Review

For reprint orders, please contact: reprints@futuremedicine.com

Personalized Medicine



Review of policies of companies and databases regarding access to customers' genealogy data for law enforcement purposes

Sevasti Skeva¹, Maarten HD Larmuseau^{2,3,4} & Mahsa Shabani*,^{1,5}

- ¹Center for Biomedical Ethics & Law, Department of Public Health & Primary Care, University of Leuven, 3000 Leuven, Belgium
- ²Department of Human Genetics, University of Leuven, 3000 Leuven, Belgium
- ³Laboratory of Socioecology & Social Evolution, Department of Biology, University of Leuven, 3000 Leuven, Belgium
- ⁴Histories vzw, 2800 Mechelen, Belgium
- ⁵Metamedica, Faculty of Law & Criminology, Ghent University, B-9000 Ghent, Belgium
- *Author for correspondence: Mahsa.shabani@ugent.be

The rapidly evolving popularity of direct-to-consumer genetic genealogy companies has made it possible to retrieve genomic information for unintended reasons by third parties, including the emerging use for law enforcement purposes. The question remains whether users of direct-to-consumer genetic genealogy companies and genealogical databases are aware that their genetic and/or genealogical data could be used as means to solving forensic cases. Our review of 22 companies' and databases' policies showed that only four companies have provided additional information on how law enforcement agencies should request permission to use their services for law enforcement purposes. Moreover, two databases have adopted a different approach by providing a special service for law enforcement. Although all companies and databases included in the study provide at least some provisions about police access, there is an ongoing debate over the ethics of these practices, and how to balance users' privacy with law enforcement requests.

First draft submitted: 2 September 2019; Accepted for publication: 20 December 2019; Published online: 3 March 2020

Keywords: consumer genomics databases • data protection • DNA data • forensic genealogy • genetic genealogy companies • long-range familial searches • nonforensic databases • privacy

The development of public or commercial genomics databases for clinical, research or recreational purposes has been growing in the last decade. Depending on the objectives of the databases, various types of genetic data are stored in them along with the relevant personal information, which can directly or indirectly identify the data subjects. As the databases contain genetic information, which may reveal vast amounts of health- and nonhealth-related information about individuals and their family members, ensuring authorized access for the intended purposes, in order to safeguard the privacy of the users and respect their consent, is of paramount importance [1,2].

To date, the concerns related to using these databases for unintended purposes have attracted a lot of attention. O' Doherty and colleagues provided an overview of potential unintended uses by third parties, which may arise when building various types of genetic databases, for example, in the context of biobanks or consumer genetic testing [1]. Among other examples of unintended uses by third parties, they highlighted the possibility of law enforcement having access to databases for forensic investigations under the issue of a court warrant or directly by submitting a formal request. For data subjects who wish to participate in building genetic databases, these practices raise several concerns [2–4].

The concerns about police access to consumer genomics databases were heightened in the aftermath of the so-called Golden State Killer arrest in April 2018 [4,5]. In solving this case, police uploaded a DNA sample collected from a crime scene into a publicly accessible genealogical website called GEDmatch and were then able to identify the suspect by matching his DNA to at least one distant relative [4,6]. It has been argued that long-range familial



searches can provide information on millions of individuals, even for those who have not undergone genetic testing [7,8]. Recently, Erlich and colleagues conducted a study on 1.28 million samples contained in a consumer genomics database and reported that up to 60% of long-range familial searches for Americans of European descent will result in a third cousin or closer relative [7].

In addition to direct-to-consumer (DTC) genetic genealogy companies, law enforcement agencies have also partnered with companies, like Parabon Nanolabs, which among others specializes in DNA phenotyping, to help them investigate unsolved cold cases. By using a DNA analysis methodology, as was described in a paper by Greytak et al., Parabon has successfully assisted law enforcement in making positive identifications and solving over 50 other high-profile criminal cases, since May 2018 [9–11]. In addition to the numerous identifications, guilty pleas and confessions, this new forensic technique has now led to a guilty verdict, making this the first case where a man was convicted after being arrested through genetic genealogy search [12,13]. Furthermore, experts in the field currently even speak of a new potential of genetic genealogy that could help exonerate individuals that have been wrongfully convicted [14–16].

The recent examples of police having access to consumer genomics databases are particularly concerning due to the fact that they did not undergo an official request [4]. Instead, undercover police agents were able to access databases by creating false profiles. They would then upload DNA samples collected from crime scenes and run them through the DNA already found in the database, looking for biological matches that would help them identify a suspect [17]. Notably, such use of databases by police departs from the conventional data requests by law enforcement agencies. Request to access data, if presented by a court-issued warrant or police direct request, allows companies to evaluate the request on a case-by-case basis. According to the privacy policy of the database, companies would share a customer's DNA information with the police if compelled to do so by law.

Despite the fact that law enforcement agencies have increasingly been using genetic genealogical services as an investigative tool, there is still great controversy over the ethics and governance of these practices, and how to balance users' privacy with law enforcement requests [12]. On the one hand, this technology has yielded results and benefits to society that cannot be understated, while on the other hand, as privacy advocates argue, users should maintain control and have the right to consent to their DNA and/or genealogical data being used in criminal investigations [18]. Given the exponential growth in this business, users participating in such genetic and/or genealogical websites need to keep in mind that their information will be shared with other users. That said, the databases set rules to define what they consider as legitimate uses and approved users. On this account, potential users should agree with those rules by accepting the privacy policy of the databases when they first sign up to use the databases or order their services. This is crucial considering the fact that, in principle, depending on the companies' policies regarding data/sample submission, anyone with a DNA sample, and/or results of a genetic or genealogy test can subscribe to these platforms. This could be cause for concern given the potential of surreptitious genetic testing, that is, the possibility to obtain genetic information of individuals by third parties in order to submit their samples or genetic data to such platforms without their knowledge or consent.

Consequently, there are some key questions deriving from police access to nonforensic databases that need to be addressed; how can databases define the scope of police access in their privacy policies and terms of services? How are these provisions being communicated to the users? In particular, the lack of a clear framework regarding police access may negatively impact the development of genomic and genealogical databases in general, including those for research purposes. In this paper, the privacy policies of some major DTC genetic genealogy companies and genealogical databases are assessed regarding how they deal with law enforcement requests.

