Trends in Data Exfiltration and Privacy Policies

Alexander Huang
alexh95@mit.edu

Alex Kimn
akimn@mit.edu

Jesse Widner
jwidner@mit.edu

Diane Zhou
dianez@mit.edu

May 13, 2020

Contents

1 Introduction

Recently, it has come to light that a tablet driver released by the graphical tablet manufacturer Wacom tracks the name of every application window opened by the user[1]. Within the privacy agreement the user sees during the driver installation process, it is mentioned that Wacom sends "aggregate usage data, technical session information, and information about [the user's] hardware device" to Google Analytics. However, the agreement fails to state exactly what is being sent and why, using only vague and uninformative language to mask a blatant example of user data exfiltration.

Unfortunately, this form of data exfiltration is now a common occurrence. Under the (often correct) assumption that users will not thoroughly read privacy agreements, well-known and trusted companies like Wacom can obtain sensitive user information without the explicit knowledge or consent of the user. These data can then be sold to data companies or other third parties for a profit, implying that unbeknownst to the user, their information could be made available to a wide range of unknown actors. Moreover, any security vulnerabilities present during the transfer of this exfiltrated data present opportunities for potentially malicious actors to obtain private user information. In this way, seemingly trustworthy actors can breach a user's privacy and data security by exploiting their naivete. Therefore, it is critically important for software users to both: (i) know if a downloaded piece of software could potentially expose them to data exfiltration and (ii) understand what the consequences of such data exfiltration are. Thus, in this work, we analyze the privacy policies of various videoconferencing services to determine potential data privacy vulnerabilities and avenues for data exfiltration. We then evaluate the privacy policies according to three metrics: specificity, responsibility, and readability to determine how well each service accounts for these data privacy issues. Finally, we provide several recommendations for users wishing to protect their data privacy.

[1]https://www.theverge.com/2020/2/6/21126245/wacom-tablet-app-tracking-google-analytics

2 Background

The historic lack of standardization in software privacy policies has conditioned most users to automatically click "yes" and agree to everything in order to access the software. A study performed in 2017 showed that 74% (N=543) of individuals did not choose to read the privacy policy when presented with the option but 97% agreed to it. Even when presented with a mandatory terms of service (TOS) policy, which should take roughly 15 minutes to read at a typical adult reading rate of 250-280 words per minute, users spent an average of 51 seconds reading it. Those who declined the TOS spent 90 seconds longer reading it[1].

This study elucidates an important point: most users do not read privacy policy and terms of service with enough time and given the obfuscated nature of data exfiltration, users are largely unaware their data is being taken from them and do not know how that data will be used.

Furthermore, privacy policies contain technical jargon that is difficult for most average users to grasp, making readability low. Even if users are proactive in reading the policies, there is a good chance they will not understand it. The readability of privacy policies is important since users tend to trust them more having understood them [2]. We argue that privacy policies serve as a contract between users and service providers. Therefore, it is important to ensure that all parties, particularly the user, are fully aware of the implications of agreeing to the contract. We will assess the readability of all privacy policies examined in this project.

2.1 Companies

Given the recent increased usage and increased scrutiny of videoconferencing platforms due to the ongoing COVID-19 pandemic, we focus on five videoconferencing platforms for our analysis: Cisco Webex, Discord, Google Hangouts, Skype, and Zoom. We provide a brief introduction and description of each of the videoconferencing services studied in the following sections.

2.1.1 Cisco Webex

Cisco Webex (originally Webex) is a videoconferencing and collaborative platform wholly owned by the technology conglomerate Cisco Systems. The primary videoconferencing service offered within the Webex platform is Webex Meetings. Thus, we primarily analyze the privacy policy of Webex Meetings [3] specifically, in addition to the general policies outlined in Cisco Systems' online

privacy policy [4] where the latter is applicable.

2.1.2 Discord

Discord is a VoIP and general communication application that features text, image, video, and audio communication over chat channels on a digital distribution platform similar to a centralized version of IRC designed primarily for gamers, but also other communities including education. We analyze the company's privacy policy [5] as it applies to the platform and software distributed by Discord, Inc.

