

Policies and Terms of Service Policies of Social Networking Services", hoping to determine the extent to which consumers sign EULAs despite the inclusion of abusive clauses in their contracts. Their experiment asked participants to join a fictitious social media service called NameDrop, which required users to agree to an EULA to use the service. NameDrop's terms of service included clauses that would allow the company to share users' data with their employers and with the National Security Agency. Signers would also be giving NameDrop legal ownership of their first-born children. Obar and Oeldorf-Hirsch found that 98 percent of subjects consented to NameDrop's terms and conditions regardless (Obar and Oeldorf-Hirsch 2018).

Consumers are well aware of the dangers of blindly agreeing to EULAs and skeptical of the true objectives of these contracts, resulting in widespread distrust of the writers of the companies requiring consent to an EULA to use their services. Returning to the 2019 Pew Research Center survey findings, only 4 percent of respondents are confident that companies use their personal information in ways they are comfortable with. Furthermore, only 5 percent are confident that companies are actually adhering to the terms of their privacy policies (Pew Research Center 2019). Consumers' obvious awareness of the dangers posed by blindly accepting the terms of a company's privacy policy points to a major contradiction: despite this awareness and near-universal feelings of wariness and suspicion, United States adults continue to agree to terms and conditions when asked without actually reading the terms in question. One likely explanation for this behavior is that privacy policies are not readable in the first place. In 2021, *Business Insider* published an article calling for government regulation of Big Tech's terms of service agreements. The author points out that many companies have adopted the EULA because of its simplicity and efficiency–a company can begin providing its services immediately after a user accepts the agreement, with all users accepting the same standard form agreement. He argues for the classification of EULAs as contracts of adhesion, or agreements constructed in a way that favors one party without giving other parties an opportunity to negotiate (Levy 2021). In other words, a contract of adhesion requires signing parties to "take it or leave it", accepting the terms exactly as presented if they wish to use the service. By making their policies unreadable, companies are able to further discourage their users from fully reviewing their terms, knowing that users will feel less inclined to do so when they ultimately will not have the power to negotiate any terms they are opposed to.

The impact of contract length on consumer behavior has received notable attention from academia. According to Pew Research Center, even excluding those survey respondents who say they never agree to privacy policies before agreeing to their terms, only 22 percent of remaining respondents reported reading all the way through a company's privacy policy before accepting its terms. The other 78 percent either read partway through or glance over the terms (Pew Research Center 2019). This is a likely result of the excessive

length of EULAs, which often use tactics like hyperlinking to covertly increase contract length. In 2008, PhD student Aleecia M. McDonald and Associate Professor Lorrie Faith Cranor of Carnegie Mellon University published their research on the use of excessive length as a way of creating unreadable contracts. The researchers set out to determine the opportunity cost of reading privacy policies at the reading speed of the average American consumer. They found that in order to read all of the privacy policies that the average consumer agrees to, it would cost the consumer approximately 201 hours each year, worth about $3,524. On the national scale, this adds up to a cost of $781 billion in time lost every year (McDonald and Cranor 2008). These companies do not expect their users to be spending nearly four hours every week reading their terms and conditions–in fact, it is far more likely they expect the opposite, making extending contract length an effective way to ensure that abusive clauses hidden deep in a terms of service agreement go unnoticed.

The impact of contract language has also been a subject of extensive study. According to Pew Research Center, excluding again the survey respondents who say they never read privacy policies before agreeing to their terms, only 13 percent of the remaining respondents reported understanding "a great deal" of the privacy policies they accept. 32 percent reported understanding very little or not understanding at all (Pew Research Center 2019). This can be attributed to language that has been intentionally overcomplicated to the point that it becomes incomprehensible to the average user. In 2010, professors Rainer Böhme of the University of California, Berkeley and Stefan Köpsell of Germany's Dresden University of Technology published their research on the use of complicated contract language as another way of creating unreadable contracts. The researchers analyzed the effect of EULA wording on a user's likelihood of accepting the agreement. Studying the behavior of 80,000 users of an online privacy tool, the researchers examined differences in participation between users receiving language resembling the more coercive wording of a typical EULA–such as button text reading "I accept" or "I decline"–and language that suggested greater user agency–such as button text reading "I take part" and "I do not take part". They found that the users asked to accept terms and conditions using the language of an EULA were 26.8 percent more likely to agree to the exact same terms and conditions as those asked using a more polite dialogue (Böhme and Köpsell 2010). These findings demonstrate how subtle changes in wording can have a significant impact on consumer behavior, making contract language a potent tool for companies to manipulate consumers who agree to their terms of service.