Methods

A list of major DTC genetic genealogy companies and genealogical databases was compiled through general Google searches including terms such as 'genealogy websites', 'direct-to-consumer (DTC) genetic testing companies' and 'genealogy and DNA testing'. Various blogs and forums, along with lists of DTC genetic ancestry companies published by Royal *et al.*, Howard *et al.* and Moray *et al.* were consulted [19–21]. We compiled a sample of 22 websites and companies providing genetic genealogy services, which have often been referred to as popular genealogy services and websites (Table 1). We selected a heterogenous sample of such services (to include both DTC genetic genealogy services and genealogical databases) that appeared most often in the results. Furthermore, we selected the most up-to-date websites that excelled at communicating information regarding their practices to their customers, provided that all the information was written in English. The primary focus was the privacy information provided in the privacy statements, terms of service and in some cases covered in other areas of the websites

Table 1. Relevant provisions	t provision	s regarding law	nent access of privacy	policies of the selected g	enforcement access of privacy policies of the selected genetic genealogy companies and databases (15
October 2019) (cont.).	pnt.).				
# DNA genealogy Country services	Country	URL	Types of companies/databases	Other relevant information included on the websites	Excerpts from privacy policies regarding LE access

GEDMarth USA www.getmarth.com/legint.jpp Genealogical database "Withen you are of the following" i "With a work of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another interview." "Public account of a violent crime against another oncopy with a violent of a violent crime against another oncopy with a violent of a violent crime against another oncopy with a violent of a violent crime against another oncopy with a violent of a violent or required and a violent crime against another oncopy with a violent another oncopy and a violent another oncopy. " I Doc on a violent another oncopy and a violent	#	DNA genealogy services	Country	URL	Types of companies/databases	Other relevant information included on the websites	Excerpts from privacy policies regarding LE access
Genes Reunited UK www.genesreunited.co.uk company - company - Law enforcement guide - Law enforcement guide - Law enforcement guide - Company - Law enforcement guide - Company - Law enforcement guide - Prior notice, only when allowed company - Company company company - Company company - Company company company company - Company compan		GEDmatch	N SA	www.gedmatch.com/login1.php	Genealogical database	ı	"When you upload Raw Data to GEDmatch, you agree that the Raw Data is one of the following: [] -DNA obtained and authorized by law enforcement to identify a perpetrator of a violent crime against another individual, where violent crime' is defined as murder, nonnegligent manslaughter, aggravated rape, robbery, or aggravated assault. -DNA obtained and authorized by law enforcement to identify remains of a deceased individual. Public + opt-in' DNA data is available for comparison to any Raw Data in the GEDmatch database using the various tools provided for that purpose. Youblic + opt-out' DNA data is available for comparison to any Raw Data in the GEDmatch database, except DNA kits identified as being uploaded for Law Enforcement purposes."
Geni UsA www.geni.com Genealogical database - Helix USA www.helix.com DTC genetic genealogy company - Law enforcement guide and prior report and prior notice, only when allowed and some accommendation and company HomeDNA USA homedna.com DTC genetic genealogy company - IGENEA Switzerland www.igenea.com/en/home DTC genetic genealogy company -		Genes Reunited	٦	www.genesreunited.co.uk	DTC genetic genealogy company		"We will disclose your personal data in order to comply with any legal obligation. This includes disclosing information to organisations for the purposes of fraud protection, credit risk reduction, or the order of a court or regulator." www.genesreunited.co.uk/contents/legalprivacypolicy
Helix USA www.helix.com DTC genetic genealogy - Law enforcement guide company - Transparency report - Prior notice, only when allowed to do so company and www.igenea.com/en/home DTC genetic genealogy - company company riGENEA Switzerland www.igenea.com/en/home company company - company company - company company - c		Geni	USA	www.geni.com	Genealogical database	1	"[We may] share your information with our regulators and with law enforcement if required to do so by law, or in the good-faith belief that such action is necessary to comply with state or federal laws or respond to a court order, subpoena, law enforcement or regulatory request, or search warrant." www.geni.com/company/privacy#full_privacy
HomeDNA USA homedna.com DTC genetic genealogy – company iGENEA Switzerland www.igenea.com/en/home DTC genetic genealogy – company	_	Helix	USA	www.helix.com	DTC genetic genealogy company		"We may also disclose your personal information, including your Genetic Information: to comply with law, a valid court order, a judicial proceeding, subpoenas, warrants, bankruptcy proceedings, or in connection with any legal process, provided that we will not disclose your Genetic Information without a valid subpoena or search warrant specific to your Genetic Information. If we are required to disclose your information, we will do our best to provide you with notice in advance, unless we are prohibited by law from doing so."
iGENEA Switzerland www.igenea.com/en/home DTC genetic genealogy – company		HomeDNA	USA	homedna.com	DTC genetic genealogy company		"[] DDC, may disclose information if it believes in good faith that such disclosure is necessary to comply with relevant laws or to respond to subpoenas or warrants served on DDC." homedna.com/privacy
	_	igenea	Switzerland	www.igenea.com/en/home	DTC genetic genealogy company		"We may also pass your personal data on if we regard this as necessary to comply with the applicable laws and regulations, for court proceedings, if required to do so by the competent courts and authorities, or under other legal obligations, in order to protect and defend our rights and/or our property." https://www.igenea.com/en/terms