2.1.3 Google Hangouts

Google Hangouts is a communications software developed by Google that supports text chats, audio chats, and video conferencing. It was originally just a feature of Google+, a discontinued social media platform, but has since become a standalone service. Hangouts does not have its own privacy policy and instead references the general Google privacy policy that is shared across many Google services.

2.1.4 Skype

Skype is a telecommunications application that allows for video chatting, voice calling, and instant messaging. Microsoft acquired Skype in 2011, so Microsoft's privacy policy now applies to Skype [6]. Skype can be used between various devices such as computers, tablets, mobile devices, the Xbox One console, smartwatches, and Amazon Echo. Specialized Skype products include Microsoft Teams (originally Skype for Business) and Skype for content creators, but we will focus on the general Skype product available for free to all consumers.

2.1.5 Zoom

Zoom is a cloud-based peer-to-peer platform for videoconferencing and online chatting developed Zoom Video Communications. The service was originally launched in 2012 by a former executive at Cisco Webex, and has since grown to become one of the most popular videoconferencing platforms. We note that due to increased media scrutiny surrounding its data privacy policies, Zoom made

significant updates to its privacy agreement on March 29th, 2020. As such, though we analyze both Zoom's current privacy agreement [7] in comparison to the other videoconferencing services, we also compare its previous privacy agreement [8], dated to December 31st, 2019 in a separate case study in Chapter 5.

3 Methodology

In analyzing the privacy policies of each company, we consider a few metrics as described in the next section. We find it important to distinguish between language and substance, so we include metrics for both. Language reflects how easy it is for a consumer to understand the privacy policy, while substance reflects what companies can actually be held accountable for.

3.1 Metrics

3.1.1 Vagueness/Specificity

The privacy policies that we have analyzed vary based on notions of specificity and vagueness related to the amount of detail put into the privacy policies by the respective companies. These notions apply to what data the company collects, when it does so, and any guarantees on what it does with such data along with the type of data that is shared with different entities (if any) and what the data is used for. If a company were to list all of this information in full detail, i.e. without some form of catch-all clause, the company would be considered fully specific in how it handles user data. If a company does not list all of this information or includes a catch-all clause in its privacy agreement, then it may be considered more vague depending on the amount of detail in the privacy policy. An example of a fully vague policy would include not disclosing what data is collected or how it is used in any detail.

3.1.2 Responsibility

Companies vary in how much responsibility they claim to take when it comes to privacy and data protection. We rate the responsibility of a company from full responsibility, i.e. the company proactively makes sure data is only collected with consent and used for certain functions, to no responsibility, i.e. it is entirely up to the user to consider what data they are sharing when they

take certain actions and should have no expectations about what that data might be used for. In the case that a privacy policy does not explicitly state what the company takes responsibility for, we assume the company to fall on the "no responsibility" side of the spectrum.

3.1.3 Readability

Readability is another metric we will use in our analysis and evaluation of each company's privacy policies. Quantifying readability is a concurrent research problem in its own right with many viable methods for producing a readability score, but we used the Flesch–Kincaid readability test in this project. The Flesh-Kincaid test takes input text and outputs a score which corresponds to the schooling level required to read the text as shown in table 1. The score S can be computed with the following formula:

$$S = 206.835 - 1.015\left(\frac{\text{total words}}{\text{total sentences}}\right) - 84.6\left(\frac{\text{total syllables}}{\text{total words}}\right)$$

We recognize that this test is limited in that it only takes into account syntactical challenges in reading (e.g. number of words, syllable of words) but not more subjective metrics such as the usage of jargon or the perceived complexity of sentences. There exists an alternative test known as the cloze test, which measures how well a test subject is able to comprehend a text by having them fill in a copy of the reading with certain words blanked out [9]. This method could be used in conjunction with the Flesh-Kincaid test, but since the cloze test requires coordination with test subjects, we did not perform it.

Score	**School Level**
> 90	5th grade
$90 - 80$	6th grade
$80 - 70$	7th grade
$70 - 60$	8th to 9th grade
$60 - 50$	10th to 12th grade
$50 - 30$	College
< 30	College graduate

Table 1: Flesch-Kincaid score and relationship to schooling required [10]

4 Results: General Trends

We analyze four specific data vulnerabilities in which data exfiltration can occur: personal data collection, third party data sharing, personal data retention, and audio/video recording.