Two of the most widely used measures of readability are the Flesch Reading Ease (FRE) and Flesch-Kincaid (F-K) metrics. Both scores are calculated using the average sentence length and average number of syllables per word of a given text, varying only by the coefficients used. In 2019, professors Uri Benoliel of the Ramat Gan Law School and Samuel Becher of Victoria University of Wellington performed these tests on the 500 most

popular websites in the United States that require users to accept an EULA. They determined that the median FRE score of these agreements to be 34.20, with 498 of the 500 websites scoring below the recommended FRE score of 60. Additionally, they determined the median F-K score to be 14.9, with 498 of the 500 websites scoring above the recommended F-K score of 8.0 (Benoliel and Becher 2019). These findings support the presumption that privacy policies are generally unreadable to the average consumer.

While there is extensive literature on the readability of EULAs and the behavior of United States adults, there is a lack of research into how this behavior varies across demographic groups. Using the raw data from Wave 49 of Pew Research Center's American Trends Panel, which was conducted in June 2019, I will examine whether the demographic qualifiers sex, race, age, education, and income identify subpopulations as being more or less likely to read and comprehend the contracts they sign, thus determining which groups are more vulnerable to exploitation by major corporations.

3 Quantifying Readability

Two of the most widely used measures of readability are the Flesch Reading Ease (FRE) and Flesch-Kincaid (F-K) metrics. Both scores are calculated using the average sentence length and average number of syllables per word of a given text, varying only by the coefficients used.

FRE scores range from 0 to 100, with higher scores indicating greater readability, and are calculated as follows:

206.835 – (1.015 * average words per sentence) – (84.6 * average syllables per word)

Texts that score below 60 are considered to be unreadable to the average consumer, with a score of 60 or above being required by United States government agencies to ensure documents are readable to the public.
F-K scores are associated with a grade level reading ability. The recommended score is 8.0 or less, indicating an eighth-grade reading level. F-K scores are calculated as follows:

(0.39 * average words per sentence) + (11.8 * average syllables per word) – 15.59

This score is required for documents published by organizations like the Department of Education, the Food and Drug Administration, and the National Institute of Health.
In 2019, professors Uri Benoliel of the Ramat Gan Law School and Samuel Becher of Victoria University of Wellington performed these tests on the 500 most popular websites in the United States that require users to accept an EULA. They determined that the median FRE score of these agreements to be 34.20, with 498 of the 500 websites scoring

below the recommended FRE score of 60. Additionally, they determined the median F-K score to be 14.9, with 498 of the 500 websites scoring above the recommended F-K score of 8.013. These findings support the presumption that privacy policies are generally unreadable to the average consumer.

4 Data & Methods

By calculating Flesch Reading Ease and Flesch-Kincaid scores, the same measures of readability used by Benoliel and Becher, I more closely examine the public-facing documents published by Big Tech companies that users are required to accept before using a product or service. By developing a tool using the Python programming language that can determine the FRE and F-K scores of a given text, I evaluate the terms of service and privacy policy agreements of the "Big Five": Amazon, Apple, Google, Meta, and Microsoft. I calculate the FRE and F-K scores for a total of 74 documents, 14 of which are currently enforced. The remainder are archived past versions of these documents that have since been updated.

To identify demographic subgroups of United States adults who are at higher risk of exploitation by companies requiring their users to agree to privacy policies, I use the raw data from Wave 49 of Pew Research Center's American Trends Panel. The survey was conducted in June 2019, and the raw data is available for download here.