future science group fsg

Tal O	Table 1. Relevant pr. October 2019) (cont.)	nt provisions ont.).	Table 1. Relevant provisions regarding law enforceme October 2019) (cont.).	ent access of privacy	policies of the selected g	enforcement access of privacy policies of the selected genetic genealogy companies and databases (15
#	DNA genealogy services	Country	URL	Types of companies/databases	Other relevant information included on the websites	Excerpts from privacy policies regarding LE access
4	insitome	USA	www.insito.me	DTC genetic genealogy company	– Prior notice, only when allowed to do so	"We may be legally required to disclose your Personal Information including your Genetic Information if such disclosure is required by subpoena, law or other legal process, necessary to assist law enforcement officials or government enforcement agencies []." www.insito.me/privacy
15	Living DNA	Ä	livingdna.com	DTC genetic genealogy company	1	"[] we may be legally required to disclose information. Examples of this include where a we are subject to a binding court order, subpoena, or a legally binding direction by a regulator []. We reserve the right to share personal information where we reasonably believe that we are legally required to do so." livingdna.com/privacy-centre/privacy
16	MyHeritage	Israel	www.myheritage.com	Genealogical database/DTC genetic genealogy company	1	"MyHeritage will not disclose any of your personal information except in very limited circumstances []: if required by law, regulatory authorities, legal process or to protect the rights or property of MyHeritage []." www.myheritage.com/FP/Company/popup.php?p=privacy_policy
17	National Geographic Society (Genographic Project)	USA	genographic.nationalgeographic .com	DTC genetic genealogy company	1	"Nothing contained in this Agreement limits National Geographic Partners' right to comply with governmental, court and law enforcement requests or requirements relating to your use of the Services or information provided to or gathered by us in connection with such use." www.nationalgeographic.com/legal/terms.html
18	ORIG3N	USA	orig3n.com	DTC genetic genealogy company	1	"ORIG3N and its affiliates are expressly permitted to use any or all of the User Information to the extent that it is required or compelled to do so by law." orig3n.com/privacy-policy-genetic-information-and-platform/
19	Oxford Ancestors	N N	www.oxfordancestors.com	DTC genetic genealogy company	1	"We may use the data you provide [], when required or permitted by law, or except as specifically detailed in this privacy policy, without your permission." www.oxfordancestors.com/content/view/26/44/
50	Vitagene	USA	vitagene.com	DTC genetic genealogy company	1	"As required by law, such as to comply with reporting requirements, a subpoena, or similar legal process in connection with an investigation of fraud, [] or as we believe reasonably necessary to protect or enforce our rights, protect your safety or the safety of others." vitagene.com/terms-use/#privacy_of_policy
21	WikiTree	USA	www.wikitree.com	Genealogical database	1	"In certain circumstances, if we believe it is reasonably necessary, we may release specific information about you or your account to comply with any valid legal process such as a search warrant, subpoena, statute, or court order, or in other special cases, such as an attempted breach of WikiTree security, without notice to you." www.wikitree.com/wiki/Help:Privacy-Policy
22	YFull	USA	www.yfull.com	DTC genetic genealogy company	1	"We make every reasonable effort to protect your privacy as described in this Privacy Policy. Nevertheless, we may be required by law to disclose your personally identifiable information." www.yfull.com/terms/
DTC	DTC: Direct to consumer; LE: Law enforcement.	; LE: Law enforcem	lent.			

(e.g., Guide for Law Enforcement, FAQs, users' blogs) that included relevant information for the purposes of this study. Content analysis of the documents has been steered with two questions; first, whether information regarding allowing access under formal law enforcement requests are being provided, and second what further information are being communicated regarding such access. We also identified two case studies, namely FamilyTreeDNA and GEDmatch, which present a different approach regarding access under formal law enforcement purposes and elaborated them in the results section.

Results

Our review showed that all 22 companies and databases included in the study communicated either the possibility of access for law enforcement purposes, or the disclosure of users' information, if requested by law. However, their privacy policies differed in how they formulated the access by law enforcement bodies, the information they provided to the users and their provisions regarding police access.

Policies of consumer DNA databases allowing access under formal law enforcement requests

All the companies and databases state that they will share their users' personal information in order to comply with the law and requests from government bodies supported by the relevant documentation, such as the issue of a court warrant or a subpoena. For example:

"To comply with law, a valid court order, a judicial proceeding, subpoenas, warrants, bankruptcy proceedings or in connection with any legal process, provided that we will not disclose your Genetic Information without a valid subpoena or search warrant specific to your Genetic Information" (Helix) [22].

These statements are followed by disclaimers saying that the companies will not argue the legitimacy of valid court orders that would lawfully allow police to look into their users' files. One company, namely, ORIG3N, states that "ORIG3N shall be under no obligation to contest a valid order of a court or other governmental body to disclose User Information. No such use or disclose will be considered a breach of this Privacy Policy" (ORIG3N) [23]. Others, like 23 and Me, Ancestry and MyHeritage highlight that they try to resist law enforcement inquiries; for example, 23 and Me in its Transparency Report states that the company will closely scrutinize all law enforcement and regulatory requests and will only comply with court orders, subpoenas, search warrants or other requests that they determine are legally valid (23 and Me) [24]. However, in all cases, it is clarified that once the validity of a court issued order is established, they cannot contest it and have to abide by it.

Furthermore, our research showed that all companies and databases have formulated their provisions regarding law enforcement access in a broad way, implying that they may provide the information beyond the formal law enforcement requests, in order to exercise their rights under their terms of service:

"[We] Share your information with our regulators and with law enforcement if required to do so by law, or in the good-faith belief that such action is necessary to comply with state or federal laws or respond to a court order, subpoena, law enforcement or regulatory request or search warrant" (Geni) [25].

Nevertheless, four companies, namely 23 and Me, Ancestry, Family TreeDNA and Helix, have provided additional information on how law enforcement agencies should request permission to use their services for law enforcement purposes. These companies have released specific guidance intended for law enforcement authorities and government agencies in the USA that are seeking users' information. In particular, they require a valid legal process in order to consider revealing any user personal and genetic information, and that all legal requests be made in writing, including all relevant information of the user account associated to a valid trial, grand jury, administrative subpoena, warrant or court order, as defined in 18 USC § 2703(c)(2) and 18 USC § 2703(d) [26–29].

Among the databases providing provisions regarding access for law enforcement purposes, two databases, namely FamilyTreeDNA and GEDmatch, have presented a different approach by providing a special service for law enforcement purposes and by revisiting their policies regarding law enforcement access beyond formal requests, respectively.

