Home is Safer than the Cloud! Privacy Concerns for Consumer Cloud Storage

Iulia Ion*, Niharika Sachdeva†, Ponnurangam Kumaraguru†, Srdjan Čapkun*
* ETH Zurich, † IIIT-Delhi
iulia.ion@inf.ethz.ch, {niharika1012, pk}@iiitd.ac.in, capkuns@inf.ethz.ch

ABSTRACT

Several studies ranked security and privacy to be major areas of concern and impediments of cloud adoption for companies, but none have looked into end-users' attitudes and practices. Not much is known about consumers' privacy beliefs and expectations for cloud storage, such as webmail, document and photo sharing platforms, or about users' awareness of contractual terms and conditions. We conducted 36 in-depth interviews in Switzerland and India (two countries with different privacy perceptions and expectations); and followed up with an online survey with 402 participants in both countries. We study users' privacy attitudes and beliefs regarding their use of cloud storage systems. Our results show that privacy requirements for consumer cloud storage differ from those of companies. Users are less concerned about some issues, such as guaranteed deletion of data, country of storage and storage outsourcing, but are uncertain about using cloud storage. Our results further show that end-users consider the Internet intrinsically insecure and prefer local storage for sensitive data over cloud storage. However, users desire better security and are ready to pay for services that provide strong privacy guarantees. Participants had misconceptions about the rights and guarantees their cloud storage providers offers. For example, users believed that their provider is liable in case of data loss, does not have the right to view and modify user data, and cannot disable user accounts. Finally, our results show that cultural differences greatly influence user attitudes and beliefs, such as their willingness to store sensitive data in the cloud and their acceptance that law enforcement agencies monitor user accounts. We believe that these observations can help in improving users privacy in cloud storage systems.

General Terms

Human Factors, Security, Privacy.

Keywords

Cloud Storage, Social Factors, Cross-Cultural, Usability.

Symposium on Usable Privacy and Security (SOUPS) 2011, July 14–16, 2011, Pittsburgh, PA, USA.

1. INTRODUCTION

Based on a recent survey by Pew Research Center, experts predict that, in the next decade, cloud computing will become more dominant for end-users than desktop computing [4]. A 2011 survey by Hosting concludes that cloud storage drives the growth of cloud computing [3]. Data is moving from user-owned desktops and laptops to dedicated online storage systems, e.g., Dropbox [9] and Google Docs [17]. This change toward cloud storage brings a number of significant benefits, such as continuous availability of data anytime, anywhere, easy sharing of picture, and documents with friends and family, and it relieves the burden of self-managing replication and data backups. By 2008, 69% of all Internet users had either stored data online or had used a web-based software application [21]. In this paper, we focus on cloud storage systems intended for private users, also known as *consumer cloud storage* systems [22].

Cloud storage poses novel security and privacy threats, which may slow down or impede its adoption. Security and privacy analysis so far has mostly focused on enterprise cloud adoption [6, 7, 15, 18]. However, clouds equally impact end-users' privacy and expose users private documents to hackers (e.g., 2009 Google cyber attack [16], bugs in access control enforcement systems [43]), or to governments [39]. While companies and governments may be able to afford to hire trained security consultants, end-users lack the necessary resources and security education to investigate the data practices of cloud storage providers. The data confidentiality, integrity, and availability risks are partly reflected by the Terms of Service (ToS) and privacy policies of consumer cloud storage companies. It is common practice for free consumer cloud storage services not to offer any service guarantees, to assume no liability for any data loss, and to reserve the right to disable accounts without reason or prior notification, as well as to change or stop providing the service at any time. Given that users don't usually read the terms of service and privacy policies, it is unclear how many users are actually aware of these conditions. Cloud reliability questions were raised when 150,000 Gmail users and 17,000 Hotmail users found decades of personal email and documents deleted from their accounts [2].

Understanding users' expectation of privacy is essential in devising appropriate laws and regulations. Governments have repeatedly demanded that companies install backdoors in security solutions and build local servers to facilitate surveillance [26, 39]. Unlike in the case of local storage, for data stored in the cloud, users do not typically know when their data is being accessed by other parties. For example, the

notice requirement for stored communications in the US is satisfied by notifying only the storage provider, not the user, of government access [39]. The issues of surveillance and notice requirement have only recently received media attention, when Twitter disclosed the U.S. government subpoena to turn over user data, including IP addresses, for a number of people connected with Wikileaks [37]. Privacy activists argue that consumers expect privacy in the cloud [19], while law enforcement agencies in United States, to which most cloud storage providers are subject, stipulate that "a person has no legitimate expectation of privacy in information he voluntarily turns over to third parties" [38].

In this paper, we analyze users' expectations of privacy in the cloud and their awareness of the terms of service agreement with cloud storage provider. We investigate how practices and concerns towards cloud storage differ from those of local storage. Through 36 interview studies with users of cloud storage systems, such as Dropbox, Google Docs or webmail, we gathered qualitative data and elicited users' current attitudes and concerns related to the security and privacy of their cloud-stored data. We tested the conclusions derived from the interviews with 402 participants in an online survey. We formulate the central research question as follows: (1) *What do users think about the security and reliability of cloud storage?* (2) *What are users' privacy concerns in cloud storage?* (3) *How do privacy concerns influence users' behavior?* and (4) *How do privacy concerns differ among cultures?*

Privacy studies so far have been mostly focused on the U.S. Instead, we compare privacy attitudes toward cloud storage in India and Switzerland, two countries with substantial cultural differences, as shown by Hofstede's cultural dimensions [13]. Switzerland has an individualistic society and India a collectivist one.[1] Indians accept that power and control in society are distributed unequally, whereas Swiss expect an equal distribution.[2] The Swiss Federal Constitution guarantees the right to privacy, but the Constitution of India does not explicitly recognize it. While in Switzerland, privacy is regulated through the Swiss Federal Data Protection Act, established in 1992 and amended in 2008 [11, 34]; in India, there is no general data protection law [30]. However, the Indian government did pass the Information Technology Act (IT Act 2000), amended in 2008 [36]. There have been efforts to introduce a data protection bill in India [25]. We expect the topic of privacy to get more momentum of discussion in India, especially with the introduction of the Unique Identification (UID) [42] and National Intelligence Grid [31] projects.

Our contributions are as follows: (1) We study users attitudes and beliefs with respect to their privacy in cloud storage systems and we observe that, despite security expertise and guarantees provided by storage providers, users still consider local storage safer than the cloud, because they believe that nothing on the Internet is safe. Users would, therefore, rather rely on physically protecting devices storing their digital data. Nevertheless, a strong feeling of security in the cloud emerges from the belief that nobody would be interested in seeing their data, because "I am not important," "not famous," or "not criminal." (2) Our results also show that users believe they have more rights and protection than the contract terms with the cloud storage provider actually grant them. The users are typically unaware of the terms and conditions, and in fact assume higher availability, integrity, ownership guarantees and privacy protection in the cloud than they actually have. Furthermore, when prompted, they agreed to pay for better privacy in their cloud storage account. (3) We study privacy concerns and expectations in populations from two distinctive cultural backgrounds and observe that their cultural differences affect their privacy concerns and expectations in the cloud. We found significant attitude differences between Swiss and Indians: Swiss store less sensitive data in the cloud than Indians do and are more aware of the lack of guarantees. Furthermore, while Swiss consider government monitoring of cloud-stored data a fundamental privacy infringement, Indians regard it as a necessary step in combating terrorism.

This paper is structured as follows: Section 2 gives an overview of previous consumer privacy studies and of cloud computing security concerns, Section 3 describes the methodology and demographics of the interview studies and online survey, Section 4 presents our main findings regarding current user practices, perceived privacy in the cloud, and awareness of terms and conditions. Finally Section 5 presents the conclusions and implications of our study results.

2. RELATED WORK

For companies, security and privacy concerns are the main issues impeding cloud adoption; as a result, major cloud adopting corporations are mostly putting only the less sensitive data in the cloud [7, 8, 35]. Many studies evaluate enterprise security risks and cloud computing adoption [24], and devise security guidelines and best practice recommendations [23], or propose instruments to assess the cloud's security [33] and to insure data [22]. For example, a study by the Data Security Council of India investigated how companies in India deal with security risks when adopting cloud computing [8]. Most of the companies mitigate risks by negotiating legal terms with the cloud provider, to share liability in case of security breaches and unavailability of data. Such risk mitigating approaches are not available in consumer cloud storage.

Several studies have analyzed the terms of usage and conditions laid down by cloud storage providers, as well as relevant national and international data protection laws [39, 41]. However, no study has explored in depth users' understanding and expectations of privacy guarantees for cloud storage. A study by the Pew Research Center surveyed levels of privacy concerns in American Internet users [21]. In the survey, 63% of participants said they would be very concerned if the cloud storage provider retained copies of files which they tried to delete. Forty nine percent of participants said it would be an issue of concern if the provider gave their files to law enforcement agencies when asked. It is not yet clear if, and to what extent, users are aware of such issues, their expectations of privacy and how these concerns would alter their behavior towards online storage services.

Hu et al. [22] evaluated four cloud storage systems – Mozy, Carbonite, Dropbox, and CrashPlan. None of these systems

[1]Switzerland has an Individualism Index (IDV) value of 68 and India of 45. The US value is 91. The European value is 61 and the world average is 43 [14]

[2]India has a very high Power Distance Index (PDI) value of 77 compared to the Switzelrland's 34 and the world average of 56.5. The European average is 45 and the US value is 40.

offered any guarantees for data integrity and availability, nor assumed any liability in case of data security breaches or data loss. Although generally viewed as safe backup solutions, online storage systems are far from the perfect solution users envisage. Hue et al. suggest that special tools are needed to make users *aware* of existent risks and to create *demand* for better data protection and privacy solutions from cloud storage companies.

