

# Design Debt and Data Governance Failure: An Analysis of the 23andMe Genetic Breach

David J. Yates  
dyates@bentley.edu

Arthur Ream III  
areamiii@bentley.edu

Department of Computer Information Systems  
Bentley University  
Waltham, MA 02452, USA

## Abstract

This paper examines the 2023 credential-stuffing attack on 23andMe, a leading Direct-to-Consumer Genetic Testing (DTC-GT) company, as an analytical case of platform design, data governance, and institutional failure. The compromise of 14,000 user accounts cascaded through the company's DNA Relatives feature to expose the genetic profiles of 6.9 million individuals, demonstrating how relational architectures magnify the consequences of security breaches. We analyze the incident across six dimensions – context, breach mechanics, organizational response, impact assessment, post-incident governance, and broader implications – to show how design debt, regulatory ambiguity, and leadership inaction amplified both technical and social harms. Our analysis contributes to cybersecurity and privacy research by reframing breach narratives that emphasize individual user error, highlighting instead the systemic risks of shared, persistent, and identity-linked data. Furthermore, we situate the 23andMe breach within emerging debates on relational privacy, consent structures, and fiduciary responsibility for biometric information, underscoring that contractual consent and weak regulatory regimes fail to protect collective data subjects. We further assess the implications of corporate restructuring and bankruptcy for genetic data stewardship, drawing attention to unresolved governance questions when sensitive datasets are treated as transferable assets. The 23andMe journey demonstrates how technical architectures, business models, and regulatory gaps converge to create privacy fragility, offering insights for scholars of cybersecurity, digital governance, and platform accountability.

**Keywords:** Privacy risk, genetic data, data governance, cybersecurity breach, relational consent, platform accountability

# Design Debt and Data Governance Failure: An Analysis of the 23andMe Genetic Breach

David J. Yates and Arthur Ream III

## 1. 23andMe AND THE DNA RELATIVES FEATURE

Founded in 2006, 23andMe became one of the most prominent direct-to-consumer genetic testing (DTC-GT) firms by marketing DNA testing kits that offered users insights into their ancestry, genetic traits, and health predispositions, without involving a clinician. Its rise coincided with a surge in public curiosity about personal genomics and the growing commercial value of large-scale genetic databases (Greely, 2020). As the first company to receive FDA authorization to offer genetic health risk reports directly to consumers, 23andMe built its brand on scientific empowerment and user-friendly data experiences.

By 2022, the platform had amassed genetic data from over 12 million users worldwide and had diversified its revenue model through research partnerships and data-driven drug discovery, notably a \$300 million agreement with GlaxoSmithKline (Greely, 2020). However, this data-centric business model also introduced new risks. As Boyd and Crawford (2012) note, when personal data becomes a commodity, user empowerment narratives can obscure the power asymmetries and opacity built into large-scale information systems.

A key feature of the 23andMe platform was its *DNA Relatives* tool. This opt-in feature allowed users to identify and connect with genetically similar individuals in the 23andMe database based on shared autosomal DNA segments. Participants could view predicted relationships, traits, and often names, locations, and photos.

While the DNA Relatives feature offered social and genealogical benefits, such as reuniting adoptees or identifying unknown relatives, it also created relational exposure. Opting in affected not only the user but also genetically similar individuals who never joined or declined the feature (Phillips, 2016; Zettler et al., 2019). These second-order exposures complicate traditional consent models by revealing how genomic data is inherently shared. A single DNA sequence conveys probabilistic insights into the ancestry, health predispositions, and identities of biological

relatives without their knowledge or consent, making genetic privacy uniquely collective and vulnerable.

Compounding these concerns were the company's evolving privacy terms. According to 23andMe's most recent Privacy Statement (Version 7.5), the platform collects registration details, genetic and sample data, self-reported traits, web behavior, and aggregate datasets described as anonymized but still susceptible to re-identification under certain conditions (Gymrek et al., 2013; 23andMe, n.d.). The company also reserves the right to update these terms unilaterally, and past versions are not always publicly archived, a practice that challenges transparency and informed consent.

