"I'm hoping they're an ethical company that won't do anything that I'll regret": Users' Perceptions of At-home DNA Testing Companies

Khadija Baig
Carleton University
Ottawa, Canada
khadija.baig@carleton.ca

Reham Mohamed
Carleton University
Ottawa, Canada
riham.mohamed@carleton.ca

Anna-Lena Theus
Carleton University
Ottawa, Canada
anna.theus@carleton.ca

Sonia Chiasson
Carleton University
Ottawa, Canada
chiasson@scs.carleton.ca

ABSTRACT
At-home DNA testing has become increasingly popular due to the ability to be able to gain both ancestry and health information, as well as connect with others who share your DNA. Do users have reasonable mental models of how these systems work? Do users have privacy concerns and what do they understand as the benefits and risks involved? We conducted 27 interviews with Canadian users of at-home DNA testing companies. Our interviews covered perceived and desired data use, data management, data sharing practices, control over data, and any regrets. Our qualitative analysis revealed that many users have inconsistencies in their mental models and liken their DNA data to their data stored with existing technologies, such as social media, rather than health data. They are generally either dismissive of privacy concerns towards themselves or their relatives or they had not considered privacy in their choice. We discuss our findings and propose possible future work in this area.

Author Keywords
Privacy; at home DNA-testing; interviews

CCS Concepts
•**Security and privacy** → **Human and societal aspects of security and privacy;**

INTRODUCTION
Direct-to-consumer (DTC) or at-home Deoxyribonucleic acid (DNA) testing has recently gained popularity. At-home DNA testing companies provide ancestry or health-related information for consumers that is thought to be beneficial [24, 40]. For instance, they may help adoptees learn about their biological families [24], or identify health concerns to prompt users to reduce risks of some diseases [17, 24, 37, 40]. On the other hand, there is no clear evidence that such results promote more positive health behavior [17]. Moreover, the emergence of at-home DNA tests has resulted in many privacy, ethical, and legal questions [1–3, 6, 9, 10, 14, 15, 21–23, 25, 33, 35, 36, 45].

The intrinsic characteristics of DNA make it personally identifiable information [2, 22, 40], even if stored anonymously [21, 40]. However, the privacy policies from at-home DNA testing companies are unclear on how they store DNA samples (e.g., [36]). Moreover, these databases are often shared with third parties [27, 28, 38, 40], which poses consent issues; DNA testing results reveal details of the individual and of other family members [3, 15] who have not consented to such testing.

Further, social media users share sensitive DNA information online while unaware of potential consequences [33]. These disclosures could be used by unsolicited third parties, such as health insurance companies, to genetically discriminate against individuals [33].

Existing laws to prevent genetic discrimination inadequately protect the privacy of individuals and better data governance is needed [1, 25, 45]. For example, some laws in the US [8] and Canada [49] prohibit genetic discrimination, but do not apply to all situations. US laws do not apply to using DNA test results for life insurance cases [25, 45] and Canadian laws do not apply to scientific or pharmaceutical research [1].

Given these privacy and legal concerns, it is important to understand individuals' mental models towards at-home DNA testing services. Previous research [5, 16, 17, 24, 37] has looked at users' motivations and desires for requesting the service, and their awareness of such services. However, there is a lack of a deeper understanding of individuals' mental models of the service and potential risks or privacy concerns. For example, there is no literature on users' perception of how companies

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376800

handle their sample, how third parties access their data, or how comfortable individuals are with such access. Moreover, there is a lack of exploration of individuals' needs, expectations, or a potential regret over requesting the service.

To address this research gap, we conducted 27 interviews in Canada with individuals who have analyzed their DNA through DTC companies. We explored users' privacy perceptions of these companies and probed aspects like data management, regret, and privacy. Users underestimated risks and were generally ambivalent towards privacy, especially if it impeded their primary goal. Users expressed desire for complete transparency over the handling of their data and wanted some level of control. We discuss implications of our results and highlight a need for transparency from DTC companies.

BACKGROUND

Over 12 million Americans have already requested tests from 23andMe and AncestryDNA and the number of consumers continues to grow [28]. Such DNA testing companies provide ethnicity, geographical ancestry, or health screening [16].

Perceived Benefits

These services have many benefits. For example, DNA databases can help families reunite, especially those searching for their biological parents [24] or extended families.

The health screening provided by these companies can also help users learn about their medical history and the probability of getting certain diseases [40]. Learning this information can in theory help users reduce risks of such diseases [16, 37]. However, there is controversy surrounding the purported health-related benefits [26, 37], and no evidence supporting a resulting positive behavior change [17]. The link between DNA and some diseases is also in question [22]. In addition, with such large and diverse samples of DNA, companies utilize their databases to conduct medical research that can contribute to the treatment of many serious diseases, such as cancer or Alzheimer [40].

These databases are useful to law enforcement in solving cold cases or identifying criminals [18, 39, 43]. However, the accuracy of DTC DNA tests is questionable, which increases chances of mistaken identity [18]. Moreover, users might have been unaware that their DNA data will be shared with law enforcement when they requested the service [18].

Privacy and Ethical Concerns

However, several privacy concerns exist [2, 3, 3, 6, 9, 10, 15, 21–23, 33, 36, 45]. First, the genomic information stored in DNA inherently introduces many privacy and ethical concerns [2, 22]. The human genetic sequences that make up the DNA is *unique* to most people, and represents a person's *identity* [2, 22, 40]. Therefore, improper use or leakage of such sensitive information can lead to irreversible consequences since a DNA cannot be replaced or revoked [2, 22]. Moreover, even if DNA samples are anonymized, individuals can be re-identified [21, 35, 40, 45]. Further, privacy policies often fail to clarify how DNA samples will be handled by the companies [35, 36].

Second, DNA databases are often used by third parties [27, 28, 38, 40] such as pharmaceutical and drug companies [28], and law enforcement [18, 39, 43]. While some of these partnerships are claimed to be for the advancement of medical research [28], sharing such sensitive data can be problematic due to the networked privacy reasons and limitations of existing legal governance, discussed next.

Third, having DTC companies analyze, store, and share DNA information with third parties introduces issues related to consent and *networked privacy* [3, 15]. DNA tests can reveal sensitive information belonging to other family members/relatives (including deceased or unborn ones) [3]. These individuals have not given permission for others to learn about their ancestry or health profiles [3, 40], or to share such information with third parties [15] or on social media [33]. Yet, many privacy policies do not articulate clear policies around consent [35].

Finally, misconceptions by the general public regarding DNA tests might lead them to inadvertently disclose sensitive information on social media [33]. In many cases, users express regret over some of their previously shared social media posts later on, especially when their posts relate to strong sentiments [50] or reveal too much information [41]; it is plausible that they may also regret disclosing sensitive DNA results.

Legal Governance and privacy policies

Unlike DNA tests requested by a doctor, at-home DNA tests lack governance from laws that adequately protect the privacy of individuals [36]. Although some laws to prevent genetic discrimination exist, they have limitations [1, 25, 45]. In the US, The Genetic Information Nondiscrimination Act [8] prohibits genetic discrimination with respect to employment and health insurance. However, it does not extend to disability and life insurance [25, 45] nor to employers with fewer than 15 employees [25]. The Genetic Non-Discrimination Act [49] in Canada similarly prohibits some companies and employers from using genetic test results. The act does not apply to those providing health-related services or to pharmaceutical or scientific researchers [1]. It is recommended that users understand the risks to their data by reading and understanding the terms and conditions/privacy policies [36]. However, the readability and clarity of these policies is low [14, 35] and they pose unresolved privacy and trust issues [14], since they are often biased in the companies' favor [35]. Moreover, the companies are free to change their privacy policies at any time [22], further raising trust concerns.