More focused on the legal issues of data confidentiality, Soghoian [39] makes a detailed analysis of threats to personal data in Web 2.0 technologies. His work emphasizes the legal and technical issues users should be aware of. Currently, inadequate data protection mechanisms expose users to hackers and excessive government prying. Not only do Web 2.0 companies have no incentives to provide better data protection as part of their free services, but their business models rely on getting large amounts of private information which can potentially be used for targeted advertisements. Soghoian argues that users are highly unaware of the privacy risks to which they are being exposed, but so far no empirical data has been collected to support this.

A number of studies on Internet privacy attitudes and social networks have been conducted. Westin designed some indices to classify people on "fundamentalist" and "pragmatists," denoting people of high and medium privacy concerns. Only around 20-25% of people are "unconcerned" [44, 28]. Hoofnagle and King [20] investigated Californians' privacy perceptions and expectations in the online world, and found that users do not read privacy policies. Furthermore, they assume that, if a website has a privacy policy, it treats data in a privacy-compliant manner and it does not sell user data to third-parties. In social networks, Acquisti and Gross [1] found that users have misconceptions about the visibility of their profiles on Facebook, and that priming about Facebook's information practices can alter some of its members' behavior.

Most privacy studies are targeted at U.S. consumers. However, there is a need for a global, technical, and legal framework for privacy protection. Furthermore, understanding consumers' privacy behavior and differences in several nations is necessary. Few studies so far have looked at privacy expectations in India and in Europe. Kumaraguru and Cranor [29] showed that Indians exhibit an overall lack of awareness of privacy issues and less concern about privacy than Americans. In a more recent study, Patil et al. [32] compared privacy attitudes of knowledge workers in India and the U.S. While their results confirmed that privacy concerns in India are lower than those in the U.S., in some regards, Indians unexpectedly expressed higher interpersonal privacy concerns compared with their U.S. colleagues. Bellman et al. [5] investigated the role of cultural differences and national regulation in Internet privacy concerns.

To fill the gap in understanding users' perceptions, in this study, we explore users' beliefs about the rights and privacy protection they enjoy in cloud storage, in particular, issues such as the right of the storage provider to disable the account at any time and with any reason, or the lack of guarantees for permanent deletion of data.

3. METHODOLOGY

To explore users' privacy practices and expectations, we conducted 36 semi-structured, in depth interview studies – 16 in Zurich, Switzerland and 20 in Delhi, India. Next, we designed an online survey to confirm our interview conclusions. In this section, we describe the structure and methodology that we used for the interviews and online studies, and present the demographics of participants.

Figure 1: Study session at the home of one of our participants in Delhi. The sessions were audio recorded for future analysis.

3.1 In-depth Interviews

Interview sessions involved one participant at a time and were run by one moderator. They took place in our offices or at the participant's home or office. Figure 1 depicts the setup of a study session. Interviews were mostly conducted in English, but also in German and Hindi. They lasted between 45 and 120 minutes (M=80min, SD=20min). The sessions were audio recorded for future analysis.

Two moderators, one living in Zurich and one in Delhi, were involved in carrying out the interviews. This ensured that the moderator understood the participant's culture, and could later provide explanations for differences in attitudes between Europe, and India. For example, events that had been featured in local press or specific services available in the region were mentioned during the interviews. Interviews in Delhi started once those in Zurich were completed. To ensure consistency of methodology and focus, the Zurich moderator travelled to India and took part in the first seven interviews in Delhi. In the course of these interviews, the Indian moderator's role changed from passive observer to main discussion leader.

We started the discussion byasking the participants about the electronic devices they use and about the types of data they store on them, as well as in the cloud. During the interviews, the moderator never used the term "*cloud*" unless the participant did first (which almost never happened). We asked participants about attachments in their webmail accounts, documents they email to themselves, picture albums on social networking sites, blogs, and personal documents in dedicated storage systems, such as Dropbox and Google Docs. A complete list of interview questions can be found in the Appendix. We then asked participants whether they currently store and where would they store – in the cloud or on their own computer: (1) digital copy of *passport* or other ID documents, (2) *financial files*, (3) *health history information*, and (4) *password list*. To avoid bias, we did not inquire about security and privacy concerns until the

participant opened up the discussion.

We next asked participants what they thought their rights were regarding country of storage, outsourcing data storage, unauthorized modification, guaranteed deletion of data, liability in case of data loss, and account disabling. We showed them a printed slide with three or four variations of statements that appear in the Google, Google Docs or Dropbox privacy policies, and asked which statement they thought was the correct. We tried to understand how much privacy participants thought they had in the cloud, as well as how safe and confidential they considered their data to be from hackers, company employees, police, and government.

The interviews involved collecting data about participants, such as password practices, where they store their sensitive data, attitudes towards police, government surveillance, and practices regarding storage of pirated music and movies. We were not required, neither in Switzerland nor in India, to go through an IRB-type approval process before conducting the interviews. However, authors of this paper have previously been involved in studies with U.S. Institutional Review Board (IRB) approvals, and have applied similar practices in this study. Prior to the interview, participants were shown a printed consent form, which they had to read and sign, if they were comfortable with it. The form stated that an audio recording would be taken, and that collected data would be anonymized and used only for the purpose of this research. Furthermore, participants were informed that they could withdraw from the interview at any point and request the deletion of the audio recording.[3] Table 1 summarizes the demographics of interview participants.

Table 1: Demographics of interview participants.

	Zurich N=16	Delhi N=20
Gender		
Male	7	12
Female	9	8
Age		
<25	8	12
25 - 30	3	3
30 - 39	1	3
40 - 49	4	2
Education		
High School	3	3
Bachelor's	8	7
Master's	5	10
Heard of encryption	6	10
Leave laptop or wallet in the car	3	8
Save credit card info on websites	6	2
Helped fix a computer	7	14
Have created a web page	4	4
Store pictures online	8	20
Use a cloud storage service	8	3

3.2 Online Study

To confirm our interview findings, we posted an online questionnaire on SurveyMonkey [40]. Some questions were multiple choice, based on frequent answers we obtained to that question during the interivews. Other questions asked respondents to specify how much they agree with certain statements, on a Likert scale from 1 to 4. An N/A option was also provided. To filter users who only click through, we included a question that tested whether participants read the question description. On average, the survey took 23 minutes to complete (excluding the largest 15 values).

3.3 Participants

Interviews: We recruited participants through flyers posted in the city and at universities, through online advertisements on a university-hosted website (at ETH Zurich), mailing lists, and word of mouth. To avoid a biased sample, the advertisement did not mention privacy or security, and only said we are looking for people who use online platforms to store data. In particular, we mentioned that they should use a webmail account, such as Gmail, Yahoo Mail, or Hotmail, or share pictures online through Picasa Web or Flickr. During recruitment, we preferred Dropbox and Google Docs users and rejected IT experts and computer science students. We offered a monetary reward of 20 Swiss Francs (aprox. USD 17) to participants in Zurich, and 250 Indian Rupees (aprox. USD 6) in India.

In total, we interviewed 19 Indians and 1 Estonian living in Delhi, and 13 Europeans (4 Swiss, 4 Germans, 2 Italians, 2 Serbians, one Austrian), one American, one Chinese and one Indian living in Zurich. Professions varied with 10 participants in business and sales, 7 in social sciences and linguistics, 5 in natural sciences such as chemistry or biology, 4 in engineering, 4 in art and design, 2 in finance, and 2 computer scientists, one economist, one urbanist.

Online Survey: Participants were recruited through Facebook postings, student mailing lists, and word of mouth. In Delhi, 100 forms were distributed as hardcopy in several universities and later collected. To incentivize participation, we offered three $100 Amazon vouchers through a lucky draw. We had 450 respondents, from which we dropped 48 based on the test question. Table 2 shows the demographics of the remaining 402 participants. 189 respondents had Indian nationality, 132 were Swiss and the other 47 were Europeans. Of the total 402, 182 participants lived in India and 160 in Switzerland.

3.4 Data Analysis

We transcribed all audio interview recordings into English. For each question in the interviews, the interview moderators identified trends and grouped answers in a few categories. If the moderators did not agree that the participant unequivocally understood the question, the answer was discarded. Throughout the interviews we received many "I don't know" answers, which we generally exclude from reporting in the results section. Finally, we formed hypotheses for the survey about current practices, perceived and expected privacy, and cultural differences based on observed trends and aggregated answers.

To analyze differences between various groups among our respondents (e.g., Swiss vs. Indians, computer scientists vs. non-computer scientists), we used the two-sample Wilcoxon rank-sum (Mann-Whitney) test for the Likert scale questions. For multiple choice questions, we applied the Fisher's exact test for each of the possible answers, to determine if a certain group (e.g., Swiss or Indians) is more likely to pro-

[3]One participant chose to stop the interview and requested the deletion of the audio file after 15 minutes, as we were asking questions about sensitive personal digital data, such as passport copy and password list. We deleted the audio file, as requested, and are not using the data in our analysis.

Table 2: Demographics of online survey participants; values presented as percentages.

	Swiss N=132	Indians N=190	All N=402
Gender			
Male	70	55	60
Female	30	45	40
Age			
18 -24	60	66	60
25 - 34	34	22	30
35 - 44	4	7	5
>45	2	5	5
Education			
High School	47	30	35
Bachelor's	30	40	35
Master's	16	25	25
PhD	4	4	5
Computer Scientists	60	21	36
Computer Skills			
Novice	2	3	3
Intermediate	26	53	40
Proficient	42	33	38
Expert	30	11	19
Platforms Used			
Google Docs	48	70	60
Dropbox	51	17	34
FolderShare	1	14	8
Gmail	61	91	77
Yahoo Mail	13	60	40
Hotmail	34	10	22

vide the respective answer.. For the Likert scale results, we discarded neutral (N/A) responses from the analysis.

