

COLLECTIVE CONSENT, RISKS, AND BENEFITS OF DNA

Xengie Doan*

Abstract

Health data is sensitive and sharing it could have many risks for personal or shared genetic data. So how can impacted individuals consent together? Collective consent has been used in person, but no digital collective consent exists yet. Challenges span legal-ethical issues and technical properties such as transparency and usability. To address these challenges, this work uses genetic data sharing as a use case to better understand what tools and methods can enhance a user-friendly, transparent, and legal-ethically aware collective consent.

Table of Contents

COLLECTIVE CONSENT, RISKS, AND BENEFITS OF DNA	161
Abstract.....	161
Keywords	162

* Xengie is a nonbinary scientist with Master's in Bioinformatics from the University of Oregon. After obtaining their Master's, they worked as a bioinformatician at the Stowers Institute for Medical Research assembling genomes, analyzing repetitive DNA sequences in cancer genomes, and creating semi-automated analysis pipelines. They then worked as a bioinformatics engineer at Sage Bionetworks helping build infrastructure, community sourced metadata models, and user-friendly dashboards and tools as part of a data coordination center for the Human Tumor Atlas Network. Now, Xengie is a PhD student with Dr. Gabriele Lenzini in the IRiSC lab at the SnT, University of Luxembourg working on transparent, private, and user-centered consent eHealth data sharing in the EU with LeADS.

This work is supported by the European Union's funded project Legality Attentive Data Scientists (LeADS) under Grant Agreement no. 956562.

xengie.doan@uni.lu

1. Understanding Genetic Data Sharing and Consent for Everyone: A Story....	162
2. So Why Should I Care?.....	164
3. Research Question.....	168
4. Results	168
4.1 Legal-Ethical Gaps.....	168
4.2 Genetic Testing Policies.....	170
4.3 User Needs and Desires.....	171
4.4 Business Testing.....	172
5. Limitations and Future Work	172
6. Conclusion.....	173

Keywords

Consent – Genetic Data – eHealth – GDPR– Business Process Improvement

1. Understanding Genetic Data Sharing and Consent for Everyone: A Story

Have you ever used a DNA test before? Maybe someone you know has. This is a common experience, as millions of people worldwide have bought and used a direct-to-consumer genetic test kit. We begin with a fictional story to illustrate the challenges in consent and sharing genetic data. The state is set in the quaint town of Genetica, where there was a family known for their unity and love, the Helixes. One day, the youngest member, Ada, driven by curiosity, decided to take a direct-to-consumer DNA test. From the website, it looks easy to do. She pays for a test and is excited to uncover the mysteries of their ancestry and share the exciting findings with her family. She follows steps outlined in the image below, which is taken from FamilyTreeDNA.com.

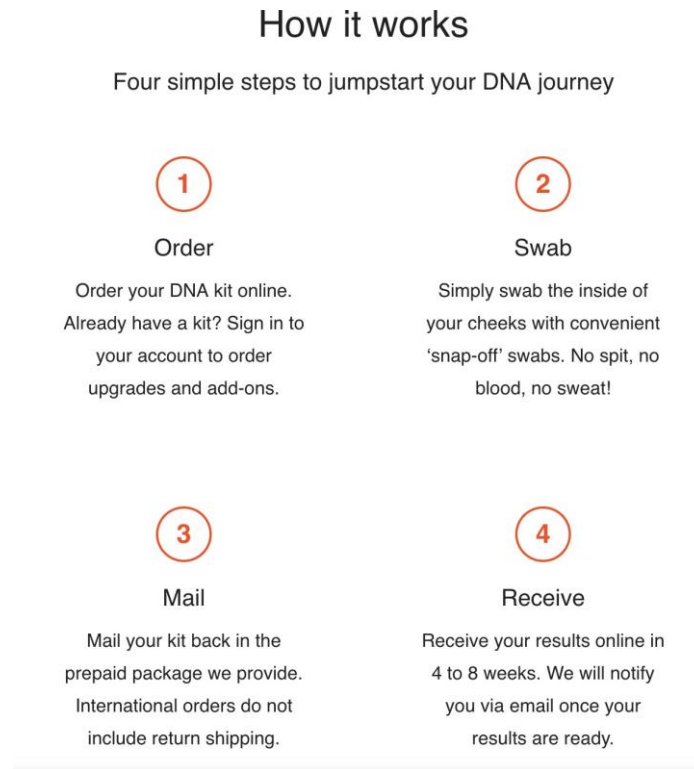


Figure 1: Steps for a DNA kit from FamilyTreeDNA.com.