Individuals' Experience, Motivations, and Concerns

Previous work [5, 16, 17, 24, 37] has explored motivations behind individuals requesting at-home DNA testing. Results from studies with early adopters [16, 37] show that individuals were motivated to learn about their health, but had misconceptions about the service. Participants were also willing to share their data with medical practitioners for recommendations on health improvement [16]. Later studies also show that most individuals' motivations centered around the wish to learn about the risk of specific diseases [17, 24, 37]. Other reasons for requesting the service include curiosity [16, 37, 47], contributing to research [47], or the desire to learn about biological families by adoptees [24]. Users' awareness of risks

related to such tests was still low, although had slightly increased in some study conditions and contexts, depending on participants' demographics [37] .

Several researchers have considered privacy issues relating to at-home DNA testing [2, 6, 9, 10, 15, 21–23], and others have proposed technical frameworks that might mitigate some of these issues [22]. However, many issues remain unresolved. Previous work considering consumers' experiences focused on understanding why they request the service and mainly consisted of survey feedback. There is a lack of qualitative research to understand consumers' recent mental models of the topic, especially to explore consumers' attitudes, concerns, or potential regrets after requesting an at-home DNA test.

METHODOLOGY

Our study explores users' privacy perceptions and mental models of at-home DNA testing companies and their services. We use an online questionnaire and an interview guide for data collection. We explore: 1) What do participants think is happening with their data? E.g., who can access it, who has control over it, and for how long is it being stored?, 2) Do participants have privacy concerns?, 3) How would they like their data to be managed?, 4) Do participants have any regrets in relation to their DNA testing?

The study was cleared by our Research Ethics Board. After iterative design of our research instruments, we pilot tested the study with two undergraduate students; one had reasonable knowledge in security. We found no major concerns that resulted in changes to the study design.

Recruitment and Participants

Recruitment notices mentioned only at-home DNA testing to avoid self-selection by users particularly interested in privacy or regrets. Participants were recruited through posters placed across campus and in public places within our city (e.g., public libraries). We also posted the study on a social media page for advertising the university's research studies and on groups dedicated to discussing this type of DNA testing service (after getting permission from the group owner). We further posted the study on online service-exchange platforms, such as Kijiji. Other recruitment methods included snowballing and email to acquaintances who further distributed recruitment materials. We recruited only participants who: 1) were over 18 years of age, 2) were residing within Canada, 3) were comfortable communicating in English, and 4) had used and received the results of at least one at-home DNA testing kit. These criteria were checked through an online screening questionnaire (described in the following section). Participants received $20.

We recruited 27 eligible participants (17 female and 10 male) with most participants (n = 15) aged between 18 – 34 years old. In all, 7 were recruited on campus. The remainder were recruited via Kijiji or Facebook groups (n=20). Participants had at least a college degree, and the majority had completed either an undergraduate or a graduate degree/certificate. Participants worked in various fields, such as business, education, medicine, science/engineering, administration, art/journalism. Others were either retired, unemployed, or students. More participants had completed at-home DNA testing for ancestry purposes (n = 27) than for health purposes (n = 9). The majority of participants (n = 26) had received their test results within the past 3 years. Ten participants had completed multiple DNA tests. Out of 27 participants, only 4 refused to have their genetic data used for research, and 12 chose not to share it with websites.

Participants read and signed a consent form that explained the study's purpose and procedure, and requested permission for audio-recording. Personally identifiable information was limited to participants' voice. Participants were assigned a pseudonym that is not linked to their identity (e.g., P1–P27); these pseudonyms were used during data analysis.

Procedure

We had two main steps to complete the study: an online screening questionnaire, and an in-person session consisting of a pre-interview questionnaire and an interview.

1. **Screening:** We emailed interested potential participants a link to a screening questionnaire (hosted on Qualtrics [1]). The questionnaire took less than 2 minutes to complete and only those who qualified were contacted for an in-person session. The questionnaire had its own consent form and collected potential participants' demographics, whether they have used at-home DNA testing services, and the number of time they used such a service for *ancestry* or for *health* purposes. It asked them to select the companies used from a pre-populated list of 13 companies; they could add other companies or select that they have not used such services.

2. **In-person session:** The sessions were conducted in-person or online using video-calling software, such as Skype or Google Hangouts, and were audio-recorded. Each session lasted approximately 60 minutes. It started with a *pre-interview questionnaire* that participants completed through Qualtrics, followed by a semi-structured *interview* where the researcher collected participants' feedback verbally.

Pre-Interview Questionnaire

The pre-interview questionnaire consisted of the following categories of questions (all questions included 'Prefer not to say' option):

Demographic questions: We collected participants' gender, age, level of education, and occupation.

At-home DNA testing history: we asked: (1) how many times they have completed an at-home DNA test for *ancestry* and *health* purposes, (2) when they took their most recent test, (3) whether they have given consent to have their genetic data used for research, (4) with which websites, if any, they have shared their raw DNA data.

Access to their most recent data: We considered *ancestry* data and *health* data separately. We asked if they believed 12 different entities had access to their data These were a 5-point Likert scale questions ranging from 'no access' to 'definitely has access'. Only the 9 participants who completed a health DNA-test answered the *health* section.

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

Preference for access to data: Participants rated their comfort with sharing their data with the same 12 entities. We considered *ancestry* and *health* data separately using 5-point Likert questions ('very comfortable sharing' to 'very uncomfortable sharing'). Only the 9 participants who completed a health DNA-test answered the *health* section.

Purposes of use: We asked participants to rate how comfortable they are with having their data used *for research* by 7 different entities. These were 5-point Likert questions ('very comfortable' to 'very uncomfortable').

Reflection on use: Using 5-point Likert questions ('very likely' to 'very unlikely'), we asked participants whether they would change anything about their decision to test their DNA if given the opportunity. We specifically asked about 9 options and included the ability to add others if needed.

Reconsidering their decision: We asked participants if they had any regrets that would cause them to reconsider their decision to take a test. Participants could select one or more of the provided options, including that they had no regrets or adding their own if it was not in the list.

Interview Questions

The semi-structured interview expanded upon the questionnaire, and consisted of questions surrounding participants' 1) DNA testing history, 2) understanding of the process, including storing, sharing, and deletion practices, anonymization procedures, and control over their data, 3) perceived risks, including third-party access, data leaks, privacy risks, and impact on others, and 4) overall experience, expectations, and regrets. The interviewer had a script with 34 main questions along with probes to elicit more detail if the participant needed prompting.

ANALYSIS

In total, we collected 15 hours of audio, which was transcribed using the online transcription software, Trint [2]. The transcripts were then edited by the research team to ensure accuracy. Interview data was analysed using inductive thematic analysis [12]. While editing the transcripts, the first author made note of any themes occurring in the data. These 24 initial themes were then used as the basis for the first code book. Three transcripts were coded with this set of codes, then the themes were revised. The second codebook had 59 finer-grained codes. The first author and a second researcher then coded 9 transcripts (including the initial 3) using the second codebook and adding any missing codes. This resulted in a final codebook of 19 themes subdivided into 85 codes. The researchers then each re-coded these 9 transcripts with this third codebook. At this point, the percentage agreement for the transcripts coded by the two researchers was 97.72% overall. The two researchers then met to clarify any misconceptions, and agreed to continue analysis using the third codebook. Finally, each researcher independently and separately coded another 9 transcripts each. In summary, one third of transcripts were coded multiple times by two researchers until agreement was reached on a suitable codebook, then the remaining two-thirds were evenly divided and independently coded by one researcher each.