4. RESULTS

In this section, we present the main findings of our study. We start by reporting on current practices,such as what kind of data users store in the cloud, and continue by presenting their mental models. Section 4.2 describes perceived privacy, and privacy expectations of consumers, and Section 4.3 discusses user understanding of key conditions stipulated in the terms of service.We refer to interview participants in Zurich as Z1, Z2, Z3, ..., Z16, and to participants in Delhi as D1, D2, D3, ..., D20.

4.1 Current Practices

Six participants in Zurich and 2 in Delhi used dedicated cloud storage systems such as Google Docs, Dropbox or FolderShare [12]. These systems were used mostly for work and collaborative projects, e.g., in school assignments and surveys. For personal data, participants made heavy use of webmail accounts. Z10 said she would rather store sensitive documents in her Gmail account than in Google Docs, because *"email feels more like your private space."* Participants emailed documents to themselves to synchronize data between computers, to back up important files, and to have documents available when needed.

Two participants in Zurich and 7 in Delhi said they have *"folders"* in their webmail account, referring to email labels. Figure 3 shows a participant's inbox. Most participants (14 in Zurich, 18 in Delhi) had several webmail accounts, to

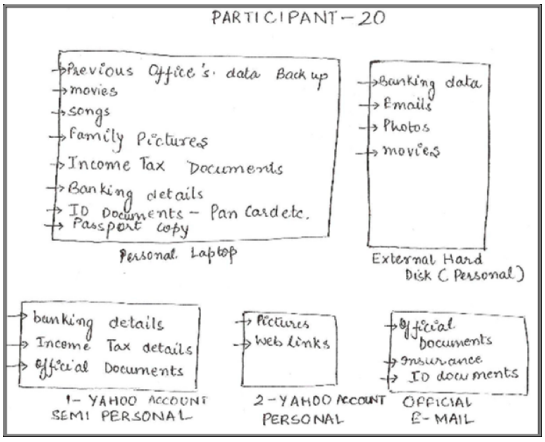

Figure 2: Participants had several webmail accounts, even with the same provider, for separate purposes: private, business use, spam, backup. As shown above, D20 stores different types of data in her accounts: from ID documents in her business webmail to pictures and web-links in her personal one.

differentiate between private, "official" and newsletter/spam use. Furthermore, some participants (7, all in Delhi) had more than one account with a single webmail provider. Figure 2 shows D20's data distribution on local devices and the cloud, as drawn and discussed during the interview. Participants stored pictures, school or project work, official letters, CVs, music files, videos, passport copies, tax, and financial files in the cloud. Complementary to email, participants made heavy use of USB and external hard drives to synchronize and back up data. For example, Z4 said that, when she creates a Word document, she stores it *"in My Documents, I back it up on a USB stick, and email it to myself for back up."* USB sticks were used not only as a data transportation device (e.g., to share files with colleagues or synchronize between computers), but also for permanent storage.

The online survey confirms that users do not use the cloud as a main storage unit. As shown in Figure 4, 83% of respondents somewhat or strongly agreed with *"I tend to keep a back up of all data I store on the Internet"* (M=1.62, SD=0.82, N=285 - where 1 is strongly agree and 4 is strongly disagree). However, Swiss agreed more strongly (M=1.5, SD=0.89, N=124) than Indians (M=1.71, SD=0.72, N=185), as shown by the Wilcoxon rank-sum test ($z=1.67$, $p<0.05$). Participants mentioned that the most annoying part about losing access to their email account would not be the loss of data, but the hassle of informing their contacts of a new email address.

Just like companies, participants were storing only the less sensitive data in the cloud. For example, Z13 said: *"If I will download a file for free, pirated, I will not put it of course on my Yahoo account. I would keep it on the laptop."* Z14 agreed: *"Would you write a diary on Google Docs, would you trust them with your secrets? I guess not."* However, what was considered *"sensitive"* differed among participants, and nationalities. Z9 considered health history more sensitive than the passport because *"a passport I show the policeman; my health card I show the doctor. The doctor is one."*

Figure 5 shows interview participants' willingness to email

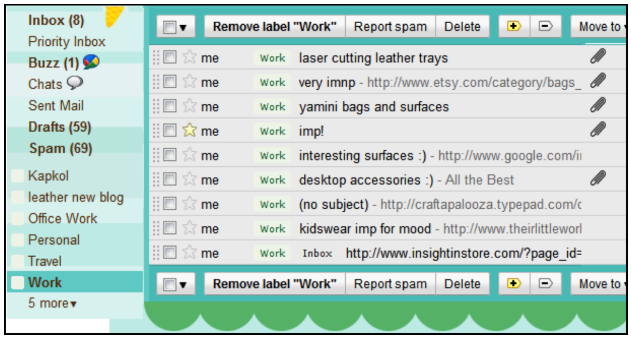

Figure 3: Most participants used their webmail accounts as a cloud storage platform. D6 regularly emails documents and links to herself and then stores them in specific "folders", by setting email labels. She lives in Delhi, is a leather designer and has 11 "folders" in her Gmail account. (Screenshot presented with the participant's written permission.)

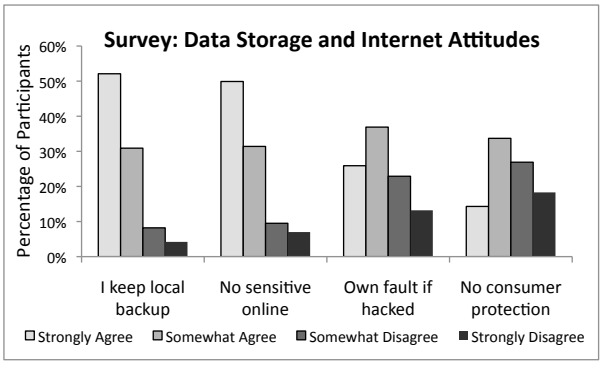

Figure 4: Users keep local backups of data they store in the cloud and try to keep sensitive data away from the cloud. Most feel it is their fault if they store sensitive data in the cloud and it gets hacked, and that there is no legal protection authority they can turn to.

some types of sensitive files to themselves. No participant said they would store sensitive data in their webmail account rather than storing it on their computer. Furthermore, for really sensitive data, like bank and tax statements, printouts were preferred to electronic copies. For example, Z1 would not keep an electronic copy of financial files: *"I don't trust myself. Sometimes my computer is kind of hectic. It happened that I sent some files to wrong people."* The online survey confirmed that users prefer to keep sensitive data on local storage. 81% of respondents (M=1.71, SD=0.90, N=287) somewhat agreed or strongly agreed with: *"I try not to store important, sensitive documents on the Internet, and instead keep them offline, on my personal computers."*

We noticed several differences between study participants in Zurich and Delhi, which were later confirmed by the online survey. While Indians did not consider health information sensitive data, Europeans were very reluctant to even store it in digital format. During our interviews, 15 people in Delhi and only 3 in Zurich said they would store financial documents in the cloud. In the online survey we asked partici-

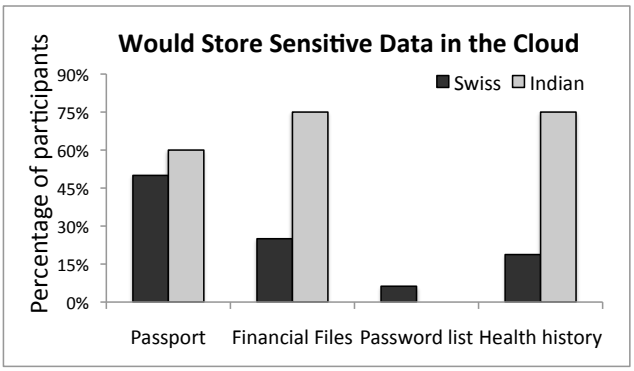

Figure 5: In interviews, Indian participants were more willing to store sensitive data in the cloud than Swiss participants.

pants to rate on a Likert scale from 1 to 7 the sensitivity of the data they have stored in the cloud. Figure 6 shows that Indians reported to have stored significantly more sensitive data than Swiss (Wilcoxon rank-sum test, $z=4.23$, $p<0.001$).

Unconcerned with identity theft, and unaware of the value of a digital copy of one's ID, participants (3 in Zurich, 1 in Delhi) considered that a passport copy is not sensitive: *"It is a copy. I think that important is the original."* Overall, users were less willing to store a password list than a copy of a passport in the cloud, though this is in large part due to perceived need. Some users have been pushed into emailing digital copies of official documents, with which they were not comfortable, by other party's requirements, e.g., when applying for a job. In Delhi, 5 users reported storing password lists on their mobile phones, which was considered more trusted, and more accessible than a computer. Some users said they do not need to store a password list, because they reuse combinations of several passwords. Only very few said they remember passwords.

4.2 Perceived Privacy

In this section, we discuss users' perceived privacy for the cloud, in particular their perception on who else, besides themselves, might be able to access their data, and what guarantees they think current technical solutions provide. Several participants in the study said privacy means *"that nobody else has access to my data."*

Participants' understanding of the cloud architecture is rather limited. They unanimously believed that their storage provider keeps one, two, or at maximum five copies of their documents. D8 said: *"I think the server needs one copy only, because from any computer in the world I can access this copy."* D17 said: *"They have so many users, [...] they would need so much space to keep multiple copies."* For Z10 *"it would be weird if they stored a backup,"* because that would mean *"they are keeping a copy for themselves."* D8 agreed: *"If it is secure and nobody can access them, why should they make more copies? One would be enough."*

Our results suggest that users consider the Internet highly insecure and feel responsible for protecting their sensitive documents themselves, instead of relying on cloud storage providers. Z16 said: *"It is still my responsibility what I upload or what I send and where it is stored."* We summarized such responses in a statement that we then included as a

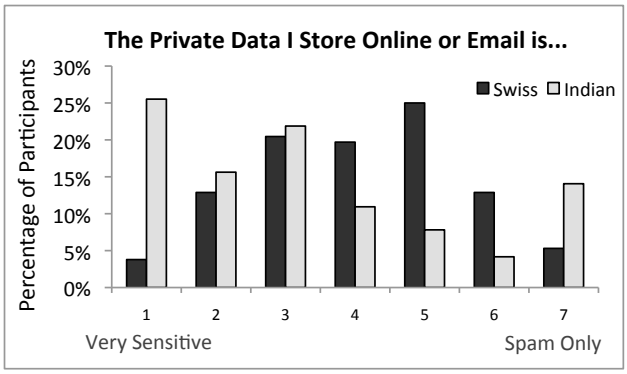

Figure 6: The online survey confirmed the interviews: Indians reported storing more sensitive data in the cloud than the Swiss.