5 Preliminary Results

Of the 14 contracts that govern the use of products and services in the present day, Microsoft's Terms of Use contract is the least readable, with an FRE score of 30.78 and an F-K score of 16.28. Apple's Website Terms of Use comes in a close second, with an FRE score of 36.20 and an F-K score of 17.37–in other words, Apple's Website Terms of Use is wrien at a graduate degree reading level. The most readable of the contracts was Microsoft's Privacy for Young People policy, which scored 65.21 on the FRE test and 8.03 on the F-K test. Given that texts scoring below 60 on the FRE test are considered unreadable to the average consumer, Microsoft's Privacy for Young People policy was the only document out of the 14 currently enforced contracts that passed this readability test. All 14 documents had F-K scores exceeding the recommended score of 8.0 or less, meaning all 14 documents are considered unreadable according to this metric.

High income adults and women are asked to agree to privacy policies more frequently

> There is no significant correlation between the frequency at which United States adults are asked to agree to the terms and conditions of a company's privacy policy and race, age, or education.

However, there are minor correlations with income and sex. Americans with higher incomes are asked to agree to privacy policies more frequently than Americans with lower incomes. Women are asked to do so more frequently than men across income brackets.

Sex, race, age, education, and income are strongly correlated with the likelihood that an American adult reads privacy policies before agreeing to them

American men are more likely to agree to privacy policies without reading them beforehand than American women, with 44 percent of men saying they never read privacy policies compared to 33.5 percent of women.

White non-Hispanic and Asian or Asian-American adults are more likely to agree to privacy policies without reading them beforehand, with 41.3 percent and 46 percent of adults saying they never read privacy policies, respectively. Only 6.8 percent and 5.6 percent, respectively, say they always read privacy policies. Black non-Hispanic adults read privacy policies at the highest rate, with 18.6 percent saying they always read them and only 22.4 percent saying they never read them.

Younger Americans are more likely to agree to privacy policies without reading them beforehand, with 50 percent of Americans between ages 18 and 29 saying they never read them, exceeding the next highest by over 10 percent. Adults who always or often read privacy policies are in the minority across all age groups, with over 75 percent of each group responding that they only sometimes read them, if not never.

Americans who have achieved a higher level of education are more likely to agree to privacy policies without reading them beforehand, with only 5.1 percent of Americans with a Bachelor's degree and 4.2 percent of Americans with a postgraduate degree saying they always read privacy policies. 16.1 percent of Americans with less than a high school education say they always read privacy policies. As seen with age, adults who always or often read privacy policies are in the minority across all age groups, with less than one-third of each group responding that they only sometimes read them, if not never.

Americans who earn higher incomes are more likely to agree to privacy policies without reading them beforehand, with less than 10 percent of adults in all income groups exceeding $30,000 annually saying they always read privacy policies. 18.6 percent of Americans who earn less than $10,000 annually say they always read

privacy policies. 53.2 percent of Americans who earn at least $150,000 annually say they never read privacy policies.

Of Americans who do read privacy policies, race, education, and income are strongly correlated with how thoroughly they do so

There is no significant correlation between how thoroughly American adults read privacy policies and sex or age. However, there are strong correlations with race, education, and income.

Only 3 percent of Asian or Asian-American adults say they read privacy policies all the way through, more than six times less frequently than the next lowest grouping. The majority of black and Hispanic adults say they read privacy policies part of the way through, while the majority of White non-Hispanic, Asian or Asian-American, and mixed race adults say they glance over privacy policies without reading them closely.

Only 14.4 percent of Americans with a Bachelor's degree and 14.1 percent of Americans with a postgraduate degree say they read privacy policies all the way through. 53.2 percent and 50.8 percent, respectively, say they glance over privacy policies without reading them closely. 33.9 percent of Americans with less than a high school education say they read privacy policies all the way through, and only 30 percent say they glance over them without reading closely.