[2] https://www.trint.com/

RESULTS

We first summarize the questionnaire data, followed by the results of our qualitative analysis by theme, noting links between themes as appropriate.

Questionnaire Results

We focus on two sections of the questionnaire, namely which entities users perceive to have access to their DNA data and with whom they would be comfortable sharing this data if they had the choice. A wording error on the Likert scale labelling for the health data questions meant that we unfortunately only collected responses for these questions from 7 of the 9 participants who had completed health testing. Due to this small number of participants for the *health* questions, we did not conduct inferential statistics between the *ancestry* and *health* results. Instead, we present descriptive statistics to provide context for the following qualitative analysis.

Figures 1 and 2 summarize this Likert scale data. The numbers in the cell represent how many participants selected each Likert scale option (grid columns) per entity (grid rows). Cells are colored to represent popularity; darker shades represent a higher number of participants selecting this option.

For example, Figure 1 shows that most users find it highly unlikely that their employer has access to their ancestry data, but many believe it likely that other users of the same ancestry service do have access. It is noteworthy that roughly half of users believe that third-parties such as insurance companies, advertisers, for-profit companies, and researchers do not have access to their data ('no access' or 'probably no access'). Similar patterns emerge for the health data — most users do not think that their health DNA data is shared.

In Figure 2, most users are very uncomfortable with advertisers and for-profit companies having access to their ancestry data; they are generally uncomfortable with giving access to most third-parties. Notable exceptions are that they are generally comfortable sharing with other users of the ancestry system and with academic researchers. For health DNA data, participants were generally uncomfortable sharing with anyone, except that they were comfortable with sharing their health data with academic and non-profit researchers.

Perceived Benefits

Greater good: Participants perceived several benefits to both themselves and for the greater good as a result of completing an at-home DNA test. Most participants thought their data could be used to advance humankind in some way via research, either because their *"DNA holds information that would be beneficial for medical research"*(P22), or *"just [for] improving knowledge of science"* (P9). By analysing DNA data, researchers could potentially *"figure out ways to cure some of these diseases that people have"* (P10), perhaps by identifying better links between illness and genes, or by improving existing medical treatment or gene therapy.

Academic institutions were most trusted for this research, with pharmaceutical companies as a close second (contrary to the questionnaire responses). That being said, it was occasionally mentioned that giving pharmaceutical companies control

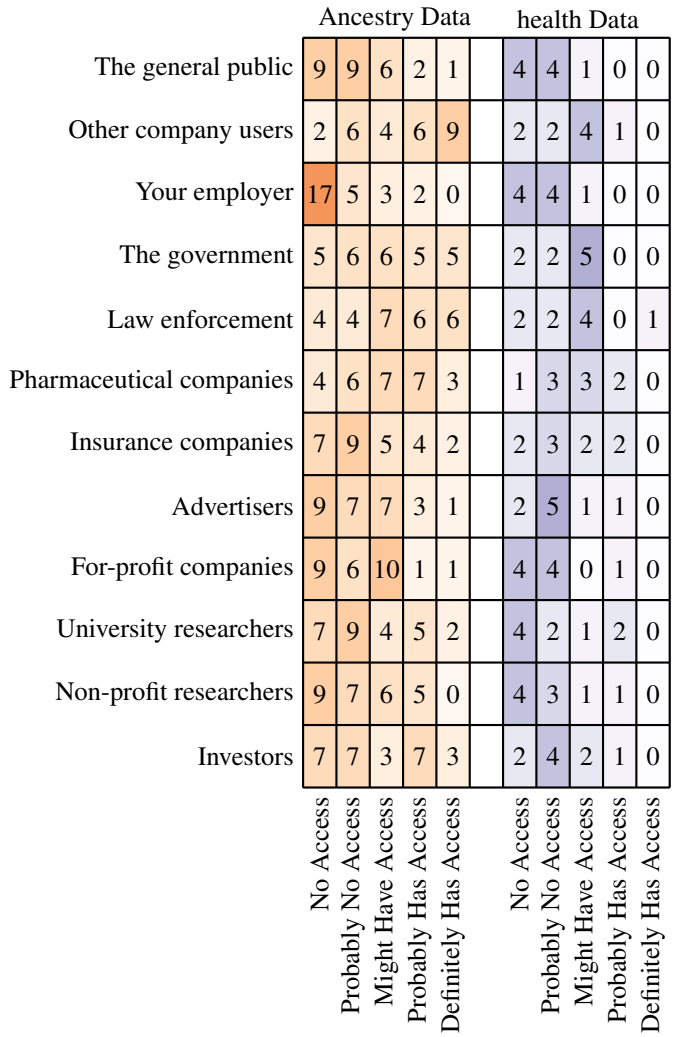

Figure 1. Perceived likelihood that various entities have access to participants' most recent ancestry and health DNA data.

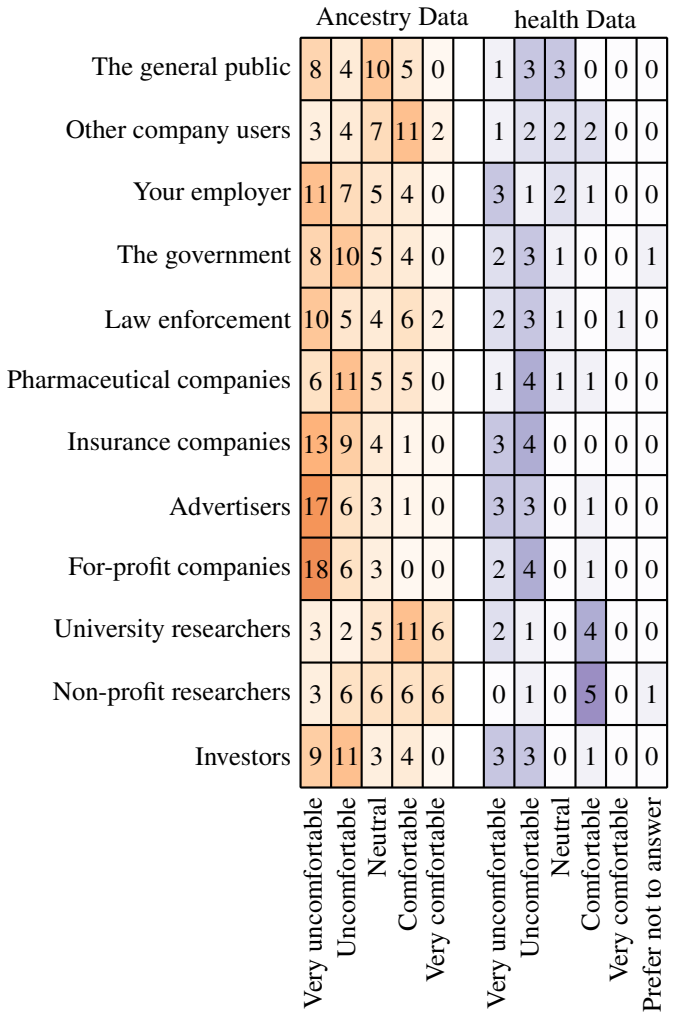

Figure 2. Participants' comfort level with sharing their most recent ancestry and health DNA data with various entities.

might backfire since they are for-profit: *"There's also a catch that a lot of pharmaceutical companies are for profit so it could make it that they create something but they won't distribute it at a fair price"*(P24). Some participants also believed the Canadian government could play a role in conducting this research, considering Canada's universal healthcare system.