As the results came in, Ada was thrilled. They were descendants of an interesting region of the world and had a possibility of diabetes! Eagerly, she shared the news at the next family gathering, expecting joy and wonder. Instead, she was met with a mixture of reactions.

Her uncle, a teacher, frowned deeply. “Did you consider the privacy risks, Ada?” he asked. “These companies collect DNA and ancestry information and don’t store it securely. Haven’t you heard of the recent data breach? They gather vast amounts of genetic data that could be used in ways we can’t even imagine yet.”

Ada's cousin, a lawyer, chimed in, "He's right. Law enforcement has used such data to track suspects without anyone's notice, consent, or knowledge ... ever heard of the Golden State Killer? Insurance companies might access it to increase premiums, and schools could use it to bar someone's admissions. The implications are far-reaching."

The family matriarch, Grandma Helix, who had seen much in her years, spoke softly, "Not all of us are equally at risk, my dear. What if someone will be discriminated against based on this information? Maybe our diabetes will be a sign of poor health and used against us. What if your future kids do not want their data shared? Your decision affects us all, yet we had no say in it."

Ada's heart sank as she realized the gravity of her action. The DNA data, like threads weaving through their lives, connected each family member, carrying potential risks they hadn't consented to.

As the discussion unfolded, the Helixes understood that while science might not yet predict their ancestry and health risk accurately, the future held possibilities that could turn their genetic information into a double-edged sword. The family decided to establish a formal process for such decisions, ensuring that everyone's voice was heard and recorded. They realized that in the tapestry of genetics, each thread was vital, and every pattern mattered.

From that day on, the Helixes approached their shared genetic legacy with caution and respect, knowing that the choices of one could unravel the fabric of many.

2. So Why Should I Care?

If you were placed in the story, how would you feel?

Researchers have been studying how we can all agree (or not) to share this kind of sensitive personal information. Personal information is anything that might identify someone under European law, and there are specific types that are called "sensitive" because there are added concerns. Genetic and health data is one of them, along with things like religion. Genetic data cannot be changed (like in the movie, *Gattaca*), and it can encode information about someone's family history, health, and more. It can be

hard to remove identifying information from DNA data, especially if there are multiple databases people can use to cross-reference. Now the world is collecting more data than ever, and it's not just about one person anymore. It can also be used for big data analytics, like machine learning algorithms or artificial intelligence. Our data can tell stories about our friends, family, and even people we don't know but who are like us in some way. It can be a big help for scientific and medical research, but it can also be leaked in data breaches or used without consent. We need a group decision-making system that's clear, fair, and respects everyone's rights—just like how indigenous groups have been doing it for a while.

As databases become more comprehensive, using a mix of public and privacy information increases the likelihood of re-identifying individuals from supposedly anonymized data. Researchers have cross-referenced datasets to uncover patterns that point to specific individuals, showing how easy it is to use public data to reveal sensitive information. This could lead to scenarios where a person's genetic traits, predispositions to diseases, or even their full names could be exposed without their knowledge. 23andMe also had a data break where millions of people's health data was leaked, which was released on the dark web and poses a privacy risk.

The traditional model of individual consent is not equipped to handle the complexities of genetic data, which inherently involves more than just the individual. This is where the concept of collective consent gains importance. It acknowledges that decisions about genetic data sharing should not be made in isolation. Instead, they should involve all those who could be affected by the data's disclosure. Collective consent allows for a more democratic approach, where the rights and concerns of everyone are considered, and decisions are made with the consensus of the group.

The implementation of collective consent would require a new framework that respects individuals while also recognizing the interconnected nature of genetic data. It would also need to have clear communication and understanding among all parties involved, ensuring that the benefits and risks of data sharing are transparent and well-understood. Under the GDPR, the key data protection regulation in the European Union (EU), this type of specific, informed consent is key.