Setting up a special Law Enforcement Matching service: the case of FamilyTreeDNA

One recently announced practice is that of FamilyTreeDNA regarding establishing a Law Enforcement Matching (LEM) service. The LEM 'is only permitted with accounts set up by FamilyTreeDNA on behalf of law enforcement authorities and their authorized representatives' [30,31]. They would then compare the DNA of profiles of unknown criminals with that of vast numbers of users, whose genetic information is on file. Law enforcement user account will have equal access to users' information as any other user would, when matched as DNA relatives, provided that

users have previously consented to participate in matching and the rest of the company's requirements are entirely

"For requests made by law enforcement and their authorized representatives that meet the requirements of our Law Enforcement Guidelines, FamilyTreeDNA may create limited access law enforcement accounts ('LE Accounts') which are permitted to upload genetic information to the database to identify the remains of a deceased individual or to identify the perpetrator of a homicide or sexual assault. FamilyTreeDNA will track Law Enforcement Accounts via an in-house identification system that will allow users to opt out of Law Enforcement Matching" [30].

For those who do not wish for their genetic information to be viewable by Law Enforcement Accounts, the company instituted a new policy on 12 March 2019 allowing its users to opt-out from matching in the relevant privacy setting. Failing to do so will be considered as providing consent and users will automatically be opted in to participate in matching [31].

However, even for individuals who have opted-out of LEM, FamilyTreeDNA may still be required to reveal their personal information in exceptional cases, in order to comply with a valid legal process and apply its terms of service. In any case, it is clarified that use of their service for law enforcement purposes is not a right, but a privilege, meaning that the company reserves the right to revoke law enforcement's account access for any reason without warning, especially for purposes not expressly agreed to by FamilyTreeDNA [30].

Moreover, in light of current events, FamilyTreeDNA explained that accounts of individuals identified as EU residents, created prior to the implementation of the LEM, are currently opted out; users who wish to opt in will have the ability to do so at any time by adjusting their *Matching Preferences* through their account settings [32].

A substantial policy change in allowing access beyond formal requests: the case of GEDmatch

GEDmatch is a public genealogical database to which people can upload DNA test results received from other consumer genetics services, like MyHeritage, Ancestry and 23andMe, in order to find genetic relatives and explore their ancestries. GEDmatch contains approximately 1.2 million profiles, and it became widely known as the 'open-source' that led to the Golden State Killer arrest [11,33–35]. Until recently, GEDmatch's privacy policy has been substantially flexible allowing law enforcement authorities to use their services as part of their criminal investigations without undergoing any formal procedure. This open policy even approved of undercover police access to its database by creating false profiles of individuals, uploading DNA samples extracted from crimes scenes and running them through the database, while looking for potential matches [4].

Privacy concerns over GEDmatch's policy to allow any data set to 'public' to be used in criminal investigations were exacerbated in May 2019, when GEDmatch was used to identify a suspect of a considerably less serious offence. Although GEDmatch's representatives claimed at the time that it was a matter of public interest, privacy advocates were particularly concerned that this could be a slippery slope for police to start using these databases to investigate not only cold, but any case in which DNA is left behind [36,37].

In response, GEDmatch changed its terms of service, and introduced a new privacy setting on 18 May 2019. According to the new policy, the list of crimes that police can now search for has been extended; particularly, DNA can be uploaded for Law Enforcement purposes in order "to identify a perpetrator of a violent crime against another individual, where 'violent crime' is defined as murder, nonnegligent manslaughter, aggravated rape, robbery or aggravated assault, or to identify remains of a deceased individual" [33,38]. However, users who do not wish to make their profiles available for comparisons to DNA kits identified as being uploaded for Law Enforcement purposes have an option to change their preference from a preselected opt-in policy to opt-out through the relevant privacy setting. Currently, out of the 1.2 million kits, 85,000 users have already opted in for law enforcement [15,38].

In addition, it is clarified that among other potential uses of uploaded raw data, results published in the database may potentially be used for "(f)amilial searching by third parties such as law enforcement agencies to identify the perpetrator of a crime, or to identify remains" [33]. In this regard, GEDmatch warns its users who do not agree with this policy and are concerned about future genealogical and nongenealogical uses of their raw data not to upload their DNA or remove DNA that has already been uploaded to the database. Lastly, any updates to GEDmatch's terms and policy will no longer be communicated via email, therefore users are expected to consult the website regularly, since continuing to use their site is equivalent to accepting the updated terms [33].

Providing further information regarding law enforcement access

Although they are not obligated to provide prior notice, six companies state that, when allowed to do so, they will try in good faith to provide notice before releasing any personal information, in order to comply with such legal or regulatory proceedings.

"Unless required to do so by law, we will not release a customer's individual-level Personal Information to any third party without asking for and receiving that customer's explicit consent" (23andMe) [24].

On the other hand, one company, namely, WikiTree, makes it explicitly clear that they would provide the requested information without giving prior notice to the user. For example:

"In certain circumstances, if we believe it is reasonably necessary, we may release specific information about you or your account to comply with any valid legal process such as a search warrant, subpoena, statute, or court order, or in other special cases, such as an attempted breach of WikiTree security, without notice to you" (WikiTree) [39].

Furthermore, for the sake of transparency, three companies, namely 23andMe, Ancestry and recently Helix, provide a transparency report, which is updated either annually or on a quarterly basis, in which they summarize all law enforcement and/or governmental requests they received:

"We publish an annual transparency report describing requests we received. When we receive a request, our team reviews it for compliance with legal requirements and our policies. If we believe a request is overly broad, we will seek to narrow it. We notify users prior to turning over any information in order to give them an opportunity to challenge the request unless it would be counterproductive, or we are legally prevented from doing so" (Ancestry) [40].

Discussion

Despite the fact that familial searching using online genetic testing services has become a frequently used practice, especially in the USA after the arrest of the alleged Golden State Killer in April 2018, these practices remain highly unregulated leaving many questions unanswered regarding the ethical and legal underpinnings of these activities [41–43].