Law enforcement have also used DNA data to aid in their investigations, which participants generally perceive as positive. As one participant describes it,*"If my DNA can solve a crime or puts someone in jail that... deserves to be there and has done the crime, or [if it] exonerates somebody that hasn't done the crime, then I'm all for it"*(P10).

Connecting families: Some participants have used at-home DNA testing kits to help them find family. Others have been contacted by users who are either building their own family tree or searching for family. The ability to answer questions relating to family history, either for themselves or for others, is generally seen as favourable. As one participant recalled, *"Someone contacted me. She's related to my mom's side but she never knew her father growing up and then she found family. She didn't know who he was. Now she's there and she knows people so it's kinda cool"* (P14). One participant did, however, comment on the privacy implications of contacting family who may not want to be contacted: *"These [stories] can be positive. But even there, did the biological mom want to be found by the biological kid? I don't know"* (P6).

Achieving goals: Participants who had completed only a single at-home DNA test had usually done so simply out of curiosity, or for fun. Participants who had completed more than one at-home DNA test generally did so to achieve personal goals. Common goals included genealogy research (building of a family tree), the search for family connections, and the desire to understanding the participant's family history. Even these participants acknowledged the entertainment value of the testing. A participant whose main goal was gathering genealogy data commented, *"It's just a different type of puzzle solving... it really is just a form of entertainment and curiosity"*(P4).

Perceived Risks

While many participants admit to not having read the terms and conditions, very few considered this lack of knowledge as a risk. Many also noted that they would find it concerning if their data was being used without their consent, or in ways in which they don't personally agree. However, most participants did not feel much risk to their privacy if their data stayed with the DNA testing company and was stripped of identifiers. Over half of participants commented that they were unsure how their DNA data could be used in a negative way.

Criminal implications: Participants also recognised that their data being used for criminal investigations was a risk. Should someone commit a crime and leave behind DNA evidence, law enforcement could access an at-home DNA testing company's database to help find the perpetrator, even if the perpetrator had not taken a test themselves. Most participants, however, did not view this as a personal risk, and, in fact, had no sympathy for individuals who had broken the law: *"If any of my relatives did anything, I would want them to be caught"* (P20). Using data for this purpose, however, could also lead to incorrect charges being filed; some participants expressed concern about being wrongly associated with a crime. DNA evidence could be incorrectly linked to them accidentally (due to inaccuracy or contamination of the data), or intentionally (having been planted by someone at the crime scene).

Misuse: Many participants also worry about data misuse by a third party, or the risk of their data being used in an unethical manner. Data being sold and the misuse of DNA data to instigate discrimination of any kind were commonly deemed unethical. Participants fear that their DNA data might result in discrimination when applying for jobs or life insurance, with one likening it to employers using Facebook as part of their hiring process (P6).

Profiteering: Some participants believe that *"pharmaceuticals have a bad rep"* (P23), speculating that the companies might do more harm than good. A participant explains their mistrust: *"Because a lot of them are for profit, they could take that [DNA] information and use it to up their prices for something that they know people will need or they could purchase the rights to some sort of medications so that they're the only ones who sell it so they can get all the money for it"* (P24).

Surveillance: Participants are also concerned that the government might use the information for surveillance purposes or to keep track of its citizens.

General data disclosure: Over half of participants expressed discomfort at the idea of their data being publicly available, or being used for purposes with which they disagree... in spite of not knowing how exactly their data could be used against them. *"Information is out there that I wasn't looking to have out [in the case of a hack]. I don't think there's anything too big. Just that information is out there that I didn't personally want people to know about"*(P24). The term "cloning" was frequently mentioned, with participants feeling like the technology isn't ready yet to allow for true misuse of data.

Trust/Mistrust

"I think there has to be a level of trust when you are using a service such as this" a participant stated, when discussing whether data is being sold or not (P19). They continue *"I'm sort of hoping that they're an ethical company that won't do anything that I would regret"*. Feelings towards the company are mixed, with some being more optimistic, hoping *"they must have like internal policies on how they would treat it (the data)"* (P17), while others lean more towards cynicism, commenting that *"you don't know who's looking at your data... you don't know who they are. Like, you've never seen them"* (P8). Most users trust the company to totally strip identifiers from their DNA data. This appears in several themes in the data; participants desire anonymity and hope that their data cannot be traced back to them. While many consider a company's terms and conditions as legally binding, there are some who are wary of a *"legal loophole"* (P13) that might result in their data being mismanaged.

Country of residence: Participants also feel relatively more secure residing within Canada. Several stated that the United States *"is far more loose with their data than Canada is"* (P27). Participants expect Canada to have privacy laws that are *"a little bit more stringent"* (P4), and commented on how *"nobody [in Canada] had ever asked yet 'have you taken a DNA test?"* (P27). Some participants expressed concern over not knowing how their data is actually handled, since most at-home DNA testing companies are not based in Canada.

Third parties: Academic and other non-profit institutions are generally viewed as altruistic. This is because *"they're searching for solutions to better conditions, you know. It's to help humanity to a certain extent"* (P13). In contrast, participants are most mistrustful of the government and law enforcement, while still acknowledging benefits to these entities having access to DNA data; attitudes towards one party is usually linked to attitudes towards the other, because participants tend to perceive that *"the government and law enforcement are little too close together"* (P14). Interestingly, many participants mistrust for-profit companies, despite identifying no specific risks associated with pharmaceutical companies when asked.

Comfort/Discomfort of Sharing Data

Benevolence: In keeping with previously noted themes, participants are happy to share their data for research purposes, providing the research is done to "help humanity". Participants thought that sharing data with academic institutions was acceptable. Most participants do not distinguish between types of research, but some specified that the research should follow a *"an ethics code"* (P21). Doing research for profit transgresses the boundaries of what is considered the "proper purposes", and many deemed it unethical. Several were very uncomfortable with their data being used for profit, *"I feel like that would be illegal"* (P23). Interestingly, only a few were uncomfortable with pharmaceutical companies; which do have a for-profit nature, suggesting that many had not fully considered the implications.

Personal sharing: Almost all participants shared their results with family and friends. Those actively involved in genealogy work were more likely to post on social media about it since

sharing the data with others on DNA-focused groups can help participants in their research. Most participants, however, have not otherwise posted on social media. Those who posted said it was because they were constantly being asked about their ethnic background and posting meant that they *"don't have to answer that question anymore"* (P22), or because they received bizarre results.

Government and law enforcement: People were less inclined to share their data with law enforcement and the government. While discussing the benefits to law enforcement having their data, a participant explained, *"I wouldn't want my DNA to be in it, but... I would want yours to be in it... like, other persons. I feel like it could help in their investigation... but... I would feel like I'm being watched all the time"* (P8).

Data Management

We asked participants about their perceptions of data storage, data access, data deletion, and control. A general theme of uncertainty is prevalent throughout this section.

Control

Control by others: Most users acknowledge that the company has the most control over their data, in terms of how long it is stored for, who gains access to it, and what it is being used for. Some participants are unsure of who might have control, but are certain that *"it's not me"* (P4, P8). Users also felt that any third parties with access also have control over the data.

Self control: In contrast, participants who considered themselves as having partial or complete control over their data did so because they believe that they have the ability to delete their data whenever they like. They also firmly believe that the company would not sell or share their data without their explicit consent, in spite of some acknowledging that they have no way of verifying that the company meets its legal obligations. Some participants also perceive the ability to take precautions, and to download their raw DNA data, as a form of control.