4.1 Personal Data Collection

Companies collect user generated content and usage statistics across their services. We examine what data is considered personal data across the different services. The relevant policy language is listed in Table 1.

Across most services, the practice of using umbrella clauses, what we define as broad statements that can be interpreted in more than one way, to justify collecting essentially all user generated data is common. For example, Google states that it can collect any content users create, including emails. This is especially concerning since sensitive information could be compromised if there is a malicious employee with enough access privilege. These umbrella clauses also fail to specify granularity. It is unclear whether the whole email in its entirety is collected or just specific parts such as the to and from fields.

Another troubling aspect that is common to all these policies is the fact that they do not mention whether the personal data collected would be anonymous or whether this is an option. This is especially concerning since it would be relatively easy for a malicious employee or attacker who has compromised the company's security measure to determine a user's identity.

Overall, personal data collection in the policies we have examined is fairly vague due to the usage of umbrella clauses and lack of granularity in specifying what data is considered personal data. There is also very little responsibility taken by the companies as there is no channel specified for handling disputes if the user were to discover their personal data has been leaked. However, we would like to recognize Zoom for their clear outline of exactly what data they collect, a number of examples that average users can understand, and reasons for why they collect the data.

4.2 Third Party Data Sharing

One of the more lucrative enterprises for technology companies today is the selling and sharing of valuable user-related data with third-parties. This third party data sharing presents a significant

Company	Privacy Statement
Cisco Webex	"'Personal information' is any information that can be used to identify an individual, and may include name, address, email address, phone number, login information (account number, password), marketing preferences, social media account information, or payment card number. If we link other data with your personal information, we will treat that linked data as personal information. We also collect personal information from trusted third-party sources and engage third parties to collect personal information to assist us
Discord	"We collect information from you when you voluntarily provide such information, such as when you register for access to the Services or use certain Services. Information we collect may include but not be limited to username, email address, and any messages, images, transient VOIP data (to enable communication delivery only) or other content you send via the chat feature."
Google Hangouts	"We also collect the content you create, upload, or receive from others when using our services. This includes things like email you write and receive, photos and videos you save, docs and spreadsheets you create, and comments you make on YouTube videos."
Skype	"Microsoft collects data from you, through our interactions with you and through our products. You provide some of this data directly, and we get some of it by collecting data about your interactions, use, and experiences with our products. The data we collect depends on the context of your interactions with Microsoft and the choices you make, including your privacy settings and the products and features you use. We also obtain data about you from third parties."
Zoom	"Technical information about your devices, network, and internet connection...Approximate Location...Information about how you use Zoom (this is NOT information or content you share in your meetings or in chats)...Setting and preferences chosen by the user...Metadata"

Table 2: Sections of the studied companies' privacy policies pertaining to the collection of personal data.

data vulnerability, as unknown to the user, their data can be shared with unknown and potentially malicious actors.

We examine the circumstances and conditions under which each company claims to share personal information to third parties. The sections of the companies' privacy policies relevant to third party data sharing are given below in Table 3

Company	Privacy Statement
Cisco Webex	"We share Registration Information, Host Information, and/or Usage Information with service providers, contractors or authorized third parties to assist in providing and improving the Service."
Discord	"Advertising platforms, which include Twitter and Facebook (and whose SDKs are integrated within our Service), may collect information for optimizing advertising campaigns outside of the Service."
Google Hangouts	"We do not share your personal information with companies, organizations, or individuals outside of Google except in the following cases: with your consent..., with domain administrators..., for external processing..., for legal reasons..."
Skype	"If you use Skype through a company other than Microsoft, that company's privacy policy governs how it handles your data... We may access, transfer, disclose, and preserve your data."
Zoom	"...they [third-party service providers] may have access to Personal Data related to the specific activity they are doing for us in the process"

Table 3: Sections of the studied companies' privacy policies pertaining to the sharing of user information with third parties.