First of all, in terms of privacy, customers have the right to know how companies handle protection of their personal data, including those deriving from their DNA, how they process it, with whom they would share it and under what circumstances. The right to privacy and data protection is endorsed by both ethical and legal frameworks. These frameworks recognize the individuals' right to be adequately informed about how their personal data are being used, and also to be allowed to object to further uses of their data, when such objection does not undermine public interest. In particular, the right to transparent information and communication is recognized by the relevant data protection regulations, such as the EU General Data Protection Regulation (GDPR) in Chapter 3 – rights of the data subject [44]. Providing "detailed transparency about how Genetic Data is collected, used, shared, and retained" is also endorsed by relevant guidelines and policies, such as those issued by the Future of Privacy Forum in 'Privacy Best Practices for Consumer Genetic Testing Services' [45]. Within this context, the Future of Privacy Forum recently removed FamilyTreeDNA from their list of supporters, due to the company's failure to provide such information to their customers in the framework of a transparency report [45].

Furthermore, in order to promote consumers' trust in using their services, it is crucial that companies adopt clear and transparent policies regarding the processing of data for law enforcement purposes. Our investigation showed that all companies and databases included in the study disclose in their privacy policies and terms of service that they will share user information for law enforcement purposes under formal requests (i.e., under court order, warrant or subpoena), or when it is believed that is reasonably necessary to do so. However, the information communicated about this possibility varied among the companies and databases. Previously, the Hazel and Slobogin study observed that 'over two-thirds of companies (38 of 55, i.e., 69%) addressed the sharing of information with law enforcement or other government authorities. However, policies varied significantly in the amount of information provided to consumers about the process [46].

Notably, a recent interim policy issued by the US Department of Justice on 'Forensic Genetic Genealogical DNA Analysis and Searching' stressed this issue by stating that "the Investigative agencies shall identify themselves as law enforcement to GG [genetic genealogy] services and enter and search FGG [Forensic Genetic Genealogy] profiles only in those GG services that provide explicit notice to their service users and the public that law enforcement may use their service sites to investigate crimes or to identify unidentified human remains" [47].

Furthermore, the adequacy of privacy policies and the terms of service in communicating such further uses of data is far from certain [1,48]. Previous studies have shown that privacy policies and the terms of service are usually seen broadly as legalese, which discourages consumers from taking the time to read them [49]. Moreover, some companies have already frequently modified their privacy statements, raising a question regarding how far the customers are truly informed of such modifications. This has been a pertinent concern regarding access for law enforcement purposes, as companies such as GEDmatch and FamilyTreeDNA have changed their policies already a number of times over the course of last year [4,32,50–52].

Most companies clearly mention that they will provide advance notice for any substantive and minor changes of the privacy policies and terms of service, so that the updated information be communicated to their users giving them the chance to decide whether to continue using their services. However, it is not made clear whether or not objecting to the changes is sufficient or whether actual consent needs to be obtained again. In particular, for privacy statements that have been substantially altered, users need to provide consent for the document to retain its original validity and for all parties involved to be legally obligated by the new terms. Currently, this was seen as the users' responsibility to stay updated, since continued use of a website's services is equivalent to providing consent to the updated statements. Moreover, as it has been reflected in existing literature, consumers need to be aware that sharing their DNA information means providing sensitive information not only for themselves, but also for close and distant relatives, and even for unborn offspring [7,17,53].

Second, the legal basis for accessing the websites data for law enforcement purposes beyond official requests has not been clear, and particularly uploading a crime scene sample by using a false profile has been a highly questionable practice. In this regard, some have argued that the samples left in the crime scene should be considered as 'abandoned samples' devoid of any ownership rights. To date, in countries such as the USA, the use of abandoned DNA collected from public places has been mainly accepted by the courts [54]. However, in some jurisdictions there are restrictions on the types of the testing that can be performed on the crime scene samples – mainly limiting it to the noncoding area of the DNA. For instance, the European Council Resolution of 30 November 2009 on the exchange of DNA analysis results states that the exchange of results should be limited to DNA analysis of noncoding parts [55]. Similarly, a recent report on forensic DNA databases showed that restricting DNA testing to noncoding regions of the DNA is a dominant approach across national legislations on forensic DNA analysis [56].

In this regard, 23andMe explicitly states that it would constitute a direct violation of 23andMe's Terms of Service "for law enforcement officials to submit samples on behalf of a prisoner or someone in state custody", and that "customers who wish to participate in the 23andMe service must guarantee that any sample they provide is either their own saliva or that of an individual for whom they have legal authorization to agree to the terms of services on their behalf" [57].

Although one can argue that there are no legitimate privacy expectations in DNA left in the crime scenes, the questions remain about the rights of the individuals who voluntarily upload their genetic data in nonforensic databases. Do they have a right to know that their uploaded genetic profile was analyzed and that they were part of a criminal investigation? [48,58]. Some have suggested that such use of genealogy websites should be limited to solving serious crimes, and to be seen in the public interest. As David Kaye suggested, such use could be limited to serious crimes, like sexual assault and homicide, and that all other investigative methods be exhausted first [42]. In light of the recent discussion, in GEDmatch's revised policy of May 2019, law enforcement access is explicitly limited to identify perpetrators of violent crimes, namely murder, nonnegligent manslaughter, aggravated rape, robbery or aggravated assault, as defined in the website, and only when users have opted in to have their DNA profiles available for Law Enforcement purposes [14,15,33].

Of interest to this discussion is the approach of the GDPR with regard to further processing of personal data in relation to identifying criminal acts. The terminology used in the GDPR is for 'indicating possible criminal acts' or 'threats to public security', in order to provide a legal basis for further processing when the original data collection has been based on consent. As stated in Recital 50 GDPR: [...] 'Indicating possible criminal acts or threats to public security by the controller and transmitting the relevant personal data in individual cases or in several cases relating to the same criminal act or threats to public security to a competent authority should be regarded as being in the legitimate interest pursued by the controller. However, such transmission in the legitimate interest of the controller or further processing of personal data should be prohibited if the processing is not compatible with a legal, professional or other binding obligation of secrecy' [44]. Notably, the impact of the GDPR's provisions on the use of nonforensic databases for law enforcement purposes has not been thoroughly investigated to date. Although the practice of forensic genealogy is largely taking place in the USA, this investigative method has captured the interest of police in other countries as well. For example, in Sweden, prosecutors have allowed Swedish murder investigators to use consumer genealogy databases to solve cold cases, as it is estimated that the DNA test results of 40,000 people are already on file [59-61]. This makes it all the more important to investigate the legal basis for access to such data of EU citizens under the GDPR and other applicable regulations such as consumer protection laws. It should be noted that regarding the use of evidence collected through such online searches and their admissibility in the courts, definitely the specific rules of each jurisdiction will apply.