Transparency: Participants want total transparency from the company, and communication that is *"user-friendly"* (P24), and *"straightforward"* (P13). One participant outlines, *"... just like a one pager, point blank, this is who we're sharing with and why. Not a 20 page 'do you agree terms and agreements' that normally happens and... no one reads them, so you scroll through to the end and hit accept"* (P24). In addition, participants desired full control over their data, being able to explicitly opt-in and opt-out of whether their data gets shared, and to whom. Some participants *"want to be able to choose each time"* (P15), while others felt that *"there's probably a lot of third party ones... it would probably be like hundreds of e-mails every day"* (P6), and would instead prefer to choose categories of acceptable third parties.

Data Storage

Many participants *"have no idea"* (P1, P7, P8, P21, P23) for how long their data is being kept, or what happens to it after the results are processed. Most think their data is stored indefinitely, while some assume it is stored for a certain number of years, as *"seems standard"* (P2). While some participants believe that their physical sample probably gets destroyed, others believe otherwise because their test results get updated: *"I know that they keep testing it over and over again. I think it gets kept in their lab."* (P14).

We saw mixed responses on how long data should be kept. Many participants are indifferent to how long their data is stored, provided it was *"stripped out of all the identifiers"* (P24). Some are only comfortable with it being stored for a certain time, while a few do not want it stored at all beyond processing their results. There was also more general discomfort around having a physical sample stored compared to digital records. Overall, the length of time was less concerning than purpose of use. One participant explains, *"I don't really care. For me it's more about consent and who it's being shared [with], what it's being used for"* (P13).

Data Access

Users had diverse perceptions about whether the companies share DNA data. Some believed that the companies are ethically bound to not sell the data, while others believe that sharing happens, either with consent or behind people's backs. The government and law enforcement are perceived as most likely recipients, with pharmaceutical companies following very closely. Some participants were unaware which third parties had access, but recognised that this information was probably listed in the terms and conditions.

Consent is key when it comes to sharing data. Many individuals specifically don't want anyone profiting from their data, while others are uncomfortable with any third party access. Trust in the company is a recurring theme, as several participants "hope" the company is not selling their data, but are unsure how to verify. One participant states, *"I hope not but I have no proof if they do or do not. So I hope they don't"* (P10).

Data Deletion

Deletion is impossible: Although uncertain, participants do not believe that complete data deletion is truly possible. *"Can you [delete your data] indefinitely? If you delete your account, probably your data is still alive, right?"* (P6). Participants likened this information to *"most information that's on the Internet"* (P4), believing that online information never dies. They believed that account deletion might remove any personally identifiable information, like name and email address, or the link between the account and the data. Some participants were even more skeptical, with one stating *"there would still be a ghost account linked to you. Without you knowing it. And even if you were to delete your account, the little tube that you shipped off will have your name on it because you wrote it and tagged it"* (P8). Because of this, some participants are less likely to delete their accounts because they *"don't really think you can delete it"* (P4). One participant who was "absolutely devastated" by their results recounts *"I tried to [delete my account] and couldn't figure it out. I work in digital [X]... I'm not an idiot, I feel like I should have been able to figure it out and it seemed overly complicated"* (P23).

Deletion is undesirable: Most participants also enjoy seeing updates to their data, *"I guess I like to see my DNA, like log [into] the Web site and see if there's an update"* (P12), and do

not see any need to delete their accounts. Only if their data was being used in ways that they considered unethical would they consider deleting their account, with a small number also likely to do so if they were highly unsatisfied with the results.

Deletion is reassuring: One participant said they might feel more secure about providing the company their data if they had the ability delete it entirely: *"Let's assume that... it would actually be deleted on all of their platforms and they would actually stop using it, then I wouldn't be worried"* (P23).

Regrets

Regrets were expressed naturally during the course of the interview, and these remarks form the bulk of our results. Users were asked specifically about regrets at the end of the questionnaire and interview, which may have biased responses; this is a limitation of this section. However, users who did not express much regret throughout the interview did not change their answers when directly asked. This is also reflected in their questionnaire answers.

Not reading the fine print: Many participants admitted to not reading the terms and conditions before their DNA test. This manifested for some as regret, as they wished they had taken the time to *"read the fine print"* (P5) and truly understand the meaning of the contract: for how long their data was stored, who had access to it, and how much control they had over it. One participant states, *"So your data is stored forever and... Like it is actually forever. It's not just like a word. It's actually forever. So I wish like I understood that before"* (P8).

Mismatched expectations: Other participants regretted not having done better research on what to expect with an at-home DNA test. This would include information on what kind of results to expect, what features are included, and accuracy of the results. *"I might not have spent the money on the test had I known that it wasn't just going to automatically populate my tree for me"* (P4).

Privacy doesn't seem to be the main cause of concern and regret for participants, with very few expressing regret at the idea of someone having access to their data without their consent. Only one participant mentioned wishing they had taken more protective measures. Curiosity is a powerful thing, and almost all participants stated that they still would have taken the DNA test for the sake of their ultimate goal, or to satisfy their curiosity. The main themes of regret revolve around having not been completely informed of the risks before taking the test, but ultimately this information would not have stopped them from using the service.

Protective Measures

A relatively simple method of taking precautions is to do some research on potential options. The scale of the research done by participants varied immensely, with some participants choosing a company that simply *"seemed the most trusty-ish out of all of them"* (P8), while others' choices reflected their privacy concerns. Most users, however, simply selected a company based on recommendations from friends or based on advertisements.

Careful choice of company: One participant only uses *"companies that are under EU controls for data"* because of the *"data protection rules in the EU"* (P15). Another participant first tested at a certain company because *"we know that company control[s] their own labs"* (P16).

Privacy is secondary: Interestingly, if participants' primary goal involved genealogy or family connections, these privacy protective measures were usually pushed aside when expanding their search to other companies. The same participant commented about two of the companies he had also later taken the test with, *"I didn't like... that companies like [X] and [Y]... don't have their [own] labs. They do it somewhere else"*(P16). The potential benefits of testing with a company participants' may trust less ultimately seems to outweigh any privacy concerns they might have. As one participant admitted, *"I was reluctant to test with [X] but I tested with them simply because of the size of their database"* (P27). One way to compensate for this perceived loss of control is to inform themselves of the terms and conditions of each company, to the best of their ability. When explaining why they decided to test with a company they don't trust, a participant mentioned, *"I fully understand what I agreed to. And that's why I specifically didn't agree to certain things on [X] where I did agree to it on [Z]"* (P27).

Fake information: A few participants used fake information and aliases to remain anonymous online and maintain their privacy. One participant even regretted not having done so, saying *"I would have signed up with totally different name from the get go. Different name, different email. Even on the package... different shipping address. I would want to know [the results as] accurate as possible but I wouldn't want it to be traced back to me"* (P8). Doing so could potentially protect them from certain negative repercussions of a data breach, as one participant articulated, *"If somebody came to me with the information that I've posted, I'd almost chuckle in a way. Because I know it's fake"* (P3).

Account deletion: Participants would also consider deleting their accounts if something goes awry. Triggers could include a data breach, bad press or a scandal on the company, or the knowledge that data is being used in an unethical way. In case of a breach, deleting your data may be a way to prevent data from being traced back to you: *"You might get scared... and in the case of a breach... you want to delete your account because you don't want it to get back to you"* (P8). Feelings of anxiety and fear are prevalent in this sort of narrative, with one participant recounting *"I was on [A] and I had this weird relative connection very close and it didn't make any sense... for a second I thought of someone hacked in. I don't know why [but] I got a little freaked out about it so I deleted it"* (P9).

Using privacy settings: To protect their privacy from other users, a few participants deliberately set their accounts to private. This change could sometimes follow the discovery of unwanted information, as one participant explains, *"when I found out the information in my family... I went private"* (P11).