Likert scale in the online survey. As shown in Figure 4, 63% of respondents agreed that: *"If people put their private data on the Internet and it gets hacked, it is their own fault. They should know that nothing is really safe on the Internet"* (M=2.24, SD=0.98, N=291).

Almost all interview participants voiced concerns about the safety of storing documents online, many even before we prompted them about possible sensitivity of the documents. According to statements by our participants, common perception as well as extensive media coverage on the subject formed their beliefs. Some participants seemed to believe that digital data cannot be contained, because the Internet is *"everywhere."* They could not imagine that it might be technically possible to have online data stored in a single country. Similarly, some believed that, once uploaded on the Internet, digital copies remain there forever. According to Z16, there is a nice saying: *"The Net will not forget."*

4.2.1 Anybody Can See My Data, If They Want To

We asked participants who else, except for themselves, might be able to see their cloud stored data. Several participants said *"anybody"* could see it. Z15 said: *"I know that when I store data [in the cloud], the data is really for more people than myself."* We inquired about *hackers*, *storage providers* and *governments*.

Hackers: Participants almost unanimously believed that it would be *"easy"* or *"really easy"* for a hacker to get their data from the cloud. Only one (in Zurich) said that it would be *"hard, but not impossible"* for a hacker to break into their account and another (in Delhi) believed that *"Google cannot be [hacked, because] they have Russian army to protect their data, but Facebook and Twitter have been [hacked]."* For example, Z11 said: *"If he is a good hacker, he can do everything."* According to D19, any measures to protect online data are useless, because ultimately *"there are supernatural hijackers who are sitting there, who can dig everything away."* Z1 agreed: *"They can even get access to the websites of governments. Why shouldn't they [be able to access my account], if they really want to."*

Storage Provider: Except for one, all interview participants were aware that their storage provider can access their data. When asked why his provider would need to see the data, Z2 said: *"To arrange it. If they are keeping an account, then they look after it."* D3 is not convinced that there is a valid reason behind access: *"They come up with all stupid, stupid excuses: security reason, we need to see it."* Except for one participant who said that it might be that every employee of the company can access customer documents, people said that only *"some"* employees would have access, most often quoted being system administrators or *"security people."* Several participants said that internal *"policies"* impede other employees'access to user data, or the fact that accessing customer data is *"taboo"* within the organization. Only one user had *"never thought about it. [...] Is an account accessible just from the user or, for example Gmail or Google [her storage provider], can have access as well? Now I am getting scared."* D7 said: *"They can but they don't."*

Governments: Only two interview participants, both in Zurich, said that the police or government cannot access their account. Z4 believed that they could not because only she knows her password, and Z5 because she is *"a normal citizen, in the sense of not criminal. [...] The state cannot access my bank account also, so I suppose it is more or less the same."* For D5, even with a paid account, *"at the end of the day, there is no gaurantee. Like the bank account, if a governmental agency wants, they can access your information."* The affirmative answers varied from *"Yes, it would be very easy," "they [the government] must have a direct access"* or *"a program,"* to *"only through Google."* We will report more on user perception of their country's and other governments rights to access their data in Section 4.2.4.

4.2.2 But I am Not Interesting to Them

Although participants believed hackers, storage providers and governments could theoretically view their data, none showed great concern about it. In practice, people did store sensitive data in the cloud and considered that the risk of somebody actually viewing *their* data was minimal or non existent. Few participants believed unauthorized access might have already taken place. The main reason given was: *"I am a normal person," "not famous," "not criminal,"* and *"not as interesting as Obama."* Such attitudes were stated by 10 participants in Zurich and 4 in Delhi. Z1 said: *"I am not interesting to them [government], because I am just a little boy somewhere in Switzerland."* Z4 agreed: *"I am a student, I don't know why a hacker should access my account."* For governments, only a couple of participants mentioned that automatic monitoring might occur, but then again: *"I don't write bomb, bin Laden."* Not storing valuable data online kept the hackers away. Z0 said: *"It is very unlikely that they [hackers] want to see my documents, as long as I don't store financial documents, access codes or passwords online. If I store my bank account access, yes, they would be interested."* Similarly, D10 said: *"There are too many documents and too few people in Gmail, [...] so not many manage to see my Gmail documents. [...] But in future, if I hold a good position, then they may."* D2 said: *"I don't think anybody has that much time. Why would someone be interested?"* It is also a matter of time. To participants, Internet attacks are targeted. Viruses targeting a bulk of random user computers or accounts were not considered by the participants.

4.2.3 Home is Safer Than the Cloud

We asked interview participants where they consider their data to be safer: in the cloud or on local storage. Participants felt that availability is better online, *"in case my com-*

puter crashes," but for sensitive documents, they strongly preferred to keep them offline. The ultimate protection against hackers is unplugging the Internet cable. For example, Z11 said: *"Hackers can access the data when we are online, not offline."* Z13 said there is a higher risk if the data is saved online compared to her own computer: *"They can try to enter on my account also if my laptop is closed, so they have more time."* Physical protection of data stored locally, i.e., by locking the disk in the cupboard, is still better than online protection of documents. Z3 said: *"There are many people online; at home it is put away."* Z5 would not store her passport copy in the cloud because *"I look after my laptop [...] and I take care of it. But on Google Docs, I just have to depend on people that program security."* Even though she knows that Google has more experts, she *"would still keep the copy on my computer.[...] It feels here and more accessible."* The USB stick is even safer than the laptop: *"I keep it always with me. Somebody has to really kidnap me to have the USB stick."* Even if she believed that it would be easier for a hacker to break into her computer than into Google systems, Z6 still considered her laptop safer than the cloud: *"Google has experts to deal with hackers, I have no one to help me,"* but *"professional hackers want to hack big companies, organizations, not individuals, because there is more value in that."* Similarly, D4 said: *"people know where Google Docs are, but it's difficult to find which connection I am using, where I am sitting!"*

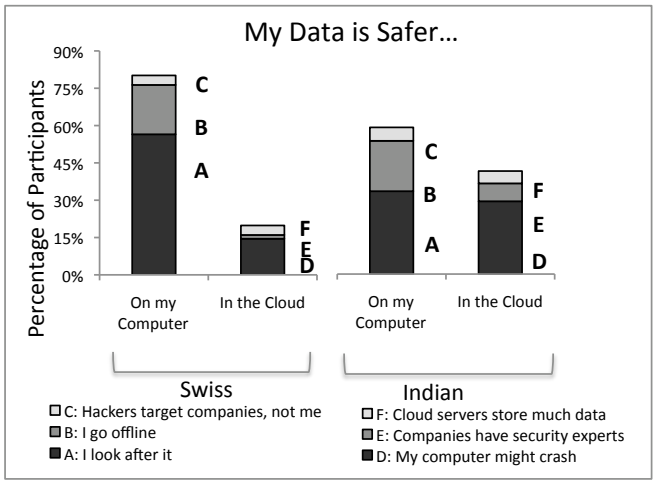

Figure 7: Indians considered cloud storage almost as secure as local storage. Swiss trust their computer much more than the cloud.

In Zurich, 2 participants said the risk is the same *"if my computer is connected to the Internet,"* and none said higher on the laptop. In Delhi, however, 13 said the risk is higher offline and 4, online. The online survey also showed differences in attitudes between Swiss and Indian participants. We asked survey respondents to rank 6 provided reasons on why local or cloud storage is safer. Figure 7 compares the choices made by survey participants of Swiss and of Indian nationality. 69% of all respondents said local storage is safer, and 31% that the cloud is. The reason to be rated the strongest was A: *"On my computer, because I can physically protect my data,"* with 44%. Next reason was availability

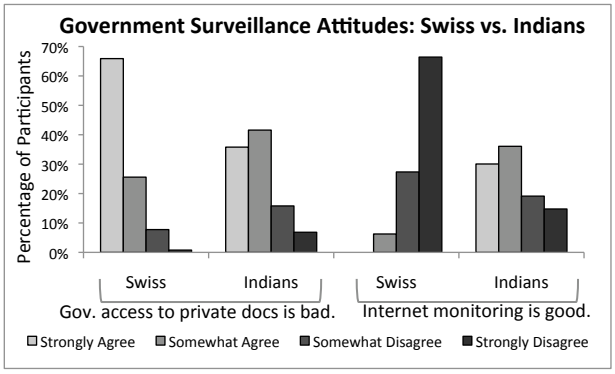

Figure 8: Indians are more accepting than Swiss of government surveillance of data stored in the cloud.

D: *"Online, because my computer may crash."* Interestingly, only 5% chose E: *"Online, because big companies have security experts,"* an argument often stated as major cloud advantage in enterprise usage.

A significantly higher percentage of Swiss respondents (82%) compared to Indians (60%), considered local storage safer than the cloud ($p<0.001$, Fisher's exact test). However, participants' background, not just nationality, might be a factor influencing the difference in perception between Swiss and Indians. The percentage of computer scientists in the Swiss group was significantly higher than in the Indian group (60% vs. 21%, see Table 2). Nevertheless, the difference persisted among the groups of Swiss computer scientists and Indian computer scientists ($p<0.01$), as well as Swiss non-computer scientists and Indian non-computer scientists ($p<0.006$).