Moreover, in terms of public opinion regarding police access and the type of crimes, the recent study conducted by Guerrini *et al.* is of interest, as it provides insights about the general public's view on this subject matter [4]. The results showed that participants were significantly more supportive of police having access to such genealogical databases or even creating false profiles of individuals, in order to identify perpetrators of violent crimes (80%), like murder, rape, arson and kidnapping, crimes committed against children (78%) or to identify missing persons (77%). When the purpose was to identify perpetrators of nonviolent crimes, like car theft or drug possession, positive response dropped in half (39%). In addition, in a study conducted by Christofides and O'Doherty in 2016 by customers of DTC genetic testing companies, the use of data for law enforcement purposes has been mentioned among potential objectionable uses by the respondents [62].

At last, predicting the consequences of unintended and unforeseen uses of large health data collections by users is a challenging task [1]. Therefore, to protect the legitimate uses of genetic data, and to adequately control whether companies lawfully apply their terms of use, sufficient safeguards must be implemented. For example, a potential solution involves third parties, such as independent ethics committees, to supervise the process and review requests from law enforcement agencies [36]. Moreover, given the identifiability of DNA data, in order to mitigate some of the risks and prevent the exploitation of long-range familial searches, Erlich *et al.* suggested that consumer genomics companies encrypt the genetic data to protect customers' personal information [7].

Conclusion

Despite the fact that there are many controversies surrounding the familial search of genealogy databases, it is undeniable that it has proven to be a powerful investigative tool. Following the arrest of the Golden State Killer, genetic genealogy has helped law enforcement agencies solve high profile cold cases in many different US states [43]. This has resulted in justice being served and has also brought closure to the families of the victims, some of which have been struggling for decades. As consumer genomics databases continue to grow rapidly and the number of voluntary samples have reached the scale of millions, there are increased opportunities to identify individuals suspected of having committed serious crimes. Consequently, genetic genealogical databases from commercial companies could prove to be a well-designed criminal justice tool leading to major breakthroughs in cold-case investigations [7,8,63].

However, considering the lack of specific legislation regarding long-range familial searches on nonforensic databases, one can argue that the consumers' DNA data are not currently effectively protected [64]. On this account, the genetic genealogy community is far from reaching an agreement, due to the raging debate over finding the balance between the ethics of these practices and the protection of users' privacy rights [12,18]. Despite the fact that there are many privacy laws in the USA (e.g., HIPAA, GINA, the Affordable Care Act), none of them provides an adequate protection against the potential misuses of DNA data, nor are the practices of consumer genomics companies fully covered by them [65]. In response, certain states have been attempting to draft legislation hoping to shed light on forensic genetic genealogy, restrict genetic surveillance or strictly limiting the practice to violent crimes [43]. It remains to be seen how current regulatory responses will be shaped in order to more fully protect the privacy rights of individuals.

Future perspective

The popularity of genetic genealogy has increased significantly over the last years. While the number of customers of DNA-testing companies is rapidly growing, it is now very likely to find far-related persons from a particular DNA-donor. Given the fact that genetic genealogy is seen as a powerful tool to solve forensic cases, the increasing use of databases for law enforcement purposes has been anticipated. Nevertheless, the scope of using genealogy databases by law enforcement is still a matter of controversy, both for the service providers and the consumers. Our paper showed that the companies and websites were not fully prepared for such use, and thus their privacy policies and terms of service are not communicating adequate information about this possibility to their users.

As this matter continues to attract public attention, it is raising awareness among current and future clients. Therefore, it is expected that in the foreseeable future, DTC genetic genealogy companies and genealogical databases shall strive to restore their clients' trust by adopting clear privacy policies that better protect their privacy rights.

Executive summary

- The popularity of direct-to-consumer (DTC) genetic genealogy companies and genealogical databases has made genetic genealogy a powerful tool for law enforcement purposes.
- Our research question is: Are users of such companies aware that their genetic and/or genealogical data can be used as means to solving forensic cases?

Methods

 We analyzed the privacy policies of 22 DTC genetic genealogy companies and genealogical databases regarding the access and use of genetic and/or genealogical data by law enforcement.

Results

- While all companies and websites included in the study provide at least some provisions about police access, only four companies provided additional information on how law enforcement agencies should request permission to use their services for law enforcement purposes.
- Two companies have adopted a different approach by providing a special service for law enforcement purposes.

Discussion

- There is still an ongoing debate over the ethics of using genetic genealogical data to solve forensic cases, and how to balance users' privacy with law enforcement requests.
- It is expected that in the foreseeable future, DTC genetic genealogy companies and genealogical databases shall strive to restore their clients' trust by adopting clear privacy policies that better protect their privacy rights.