Collective consent would expand individual ethics to include collective ethical approaches, resulting in a more equitable approach that empowers individuals and groups to have a voice in decisions that could impact their privacy and well-being. This should be clearly communicated and user-friendly to foster a sense of shared responsibility and mutual respect in the management of sensitive genetic information.

The Sensitivity of Genetic Data: Sharing health data, particularly genetic information, carries significant risks. Even when anonymized, genetic data can potentially be traced back to the individual it came from (Erich et al., 2018). This could reveal not just personal health risks, but also familial connections and population level health risks. Some benefits could be that health information about diseases helps to improve the family, but there are also risks. Such information could be used by the police to find suspects or adjust insurance premiums (Joly et al., 2013), affecting not only the person who agreed to share their data but also their relatives and others with similar genetic markers. In addition, there can be informal sharing of genetic results, like those from consumer DNA tests. But what if family members disagree with this decision? They have no formal say in the matter, even though they could be impacted. This may be especially complex when in the future, those generations wish they could track down how DNA data was shared and delete it. If there is no formalized tracking of the consent and data, it would become impossible to manage personal” data after a few generations.

Collective Consent: This is where collective consent comes in, which can help to protect people’s privacy and create a formal system for shared notice, consent, decision-making. Collective consent is a way for groups to agree on sharing genetic data. This idea comes from indigenous communities, who have asserted their right to approve biomedical research collectively, rather than individually (Hudson, 2009). This approach requires engagement with research partners and respect for the community as stakeholders rather than mere subjects, including considering non-Western cultural beliefs. To adapt it to everyone, collective consent would involve shared decision-making and keeping records of who agrees or disagrees with sharing genetic information. This is not just about being fair; it’s about respecting each person’s autonomy and rights. A system that formalizes decision-making and consent could protect privacy and individual rights. Collective consent works for indigenous

groups anymore; and it could be adapted for broader use in genetic data sharing. While the European Data Protection Board, an independent European Union body that is in place to ensure a consistent application and enforcement of data protection law, stated in their guidelines for DNA data recognize genetic data as personal data that may affect more than one person (e.g., may identify all members of a family because they are all genetically related), the current laws are designed for individuals, not groups, making enforcement tricky (*12178/03/EN WP 91 Working Document on Genetic Data*, 2004, Kuru and de Miguel Beriain, 2022).

A Framework for Digital Collective Consent: Collective consent has been effective for ensuring autonomy and rights for indigenous groups, where decisions are made by community leaders. However, in the era of extensive data collection, similar approaches may need to be digitalized to operate the new era of digital data sharing. While the goal is to develop a framework for digital collective consent, many challenges must be addressed first, including the transition from individual to collective models and the integration of user needs and risks. While collective consent is recognized in traditional settings (e.g., for physical consent where the participants are in-person), its digital counterpart is still in development. There are complex issues to resolve, such as balancing the right to privacy against the need to inform, defining decision-making processes, and establishing governance structures for the collective. Legal, ethical, technical, and trust-related factors all play a role in shaping digital collective consent. Consent itself is also complicated. Things like digital literacy, reading level, and data management preferences vary among users (Niemiec et al., 2018). Systems must consider informational transparency, storage and access controls, and user interface design. Trust in the institution requesting consent is also crucial and can be influenced by past events, such as data breaches.

A promising solution is dynamic consent, a flexible model that allows individuals to manage their consent and interact with research projects (Haas et al., 2021, Mascalzoni et al., 2022). It could potentially align with collective consent, but research on group decision-making in a digital context for genetic data is lacking. Questions remain about the appropriate level of transparency and the most relevant user attributes. This research looks to find a framework for digital collective consent that can address the needs that people have, help businesses improve their services, and share information

clearly. This leads to the research questions, which will not address the full problem, but try to answer a small part of it.

3. Research Question

The journey towards a formalized, legal-ethical framework for collective consent in genetic data sharing is filled with challenges. This work will explore the challenges of collective consent, then research user needs and the potential for tools and methods to increase transparency of the systems and address user needs. Legal, ethical, business, and stakeholder considerations must be balanced to create a consent process that is both effective and respectful of individual and collective rights.