Fisher's exact test showed no significant differences between the scores for Indian, non-computer scientists group and Indian, computer scientists ($p=1.0$). Similarly, we obtained no significant difference between Swiss non-computer scientists and computer-scientists ($p=1.0$). A failure to see a difference between computer scientists and non-computer scientists might be also attributed to the young age of participants. (Computer scientists in our study are mainly students; their attitudes towards cloud security might be different from those of experienced professionals.) Differences might be noticeable in a future study among higher age groups.

4.2.4 Government Surveillance

Throughout our interviews, Indians showed very different attitudes towards government surveillance than Europeans. For example, we asked participants whether it is their right to protect the privacy of their data and communications, followed by whether everybody should be able to, and then by: even terrorists? In Zurich, 6 participants said everybody should have the right, including terrorists. In Delhi, 11 people said that terrorists should not have the right to privacy and only 3 said that everybody should. For example, Z13 said: *"There are terrorists, but it is not because of them all the people cannot have their privacy.[...] I think this is an excuse to control everything."* Z14 agreed: *"Who defines who is terrorist?"* Only one person answered that *"the police from all states"* should be able to access any data. Participants in Delhi showed a much stronger acceptance of government surveillance. They felt that *"national security comes first."*

Furthermore, we asked interview participants if a communication technology currently exists, through which they could talk to a friend, for example over the Internet, and nobody, not even the government, could listen in to their communication. Overall 13 people said such a technology is technically possible, and 13 said it is not. While among Zurich participants, the general trend was that this technology is not currently being deployed for surveillance and security reasons, in Delhi people felt that such a technology should not exist, because it would be misused: *"then terrorists will enjoy themselves."*

In the online survey, we asked respondents to rate on a 4 point Likert scale, with 1 for strongly agree, two statements which we received from our European and Indian participants in the interviews. Figure 8 shows the answers for respondents of Swiss and of Indian nationalities. For the statement *"If the government had access to every document users store on the Internet, that would be a major violation of individual privacy,"*, Swiss (M=1.43, SD=0.67, N=129) agreed more strongly than Indians (M=1.94, SD=0.89, N=190): Wilcoxon rank-sum test, z=4.96, p<0.001. For the statement *"It is good if the government monitors every Internet communication and all user accounts. National security comes first,"* Indians (M=2.18, SD=1.03, N=193) agreed more strongly than the Swiss (M=3.60, SD=0.61, N=128): Wilcoxon rank-sum test, z=10.56, p<0.001.

4.2.5 I Would Pay for Privacy

During the interviews we asked 8 Dropbox users and some non-users to identify the statement that appears in the Dropbox privacy policy from three statements. None chose the correct variant from the tree possible choices: *"Dropbox may sell, transfer or otherwise share some or all of its assets, including your Personal Information, in connection with a merger, acquisition, reorganization or sale of assets or in the event of bankruptcy."* All participants said that this should not be, and non-users said they would not open an account with such a company. Motivated by this finding, we used the online survey to see if respondents would be willing to pay for privacy guarantees.

In the online survey, respondents had to choose between two companies with which they could open an online data storage account. Company A offered the service for free, but said that they may sell documents of personal information. Table 3 shows the results of the online survey. 79% of respondents agreed that they would pay 20 USD per year for Company B, whose policy says that they will not sell any personal information. This amount would be enough to buy twelve 2 GB of increased redundancy storage on Amazon S3 for a year.[4] We did not notice a strong difference between Swiss and Indian respondents: 81.6% of Indians and 78.5% of Swiss chose Company B. While the statement used is much stronger than the Dropbox policy, which may sell only in connection with a merger, our survey does show a strong user response toward privacy protection.

During our interviews, we asked participants if they would be interested in purchasing insurance for their cloud-stored data, the same way they have for cars and houses, so that if a hacker breaks in and data is lost they would receive some compensation. Half of the interview participants, split evenly among Delhi and Zurich, said they are interested. For example, D6 said he would pay 1000 Indian Rupees (approx. 20 USD) per year for data insurance, while D10 said he would pay 60 USD per year. Z5 would pay 50 Swiss Francs per year and Z6 would pay *"several hundred Swiss Francs."* Others said what matters is the data, and they would instead prefer investing in an additional back up system.

[4]Amazon S3 charges $0.140 per GB per month: http://aws.amazon.com/s3/pricing/.

Table 3: Willing to pay for privacy: *"Which company would you choose to store your data, and why?"*

CompanyA: free, may sell user data	
– It is free.	3.0%
– I don't have sensitive data anyway.	11.4%
– I never know what they do with my data.	6.5%
	Total: 20.9%
CompanyB: costs $20, won't sell data	
– I value my privacy.	37.3%
– If the price was lower.	9.7%
– If they are trustworthy.	32.1%
	Total: 79.1%

4.3 Terms and Conditions

Unsurprisingly, our results confirm that users do not read Terms of Service and Privacy Policies. D14 said: *"It's massive! It's just in five Arial font and it's massive! It's ten pages!"* A few participants said they skim through the text. Although they do not read them, participants believed strongly that these documents are legally binding and valid contracts in court. Z8 said: *"It is your fault if you did not read it."* D14 said: *"You should be smart enough not to do all that stuff [store confidential customer information]. And if you've done it, then welcome to the world, wherein you had said, 'I accept'. So, if you have accepted it, you have to take it."* Only one user said that some of the things the company claims in the terms might not be legal.

Several participants said they do not read these documents because *"I don't think they [Terms of Service] would have an impact."* However, the Terms of Service and Privacy Policy documents explain conditions such as: Google has the right to disable the account at any time and without notice, to read, delete and modify their data; Dropbox may sell user data, the storage provider assumes no liability in case of data loss. We explored in detail users' awareness of these terms.

4.3.1 Country of Storage and Storage Outsourcing

We asked participants where they thought their online data was being stored and whether the country of storage was important to them. We then asked them to imagine that their storage provider contracted a third-party company to store their data[5]. Only 7 out of 36 interview participants said the country of storage is important or they care about storage outsourcing. Another 4 participants said it might matter if they were storing more sensitive data. Reasons given for not caring were *"I trust the company," "maybe if I had more sensitive data,"* but also *"as long as security is guaranteed"* and *"if they [the third-party company] have*

[5]This is, for example, the case with Dropbox, who is using Amazon S3 to store user files.

the same privacy policy." All participants said the data is safer in their own country, except for one Indian who said in India there are more hackers. Only 2 participants, both in Zurich, mentioned country-specific data protection laws to be a factor in data security. Fourteen participants said they care if their data storage is outsourced by their storage provider, and 19 said they should be informed. Two participants said they would close their account if data storage was outsourced, and one that he would sue the company.

4.3.2 Unauthorized Modification

We showed participants a slide with three variants of the policy *"Google reserves the right (but shall have no obligation) to pre-screen, review, flag, filter, modify, refuse or remove any or all Content from any Service."* One variant did not grant the right to modify, and one said *"except for personal documents in user accounts."* Only 4 participants chose the correct variant. Like 13 other participants, Z10 said: *"They should have the right to review but I don't think they should have the right to modify."* Similarly, D5 said: *"No, not modify it, but this if it's illegal then they [Google] can delete it."*

Participants felt a strong ownership right over the data, even if stored in the cloud. They accepted that their data might not remain confidential and that the provider might choose to delete it, but expected full data integrity. For example, D17 said: *"It's my personal data they have to respect this thing"* or *"they are my personal documents [...] even if I put them on Google."* Z13 thinks she still owns the copyrights of whatever data she uploads: *"They give me a space on their system. They don't say put the stuff here and everything gets mine."* D3 said that the documents she stores in the cloud: *"are my things. [...] It was created by me. [..] They might delete it, they might remove it, but they cannot make changes themselves."* For security reasons, many participants accepted that the provider might need to look at their data. For example, Z12 said: *"I can understand that Google wants to be able to look at the data that is stored. In case it is criminal data, they could inform the police or delete it."*

We summarized participants' views in four choices in the online survey. Table 4 shows the results for Swiss and Indian nationalities. From all respondents in the survey, only 8% answered "Yes"; 77.3% were using Gmail and 50% named Gmail as their main email account. We applied the Fisher's exact test for each of the multiple choice answers of the survey question, and, except for the willingness to answer "I don't know," observed no significant difference between Swiss and Indians.

4.3.3 Guaranteed Deletion of Data

The Google Docs policy states that "residual copies of your files may take up to 30 days to be deleted from our active servers and may remain in our offline backup systems for up to an additional 60 days." We asked participants to identify the correct statement among three other variations: one saying that data can never be deleted, one saying it gets deleted within 24 hours, and one that it gets immediately deleted. The correct variant was chosen by 14 participants. Five participants said the data never gets deleted and 4 said deleted data resides for 24 hours. The most mentioned source of information was the media: *"Probably there are traces still there. I heard in the media, television, advertisement in journal."* No participant said the data would be immediately deleted.

Table 4: Unauthorized modification: *"Does your webmail provider have the right to see or modify the documents you have as attachments in your email account?"*

Response	Swiss	Indians
No.	22.3%	28.4%
They can see, but not modify my files.	12.2%	26.8%
They have the right to see and modify only in criminal or terrorists cases.	7.2%	21.1%
Yes.	10.1%	6.8%
I don't know.	48.2%	16.8%

Few participants felt strongly that when they delete data it should get deleted. For example, Z8 said: *"It is the private right that when it is deleted it actually is, and if somebody uses it nevertheless it is infringing my privacy."* Others said they would care about copies of sensitive data such as online banking transactions, but not about advertisement emails. Overall, participants did not show great concern: *"If somebody is storing important stuff like ID, official documents then they should be deleted immediately."* Other participants regarded this as a good feature. D10 believed that the data would still remain on the cloud *"because it is a very good system. If I delete my document, there must be some technology through which I can retrieve my data back."*

Table 5: Guaranteed deletion of data: *"When you delete a file stored on the Internet or an email in your Webmail account, what do you think happens?"*

Response	Swiss	Indians
The file gets permanently deleted.	2.9%	15.3%
Some copies still exist for a few weeks.	34.5%	38.9%
Copies are kept, for security reasons.	36.7%	25.3%
I don't know.	25.9%	20.5%

We followed up these findings in our online survey. As Table 5 shows, very few respondents believed data gets immediately deleted. Only 15% of Indians and 3% of Swiss chose: *"The data gets permanently deleted, just as when I deleted it from my computer"* (significant according to the Fisher's exact test, $p<0.001$).