Most participants, however, did not read through the terms and conditions, and continue to not read any updates they receive from the companies on their policy. Only one participant

mentions regularly reading through updates, saying *"every time there's an update on the use, I go take a look at it just to see exactly – if I can figure out what's been changed"* (P15).

Attitudes Towards Risk/Privacy

A recurring theme of resignation exists amongst users, with many participants believing that they no longer have any control over their data because they chose to give it away.

Trading privacy for health: Some have had to make the decision between potentially life-altering information and the right to their privacy. A participant who chose to undergo health-related at-home DNA testing describes having to put their privacy aside, quite literally, for the sake of their life, *"If it means that I have to share my privacy, fine. What can you do? You can hide all the time, you can hide under a rock and you don't get help"* (P1). Going through a doctor would have taken much longer, and the at-home DNA test was a faster, more convenient option.

Resignation: Others believe they no longer have a right to their privacy now that they've signed on the dotted line: *"I'm OK with that because it is my choice, I volunteered to give my DNA"* (P24). Pairing this with the trend of not reading the privacy policy means that although participants consented, they may not know what they consented to nor the associated risks. In addition, by drawing a parallel to the use of other technology, it is easy for users to normalise and justify the risk, saying *"Same thing with Facebook. You know that your photo that you deleted from three years ago is still around, right? And we still do it"* (P6). Others see no difference between DNA data and any other data, and some, in fact, may even expect their data to be sold, *"Everybody sells data, in essence that's [the] norm... It's something you can't avoid. It's become a part of life now"* (P1).

DNA is non-sensitive data: Many have no idea how their data could be used to affect them negatively because *"it's not, like, a super sensitive information... it doesn't really do any harm"* (P17). This feeling is usually compounded if they perceive themselves as less of a target (which many do), *"I'm extremely healthy. I'm retired. I'm not a criminal. I don't do anything bad so I don't have a problem"* (P7).

Disregard for others' privacy: Another interesting attitude observed is dismissal of others' privacy. Participants who perceive to have high control over their data, or have done multiple tests, are generally more dismissive of other people's concerns. The word *"paranoid"* was used by a participant (P10) to describe a family member's unwillingness to do a test. Other participants also disregarded other people's privacy in the face of their own research, *"I get mad when people have made their trees private... they're mining other people's information and they're being very selfish about the fact that they're, you know, it's just, it has to go both ways"* (P4). Participants may even purchase kits for their family to help with the genealogy research, even after acknowledging several privacy risks. Those with this sort of attitude tend to either have no privacy concerns of their own, or are highly aware but have decided to trade privacy in exchange for achieving their own personal goals.

DISCUSSION

Mental Models

Although it is a complex subject, a reasonable mental model by users would encompass elements from our Background section: understanding of the benefits, privacy concerns to themselves and others, the legal protections and their limitations. Our users could be broadly classified into 3 categories: those who were entirely uncertain of how their data was being handled, those who were certain their data was being sold and they had no control of it, and those who felt comfortable with the amount of control they perceived to have. We discuss these mental models below.

Uncertain: These users frequently responded "I don't know" or "whatever the contract says" to our interview questions. Several users were unsure of whether the data was being sold, how long it was being kept for, and who had access to it. For example, several "hoped" that their data was not being sold, while others would speculate as to what was happening with their data, often adding a disclaimer that they *"have no proof"* (P10) as to what is actually being done. Some participants even admitted to intentionally avoiding too much knowledge about the process because it might make them regret their decision. Other users expressed conflicting understanding at different points in the interview. For example, a user mentioned they *"hope [the companies] burn"* physical samples after testing, but later said they'd want their sample stored for *"as long as [the company's] fridges don't break"* (P15).

Powerless: Some acknowledged they were *"aware of the risk that data could be shared"* (P6), but this peripheral awareness of risks leads users to believe that they no longer have the right to privacy and/or control because they *"volunteered to give [their] DNA"* (P24). Due to this resignation, these participants were not as concerned about their privacy because they believe they don't have any control it.

Unconcerned: These users believed that their data could, in no way, be traced back to them. Several even outlined this as a precursor to sharing their data with third parties, and were confident that nothing could be linked to them, indicating an implicit trust in the company. A small number were also certain that they were in full control over their data. Many unconcerned users did not differentiate their raw DNA data from data stored by current digital technologies (e.g., social media, browsers). These users believe that nothing negative has come out of the large tech companies having access to their data, and extended this view to their raw DNA data: *"Google is already using all our data and it doesn't hurt me in any way"* (P22). This sense of security also potentially impacted others related to them. Many were entirely unaware that in providing their own DNA sample, they were also providing a chunk of DNA shared with their family. Some would not object to family doing the test because they themselves have already done it and *"nothing crazy happened"* (P14).

Underestimated consequences

Misconceptions: A common trend observed in the data is that many users underestimated the possible consequences and implications of having their DNA data and test results

available. Most users do not view that this data as especially sensitive information. They believed that anything shared with third-parties was de-identified, but recent literature [21, 36, 40] indicates that this is not the case. Privacy policies do not always specify whether the shared data will contain explicit identifiers (e.g., name) [14, 35, 36]; but in either case, it could still be traced back to users [21, 35, 40] since DNA is inherently identifiable. Instances of genetic discrimination have appeared in media [51], and Canadian border agencies have also made use of DTC genetic kits to assist in deportation investigations [4].

Canadian laws: Laws in Canada prohibit insurance companies and employers from using a DNA test to discriminate against individuals [49]. The Genetic Non-Discrimination Act was adopted in 2017 [34]. However, at the time of writing of this paper, there are legal challenges to this law [30]. It has been ruled unconstitutional by a provincial court and currently awaits decision from the Supreme Court of Canada [31]. Citizens of countries with weaker privacy laws may be at even greater risk.

In addition, making DNA test results public (e.g., by posting on social media) still poses a risk despite relevant law. For example, employers regularly do an online search for people they are interviewing [20,42,46]. These unofficial background checks can negatively impact hiring decisions [7,44,48]. These public results could also be used as part of profiling by advertisers or social media.

Adversarial attacks: Data breaches have the potential for the large-scale exposure of sensitive DNA information of individuals. While no such breaches have been reported thus far, breaches of other types of sensitive user data have occurred (e.g., [19, 32]. It is reasonable to assume that DNA databases would be prime targets. Attacks on DNA databases affect not only the consumer, but their genetic relatives as well.

It is possible to exploit certain DTC DNA company APIs and websites to compromise user data [11, 29]. The compromised information could then be used for impersonation of the user's relatives [29]. Techniques to identify individuals who have **not** taken an at-home DNA test also exist [13].

User perception: 'Bad things' could be happening without users' knowledge. We would not see this reflected in users' *perception* of security and privacy since they are unaware that anything went wrong in the first place. However, this false sense of security is worrisome since it suggests that users are largely unaware of the risks, are operating in ways that may further increase risks, and have no voice in what is happening to their sensitive data.

Transparency and Control

This study has uncovered the glaring need for transparency from at-home DNA testing companies. The most common desire expressed by users was for companies to be *"totally transparent"* and use *"user-friendly language"* (P6, P14). Participants expressed their dislike of the *"20 page 'do you agree' terms and agreements that normally happens"* (P24), because when they see *"the lawyer talk"* (P1), they feel distrust towards the companies.

Users should have the option to choose their preferred level of control. Some users wanted to provide explicit consent for sharing their information with each third party. One participant very aptly noted: *"That would be scary though, the number of companies... to see all the e-mails... that would be scary"* (P6). Others would like to provide access to certain *types* of third parties rather than being consulted at every instance.