Financial & competing interests disclosure

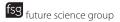
The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

- O'Doherty K, Christofides E, Yen J et al. If you build it, they will come: unintended future uses of organised health data collections. BMC Med. Ethics 17(1), 54 (2016).
- Curtis C, Hereward J, Mangelsdorf M, Hussey K, Devereux J. Protecting trust in medical genetics in the new era of forensics. Genet. Med. 21(7), 1483–1485 (2018).
- 3. Ram N, Guerrini C, McGuire A. Genealogy databases and the future of criminal investigation. Science 360(6393), 1078–1079 (2018).
- Guerrini C, Robinson J, Petersen D, McGuire A. Should police have access to genetic genealogy databases? Capturing the Golden State Killer and other criminals using a controversial new forensic technique. PLoS Biol. 16(10), E2006906 (2018).
- 5. Hesman Saey T. New genetic sleuthing tools helped track down the Golden State Killer suspect. *Science News*(2018). www.sciencenews.org/article/golden-state-killer-suspect-dna-genetics-genealogy
- 6. Selk A. The ingenious and 'dystopian' DNA technique police used to hunt the 'Golden State Killer' suspect. Washington

 Post (2018).www.washingtonpost.com/news/true-crime/wp/2018/04/27/golden-state-killer-dna-website-gedmatch-was-used-to-identi
 fy-joseph-deangelo-as-suspect-police-say/?outputType=amp
- Erlich Y, Shor T, Pe'er I, Carmi S. Identity inference of genomic data using long-range familial searches. Science 362(6415), 690–694
 (2018).
- 8. Regalado A. More than 26 million people have taken an at-home ancestry test. MIT Technology Review (2019). www.technologyreview.com/s/612880/more-than-26-million-people-have-taken-an-at-home-ancestry-test/
- Parabon Nanolabs. Parabon® Snapshot® DNA analysis service helps investigators solve 1,000 years of cold cases in 9 months (2019). https:
 - //parabon-nanolabs.com/news-events/2019/03/parabon-helps-investigators-solve-1000-years-of-cold-cases-in-9-months. html
- 10. Greytak E, Moore C, Armentrout S. Genetic genealogy for cold case and active investigations. Forensic Sci. Int. 299, 103-113 (2019).
- 11. Kennett D. Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes. *Forensic Sci. Int.* 301, 107–117 (2019).
- Murphy H. Genealogy sites have helped identify suspects. Now they've helped convict one. NYTimes (2019). www.nytimes.com/2019/07/01/us/dna-genetic-genealogy-trial.html
- 13. Shapiro E. Signaling a 'new era' in forensic investigation, a man caught through genetic genealogy gets life in prison for 1987 double murder. ABC News (2019).
 - https://abcnews.go.com/US/signaling-era-forensic-investigation-man-caught-genetic-genealogy/story?id=64540673



- 14. Hernandez S. It helped lock up killers and rapists. Now investigative genealogy has cleared an innocent man of murder. *BuzzFeed News* (2019). www.buzzfeednews.com/article/salvadorhernandez/dna-investigative-genealogy-wrongful-conviction-angie-dodge.
- 15. Armstrong M. In an apparent first, genetic genealogy aids a wrongful conviction case. *The Marshall Project* (2019). www.themarshallproject.org/2019/07/16/in-an-apparent-first-genetic-genealogy-aids-a-wrongful-conviction-case
- 16. Shapiro E. A new beginning: man convicted of murder in the 90's exonerated thanks to genetic genealogy. ABC News (2019). https://abcnews.go.com/US/90s-christopher-tapp-convicted-rape-murder-today-set/story?id=64339196&cid=clicksource_4380645_null_card_image
- 17. Garciía Ó, Crespillo M, Yurrebaso I. Suspects identification through "familial searching" in DNA databases of criminal interest. Social, ethical and scientific implications Span. J. Leg. Med. 43(1), 26–34 (2017).
- Informed Consent: what it is, what it isn't, and why it's necessary. The DNA Geek(2019).
 https://thednageek.com/informed-consent-what-it-is-what-it-isnt-and-why-its-necessary/
- 19. Royal CD, Novembre J, Fullerton SM *et al.* Inferring genetic ancestry: opportunities, challenges, and implications. *Am. J. Hum. Genet.* 86(5), 661–673 (2010).
- 20. Howard HC, Avard D, Borry P. Are the kids really all right? Direct-to-consumer genetic testing in children: are company policies clashing with professional norms? Eur. J. Hum. Genet. 19(11), 1122–1126 (2011).
- Moray N, Pink K, Borry P, Larmuseau M. Paternity testing under the cloak of recreational genetics. Eur. J. Hum. Genet. 25(6), 768–770 (2017).
- 22. Helix. Helix privacy policy (2019). www.helix.com/pages/privacy-policy
- ORIG3N. Orig3n's privacy policy genetic information and platform (2019). https://orig3n.com/privacy-policy-genetic-information-and-platform/
- 24. 23andMe. 23andMe Transparency Report (2019). www.23andme.com/transparency-report/
- 25. Geni. Geni your privacy (2019). www.geni.com/company/privacy#full_privacy
- 26. govinfo. Stored Communications Act. 18 U.S.C. 2703(c)(2) and (d) required disclosure of customer communications or records (2018). www.govinfo.gov/app/details/USCODE-2018-title18/USCODE-2018-title18-partI-chap121-sec2703/context
- 27. 23andMe. 23andMe guide for law enforcement (2019). www.23andme.com/law-enforcement-guide/
- 28. FamilyTreeDNA. FamilyTreeDNA guide for law enforcement (2019). www.familytreedna.com/legal/law-enforcement-guide
- 29. Helix. Helix law enforcement guide (2019). https://blog.helix.com/law-enforcement-guide/
- 30. FamilyTreeDNA. FamilyTreeDNA privacy statement (2019). www.familytreedna.com/legal/privacy-statement
- 31. FamilyTreeDNA. FamilyTreeDNA consent to participate in matching (2019). www.familytreedna.com/legal/consent/matching
- 32. FamilyTreeDNA. We are updating our terms of service and privacy statement regarding law enforcement matching preferences (2019). https://mailchi.mp/familytreedna/updates-to-our-terms-of-service-and-privacy-policy
- 33. GEDmatch 2019. GEDmatch.com terms of service and privacy policy (2019). www.gedmatch.com/tos.htm
- 34. Gafni M, Krieger L. Here's the 'open-source' genealogy DNA website that helped crack the Golden State Killer case. *The Mercury News* (2018). www.mercurynews.com/2018/04/26/ancestry-23andme-deny-assisting-law-enforcement-in-east-area-rapist-case/
- Molteni M. The future of crime-fighting is family tree forensics. Wired (2018).
 www.wired.com/story/the-future-of-crime-fighting-is-family-tree-forensics/
- 36. Aldhous P. The arrest of a teen on an assault charge has sparked new privacy fears about DNA sleuthing. *BuzzFeed.News* (2019). www.buzzfeednews.com/article/peteraldhous/genetic-genealogy-parabon-gedmatch-assault
- 37. Reavy P. Plastic milk container, genealogy helped Utah police crack church assault case. *Deseret News* (2019). www.deseretnews.com/article/900070489/plastic-milk-container-genealogy-helped-utah-police-crack-church-assault-case.html
- 38. Aldhous P. This genealogy database helped solve dozens of crimes. But its new privacy rules will restrict access by cops. BuzzFeed News (2019). www.buzzfeednews.com/article/peteraldhous/this-genealogy-database-helped-solve-dozens-of-crimes-but
- 39. WikiTree. WikiTree Help: Privacy Policy (2019). www.wikitree.com/wiki/Help:Privacy_Policy#Links_to_Third-Party_Sites
- 40. Ancestry. Ancestry 2018 Transparency Report (2019). www.ancestry.com/cs/transparency
- Piquado T, Matthies C, Strang L, Anderson J. Forensic familial and moderate stringency DNA searches: policies and practices in the United States, England, and Wales. RAND Corporation, Santa Monica, CA, USA (2019). https://www.rand.org/pubs/research_reports/RR3209.html
- 42. Pauly M. Police are increasingly taking advantage of home DNA tests. There aren't any regulation to stop it. *Mother Jones* (2019). www.motherjones.com/crime-justice/2019/03/genetic-genealogy-law-enforcement-golden-state-killer-cece-moore/?fbclid=IwAR3lOZl 6fAtkNdgYE5umtSgxz1qbbaS5_aOU9j7eGJHv68UY_JOX2MRw6EM
- Molteni M. What the golden state killer tells us about forensic genetics. Wired (2019).
 www.wired.com/story/the-meteoric-rise-of-family-tree-forensics-to-fight-crimes/