4.3.4 Account Disabling

Google's Terms of service state: *"You acknowledge and agree that if Google disables access to your account, you may be prevented from accessing the Services, your account details or any files or other content which is contained in your account."* Other storage providers follow a similar policy. For example, *"Dropbox reserves the right to terminate Free Accounts at any time, with or without notice."* Eight people in Zurich and 7 in Delhi said their service provider (mostly referring Google) has the right to disable their account. Seven participants in Zurich and 4 in Delhi said they do not. For example, D2 said: *"There is some trust that we*

have put in, they should take care of that. Not without my consent." Doing so *"is not ethically correct."* D5 said they do not have the right *"because all my data is there, they should inform me before."* Other participants said Google may disable their account, but only with prior notice, only with a reason of *"if you do not access it anymore."* Several participants said they would sue Google if they disabled their account. Many participants accepted security reasons such as *"if I am a terrorist"* or *"use it for criminal purposes," "if I have done something and that is against their rule,"* or *"if they get complaint from other people."* A paid-for service was not always regarded as a guarantee of having more rights. For example, Z14 said: *"you paid for the storage, not the privacy."*

Table 6: Account disabling: *"Does your Webmail provider have the right to disable your account?"*

Response	Swiss	Indians
Yes, at any time, without advanced notice and without explanation.	34.5%	15.3%
Yes, but only with advanced notice and a valid reason.	21.6%	48.4%
Only if I use it for criminal purposes.	10.1%	13.7%
No.	4.3%	8.4%
I don't know.	25.2%	12.6%

We asked the same question in the online survey, with four multiple choice answers. Table 6 shows the responses, which confirm limited user awareness. The Fisher's exact test confirmed that Swiss are more aware than Indians of the fact that their storage provider has the right to disable their account without advanced notice and without explanation ($p<0.001$). Indians, on the other hand, assume that this can happen only with advance notice and a valid reason ($p<0.001$). The difference persisted between Swiss non-computer scientists and Indian non-computer scientists ($p<0.001$ and $p<0.02$ respectively).

4.3.5 Data Loss Liability

We asked participants what their rights would be if their storage provider lost some of their data (e.g., due to accidental deletion or server crash). The terms of service of all important online storage and Webmail companies dismiss any liability for data loss. For example, *"Google [...] shall not be liable to you for [...] the deletion of, corruption of, or failure to store, any content [...] whether or not Google has been advised of or should have been aware of the possibility of any such losses arising."* Participants had diverse views on companies' liability and their rights. For instance, D14 said: *"They're already giving you a service. [...] if you're stupid enough to keep your important documents there as a storage device and not use your external hard disks and stuff, then it's not their liability."* D5 disagreed: *"I didn't ask them to give a free service, they decided this. [...] They should pay me a large sum."*

Five participants in Zurich believed that the storage provider would be liable to them in case of data loss, and 4 participants said they would not have any rights. In Delhi, 14 participants said the storage provider is liable, and 5 participants said the consumer would have no rights to claim compensation. Several participants said they would not care about money, because the data is lost anyway, or that even if they had rights the company would not pay or there is nobody they could contact to make the claim. A few said they would sue the company. We investigated these issues further in the online survey. Table 7 shows the results.

Table 7: Data loss liability: *"If your Webmail provider lost some of the data you store with them, what would your rights be?"*

Response	Swiss	Indians
They should pay me for the damages.	10.8%	15.3%
If it is a free service, I have no rights, otherwise they have to pay me.	27.3%	48.4%
I have no rights even if paid-for service.	20.9%	13.7%
Don't care, my data is lost anyway.	8.4%	7.2%
I don't know.	33.8%	12.6%

Our results show that Indians are more prone than Swiss to expect liability from their service provider. They are more likely to expect the provider to pay them for damages (Fisher's exact test, $p<0.001$), whereas Swiss are less prone to believe that they do not have any rights, even if it is a paid-for service ($p<0.003$).

Interview participants said that, if their account was hacked, disabled, or if the provider lost some of their data, they would: *"change the password," "delete all my emails," "close my account,"* or *"write to Google and ask why."* Some participants (8 in Delhi and 1 in Zurich) said they would sue the company if they felt their rights had been infringed. Others said they would not because their data is not that important, lawyers are expensive, or they don't have the time. Generally, participants did not know with whom to file a complaint: *"I don't know whom I should go to. [..] I don't think you can contact anybody."* Participants felt stronger about complaining in case of account disabling than in case of unauthorized data modification. If his account got disabled, D15 said: *"I will shut down my computer,"* because it must have been because of *"Virus attack or some hijack."* No participant said they would go to police if their account had been hacked. Some said there is no possibility to complain, or that they do not know how to contact the company. Most were not aware of laws or agencies protecting their rights, but *"would like to have some laws so that I can complain."* nor of data protection agencies. D3 mentioned *"Cyber client court"* and Z4 the *"Postal police."* We followed up in the online survey. Figure 4 shows that 58% percent of all respondents agreed with the statement *"There is no such thing as consumer protection service or police on the Internet, whom I could turn to, if I felt that my rights were violated"* (M=2.53, SD=0.98, N=279). The Wilcoxon rank-sum test showed no statistical difference between Swiss and Indian participants.

5. CONCLUSIONS

In this paper, we explored end-users' privacy expectations and assumptions about cloud storage, as well as their awareness of risks, terms and conditions. We conducted 36 in-depth interviews in Switzerland and India, and followed up

with an online survey with 402 participants. Our results suggest that users make heavy use of free webmail accounts as cloud storage drives. However, instead of relying on the cloud as a main storage unit, users keep local backups of cloud-stored data. Our results show that end-users have a strong belief, fueled by media stories and hacker stereotypes, that the Internet is intrinsically insecure. The loss of control over where their data is stored, and the inability to physically protect it prevent them from storing sensitive data in the cloud. Our results suggest that users' mental models of cloud storage are very different from that of banks. Unlike money (people trust banks to protect their savings), personal documents are still perceived to be safer at home, regardless of how many security experts the cloud storage providers hire.

Unlike data stored locally, consumers accept that cloud-stored data might be viewed by other parties, such as hackers, cloud storage providers, or law enforcement agencies. However, they believe this privacy breach would only happen to famous people or criminals, not to themselves. Users don't read privacy policies or and terms of service, and believe they have more rights and guarantees than what these documents actually grant them. For example, an alarmingly high percentage of users are unaware that their storage provider reserves the right to modify user data and disable user accounts at any time. Consumers assume they have the same ownership rights over their data stored in the cloud as if stored on their personal devices.

Clearly, there is a great mismatch between users' expectations of privacy and the actual rights and guarantees they enjoy for their data in the cloud. To foster business and cloud adoption and to protect consumers, regulation bodies and cloud storage companies alike should try to close this gap by meeting users' expectations and/or educating consumers on the risks they face. Possible measures to take include: (1) changing the content and the presentation of privacy policies and Terms of Service agreements to make it easier for users to read and understand, (2) offering better visibility into security settings by adopting stronger authentication mechanisms such as two-factor authentication, access log visualization, etc, and (3) accounting for internationalization. The latter involves going beyond just translating the service interface and privacy policy. Companies should keep in mind that users from different countries may have different privacy expectations and understanding of privacy guarantees offered by the cloud storage system.

Our results show that cultural differences and local events influence users' expectation and perception of cloud storage privacy. Furthermore, our results imply that certain countries place a much greater emphasis on individual privacy, whereas others prioritize national security over privacy, differences which companies and international cloud privacy bodies should keep in mind when designing global policies and services. For example, Swiss respondents were more aware of the lack of guarantees and stored less sensitive data in the cloud than Indians. While Indians considered government monitoring of users accounts to be a good thing because *"national security comes first,"* to Swiss government surveillance was a great violation of individual privacy. This is not surprising considering the two countries' political situations and cultural attitudes towards privacy. First, Switzerland is considered a safe haven of stability, whereas India is increasingly dealing more with terrorist attacks and violence [27]. Switzerland receives a score of 1.39 for the factor "Political Stability & Absence of Violence/Terroris" in the World Bank's Governance Indicators for 2008, whereas India receives a score of -0.99. The United States is scored at 0.59 [10]. Second, we, as members of Swiss and India society, have observed that, while privacy is deeply rooted into the Swiss culture, in India the social and family structures place much less importance on privacy. Differences in perceptions of guarantees and privacy in the cloud suggest that the cloud storage policy and system level designers cannot expect one-size-fits-all solution that can accommodate different cultures.

Participants in our study were mostly young. Although young people are a major target group for consumer cloud storage systems, they are not representative of the entire world population. However, young people tend to be more technically savvy than the general population, and likelier to use such cloud storage systems and understand how they work. The general population is, therefore, likely to have an even stronger mistrust of the cloud and a higher misunderstanding of the privacy guarantees it offers than our study participants. Future work could look into privacy attitudes and differences among older age groups and compare awareness of privacy policies among technical and non-technical users.

Furthermore, future work should explore consumer perception of international laws and regulation, as well as data protection authorities they could turn to. Finally, novel, usable mechanisms are needed to educate users and provide them with visibility and control over personal data in the cloud.

Acknowledgements

We would like to thank Prof. Stefan Bechtold for his input on current legislation, Prof. Marc Langheinrich, Aleecia McDonald, Rob Reeder, Martin Ortlieb for input on the study design, and Deepansha Sachdeva for helping in collecting the surveys. We would like to thank all participants in the study. The authors would also like to thank International Development Research Centre (IRDC) and all members of PreCog research group at IIIT-Delhi.