To answer this, here's what I did:

- **Looked at the Gaps:** I worked with legal and ethical experts to check the ongoing challenges in consent for sharing health data in the EU.
- **Checked the Fine Print:** I worked with another lawyer to also read through the privacy policies of companies that test your DNA to see if they're clear, relevant, and fair.
- **Asked People What They Want:** Then, we asked regular adults in Germany what they wanted from consent and what format (video, text, newsletter, comic, infographic) they would prefer.
- **Tried It Out:** We tested methods to improve information transparency and clarity with a company in Norway to see if they could make better privacy policies and consent processes.

4. Results

4.1 Legal-Ethical Gaps

Self Determination: Imagine you're signing up for a new health app on your phone. It asks you to agree to terms and conditions that are really long and full of complicated words. This is about giving your permission, or consent, to let the app use your health

information. But it's not just for your doctor anymore; now, this info might be used for other things like ads, which are not necessary for your care. The problem is, that these apps don't always make it clear what you're agreeing to. Sometimes they don't even have a privacy policy that's easy to understand. So, when you just tap "agree," you might not really know what you're getting into. This makes people worry that we're not really in control of our own information, which is super important when it comes to our health.

So, are we really making our own choices if we don't fully understand what we're agreeing to? It's important that we can make our own decisions, which is called *self-determination*. This right is reflected in legal and ethical guidelines and regulations in Europe, and consent is a key part of lawful data processing and biomedical research.

Genetic Data is Shared: In Europe, the data protection board guidelines explore the way genetic data is shared between people. However, it's just a guideline, so there are no laws that force people to address collective data privacy. This is because the rules, if they went into action, would be very complicated. The current system is based on individual rights, and there is no precedent or framework for how to manage multiple people's rights together. There are no rules for agreements or disagreements. What if one person doesn't want to share the data, but another person does? Or what if someone doesn't even want to know the results or risks at all? Some think we can sort out these tricky situations by starting with the rules we already have and working from there. Different countries are also trying their own ways to balance everyone's rights. For example, sometimes they decide that knowing about health is more important than keeping a secret and a doctor can share important health information with relatives.

Specific Consent: The General Data Protection Regulation (GDPR) says that consent should be clear and specific. This means you should know exactly what you're agreeing to and why. But sometimes, it's hard to be specific. Like when scientists collect data for research, they might not know all the ways they'll use it in the future. So, they ask for a "broad consent," which is like saying, "I trust you to use my data for good things later on." The law tries to protect us by saying we need to give specific consent. But if we're too strict about this, we might end up having to say "yes" over

and over again for each little thing, which can get really tiring. There's a part of the law (a Recital that tries to expand on the main body) that tries to help by saying it's okay to give broad consent for science, as long as it follows good ethical standards. But this isn't supported in the core regulation, which still talks about being specific. So, it's confusing in practice and people can interpret it in different ways.

Consent is Relevant Even when its not required by the law, ethical consent plays a crucial role in data processing. First, what is the difference? Legal vs. Ethical Consent:

- Legal consent refers to obtaining permission from individuals to process their personal data based on legal requirements (such as the General Data Protection Regulation, GDPR).
- Ethical consent, on the other hand, goes beyond legal obligations. It ensures that data processing respects human dignity, autonomy, and privacy.

Second, ethical consent is relevant as a legal safeguard to help make sure people's right are taken care of. Some researchers (Staunton et al., 2019) argue that ethical requirements such as consent and transparency could serve as safeguards to help inform the data subject of their rights. Third, to uphold self-determination, consent should be asked. Ethics scholars have prioritized this decision making, or autonomy, and they are part of the internationally recognized guidelines regarding people's rights for medical research, like the Belmont Report, to ensure people's rights are safe in biomedicine. It should also be considered in a collective sense,

when multiple people's autonomy is involved.

Without consent, there would be less self-determination and safe- guards for data collection, sharing, and processing activities.

4.2 Genetic Testing Policies

Privacy policies are like public explanation of the internal rulebooks that companies follow when handling your personal data. They are required to write their data processing activities so you can be informed of what is happening, how they store the data, and more. Transparency means being open and clear about what they're doing

with your information. While the companies should write everything they are doing and have it been clear what is happening, many privacy policies fall short, causing confusion and misaligned expectations. Maybe you think they are using the best privacy techniques, but your data can be easily hacked. Companies that test DNA data are dealing with sensitive genetic data (like DNA, your family tree, health and wellness information, and more). They have millions of users, but how well do they communicate their practices to users?