Limitations and Future Work

The study could have a self-selection bias, since participants were interested in the subject. In addition, social desirability effect may be present due to the delicate nature of the topic; participants may have provided answers that they think the researchers wanted to hear. Our pre-interview questionnaire may have primed participants to respond in a way more geared towards privacy, especially about regrets. Because of these limitations, generalising the findings to the general community should be done with caution.

Our results highlight the need for research into stronger, more stable policy, and for better transparency and control of data practices for individuals. In addition, research must be done to address how to protect genetic relatives of consumers from related disclosure.

It would be interesting to further explore users' understanding of the legal protections surrounding DNA data, particularly in an international context.

We prioritized interviewing current users of at-home DNA testing since these users have already engaged with the system and provided their (and their relatives') sensitive data, and we wanted to understand why. However, in obvious contrast to this study, it would be interesting to hear the perspectives of individuals who have considered doing at-home DNA testing, but ultimately decided against it, and those who would never consider doing an at-home DNA test.

CONCLUSION

We conducted an interview study with 27 Canadian users of at-home DNA testing companies to gather their privacy perceptions and mental models of at-home DNA testing companies and their services. We explored users' expectations and needs in relation to data use, data management, data sharing, control, and regrets. Our analysis uncovered inconsistent, and sometimes contradictory, mental models that generally fell into three categories: those who were uncertain about most of the process, those who felt powerless, and those who are unconcerned about the risks. Participants largely underestimated the implications of sharing their DNA data. Privacy is often not the users' first priority, and most users admitted to being unaware of the terms they agreed to. They were often unaware of or ignored the privacy impact on family members. We further found that users compare their DNA to other digital data (for example, browser history, location), believe it to be relatively secure, and some evaluated risk based on their country of residence or perceived company location. Overall, users were mostly unaware of the risks and may further increase these risks through their own actions. Not reading the fine print is generally what they regretted most; and most users desired complete transparency and control over their data.

ACKNOWLEDGMENTS

We thank the participants for their time, and the reviewers for their feedback in improving the paper. We also acknowledge research funding from an NSERC Discovery Grant and the Canada Research Chairs program.

REFERENCES

[1] Canadian Civil Liberties Association. 2018. The Genetic Non-Discrimination Act - An Overview. https://ccla.org/genetic-non-discrimination-act-overview/?fbclid=IwAR3iNOmCaE--e6f_EMzhyGkYZCC3Ct7IbKri_dnri1djiHHO9DZhlZgnqTU#_ftnref1. (2018).

[2] Erman Ayday, Emiliano De Cristofaro, Jean-Pierre Hubaux, and Gene Tsudik. 2015. Whole Genome Sequencing: Revolutionary Medicine or Privacy Nightmare? *Computer, publication of the IEEE Computer Society* 48, 2 (Feb 2015), 58–66.

[3] Danah Boyd. 2012. Networked Privacy. *Surveillance & Society* 10, 3 (2012), 348–350. http://proxy.library.carleton.ca/login?url=https://search.proquest.com/docview/1314733037?accountid=9894

[4] Emanuela Campanella. 2018. Government using ancestry websites to deport immigrants: lawyer, court documents. https://www.wired.com/2016/02/schools-kicked-boy-based-dna/. (2018).

[5] Anna Childers. 2017. *Adoptees' Experiences with Direct-to-Consumer Genetic Testing: Emotions, Satisfaction, and Motivating Factors*. University of South Carolina, Scholar Commons.

[6] Peter A. Chow-White, Maggie MacAulay, Anita Charters, and Paulina Chow. 2015. From the bench to the bedside in the big data age: ethics and practices of consent and privacy for clinical genomics and personalized medicine. *Ethics and Information Technology* 17, 3 (Sep 2015), 189–200. https://doi.org/10.1007/s10676-015-9373-x

[7] Leigh A. Clark and Sherry J. Roberts. 2010. Employer's Use of Social Networking Sites: A Socially Irresponsible Practice. *Journal of Business Ethics* 95, 4 (2010), 507–525.

[8] U.S. Equal Employment Opportunity Commission. 2008. The Genetic Information Nondiscrimination Act of 2008. https://www.eeoc.gov/laws/statutes/gina.cfm. (2008).

[9] Emiliano De Cristofaro. 2014. Genomic Privacy and the Rise of a New Research Community. *IEEE Security Privacy* 12, 2 (Mar 2014), 80–83.

[10] Emiliano De Cristofaro, Sky Faber, Paolo Gasti, and Gene Tsudik. 2012. Genodroid: Are Privacy-preserving Genomic Tests Ready for Prime Time?. In *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society (WPES '12)*. ACM, New York, NY, USA, 97–108. http://doi.acm.org/10.1145/2381966.2381980

[11] Michael D Edge and Graham Coop. 2019. Attacks on genetic privacy via uploads to genealogical databases. *BioRxiv* (2019), 798272.

[12] Satu Elo and Helvi Kyngäs. 2008. The qualitative content analysis process. *Journal of Advanced Nursing* 62, 1 (2008), 107–115.

[13] Yaniv Erlich, Tal Shor, Itsik Pe'er, and Shai Carmi. 2018. Identity inference of genomic data using long-range familial searches. *Science* 362, 6415 (2018), 690–694.

[14] Tatiana Ermakova, Annika Baumann, Benjamin Fabian, and Hanna Krasnova. 2014. Privacy Policies and Users' Trust: Does Readability Matter?. In *20th Americas Conference on Information Systems (AMCIS '14)*.

[15] Julie Frizzo-Barker, Peter A. Chow-White, Anita Charters, and Dung Ha. 2016. Genomic Big Data and Privacy: Challenges and Opportunities for Precision Medicine. *Computer Supported Cooperative Work (CSCW)* 25, 2 (Jun 2016), 115–136. https://doi.org/10.1007/s10606-016-9248-7

[16] Sarah Gollust, Erynn Gordon, C Zayac, G Griffin, M.F. Christman, R.E. Pyeritz, L Wawak, and Barbara Bernhardt. 2011. Motivations and Perceptions of Early Adopters of Personalized Genomics: Perspectives from Research Participants. *Public health genomics* 15 (06 2011), 22–30. DOI:http://dx.doi.org/10.1159/000327296

[17] Stacy Gray, Sarah E Gollust, Deanna Alexis Carere, Clara A Chen, Angel Cronin, Sarah Kalia, Huma Q Rana, Mack Ruffin IV, Catharine Wang, J Roberts, and Robert C Green. 2016. Personal Genomic Testing for Cancer Risk: Results From the Impact of Personal Genomics Study. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 35 (12 2016), 636–644. DOI: http://dx.doi.org/10.1200/JCO.2016.67.1503

[18] Michael Grothaus. 2018. Police are using genetic testing companies to track down criminals. https://www.fastcompany.com/40564861/police-are-using-genetic-testing-companies-to-track-down-criminals. (2018).

[19] Alex Hern. 2018. Fitness tracking app Strava gives away location of secret US army bases. https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases. (2018).

[20] Debora Jeske and Kenneth S Shultz. 2016. Using social media content for screening in recruitment and selection: pros and cons. *Work, Employment and Society* 30, 3 (2016), 535–546.

[21] Bartha Maria Knoppers. 2010. Consent to 'personal' genomics and privacy: Direct-to-consumer genetic tests and population genome research challenge traditional notions of privacy and consent. *EMBO reports* 11, 6 (2010), 416–419. https://www.embopress.org/doi/abs/10.1038/embor.2010.69

[22] Jingquan Li. 2016. Genetic Information Privacy in the Age of Data-Driven Medicine. In *2016 IEEE International Congress on Big Data (BigData Congress)*. 299–306.