6. REFERENCES

[1] A. Acquisti and R. Gross. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *Proceedings of the 6th Workshop on Privacy Enhancing Technologies (PET 2006)*, pages 36–58, 2006.

[2] C. Albanesius. Google: Software bug caused Gmail deletions. PCMAG News, http://www.pcmag.com/article2/0,2817,2381168,00.asp, March 1 2011.

[3] J. Q. Anderson and L. Rainie. The 2011 cloud trends and best practices report. Hosting.com. http://www.hosting.com/resources/ebooks/2011-cloud-computing-trends-report.

[4] J. Q. Anderson and L. Rainie. The future of cloud computing. Pew Research Center, June 2010. http://pewinternet.org/Reports/2010/The-future-of-cloud-computing.aspx.

[5] S. Bellman, E. J. Johnsonb, S. J. Kobrinc, and G. L. Lohse. International differences in information privacy

concerns: A global survey of consumers. *The Information Society*, 20:313 – 324, 5 November 2004.

[6] R. Chakraborty, S. Ramireddy, T. Raghu, and H. Rao. The information assurance practices of cloud computing vendors. *IT Professional*, 12:29–37, 2010.

[7] R. Chow, P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masuoka, and J. Molina. Controlling data in the cloud: outsourcing computation without outsourcing control. In *Proceedings of the 2009 ACM workshop on Cloud computing security (CCSW 2009)*, Chicago, IL, USA, November 2009.

[8] Data Security Council of India (DSCI). Data protection challenges in Indian cloud computing. an Indian perspective. Special Report, December 2010.

[9] Dropbox. https://www.dropbox.com.

[10] eStandards Forum, Financial Standards Foundation. Business indicators. http://www.estandardsforum.org/switzerland/business-indicators?id=166, 2009.

[11] The Swiss Federal Data Protection Act. http://www.edoeb.admin.ch/org/00828/index.html.

[12] FolderShare. https://www.foldershare.com/.

[13] H. Geert. Geert hofstede cultural dimensions. http://www.geert-hofstede.com.

[14] H. Geert. *Cultural and Organizations - Software of the Mind - Intercultural Cooperation and its importance for survival.* McGraw-Hil, 1991.

[15] R. Gellman. Privacy in the clouds: Risks to privacy and confidentiality from cloud computing. World Privacy Forum, February 2009.

[16] Google 'may pull out of China after Gmail cyberattack'. BBC News, http://news.bbc.co.uk/2/hi/8455712.stm, January 13, 2010.

[17] Google Docs. https://docs.google.com.

[18] B. Grobauer, T. Walloschek, and E. Stocker. Understanding cloud-computing vulnerabilities. *IEEE Security & Privacy*, PP(99):1 –1, 2010.

[19] G. Gross. Cloud computing may draw government action. PCWorld, September 12, 2008, http://pcworld.about.com/od/businesscenter/Cloud-Computing-May-Draw-Gover.htm.

[20] C. J. Hoofnagle and J. King. Research report: What Californians understand about privacy online. Available at SSRN: http://ssrn.com/abstract=1133075, September 3, 2008.

[21] J. B. Horrigan. Use of cloud computing applications and services. Pew Research Center, September 2008.

[22] W. Hu, T. Yang, and J. N. Matthews. The good, the bad and the ugly of consumer cloud storage. *ACM SIGOPS Operating Systems Review*, 44(3):110–115, 2010.

[23] W. Jansen and T. Grance. Guidelines on security and privacy in public cloud computing. NIST Special Publication, January 2011.

[24] A. Joint, E. Baker, and E. Eccles. Hey, you, get off of that cloud? *Computer Law & Security Review*, 25(3):270 – 274, 2009.

[25] M. N. Khan. Does India have a data protection law? http://www.legalserviceindia.com/article/l406-Does-India-have-a-Data-Protection-law.html, November 2009.

[26] E. Kinetz. Google, skype targeted in india security crackdown. The Huffington Post, February 9, 2011, http://www.huffingtonpost.com/2010/09/02/google-skype-targeted-in-_n_703198.html.

[27] M. Kumar. List of terrorist attacks on India. http://kumarmohit.wordpress.com/2009/07/29/list-of-terrorist-attacks-on-india/, July 29, 2009.

[28] P. Kumaraguru and L. Cranor. Privacy Indexes: A Survey of Westin's Studies. Technical Report CMU-ISRI-05-138, Carnegie Mellon University, 2005.

[29] P. Kumaraguru and L. F. Cranor. Privacy in india: Attitudes and awareness. In *Privacy Enhancing Technologies*, pages 243–258, 2005.

[30] H. Michael. *International Privacy, Publicity and Personality Laws.* Reed Elsevier, 2001.

[31] National Intelligence Grid (NATGRID). http://currentaffairs.gktoday.in/2010/02/national-intelligence-grid-natgrid.html.

[32] S. Patil, A. Kobsa, A. John, and D. Seligmann. Comparing privacy attitudes of knowledge workers in the U.S. and India. In *Proceedings of the 3rd international conference on Intercultural collaboration*, ICIC '10, pages 141–150. ACM, 2010.

[33] W. Pauley. Cloud provider transparency - an empirical evaluation. *IEEE Security & Privacy*, PP(99):1 –1, August 2010.

[34] D. Rosenthal. New data protection act in switzerland: More transparency, additional costs. *Privacy Laws & Business International Newsletter*, 2006(3), December 2007.

[35] E. Schindler. Cloud development survey. Evans Data Corporation, Strategic Reports, July 2010. http://www.evansdata.com/reports/view Release.php?reportID=27.

[36] R. Singel. Information technology act 2000. Department of Information Technology. http://www.mit.gov.in/content/it-act-2000-dpl-cyber-laws.

[37] R. Singel. Twitter's response to wikileaks subpoena should be the industry standard. WIRED, January 10, 2011, http://www.wired.com/threatlevel/2011/01/twitter/.

[38] Smith v. Maryland. 442 U.S. 735 (1979): http://laws.findlaw.com/us/442/735.html.

[39] C. Soghoian. Caught in the cloud: Privacy, encryption, and government back doors in the Web 2.0 era. *Journal on Telecommunications & High Tech. Law 359*, 8(2), 2010.

[40] SurveyMonkey. http://www.surveymonkey.com.

[41] D. Svantesson and R. Clarke. Privacy and consumer risks in cloud computing. *Computer Law & Security Review*, 26(4):391–397, July 2010.

[42] Unique Identification Authority India (UID). http://uidai.gov.in/.

[43] J. E. Vascellaro. Google discloses privacy glitch. WJS Blogs, March 8, 2009. http://blogs.wsj.com/digits/2009/03/08/1214/.

[44] A. F. Westin. *Privacy and Freedom*. Atheneum, New York, first edition, 1967.

APPENDIX

Interview Study Script

1. INTRODUCTION

Small description, introduction about the scope of the study, will be told to the participant when starting the interview. We will refer to people and organizations whom you have not explicitly given permission to see your files as "unauthorized parties", and to online websites where you put your files like Dropbox, Google or Facebook as "online site" or "online storage."

2. CURRENT PRACTICES

1. Please draw a diagram showing what kind of files you store on your own computers (laptops, desktops) and what you store on online services such as Google Docs, Dropbox, Facebook, Flickr, Picasa Web.
2. What data do you upload online? Since when? Why? (e.g., to share with friends, for backup, to be able to access it from other computers, etc)
3. What data do you store in more than one location?
4. What documents do you still keep only on your computer and why?

3. WHERE WOULD YOU STORE THE FOLLOWING DOCUMENTS?

1. Where do you store/would you store...
   (a) Financial files such as bank transactions, income, or your tax documents.
   (b) ID documents such as copy of your passport or copies of passports of your family members, scanned visa application forms, in case somebody steals your documents while you travel.
   (c) Your password list or bank login information and credit card number so you can log in from anywhere.
   (d) Health history so that your doctors can access your entire profile fast when you go to a new hospital.

4. EXPECTATION OF PRIVACY

4.1 Physical location:

1. When you write a Word document, where is it stored?
2. How about your email attachments or document in Google Docs, where are they stored?
3. How many copies of your online data are out there?
4. In which country? Does it matter to you?
5. Would you be willing to pay extra to have the guarantee that your data is stored in a specific country, like in Switzerland? For which documents?

4.2 Data protection:

1. How do you think your online data is being protected?
2. Have you heard of data encryption?
3. Do you think your data is safer on an online storage than on your computer? Why?
4. When do you think the risk is higher of somebody obtaining unauthorized access to your files: when stored locally or online?
5. Would you like to be able to request higher protection levels for more sensitive data? By what means? Would you pay?

4.3 Unauthorized access:

1. Who else, accept for you, might be able to see the private data you store online [pick an example: in your Dropbox, Gmail inbox]?
2. How easy would it be for ...
   (a) Hackers
   (b) Employee of your online storage provider
   (c) Your government
   (d) US government
3. How likely do you think it is, the above entities would access some of your online stored documents intentionally or maliciously?
4. Do you think that any of these parties have already accessed your documents?
5. Do you think you would be informed, if an unauthorized party/person accessed your data?
6. Do you think you should be informed?

4.4 Third-parties:

1. Imagine that instead of storing your data on their own servers, your storage provider (e.g., Google/Dropbox) hired another company to store your data, on their servers
   (a) How concerned would you be if this happened?
   (b) Do you think this might currently be the case?
   (c) How likely do you think this is to happen?
   (d) How upset would you be if they did?
   (e) Do you think you would be informed? Should you be informed? Through which means?
   (f) How upset would you be if you were not informed?
2. Have you heard of Terms of Service and Privacy Policies? Did you read them? What do you think they say?
3. Are the Terms of Services and Privacy Policies legal contracts, enforceable in court?
4. If your data is stored in another country (e.g. the U.S.), do you think that Swiss/Indian or U.S. regulations apply?
5. For Dropbox users: Which one of the following four statements do you think apply to your contract? ... Discussion.
   (a) Dropbox may sell, transfer or otherwise share some or all of its assets, including your
   (b) Personal Information, in connection with a merger, acquisition, reorganization or sale of assets or in the event of bankruptcy.
   (c) Dropbox will not sell, transfer nor otherwise share any of your Personal Information to another party.
   (d) Dropbox may only sell, transfer or otherwise share some or all of its assets, including your Personal Information, in the event of bankruptcy.
   (e) Dropbox may sell or otherwise share some or all of its assets, including your Personal Information, in connection with a sale of assets.
6. Do you think you are allowed to store third-party data (music, videos, photos, text etc.) on your Google Docs/Dropbox account?