In general, the services are fairly explicit as to exactly when data sharing occurs. All five services enumerate a list of specific situations during which data sharing regularly occurs.[2]. Google's list is by far the most comprehensive, and also details the type of information shared in some of the scenarios.

However, one concerning observation is the lack of any guarantees from any of the services that a user's information will not be shared with third parties. Essentially, all five services tacitly acknowledge that regardless of the knowledge or consent of the user, at least some user information is always shared with third parties.

A common strategy used to accomplish this is the idea of opt-out, rather than opt-in, consent. Under this idea, assent to the privacy agreement itself is equivalent to the user's consent to the release of their data to third parties as stipulated in the policy. To avoid this, the user must make additional effort to then opt-out of some of the data releases. For example, Cisco Webex asks users to: "...To opt-out of Cisco sharing with third parties for their marketing purposes, please send an email to privacy@cisco.com." Discord and Skype list almost identical opt-out policies for marketing communications.

[2]Note that Google frames this as a list of exceptions with the default being data not being shared.

This places significant onus on the user to be informed and aware about which third parties their personal information is being shared with and where that data ends up. However, this is made even more difficult that many of the services, in particular Skype and Zoom, appear to transfer the responsibility of handling the shared data to the third parties themselves. Discord and Google go a step further and give third parties agency to collect information on their services. This makes it almost impossible to track who has access to a user's personal information.

Thus, overall, while third party data sharing is generally defined in detail by the five services, there is very little responsibility taken by the companies for the privacy of the shared information.

4.3 Personal Data Retention

Companies often retain personal user information long after it is originally obtained. This presents a potential privacy vulnerability, as the users typically do not have access control over service databases.

We analyze the reasons the five services use to justify personal data retention as well, the contents of retained information, and the protocols each service has in place to dispose of personal information once it is no longer needed. The relevant excerpts from the services' privacy policies are given below in Figure 4.

It is immediately apparent from the excerpts that the language that all of the services use to describe their data retention policies is exceptionally vague. In particular, Cisco Webex, Discord, and Zoom frame the basis for their data retention policies as the privacy policy itself, creating circular logic equivalent to the policy "personal data is stored for the reasons it is stored". Discord and Zoom elaborate no further on this and are therefore fully vague with respect to the justification for personal data retention. Google and Cisco further reference that data is stored for their personal interests but otherwise provides no specific justification. Microsoft at least includes a reference to the user's interests in it's justification; however, even this is fairly vague as it makes no reference to specific user activities.

Moreover, none of the services are very clear on the exact contents of the retained information. Both Skype and Google are fairly vague as they do provide a few indirect examples. The two services imply that any uploaded content is retained, while Skype further mentions that personal information is retained for shorter a duration than anonymized data. The remaining three services are fully vague as they do not elaborate on the data contents whatever.

By contrast, most of the companies provide specific guidance as to how users can delete their

Company	Privacy Statement
Cisco Webex	"We will retain your personal information as needed to fulfill the purposes for which it was collected. We will retain and use your personal information as necessary to comply with our business requirements, legal obligations, resolve disputes, protect our assets, and enforce our agreements."
Discord	"We generally retain personal data for so long as it may be relevant to the purposes identified herein. To dispose of personal data, we may anonymize it, delete it or take other appropriate steps. Data may persist in copies made for backup and business continuity purposes for additional time."
Google Hangouts	"We retain the data we collect for different periods of time depending on what it is, how we use it, and how you configure your settings...Some data you can delete whenever you like, such as the content you create or upload...We keep some data until you delete your Google Account, such as information about how often you use our services...And some data we retain for longer periods of time when necessary for legitimate business or legal purposes..."
Skype	"Microsoft retains personal data for as long as necessary to provide the products and fulfill the transactions you have requested, or for other legitimate purposes such as complying with our legal obligations, resolving disputes, and enforcing our agreements. Because these needs can vary ... actual retention periods can vary significantly."
Zoom	"We will retain personal data collected for as long as required to do what we say we will in this policy, unless a longer retention period is required by law. Customers can delete their own content."