We looked at privacy policies from the 6 top companies chosen and analyzed the sections that talk about how they share genetic data. Shockingly, 81% of these explanations were vague, using terms like may, possibly, perhaps, etc. 37% were confusing with more than 2 distinct subjects or purposes for processing the data (some had 10+ unique purposes in one section), making it hard for users to understand. The GDPR requires clear, direct, and transparent information so they might not fully meet the legal requirements. In addition, the way the information is framed is tailored for legal experts, so it makes it hard for a normal person to understand what is happening. They also don't address the collective responsibility of sharing DNA data, which can affect your family. They just say how it will affect their single customer, which makes it seem like it's not a big issue. Some companies only share one risk of anything happening, while others detail about the possible risks. Some could be great and affect your family. So, we suggest making policies more user-centered by framing it in an useful way to general users and more risk aware.

4.3 User Needs and Desires

Consent forms for individuals are complicated. So how can we make them better and try to apply them to collectives?

We looked at what people wanted from consent and the way they want it presented. We took a portion of the consent process and translated it into an infographic, video, text, newsletter, and comic. Then we interviewed 24 German adults about their expectations and experiences with consent forms.

We found that people have different goals when reading consent forms, and while everyone wanted easy to understand information, some relied more on trust (e.g., how

much they trusted their doctor or the organization) or placed a burden on themselves to record the consent decision.

The infographic was the top voted way to receive consent information because it helped understanding and they could focus on important info. It suited serious scenarios (like health data consent), where comics were seen as childish or unserious for health data consent. Things like structure and readability were also important because they helped make the form engaging and easier to follow.

The participants also wanted a centralized digital platform to manage consent over time, and possibly to revoke consent.

4.4 Business Testing

Following the previous results, we were interested in how to improve privacy policies before they go public. They should reflect the company's own actions for how they collect, use, and share your data. To create better privacy policies, we need a solid approach, or methodology. While there are many methods, none considers the context. In this ongoing work, I tested the usefulness of methods for clarifying privacy policies and consent processes with 13 employees at a company in Norway to see if they thought it would be useful for their jobs.

Results are still being published, but it seems that the employees do appreciate the methods for giving a more structured, visual, and understandable way to process very complex information.

5. Limitations and Future Work

First, these studies cover many different areas with a limited scope. Many of the studies mentioned were pilot projects to check the feasibility and interest in the method, idea, or framework. While useful, they might not cover all aspects or scenarios. More in-depth, larger-scale studies could provide a more comprehensive view. Second, there may be sampling bias. For different studies, the way participants were chosen for the study could introduce bias. This is especially true for the testing in Norway because we were constrained by the availability of people on different teams. In the German

study, we looked for an equal number of people of different ages, education levels, and sexes. A more diverse participant pool would enhance the study's validity. We also only looked at one company, and it might not be as reliable. More deliberate sampling and larger and more varied samples would strengthen the findings.

The work was also around a specific domain, so it would not be applicable to others. The work focused on health and DNA data and applying these findings to other fields (like finance or social media) might not work because consent norms can differ across domains. Consent has a strong history in biomedicine and people are used to it in health-related scenarios. From the study with German participants, they took it more seriously than cookie consents.

The studies also had a short-term focus. We looked at privacy policies and laws at one point in time, and the user studies were for immediate perceptions, understanding, or engagement. We did not explore how perceptions might change over time, or measure beyond people's perceptions. Maybe people will have different ideas after working with the concepts for longer, or putting it into practice.

Future work includes repeating and extending the studies. They should be repeated, and if people were interviewed, more people should be interviewed. It should try to get people from many different demographics, like country, age, sex, etc. to be more generalizable, so the results could apply to more people. For the user needs, the study should be repeated in different countries to get a better understanding of how the context (like how comics are not serious enough) might change from country to country. It would be interesting to have a prototype of collective consent using all the suggestions from this work to see how people react to it.

6. Conclusion

From this, we have a better understanding of what users want and can try to translate the findings for collective consent design. This can tie into previous methods such as the analysis of privacy policies, which can be shown before the consent form, and be prototyped all together into a dynamic consent system. Altogether, it is a first step into building a digital collective consent, deeply considering the gaps, user-needs, and possible implementations in small businesses. As shown in Figure 2, the dynamic

consent prototype should be built and tested based on a specific use-case for collective consent on the findings from this work, then be used iteratively to build better systems from the ground up.