In technical terms, DNA Relatives effectively transformed the 23andMe platform into a relational network. This design choice increased the volume of linkages and the potential attack surface. Even though only a subset of users enabled the feature, each new participant expanded the graph of visibility across the system. As Erlich and Narayanan (2014) showed in their foundational study on genetic re-identification, even anonymized genomic data can be traced back to named individuals through cross-referencing with public records, genealogical databases, and auxiliary datasets. The addition of names, photos, and family trees through DNA Relatives only lowers the threshold for such inference attacks.

These risks became alarmingly real in late 2023 when 23andMe was breached by an attacker known as *Golem*, who exploited reused credentials from external data breaches to gain access to approximately 14,000 user accounts (Holthouse, Owens & Bhunia, 2025). However, due to the structure of DNA Relatives, the attacker accessed profile data of approximately 6.9 million other users linked to those accounts (23andMe, 2024). The attack highlighted the vulnerability of relational features that lacked containment boundaries, allowing one compromised account to reveal sensitive information about many others.

23andMe initially framed the breach as due to user error – specifically password reuse – rather than a platform failure (23andMe, 2023). Yet cybersecurity experts argued that failure to mandate two-factor authentication (2FA) or implement automated login monitoring contributed significantly to the breach's scope (Florêncio & Herley, 2010). In turn, privacy scholars noted that relational data exposures through DNA Relatives magnified the harm beyond what users might have reasonably anticipated or consented to (Nissenbaum, 2004).

Beyond the technical and consent issues, the 23andMe breach raises broader questions about data governance. As Bradshaw, Millard, and Walden (2011) warn, cloud-based platforms that collect sensitive data often embed complex third-party arrangements and variable retention policies that create uncertainty for users. In the case of 23andMe, the company's Terms of Service explicitly allow for the transfer of user data in the event of a merger, acquisition, or bankruptcy – provisions that would become central when the company filed for Chapter 11 bankruptcy and was sold in 2025 (Gerke, Jacoby & Cohen, 2025; Hernandez, 2025).

In sum, by 2023, the DNA Relatives feature exemplified the trade-offs embedded in genetic platform design; it enabled connection but also catalyzed cascading data exposure. It personalized the user experience while externalizing privacy risk to genetic relatives. It emphasized empowerment through data but blurred the boundaries of meaningful consent. For instructors and students, the 23andMe context offers a critical opportunity to examine how privacy risks arise not just from singular failures, but from the cumulative effects of design decisions, governance choices, and legal ambiguity in a data-intensive business model.

The remainder of our analysis is organized into five sections. Section 2 outlines the immediate trigger for the breach and describes the initial attack vector and user impact. Section 3 details the company's short-term incident response, public communications, and internal mitigation steps. Section 4 analyzes how platform design decisions, visibility defaults, and consent architectures contributed to the scale and persistence of harm. These architectural choices and policy provisions represent a form of design debt, where short-term emphasis on engagement and growth created long-term vulnerabilities in privacy and governance. Section 5 examines broader governance and accountability failures, including the ethical implications of asset transfer

during bankruptcy. Finally, Section 6 reflects on limitations, highlights lessons for cybersecurity education, and proposes directions for reform in data governance and platform design.

## **2. CREDENTIAL STUFFING ATTACK AND SCOPE OF EXPOSURE**

In October 2023, 23andMe publicly confirmed a major security incident that compromised genetic profile data linked to 6.9 million users. At the core of the breach was a method known as credential stuffing, a technique in which attackers use previously leaked username-password combinations, often gathered from unrelated data breaches, to gain unauthorized access to user accounts. Credential stuffing is particularly effective against platforms that do not enforce multi-factor authentication (MFA) or monitor for high-volume login attempts, both of which were lacking on 23andMe's platform at the time (Holthouse, Owens & Bhunia, 2025).

The attacker, using the alias Golem, exploited this weakness to access approximately 14,000 user accounts directly through reused credentials. Many of these accounts, in turn, had access to the DNA Relatives feature, which allowed the attacker to view profile data linked to an additional 6.9 million users (23andMe, 2024; Holthouse, Owens & Bhunia, 2025). The breadth of the breach, then, was not due to a deep compromise of internal systems or unauthorized access to databases, but to an authentication architecture that failed to prevent (or even detect) credential-stuffing, and to a design architecture that facilitated relational visibility. A single compromised account had the potential to reveal names, ancestry results, locations, photos, and family connections of numerous relatives – many of whom had not themselves been directly breached (Erllich & Narayanan, 2014).