4.5 Data Integrity:

1. Imagine that you are accessing your online documents and notice that somebody modified or deleted some of your data (e.g., emails you know you sent now contain a different text).
   (a) If this happened, what would you do?
   (b) How likely do you think this is to happen?
   (c) Whom would you suspect to have modified your data?
   (d) Do you think anybody has the right to modify or delete your data?
2. Which of the following statements do you think is in the Google privacy policies document?
   (a) Google reserves the right (but shall have no obligation) to pre-screen, review, flag, filter, modify, refuse or remove any or all Content from any Service.
   (b) Google reserves the right (but shall have no obligation) to pre-screen, flag, filter, refuse or remove any or all Content from any Service.
   (c) Google reserves the right (but shall have no obligation) to pre-screen, review, flag, filter, modify, refuse or remove Content from any Service, except for personal documents in user accounts.
3. The first option is the correct one. Do you think there might be a reason for this policy?

4.6 Controls:

Imagine that you could set a lock on your data (set a flag) before you upload it, or when it is already uploaded. If you set this lock then nobody can modify your data.

1. Do you think such a technology could be possible?
2. If your provider offered it to you, would you use it?
3. How much would you be willing to pay for this feature?

4.7 Guaranteed Deletion of Data:

1. Can your data still be recovered after you delete it from your computer?
2. How about after you deleted from your email account?
3. Which one of the following statement is correct?
   (a) You may permanently delete any files you create in Google Docs. Residual copies of your files will be deleted within 24 hours.

(b) Because of the way we maintain this service, residual copies of your files reside on several active servers and offline backup systems. We therefore do not guarantee permanent deletion of files you create in Google Docs.

(c) You may permanently delete any files you create in Google Docs. Once you do, all copies of your files will be deleted from all of our servers.

(d) You may permanently delete any files you create in Google Docs. Because of the way we maintain this service, residual copies of your files may take up to 30 days to be deleted from our active servers and may remain in our offline backup systems for up to an additional 60 days.

4.8 Lock-out/Data Migration:

Imagine that in the future you will decide to abandon Dropbox/Gmail/Yahoo and move to a new system that is gaining popularity. Perhaps these companies are going bankrupt.

1. What data would you save?
2. Do you know how to get your data out of the system easily?

4.9 Account Disabling:

Imagine that tomorrow when you are trying to access your Google/Dropbox account you are being informed that your account has been disabled and you may no longer log in.

1. How likely do you think this is to happen?
2. Do you think your provider has the right to disable your account?
3. What would be the worse/irepleaceable thing/data to lose?
4. Whom would you turn/complain to?

4.10 Liability in case of failure:

Imagine that your storage provider lost some of your data, perhaps an administrator accidently deleted it or there was a server crash.

1. What do you think your rights are in such a case? What actions would you take?
2. What if you paid for the service? Does it change your rights?
3. Do you think you would have to file a lawsuit?
4. Or a complaint with a privacy protection authority? In which country?

4.11 Government, surveillance and coercion:

1. Do you think the Swiss/Indian police or government can access the data you store online?
2. Would they need a court order?
3. Would you be informed if this happened? Should you be informed?

4.12 Coercion (US vs. Swiss, local vs online):

1. Could you be forced by the Swiss/Indian police to give your Gmail/Yahoo password? How about the password of your laptop? Would they need a court oder?
2. How about the US police or government? When could they access your data? Would they need a court order?
3. Do you think the technology exists for you and a friend to communicate electronically and exchange data without any other party being able to decrypt or see your communication?
4. Is such a technology possible? Why not?

4.13 The right to privacy:

1. Do you think YOU should be able to protect the privacy of your data and communications, whether stored locally or online?
2. Do you think EVERYBODY should?
3. Do you think TERORISTS should?

4.14 Regulation:

1. Do you think somebody is responsible to check that your online storage provider does not sell your data and that they apply appropriate data protection levels?
2. What data protection laws do you think apply to the personal data you store online? (e.g., Swiss, internationals, EU, US?)

3. If your data is stored in another country, e.g. the U.S., do you think that you will have the same rights & privileges as U.S. citizens or do you think special rules apply to you because you are located in Switzerland?
4. Would you like to be able to insure the data you store online, in a similar way that you insure your car or the assets in your home? If something bad would happen and you would lose your data you would be reimbursed by your insurance company.

Online Survey Questionnaire

1. **Where do you consider your private data to be safer: on your computer or stored online (for instance as email attachment)? Order the following arguments from 1 to 6, according to their relevance for you, where 1 is the one you most agree with.**
   On my computer, because I can look after it and physically protect my data, whereas online I cannot see where it is actually stored or who has access to it.
   On my computer, because I can disconnect it from the Internet, whereas online it is always exposed to hackers.
   On my computer, because hackers target big companies. They would need to identify my computer first, and they don't know where I am.
   Online, because my computer might crash or somebody might steal it and then I would lose all my data, but if I put it online I can always access it.
   Online, because big companies have more security experts and can guarantee better protection than what I could do for my laptop.
   Online, because on those servers there are many documents, from many users. Nobody would have the time to look at mine.
   Select the correct option
2. **When you delete a file stored on the Internet or an email in your Webmail account, what do you think happens?**
   The file gets permanently deleted just as when I would delete it from my computer.
   Copies will still be kept for security reasons, in case they are ever needed in criminal investigations.
   Some copies might still exist, but only for a few weeks, until the company manages to delete all of them.
   I don't know.
   Other (please specify)
3. **Google began in January 1996 as a research project by Larry Page and Sergey Brin. Its initial public offering took place on August 19, 2004. In which year did the initial public offering of Google take place?**

   1996 1998 2004 2006 2011
4. **You want to open a new account with a company that provides storage space for personal documents on a server on the Internet. You have come across these two companies. Which one do you choose and why?**
   Company A: Offers the service for free, but their privacy policy says that they may sell, transfer or share your personal information and documents to another company.
   Company B: Asks you to pay $20 per year. Their privacy policy says that they will not sell, transfer nor share any of your personal information to another companies.
   Company A, because it is free.
   Company A, because I don't have sensitive data anyway.
   Company A, because I can never be sure what they do with my data anyway.
   Company B, because I value my privacy.
   Company B, if the price was lower.
   Company B, if I am sure they are trustworthy.
   Other (please specify)
5. **Does your Webmail provider have the right to see or modify the documents you have as attachments in your email account?**
   They don't have the right to look at nor modify any of my documents.
   They can see them, but not modify them, because these are my documents and they belong to me, even if I store them there.
   They have the right to see and modify my documents only in criminal or terrorists cases.
   They have the right to see and modify any of the documents I store.
   I don't know.
   Other (please specify)

6. **Does your Webmail provider have the right to disable your account?**
   Yes, at any time, without advanced notice and without explanation.
   Yes, but only with advanced notice and a valid reason.
   Only if I am using it for criminal purposes.
   No.
   I don't know.
   Other (please specify)

7. **If your Webmail provider lost some of the data you store with them, what would your rights be?**
   They should pay me for the damages, regardless whether it was a paid for or free service. We had a contract.
   If it is a free service, I have no rights, but if I paid for it, they would have to pay me for the damages.
   I have no rights even if it is a paid-for service. There are no guarantees.
   My data is lost anyway. I wouldn't care about money. An apology would be enough.
   I don't know.
   Other (please specify)

8. **How much do you agree with each of the following statements?**
   Mark on the likert scale(Strongly agree, somewhat agree, somewhat disagree, strongly disagree, N/A)
   I try to keep local backups of every important document I store on the Internet.
   I try not to store important, sensitive documents on the Internet, and instead keep them offline, on my personal computers.
   Most businesses handle the personal information they collect about customers in a proper and confidential way.
   If people put their private data on the Internet and it gets hacked, it is their own fault. They should know that nothing is really safe on the Internet
   There is no such thing as consumer protection service or police on the Internet whom I could turn to, if I felt that my rights were violated.
   If the government had access to every document users store on the Internet, that would be a major violation of individual privacy.
   Consumers have lost all control over how personal information is collected, circulated and used by companies.
   Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today.
   It is good if the government monitors every Internet communication and all user accounts. National security comes first.

9. **How important do you consider the data you store online (data that is private, not for public access on the Internet)?**
   Mark it on a likert-scale one to seven:
   Very Important to Only spam or things I can live without
10. **What is your age?** 18-24 25-34 35-44 45-54 55+
11. **What is your gender?** Male Female
12. **What is your nationality?**
13. **What country do you live in?**
14. **What is the highest education degree that you completed?**

    High school Bachelor Masters PhD Other (please specify)

15. **How would you rate your computer skills?**

    Novice Intermediate Proficient Expert Comment

16. **What Webmail accounts do you use?**

    Yahoo Mail Gmail Hotmail AOL Other (please specify)

17. **Which is the main Webmail account you use?**
18. **Do you use any of the following systems for storing your documents online?**

    Dropbox FolderShare GoogleDocs Other (please specify)

19. **Three survey participants will be randomly selected to win a USD 100 Amazon vouchers. If you want to take part in the lucky draw, please specify an email address or phone number where we could contact**

you. All data collected during this survey will be anonymized and aggregated. Your answers are treated confidentially and used for research purposes only. We will not use your contact information for any other purposes but to contact you to collect