Table 4: Sections of the studied companies' privacy policies pertaining to user data retention.

stored personal data. For Skype, users are directed do delete their information using the Microsoft privacy dashboard, while Google directs its users to go to their Google accounts and either delete specific Google products or remove uploaded content. Both of these companies also state the speed at which these data deletion requests are applied and under what specific circumstances they may be rejected. By contrast, Discord and Cisco place additional responsibility on the user by asking users to directly contact them regarding specific requests for data disposal. Moreover, both of these companies only provide a fairly vague set of conditions under which these requests may be denied. However, even this is far better than the data disposal options given by Zoom. Outside of simple content deletion, the only way that Zoom appears to allow users to delete any personal information is through account deletion.

Overall, data retention is only vaguely justified and defined in the analyzed privacy policies.

However, there is a significant amount of responsibility taken by the services with regards to personal data disposal. All of the services except for Zoom provide a specific means by which users can control the personal information stored by the services.

4.4 Audio/Video Recording

We particularly consider audio and video recording because the products whose privacy policies we are analyzing specialize in video chats and audio calls. The ability to record video and audio presents additional potential privacy and security vulnerabilities as multiple parties, including some who may not even have an account with the company, are involved.

There are generally two models of video and audio communication:

- Host and guests - The host initiates the video or audio conference session. Guests can join the session. The host has privileges that generally include who to allow as guests of the session and whether to record the session.
- Group chat / call - A predetermined set of users all join a session of video chatting or audio calling.

Cisco Webex, Skype, and Zoom fall under the host and guests model, while Discord and Google Hangouts fall under the group chat / call model. Discord and Google Hangouts do not provide the option of recording sessions, so we focus on Cisco Webex, Skype, and Zoom in this section. Excerpts of the privacy policies related to recording are shown in Table 5.

For all three products, the host has full control and responsibility over the recording of a meeting. Companies take no responsibility in ensuring that guests are notified about recording, guests who are part of the recording have given consent, and that the recording and its distribution are legal. Zoom is the only company that mentions that the product itself will usually issue a notification about a session being recorded.

The language is very specific across all three companies with respect to the expectation that the host is fully responsible for all logistics and enforcement of policies related to recordings. However, it is not always clear what recording data the companies might collect and retain, as well as how they may use the data. Cisco Webex does not address the issue of what data Cisco may collect or retain with respect to recordings at all, so it is fully vague. Skype mentions that the recording will be stored and shared as part of the conversation history, which implies that Microsoft retains the

Company	Privacy Statement
Cisco Webex	"The meeting host will control any meeting recordings or files shared during the meeting, which will be subject to the host's corporate policies regarding access, use, monitoring, deletion, preservation, and export of information."
Skype	"Some versions of Skype have a recording feature that allows you to capture and share all or part of your audio / video call. The recording will be stored and shared as part of your conversation history with the person or group with whom the call occurred. You should understand your legal responsibilities before recording any communication. This includes whether you need to get consent from all parties to the communication in advance. Microsoft is not responsible for how you use your recordings or the recording features."
Zoom	"If you participate in a Recorded Meeting or you subscribe to Zoom cloud recording services, we collect information from you in connection with and through such Recordings. This information may include Personal Data. Meeting hosts are responsible for notifying you if they are recording a meeting, and you will generally hear a notice or see an on-screen notification when recording is in progress."

Table 5: Sections of the studied companies' privacy policies pertaining to video/audio recording.

data indefinitely. Zoom explicitly mentions that they may collect personal data and any other data in connection with recordings, but it is unclear what they do with such data.

Overall, there is much emphasis on the responsibility of recording legally and with consent falling on the meeting host, and not much onus on the side of the companies to build in features to help guarantee consent and legality. Companies also make no guarantees about what happens to the recording data, but based on the previous sections we would assume that the companies treat recording data similar to other data that they may collect. Unfortunately, companies also do not address the security issue of unwarranted snooping on sessions that may lead to recordings being made without the knowledge of anyone else in the session.