The DNA Relatives feature, while marketed as a tool for discovery and connection, became a vector for cascading exposure. As critics later noted, the platform's privacy architecture lacked internal boundaries to prevent second-order disclosures. Unlike data minimization approaches where access is limited to essential attributes, 23andMe's model prioritized user engagement and discovery, allowing wide visibility into linked profiles through shared genetic segments.

The breach was first acknowledged by 23andMe in a public statement on October 6, 2023, in which the company emphasized that its internal systems had not been "hacked" in the traditional sense. Instead, it framed the incident as the

result of users failing to protect their own credentials (23andMe, 2024). This framing deflected responsibility onto individuals, even though research in security usability has consistently shown that password reuse is endemic and foreseeable, particularly in systems lacking proactive defenses (Florêncio & Herley, 2010).

Following the attack, Golem began advertising ethnicity-specific datasets on dark web forums, beginning with profiles of over 1 million Ashkenazi Jewish users and over 100,000 users of Chinese descent (Carballo, Schmall & Tumin, 2024). These datasets were offered for sale at \$1–\$10 per profile, raising immediate concerns about ethnic targeting and genetic discrimination. The sale of curated, population-specific data introduced not only privacy harms but also potential threats to civil liberties, echoing past warnings from scholars about the discriminatory potential of genomic data misuse (Macdonald, 2024; McGuire, Caulfield & Cho, 2008).

While the breach involved no direct intrusion into the company's servers, its scale and sensitivity far exceeded typical credential-stuffing attacks. What made the 23andMe incident extraordinary was the way a relatively small number of account compromises unlocked access to a vast network of sensitive profiles, enabled by poor product design rather than technical exploit. This architectural vulnerability illustrates a form of privacy fragility (Kotlan, Magoon & Yates, 2026), in which the failure of a single account unlocks cascading exposures across an entire relational network.

The exposed data included not just genetic ancestry results, but also inferred haplogroups (genetic population group consisting of individuals who share a common ancestor through a specific set of mutations in DNA), relative matches, photos, names, user-submitted traits, and geographic locations (Wikipedia, n.d.) – constituting a highly identifiable composite of personal and biometric data. Research shows anonymized genetic datasets can often be re-identified, especially when paired with auxiliary data (Bampoulidis & Lupu, 2019; Narayan, Kohli & Martin, 2025).

From a governance perspective, the breach reflected multiple systemic oversights:

- MFA was optional rather than mandatory, despite the sensitive nature of the data;
- Rate-limiting or anomaly detection was insufficient to flag automated login

attempts; and

- Users were not alerted that their DNA Relatives settings could expose others, nor were they informed of the full consequences of a breach.

Such omissions undermined cybersecurity best practices and reflected a fiduciary lapse, as the company failed to exercise care proportional to the sensitivity of genetic data. These choices reveal a misalignment between the permanence of genetic information (its enduring, immutable, and unchanging nature), predictive potential, and shared nature of DNA relative to the safeguards applied to it. Treating genetic data like ordinary consumer information leaves organizations unprepared for the scale of harm possible when such enduring and identity-linked information is exposed.

Regulatory scrutiny quickly followed, as state attorneys general opened investigations and lawsuits alleged deceptive practices and weak protections (Hernandez, 2025; Kirk, 2025). Many users never opted into DNA Relatives yet were profiled and exposed, raising new legal and ethical issues about relational consent and platform accountability (Gerke, Jacoby & Cohen, 2025). These concerns reflect how fiduciary responsibility for biometric data was reduced to contractual formality rather than enforceable obligation.

In sum, the breach exposed a design flaw where one user's compromise endangered many others, challenging breach narratives centered on individual error. It underscored the need for stronger safeguards, realistic user expectations, and recognition of shared data risks in consumer genomics. Ultimately, privacy fragility stems less from isolated mistakes than from accumulated design debt and weak governance that expose entire families and populations.