- 44. EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (2016). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN
- 45. Future of Privacy Forum. Privacy best practices for consumer genetic testing services. (2018). https://fpf.org/2018/07/31/future-of-privacy-forum-and-leading-genetic-testing-companies-announce-best-practices-to-protect-privacy-of-consumer-genetic-data/
- Hazel J, Slobogin C. Who knows what, and when?: A survey of the privacy policies proffered by U.S. direct-to-consumer genetic testing companies. J.L. & Pub. Pol'y. Vanderbilt Law Research Paper No.18-18 (2018). https://papers.ssrn.com/sol3/papers.cfm?abstract.id=3165765
- 47. The US Department of Justice. Department of justice announces interim policy on emerging method to generate leads for unsolved violent crimes. Office of Public Affairs (2019).
 www.justice.gov/opa/pr/department-justice-announces-interim-policy-emerging-method-generate-leads-unsolved-violent
- 48. Berkman B, Miller W, Grady C. Is it ethical to use genealogy data to solve crimes? Ann. Inter. Med. 169(5), 333-334 (2018).
- Vu K, Chambers V, Garcia F et al. How users read and comprehend privacy policies. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 4558(2), 802–811 (2007).
- Brown K. Major DNA testing company sharing genetic data with the FBI. Bloomberg (2019).
 www.bloomberg.com/news/articles/2019-02-01/major-dna-testing-company-is-sharing-genetic-data-with-the-fbi
- 51. Augenstein S. GEDmatch Changes: 'Blow' to Forensic Genealogy? *Forensic Magazine* (2019). www.forensicmag.com/news/2019/05/gedmatch-changes-blow-law-enforcement-and-forensic-genealogy
- More Privacy Concerns at Family Tree DNA. The DNA Geek (2019). https://thednageek.com/more-privacy-concerns-at-family-tree-dna/
- Raeburn P. How to identify almost anyone in a consumer gene database. Scientific American (2018).
 www.scientificamerican.com/article/how-to-identify-almost-anyone-in-a-consumer-gene-database/
- 54. Joh E. Reclaiming "abandoned" DNA: the fourth amendment and genetic privacy. Northwest. Univ. Law Rev. 100(2), 857-884 (2006).
- Council Resolution of 30 November 2009 on the exchange of DNA analysis results. OJ C 296, 5.12. (2009). https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:296:0001:0003:EN:PDF
- 56. Samuel G, Howard HC, Cornel M et al. A response to the forensic genetics policy initiative's report "Establishing Best Practice for Forensic DNA Databases". Forensic Sci. Int. Genet. 36, e19–e21 (2018).
- 57. 23andMe 2019. 23andMe Full Privacy Statement. www.23andme.com/about/privacy/
- 58. Scudder N, McNevin D, Kelty S, Funk C, Walsh S, Robertson J. Policy and regulatory implications of the new frontier of forensic genomics: direct-to-consumer genetic data and genealogy records. *Curr. Issues Crim. Justice* 31(2), 194–216 (2019).
- Linter S. Genealogy websites used to track down murderer. Radio Sweden (2019). https://sverigesradio.se/sida/artikel.aspx?programid=2 054&artikel=7276285&fbclid=IwAR19v18AiIIZqJRso2uJCmxSLYJIGJRW1FdHC5cACkWz1YQg1rLiQOf_Gs0
- Swedish police mull using DNA family tree websites to catch killers. The Local (2018).
 www.thelocal.se/20180507/swedish-police-mull-using-dna-family-tree-websites
- 61. Phillips C. The Golden State Killer investigation and the nascent field of forensic genealogy. Forensic Sci. Int.: Genetics 36, 186–188 (2018).
- 62. Christorides E, O'Doherty K. Company disclosure and consumer perceptions of the privacy implications of direct-to-consumer genetic testing. *New Genet Soc* 35(2), 101–123 (2016).
- 63. Kayser M. Forensic use of Y-chromosome DNA: a general overview. Human Genet. 136(5), 621-635 (2017).
- 64. Greytak E, Kaye D, Budowle B, Moore C, Armentrout S. Privacy and Genetic Genealogy Data. Science 361(6405), 857 (2018).
- Molteni M. The US Urgently Needs New Genetic Privacy Laws. Wired (2019). www.wired.com/story/the-us-urgently-needs-new-genetic-privacy-laws/