4.5 Overall Readability

We computed the Flesch-Kincaid score of each privacy policy as shown in Table 6. Our subjective evaluation agrees with the Flesch-Kincaid score in terms of readability. We note that most privacy policies require a college reading level. The most readable privacy policy is Zoom's current policy, only requiring a 9th grade reading level. However, we also note that Zoom's 2019 policy was actually

much less readable than the current version. This is an interesting case study in and of itself since the rewrite to be more readable may have been a result of negative press coverage over security and privacy concerns with Zoom. At the time of writing, Zoom's user base has surged to over 300 million due to an unprecedented shift to remote work due to the COVID-19 pandemic [11]. Given the scrutiny, simpler language that users can understand and trust may have been the goal for Zoom.

Company	Score	Grade Level
Cisco Webex	40.9	College
Discord	49.4	Early College
Google Hangouts	52.6	12th grade
Skype	44.2	College
Zoom 2020	60.4	9th grade
Zoom 2019	40.2	College

Table 6: Flesch-Kincaid score of each company's privacy policy

5 Results: Case Studies

5.1 Zoom 2019 vs. Zoom 2020

We consider differences in the Zoom privacy policy as of December 31, 2019 [8] and the Zoom privacy policy as of March 29, 2020 [7]. The updated Zoom privacy policy has become significantly more readable according to our metrics, requiring only a 9th grade reading level as opposed to the college grade reading level that it possessed in 2019. One subjective advantage in readability that the new privacy policy has is the addition of a tabular format that is used to list all of the types of data that are collected along with what the data is used for.

Arguably the privacy policy has become more vague over time. One major aspect that has made it more vague is the use of the word "sell". In the old privacy policy, Zoom was up front with the fact that they were using a definition of sell that may or may not line up with a layperson's meaning of the word stating, "Depends what you mean by 'sell.'" [8]. In the new privacy policy, Zoom states outright that it does not sell your data. However, Zoom's practices have not changed. To illustrate this point, we note that in the new privacy policy, Zoom says, "It is only with the recent developments in data privacy laws that such activities may fall within the definition of a

'sale'" [7], and this comes in the same section that Zoom says it does not sell your data. Overall, both versions of the privacy policy seem to do a comparable and reasonably good job of listing the types of data that are collected. The reasons for collection are either stated explicitly or lumped under the umbrella of improving user experience.

When it comes to responsibility, the new Zoom privacy policy has more guarantees than the old one. These guarantees are mainly related to the contents of Zoom meetings themselves. In particular, the new Zoom privacy policy guarantees that the company does not monitor or collect data from Zoom meetings, although it still collects metadata about such meetings. One new privacy feature that Zoom added and mentions in the new privacy policy is that the user is notified if the host of a meeting decides to record the meeting. In the old version of Zoom (and its privacy policy), the responsibility fell entirely on the host of the meeting to notify participants that the meeting was being recorded.

5.2 Jitsi Meet

Finally, we consider Jitsi Meet, a videoconferencing option that differs from the five products we previously analyzed in that it is an open source project. Jitsi touts Jitsi Meet as "fully encrypted" and anonymous due to the fact that you do not need an account to join a meeting [12].

Jitsi Meet's privacy policy wins by far in our three metrics of specificity, responsibility, and readability over the other privacy policies we analyzed. The caveat is that there may be multiple implementations and instances of Jitsi Meet since it is an open source project. We did not analyze any actual code and implementation to see whether Jitsi Meet can follow through on the guarantees stated in its privacy policy, but given that it is an open source project we assume that there are developers who are actively scrutinizing the product for security loopholes. One such example is that Jitsi Meet is not yet truly "fully encrypted" as there is no end-to-end encryption, but work is underway to make this possible. A demo version of Jitsi Meet on which end-to-end built was released on April 22, 2020 [13].

Jitsi Meet is hosted by a company called 8×8, so its privacy policy applies to Jitsi. 8×8 claims that personal data related to a meeting is only temporarily stored for the duration of the meeting and never retained afterward. The company also explicitly states that it is "not in the business of selling personal information to third parties" and lists what they do use the data for, i.e. delivering the service, identifying and troubleshooting problems the service, improving the service, and investigating fraud or abuse [14]. Jitsi Meet's privacy policy has a section dedicated to addressing what data Jitsi uses for analytics, which supports the statement that there is no selling,

directly or indirectly, of personal information. Jitsi promises that even if participants provide their names, e-mail addresses, or profile pictures, the data is not retained beyond the duration of the meeting. Examples of data that might be used for analytics include "an anonymous identifier, bitrate, available bandwidth, SDP offers and answers, product utilization events, mobile app crash dumps" [13].