### **3. IMMEDIATE RESPONSE: COMMUNICATIONS, REMEDIATION, AND GOVERNANCE BREAKDOWN**

23andMe's response to the October 2023 credential-stuffing breach was widely criticized as fragmented, slow, and overly focused on user blame. The firm's initial public acknowledgment of the breach, posted on October 6, 2023, stated that its systems had not been "hacked" in the conventional sense and instead framed the incident as caused by users reusing passwords from prior data breaches (23andMe, 2024). This framing, while technically accurate, reflected a deflection strategy that has drawn increasing

scrutiny in cybersecurity governance literature (Florêncio & Herley, 2010).

Rather than a platform-wide alert or full forensic audit, 23andMe relied on staggered disclosures. The company's first breach notification to the California Attorney General did not occur until January 2024, nearly three months after the public confirmation of the attack and four months after initial breach activity was observed (23andMe, 2024). The delay was justified by the company as a function of investigative complexity, yet it undermined public trust and raised questions about the company's readiness to handle sensitive data at scale (Hernandez, 2025).

When 23andMe eventually forced password resets for affected accounts in December 2023, it did so *without requiring* two-factor authentication (2FA), a critical omission given the nature of the attack. Optional 2FA had long been a vulnerability in the platform's design, particularly in a context where the data at stake – genetic, relational, and identity-linked – was uniquely sensitive and non-revocable.

The breach also revealed a governance breakdown across internal and external response layers. Internally, there was no public evidence of a structured incident response plan aligned with frameworks such as NIST SP 800-61 Rev. 3 (Nelson et al., 2025), which emphasizes preparation, communication, and post-incident review. Externally, the company delayed communication not only with regulators and affected users, but also with research partners and public stakeholders. By the time the breach was formally disclosed to users in January 2024, a version of the compromised data had already circulated widely in dark web markets.

23andMe's breach FAQ and blog communications adopted a minimization strategy, emphasizing that only users who reused passwords were directly compromised and that DNA Relatives participation was voluntary (23andMe, 2023). Yet this response sidestepped a critical architectural issue. The relational exposure model embedded in DNA Relatives enabled cascading harm (Erllich & Narayanan, 2014). According to user lawsuits filed in early 2024, many plaintiffs had not opted into DNA Relatives, but were nonetheless exposed through genetic linkages to those who had (Kirk, 2025). These claims underscore how second-order harms – those suffered by individuals not directly involved in a breach mechanism – challenge traditional notions of consent and legal standing, as they violate

contextual norms of information flow (Nissenbaum, 2004) and expose the limits of personally identifiable information as a legal category (Schwartz & Solove, 2011).

Adding to concerns was the lack of individualized notification to users whose data had been accessed indirectly. While 23andMe eventually offered free credit monitoring to a subset of affected users, the offer was narrow and did not extend to relatives whose data may have been linked or downloaded through matching profiles. This narrow remedial approach is inconsistent with modern interpretations of data protection best practices, including those grounded in Fair Information Practices (FIPs) and the GDPR's emphasis on data subject rights (Gellman, 2025; Greenleaf, 2012; Shabani & Borry, 2015).

Further eroding public confidence, 23andMe's board and executive leadership remained relatively silent during the crisis. CEO Anne Wojcicki did not issue a personal statement until several months after the breach. This silence contrasted sharply with the company's pre-IPO branding as a science-forward, ethics-conscious platform (Carballo, Schmall & Tumin, 2024; Rutherford, 2025). Wojcicki eventually testified before the U.S. House Oversight Committee on June 10, 2025, at a hearing titled *Securing Americans' Genetic Information: Privacy and National Security Concerns Surrounding 23andMe's Bankruptcy Sale*.

The governance failures were magnified by conflicts between regulatory frameworks and platform policy. U.S. law does not require immediate breach notification to indirectly affected parties. Genetic privacy is not fully protected by HIPAA, and while the Genetic Information Nondiscrimination Act (GINA) prohibits use of genetic data in employment and health insurance contexts, it does not apply to consumer data usage or data sales (Rothstein, 2008; Terry, 2012). As Gerke, Jacoby, and Cohen (2025) note, this regulatory ambiguity becomes especially concerning when platforms like 23andMe pursue bankruptcy, asset transfers, or sale – events that introduce risk of data commodification during restructuring.