Americans who earn lower incomes are more likely to thoroughly read privacy policies, with 35.3 percent of adults earning less than $10,000 annually and 31 percent of adults earning between $10,000 and $19,999 annually saying they read privacy policies all the way through. Less than 20 percent of all groups earning over $50,000 annually say they read privacy policies all the way through, with the majority of these groups saying they glance over privacy policies without reading them closely.

Sex, race, age, education, and income have little bearing on how well Americans understand the privacy policies they read

There is no significant correlation between the frequency at which United States adults are asked to agree to the terms and conditions of a company's privacy policy and sex, race, age, education, or income.

To view the data in more detail, or to interact with the readability calculator I developed for this project, please refer to my interactive web app here.

6 Discussion & Limitations

Having determined that Big Tech terms of service are unreadable to the average American consumer, the question naturally follows of why companies are able to avoid facing consequences for this behavior. The economic and political climate of the United States makes it especially challenging to regulate these contracts.

One of the primary reasons, as illustrated by the tighter regulations faced by Big Tech in Europe, is the strength of laissez-faire sentiments in the United States. Americans tend to have greater faith that free market principles will protect consumers without government interference–they believe that a company enforcing abusive terms will drive users away and adjust those terms accordingly. There is also a precedent of leniency when it comes to regulation of the Internet. When the Internet became more widely available in the United States in the 1990s, the Federal Trade Commission chose not to impose strict regulations on its use, arguing that legislation would only hinder its rapid growth. This argument is commonly used to justify the use of EULAs, as the standard form, "take it or leave it" contract minimizes transaction costs and is far more efficient than allowing users to negotiate a terms of service agreement. While regulation of the Internet has become more necessary in recent years, historical policy still poses a notable challenge for legislators.

Likely a result of this historical leniency towards regulation of the Internet, Big Tech companies are able to take advantage of a number of legal loopholes related to the nature of their products. First, because companies have no way of forcing their users to comply with their terms and conditions, there is less of a pressing need for the government to ensure that those terms are not abusive. Companies can prevent users from accessing their services if they violate or decline their terms, but because use of those services is a choice, user consent is seen as purely voluntary (Levy 2021). A second major loophole concerns the Uniform Commercial Code, the consumer-protection laws that are enforced in all 50 states. Because most Big Tech companies provide services, not products, they are exempt from the protections detailed in the code. The third loophole is related to the duty to read doctrine, a foundational pillar of contract law in the United States. Under the duty to read doctrine, all participating parties are held responsible for the written terms of a contract, regardless of whether or not they have actually read those terms. However, nowhere in the duty to read doctrine are those contracts required to be readable. Some states have enacted their own "plain language laws", but these laws are often limited and lack an objective standard for readability (Benoliel and Becher 2019).

As of 2019, white, highly-educated, and high income Americans were most likely to feel as if their personal information was less secure than it had been five years prior. White, highly-educated, and high income Americans are also more likely to follow privacy news closely (Pew Research Center 2019). However, the responses to the questions surrounding privacy policies indicate that marginalized subpopulations in the United States are more proactive in efforts to protect their personal information, being more likely to read privacy policies than their counterparts, as well as being more likely to read them thoroughly.

A possible explanation for this contradiction is that United States adults who belong to a privileged demographic subgroup are less concerned about the possible consequences of agreeing to a privacy policy or terms of service agreement without reading it beforehand. On the other hand, groups that have historically experienced negative prejudice in the legal space may feel a greater responsibility to protect themselves from abusive and exploitative practices. Underprivileged Americans are victims of privacy-related crime more often than members of more privileged demographic subgroups, despite feeling less concerned about the security of their personal information. Black adults are around three times as likely as white adults to experience social media or email breaches. Black adults are also more likely to experience identity theft, with 12 percent of adults saying that someone attempted to use their name to open a line of credit or apply for a loan in the last year (Pew Research Center 2019).