Jitsi also explicitly addresses what happens with recordings: "They are kept on our servers until we can upload them to the place you indicated (currently Dropbox). If we haven't managed to do that in 24 hours we still delete them and they are gone forever (so make sure you have enough space in your Dropbox folders)" [13].

Thus, in all four data vulnerability categories of personal data collection, third party data sharing, personal data retention, and audio/video recording, Jitsi Meet scores well on our metrics of specificity and responsibility. We also computed the readability of Jitsi Meet's privacy policy to be 63.9, which is a 9th grade level, on par with the readability of Zoom's privacy policy in 2020 and much higher in readability than all of the other policies.

If Jitsi Meet is superior in its privacy and security guarantees, why is it not more popular? Jitsi Meet does not allow for the host and guests model and is more similar to a group chat / call model, with anyone being able to mute or remove anyone else. In order to enable a host with privileges, a custom version must be constructed and deployed. There may also be more bugs and issues with scalability.

6 Conclusion

Perhaps not surprisingly, we conclude that companies generally have privacy policies with vague language and difficult readability, and that they avoid responsibility in guaranteeing anything about what happens to personal data collection, retention, and sharing. Two companies that do stand out positively for more specificity and easier readability in their privacy policies are Zoom and Jitsi. In addition, Jitsi claims much more responsibility as well. These are two interesting cases as Zoom has been under particular scrutiny due to its increased usage during the COVID-19 pandemic and Jitsi Meet is an open source project. We gather that when more users are able to have more input into a product's implementation, whether via increased scrutiny or open sourcing the software, the product's privacy and security guarantees become higher and more transparent. Further work on more generally comparing whether open source projects fare better than closed source products in privacy and security guarantees would be interesting.

References

[1] Jonathan A Obar and Anne Oeldorf-Hirsch. "The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services". In: *Information, Communication & Society* 23.1 (2020), pp. 128–147.

[2] Tatiana Ermakova et al. "Privacy policies and users' trust: does readability matter?" In: (2014).

[3] Cisco. *Cisco Webex Meetings Privacy Data Sheet; Version 4.1, May 2020*. https://trustportal.cisco.com/c/dam/r/ctp/docs/privacydatasheet/collaboration/cisco-webex-meetings-privacy-data-sheet.pdf. 2020.

[4] Cisco. *Cisco Online Privacy Statement; 23 December 2019 Revision*. https://www.cisco.com/c/en/us/about/legal/privacy-full.html. 2019.

[5] Discord, Inc. *Discord Privacy Policy*. https://discord.com/privacy. 2020.

[6] Microsoft. *Microsoft Privacy Statement*. https://privacy.microsoft.com/en-us/privacystatement. 2020.

[7] Zoom. *Privacy Policy*. https://zoom.us/privacy. 2020.

[8] Zoom. *Privacy Policy*. https://web.archive.org/web/20200119034606/https://zoom.us/privacy. 2020.

[9] Ravi Inder Singh, Manasa Sumeeth, and James Miller. "A user-centric evaluation of the readability of privacy policies in popular web sites". In: *Information Systems Frontiers* 13.4 (2011), pp. 501–514.

[10] Rudolf Flesch. *How to Write Plain English*. URL: https://web.archive.org/web/20160712094308/http://www.mang.canterbury.ac.nz/writing_guide/writing/flesch.shtml.

[11] Nico Grant. *Zoom Daily Users Surge to 300 Million Despite Privacy Woes*. Apr. 2020. URL: https://www.bloomberg.com/news/articles/2020-04-22/zoom-daily-users-surge-to-300-million-despite-privacy-woes.

[12] Jitsi. *Jitsi Meet*. https://jitsi.org/jitsi-meet/. 2020.

[13] Jitsi. *Jitsi Meet Security Privacy*. https://jitsi.org/security/. 2020.

[14] Jitsi. *meet.jit.si Privacy Supplement*. `https://jitsi.org/meet-jit-si-privacy/`. 2020.