Indeed, 23andMe's initial response offered no binding assurances that breached or collected data would be deleted, sealed, or constrained in future transactions. Only after increased media and legal pressure did the company release a limited action plan, pledging to strengthen authentication, expand security auditing, and provide more granular consent options

(23andMe, 2023). However, these commitments were voluntary and non-enforceable, and critics noted that they appeared to be more timed for litigation strategy and public relations than for systemic reform.

In the months that followed, the firm faced class-action litigation, multiple state investigations, and rising scrutiny from data protection authorities in the U.S., U.K., and EU. Despite these pressures, 23andMe remained legally within its Terms of Service in most jurisdictions – a fact that highlights how contractual models of consent can undercut substantive privacy protections, especially in the absence of strong sectoral or omnibus regulation.

In short, 23andMe's response to the breach was technically incomplete, procedurally delayed, and strategically defensive. It failed to meet core expectations of cybersecurity readiness, transparency, and ethical stewardship (Barocas & Nissenbaum, 2009). The company's attempt to blame individual users obscured the broader institutional, architectural, and legal contributors to the incident. This response phase offers a case study in how insufficient preparation and minimization rhetoric can intensify the long-term damage of a breach – reputationally, legally, and operationally.

#### **4. IMPACT ASSESSMENT: DATA BREACH SCALE, SENSITIVE DATA AT RISK, AND LEGAL/FINANCIAL FALLOUT**

The 2023 breach of 23andMe exposed a profound misalignment between the scale of sensitive data collected and the strength of the privacy and security infrastructure protecting it. Although the attacker directly accessed approximately 14,000 user accounts using credential stuffing techniques, the design of the platform's DNA Relatives feature allowed data associated with an estimated 6.9 million additional users to be indirectly accessed and scraped (23andMe, 2024; Lanzing, 2016; Holthouse, Owens & Bhunia, 2025). This amplification of impact, due not to malware or insider threat but to architectural design, significantly heightened the breach's scope and long-term consequences.

The exposed data included names, ancestral origins, haplogroup classifications, familial matches, profile photographs, geographic locations, and shared DNA segment information. These biometric and familial attributes are difficult to revoke or secure (Erich & Narayanan, 2014; Wikipedia, n.d.). Genomic data is immutable. Its value also grows over time due to

its predictive, inferential, and relational qualities, which can affect not only users but their relatives and descendants (Narayan, Kohli & Martin, 2025).

The breach triggered not only technical containment challenges but also legal, regulatory, and reputational fallout. In early 2024, affected users in the U.S. and U.K. filed class-action lawsuits against 23andMe, alleging deceptive business practices and negligent data protection. Plaintiffs emphasized that even users who had not opted into the DNA Relatives feature were exposed, since their genetic information was visible to relatives who had; an allegation that underscored the shared nature of genetic data and the inadequacy of traditional, individual-centered consent models (Kirk, 2025).

Regulators in the U.S., Canada, and the U.K. initiated formal inquiries. In the U.K., the Information Commissioner's Office (ICO) levied a £2.31 million fine for GDPR violations, citing insufficient organizational and technical safeguards to protect special-category data (see <https://ico.org.uk/media2/kcbljpo/23andme-penalty-notice.pdf>). The filing associated with the fine identified evidence of raw genetic data downloads for four customers (i.e., a small number). In the U.S., the Federal Trade Commission (FTC) signaled Section 5 concerns about privacy promises and data transfer during bankruptcy, particularly around consent to data sharing and notice of downstream risks (Lee, 2025).

A \$30 million settlement agreement was preliminarily approved in 2024, though it covered only certain direct users and did not extend to relatives affected through DNA Relatives (Hernandez, 2025). Plaintiffs and privacy scholars alike pointed out that these legal remedies remained narrow, largely because U.S. privacy law does not yet recognize shared data risk or second-order data subjects (Nissenbaum, 2004; Schwartz & Solove, 2011).