When it comes to data privacy, Americans belonging to underprivileged racial subgroups tend to be more concerned about being exploited by the government than by corporations. Only 56 percent of white adults report being concerned about how much information law enforcement may know about them, compared to 67 percent of Hispanic adults and 73 percent of black adults. In addition, 60 percent of black adults and 56 percent of Hispanic adults believe the government is tracking their online activities, compared to only 43 percent of white adults. Greater skepticism of government activities may explain why Americans belonging to underprivileged demographic groups tend to read privacy policies more thoroughly before agreeing to them, as members of these groups perceive the government as an even greater threat than private companies rather than a line of defense against their exploitative practices.

7 Conclusions

Having determined that the terms of service and privacy policy agreements of Big Tech companies are unreadable to the average consumer, it is clear that stricter regulation is necessary to close the legal loopholes that allow major corporations to exploit their customers. However, despite a clear need for revised legislation, a lack of consumer awareness diminishes the efficacy of existing protections. Returning again to American

Trends Panel survey findings, 63 percent of respondents reported having very little or no understanding of the laws and regulations currently in place to protect their data privacy (Pew Research Center 2019). This challenges the argument that government intervention is the most effective way to prevent Big Tech companies from manipulating their users with unreadable contracts, if the policies meant to protect consumers are just as unreadable. The demographic disparities in responses to the American Trends Panel, which indicate that marginalized populations in the United States feel a need to greater lengths to protect themselves from abuse and exploitation from major corporations, signal that unreadable privacy policies and terms of service agreements are an issue of equity that goes much deeper than companies versus their customers.

8 References

Auxier, Brooke, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar AND Erica Turner. 2019. "4. Americans' attitudes and experiences with privacy policies and laws." *Pew Research Center*. Retrieved May 9, 2023 (https://www.pewresearch.org/internet/2019/11/15/americans-attitudes-and-experiences-with-privacy-policies-and-laws/).

Benoliel, Uri and Samuel Becher. 2019. "The Duty to Read the Unreadable." *Boston College Law Review* 2255. Retrieved May 9, 2023 (hp://dx.doi.org/10.2139/ssrn.3313837).

Böhme, Rainer and Stefan Köpsell. 2010. "Trained to accept?: a field experiment on consent dialogs." *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 2403-6. Retrieved May 9, 2023 (hps://doi.org/10.1145/1753326.1753689).

Frier, Sarah. 2019. "Facebook Paid Contractors to Transcribe Users' Audio Chats." *Bloomberg*. Retrieved May 9, 2023 (https://www.bloomberg.com/news/articles/2019-08-13/facebook-paid-hundreds-of-contractors-to-transcribe-users-audio?leadSource=uverify%20wall&in_source=embedded-checkout-banner).

Levy, Bob. 2021. "Big Tech's head-scratching terms of service agreements need simple, specific government regulation." *Business Insider*. Retrieved May 9, 2023 (https://www.businessinsider.com/free-market-government-regulation-big-tech-terms-of-service-agreements-2021-3).

"Licensed Application End User License Agreement." *Apple*. Retrieved May 9, 2023 (https://www.apple.com/legal/internet-services/itunes/dev/stdeula/).

McDonald, Aleecia M. and Lorrie Faith Cranor. 2008. "The Cost of Reading Privacy Policies." *I/S: A Journal of Law and Policy for the Information Society* 4(3):543-68. Retrieved May 9, 2023 (hp://hdl.handle.net/1811/72839).

Obar, Jonathan A. and Anne Oeldorf-Hirsch. 2018. "The Biggest Lie on the Internet: Ignoring the Privacy Policies of Social Networking Services." *Information, Communication, & Society* 1-20. Retrieved May 9, 2023 (hp://dx.doi.org/10.2139/ssrn.2757465).

Volokh, Eugene. 2016. "Amazon Web Services and the zombie apocalypse." *Business Insider*. Retrieved May 9, 2023 (https://www.businessinsider.com/amazon-zombie-clause-2016-2).

2013. "A case for reading the small print." *BBC*. Retrieved May 9, 2023 (https://www.bbc.com/news/blogs-magazine-monitor-24992518).