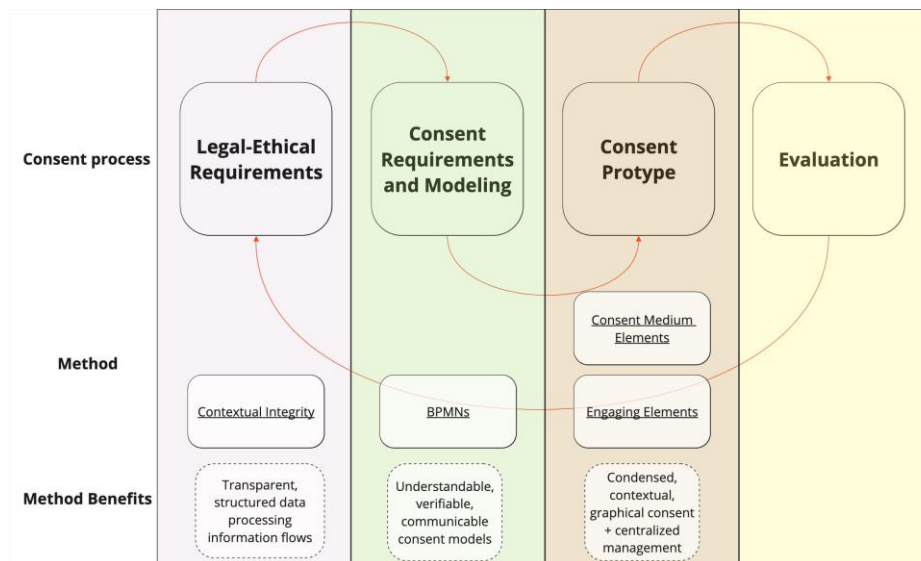


Figure 2: A visual summary of the work as it targets different parts of the consent process, the methods tested, and the top findings

References

- 12178/03/en wp 91 working document on genetic data. (2004). https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2004/wp91_en.pdf
- Erich, Y., Shor, T., Pe'er, I., & Carmi, S. (2018). Identity inference of genomic data using long-range familial searches. *Science*, 362 (6415), 690–694.
- Haas, M. A., Teare, H., Prictor, M., Ceregra, G., Vidgen, M. E., Bunker, D., Kaye, J., & Boughtwood, T. (2021). ‘ctrl’: An online, dynamic consent and participant

engagement platform working towards solving the complexities of consent in genomic research. *European Journal of Human Genetics*, 29 (4), 687–698.

Hudson, M. (2009). Think globally, act locally: Collective consent and the ethics of knowledge production. *International Social Science Journal*, 60 (195), 125–133

Joly, Y., Ngueng Feze, I., & Simard, J. (2013). Genetic discrimination and life insurance: A systematic review of the evidence. *BMC medicine*, 11, 1–15.

Kuru, T., & de Miguel Beriain, I. (2022). Your genetic data is my genetic data: Unveiling another enforcement issue of the gdpr. *Computer Law & Security Review*, 47, 105752.

Mascalzoni, D., Melotti, R., Pattaro, C., Pramstaller, P. P., Gögele, M., De Grandi, A., & Biasiotto, R. (2022). Ten years of dynamic consent in the chris study: Informed consent as a dynamic process. *European Journal of Human Genetics*, 30 (12), 1391–1397.

Niemiec, E., Vears, D. F., Borry, P., & Howard, H. C. (2018). Readability of informed consent forms for whole-exome and whole-genome sequencing. *Journal of Community Genetics*, 9 (2), 143–151. <https://doi.org/10.1007/s12687-017-0324-6>

[//doi.org/10.1007/s12687-017-0324-6](https://doi.org/10.1007/s12687-017-0324-6)

Staunton, C., Slokenberga, S., & Mascalzoni, D. (2019). The GDPR and the research exemption: Considerations on the necessary safeguards for research biobanks [Number: 8 Publisher: Nature Publishing Group]. *European Journal of Human Genetics*, 27 (8), 1159–1167. <https://doi.org/10.1038/s41431-019-0386->