In financial terms, the breach hastened 23andMe's decline. Valued at about \$3.5B at IPO, its shares fell below \$1 by 2023, and the firm entered Chapter 11 in March 2025 (Greely, 2020; Hernandez, 2025). By May 2023, it reported over 14 million genotyped customers. By bankruptcy filing, the company reported more than 15 million customers, alongside over \$300 million in cumulative losses. Appendix B shows the 23andMe share price over time and presents a timeline of some of the firm's missteps.

The bankruptcy itself raised further concerns about the disposition of genetic data. As Gerke, Jacoby, and Cohen (2025) argue, the U.S. bankruptcy code offers limited safeguards for personal data, including sensitive genetic data, once it becomes a commercial asset. Although 23andMe's privacy policy claimed that personal information would not be sold without user consent, it also reserved the right to transfer data in the event of corporate restructuring – a contradiction that became material during the bankruptcy auction (23andMe, n.d.; Bradshaw, Millard & Walden, 2011).

On July 14, 2025, Anne Wojcicki, 23andMe's co-founder and former CEO, successfully repurchased the company's assets, including its vast genetic data trove, through a new entity, TTAM Research Institute. Competing bidders, including Regeneron Pharmaceuticals, had reportedly sought access to the dataset for research and commercial applications (Saey, 2025; Kirk, 2025). The sale raised alarms among privacy advocates, who questioned whether the original terms of user consent extended to such post-bankruptcy transfers. While the company asserted that its new owner would honor prior privacy terms, those terms had been updated multiple times since 2008, often without requiring affirmative opt-in from legacy users (Gellman, 2025).

The long-term impact of the breach also included the chilling of public trust in direct-to-consumer genetic testing. Media coverage from *The Guardian* and *Science News* noted a growing movement among users to delete their profiles, request raw data downloads, and disengage from the platform altogether (Saey, 2025; Rutherford, 2025). Yet for many users, the ability to fully retract genetic data was limited by the company's retention policies and data-sharing commitments already in place with research partners.

Ethically, the incident reinforced what privacy scholars have long argued – that privacy harms in data-driven systems are often cumulative, systemic, and difficult to trace to a single bad actor (Boyd & Crawford, 2012). The breach exposed how platform design, business strategy, legal gaps, and relational data architecture can converge to create privacy fragility. Privacy fragility is a state in which the failure of even one node in the system can result in widespread and irreversible harm (Phillips, 2016).

For students of cybersecurity, law, and digital governance, the 23andMe breach offers a rare and instructive convergence of breach mechanics,

shared risk, governance breakdown, and regulatory ambiguity. Its impact is still felt and it has already prompted new calls for reform, including:

- Stronger breach notification standards that include indirect victims;
- Clearer limits on genetic data reuse in bankruptcy or acquisition; and
- Greater accountability for platform decisions that affect privacy beyond the individual.

## **5. POST-INCIDENT GOVERNANCE: ACCOUNTABILITY, ACQUISITION, AND LESSONS FOR FUTURE PLATFORMS**

In the aftermath of the 2023 breach, 23andMe entered a prolonged phase of legal, operational, and reputational crisis that included its Chapter 11 bankruptcy filing in March 2025 (Hernandez, 2025). Rather than restoring public trust through transparent remediation and systemic governance reform, the company adopted a narrow incident framing and defensive communications strategy, undermining its credibility. As litigation, regulatory scrutiny, and user attrition accelerated, 23andMe became emblematic of how digital platforms, especially those dealing in biometric data, can succumb to cascading governance failures when architectural risk, legal ambiguity, and public trust intersect.

### **Governance and Fiduciary Shortfalls**

Throughout 2024, 23andMe's board of directors and executive leadership failed to issue a comprehensive public review of the breach or its systemic causes. CEO Anne Wojcicki made only limited public comments, and no member of the executive team appeared before Congress or in regulatory hearings to explain the company's handling of the incident (Rutherford, 2025) until June 2025. Nor did the company commission an independent audit or breach report (Schwartz & Solove, 2011).

The lack of transparency and third-party oversight drew criticism from privacy experts and investor advocacy groups, many of whom called attention to the conflict between fiduciary duties to shareholders and ethical obligations to users. The lack of internal reform or institutional accountability revealed a broader failure to treat genetic data governance as a matter of public trust rather than proprietary control (Nissenbaum, 2004; Barocas & Nissenbaum, 2009; Kawaguchi & Lee, 2025).