[23] Jingquan Li. 2017. Open Medical Big Data and Open Consent and Their Impact on Privacy. In *2017 IEEE International Congress on Big Data (BigData Congress)*. 511–514.

[24] Natalie M Baptista, Kurt Christensen, Deanna Alexis Carere, Simon Broadley, J Roberts, and Robert C Green. 2016. Adopting Genetics: Motivations and Outcomes of Personal Genomic Testing in Adult Adoptees. *Genetics in medicine: official journal of the American College of Medical Genetics* 18 (01 2016). DOI:http://dx.doi.org/10.1038/gim.2015.192

[25] Thomas May. 2018. Sociogenetic Risks — Ancestry DNA Testing, Third-Party Identity, and Protection of Privacy. *New England Journal of Medicine* 379, 5 (2018), 410–412. DOI: http://dx.doi.org/10.1056/NEJMp1805870

[26] Colleen M. McBride. 2015. Personal Genomic Tests for Healthy Aging: Neither Feast nor Foul. *Generations - Journal of the American Society on Aging* 39, 1 (04 2015), 41–45. http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/1752994685?accountid=9894

[27] Sarah C. Nelson, Deborah J. Bowen, and Stephanie M. Fullerton. 2019. Third-Party Genetic Interpretation Tools: A Mixed-Methods Study of Consumer Motivation and Behavior. *The American Journal of Human Genetics* 105, 1 (2019), 122–131. DOI:http://dx.doi.org/https://doi.org/10.1016/j.ajhg.2019.05.014

[28] CBS News. 2018. Behind at-home DNA testing companies sharing genetic data with third parties. https://www.cbsnews.com/news/dna-privacy-at-home-tests-23andme-ancestrydna-sell-data-to-third-parties/. (2018).

[29] Peter Ney, Luis Ceze, and Tadayoshi Kohno. Genotype extraction and false relative attacks: Security risks to third-party genetic genealogy services beyond identity inference. Network and Distributed System Security Symposium (NDSS). *Preprint Posted* 10, 29 (????), 19.

[30] Huntington Society of Canada. 2019. https://www.huntingtonsociety.ca/gna/. hhttps://www.huntingtonsociety.ca/gna/. (2019).

[31] Office of the Privacy Commissioner of Canada. 2017. Policy statement on the collection, use and disclosure of genetic test results. https://www.priv.gc.ca/en/privacy-topics/health-genetic-and-other-body-information/s-d_140710/. (2017).

[32] Kate O'Flaherty. 2019. Facebook Exposed Up To 600 Million Passwords –Here's What To Do. https://www.forbes.com/sites/kateoflahertyuk/2019/03/21/facebook-has-exposed-up-to-600-million-passwords-heres-what-to-do/#d5fc8abc90d9. (2019).

[33] Lukasz Olejnik, Agnieszka Kutrowska, and Claude Castelluccia. 2014. I'm 2.8% Neanderthal: The beginning of genetic exhibitionism?. In *Proceedings of the PETS Workshop on Genome Privacy, 14th Privacy Enhancing Technologies Symposium (PETS '14)*.

[34] openparliament.ca. 2019. Bill S-201 (Historical). https://openparliament.ca/bills/42-1/S-201/. (2019).

[35] Andelka Phillips. 2015. Genomic Privacy and Direct-to-Consumer Genetics: Big Consumer Genetic Data – What's in that Contract?. In *2015 IEEE Security and Privacy Workshops*. 60–64. DOI: http://dx.doi.org/10.1109/SPW.2015.19

[36] Eric Ravenscraft. 2019. How to Protect Your DNA Data Before and After Taking an at-Home Test. https://www.nytimes.com/2019/06/12/smarter-living/how-to-protect-your-dna-data.html. (2019).

[37] J. Scott Roberts and Jenny Ostergren. 2013. Direct-to-Consumer Genetic Testing and Personal Genomics Services: A Review of Recent Empirical Studies. *Current Genetic Medicine Reports* 1, 3 (01 Sep 2013), 182–200. DOI: http://dx.doi.org/10.1007/s40142-013-0018-2

[38] Denise Roland. 2019. 23andMe and GSK are mining customers' DNA data in a hunt for new drugs. https://www.marketwatch.com/story/23andme-and-gsk-are-mining-customers-dna-data-in-a-hunt-for-new-drugs-2019-07-23?rss=1&siteid=rss. (2019).

[39] Dr.Chuck Russo and Katherine C Russo. 2018. How Law Enforcement is Using Genealogy Testing Services to Solve Cold Cases. https://inpublicsafety.com/2018/07/how-law-enforcement-is-using-genealogy-testing-services-to-solve-cold-cases/. (2018).

[40] Julian Segert. 2018. Understanding Ownership and Privacy of Genetic Data. http://sitn.hms.harvard.edu/flash/2018/understanding-ownership-privacy-genetic-data/. (2018).

[41] Manya Sleeper, Justin Cranshaw, Patrick Gage Kelley, Blase Ur, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. 2013. "I Read My Twitter the Next Morning and Was Astonished": A Conversational Perspective on Twitter Regrets. In *CHI*. ACM.

[42] William P. Smith and Deborah L. Kidder. 2010. You've been tagged! (Then again, maybe not): Employers and Facebook. *Business Horizons* 53, 5 (2010), 491–499.

[43] Kate Snow and Jon Schuppe. 2018. 'This is just the beginning': Using DNA and genealogy to crack years-old cold cases. https://www.nbcnews.com/news/us-news/just-beginning-using-dna-genealogy-crack-years-old-cold-cases-n892126. (2018).

[44] Robert Sprague. 2011. Invasion of the Social Networks: Blurring the Line between Personal Life and the Employment Relationship. 50 (02 2011).

[45] Nicolle K. Strand. 2016. Shedding Privacy Along with our Genetic Material: What Constitutes Adequate Legal Protection against Surreptitious Genetic Testing? *The American Medical Association Journal of Ethics* 18, 3 (2016), 264–271. https://journalofethics.ama-assn.org/sites/journalofethics.ama-assn.org/files/2018-05/pfor2-1603.pdf

[46] Steven L. Thomas, Philip C. Rothschild, and Caroline Donegan. 2015. Social Networking, Management Responsibilities, and Employee Rights: The Evolving Role of Social Networking in Employment Decisions. *Employee Responsibilities and Rights Journal* 27, 4 (2015), 307–323.

[47] Effy Vayena, Elli Gourna, Jürg Streuli, E Hafen, and Barbara Prainsack. 2012. Experiences of Early Users of Direct-to-Consumer Genomics in Switzerland: An Exploratory Study. *Public health genomics* 15 (11 2012), 352–362. DOI:http://dx.doi.org/10.1159/000343792

[48] Margaret Vroman, Karin Stulz, Claudia Hart, and Emily Stulz. 2016. Employer Liability for Using Social Media in Hiring Decisions. *Journal of Social Media for Organizations* 3 (2016), 1–13.

[49] Julian Walker. 2016. Bill S-201: An Act to prohibit and prevent genetic discrimination. http://publications.gc.ca/collections/collection_2017/bdp-lop/ls/YM32-3-421-S201-eng.pdf. (2016).

[50] Yang Wang, Saranga Komanduri, and Pedro Giovanni. 2011. “I Regretted the Minute I Pressed Share: A Qualitative Study of Regrets on Facebook". In *Symposium on Usable Privacy and Security (SOUPS)*. ACM, 10:1–10:16.

[51] Sarah Zhang. 2016. DNA Got a Kid Kicked Out of School—And It’ll Happen Again. https://www.wired.com/2016/02/schools-kicked-boy-based-dna/. (2016).