### **Legal Ambiguity and Platform Leverage**

23andMe operated in a legal gray zone. As a

consumer genetic testing company, it was not covered by HIPAA, and GINA's protections were limited to preventing discrimination in employment and health insurance, not governing data processing or resale (Rothstein, 2008; Terry, 2012). Its Terms of Service permitted transferring user data in mergers, acquisitions, or bankruptcy – a clause that became critical when the company's assets, including its genomic database, were auctioned during bankruptcy proceedings (Bradshaw, Millard & Walden, 2011).

In June 2025, TTAM Research Institute, a new company founded by Anne Wojcicki, emerged as the winning bidder for 23andMe's assets in a deal worth \$305 million, surpassing rival bids from entities including Regeneron Pharmaceuticals (Saey, 2025; Kirk, 2025; Herper, 2025). While the sale, which was finalized on July 14, has promised continuity for existing users and partners, it also raised urgent concerns about data portability, retroactive consent, and platform self-acquisition. Critics questioned whether consent agreements made years earlier – often under older privacy policies and opt-in frameworks – could legally or ethically support the wholesale transfer of sensitive biometric and relational data to a newly formed entity, even one helmed by the company's former CEO (Gerke, Jacoby & Cohen, 2025; Gellman, 2025).

The absence of a user re-consent process before or after the transfer further weakened the credibility of 23andMe's governance claims. Privacy scholars have noted that the transfer of data under these conditions effectively converted consent into a one-time contractual event, rather than an ongoing, contextual process – a practice incompatible with modern interpretations of informed consent and privacy-by-design principles (Greenleaf, 2012; Shabani & Borry, 2015).

### **Design Debt and Institutional Inertia**

The breach and its aftermath have illustrated how design debt – the accumulation of risky architectural decisions made for convenience or growth – can compound over time into governance crises. The DNA Relatives feature, though widely used, lacked containment safeguards or explicit disclosure about secondary exposures. Despite repeated updates to its privacy policy and a growing user base, the company failed to implement system-level controls that would prevent cascading visibility through kinship graphs (Erich & Narayanan, 2014; Phillips, 2016).

Nor did 23andMe establish a governance

mechanism for shared data risk, such as providing control panels that allowed users to restrict how their data appeared in others' match results or setting visibility defaults to "off" for new participants. Optional privacy controls reflected a philosophy that externalized risk and underestimated the familial implications of genetic visibility (Nissenbaum, 2004).

### **Lessons for Future Platforms**

The 23andMe breach offers four key governance lessons for data-centric platforms managing sensitive user information:

1. Breach minimization is not sufficient without systemic accountability. Merely blaming users for credential reuse obscures structural vulnerabilities in platform architecture, access controls, and monitoring capabilities (Florêncio & Herley, 2010; Holthouse, Owens & Bhunia, 2025).
2. Relational data requires relational governance. Platforms must recognize that users often expose others through their participation, and consent models must reflect that complexity (Narayan, Kohli & Martin, 2025).
3. Contractual privacy terms do not ensure ethical legitimacy. The ability to transfer sensitive data through bankruptcy or sale does not mean such transfers are aligned with user expectations or ethical best practices (Bradshaw, Millard & Walden, 2011).
4. Leadership silence undermines trust. In the wake of breaches, organizational leaders must visibly engage in the response – not only with shareholders, but with users, regulators, and the public (Rutherford, 2025).

## **6. DISCUSSION, LIMITATIONS, AND CONCLUDING REMARKS**

The 23andMe breach and its aftermath offer a cautionary tale for digital platforms that collect, process, and monetize genomic and relational data. Our analysis highlights how architectural decisions, especially those concerning visibility and consent, can create privacy risks that extend far beyond individual users to affect entire genetic networks (Phillips, 2016). While the credential-stuffing attack may appear to be a technical failure stemming from weak user passwords, deeper flaws in the DNA Relatives feature, a lack of strong access controls, and opaque consent practices significantly magnified the breach's impact. Appendix C summarizes the technical vulnerabilities and security failures that



contributed to the 2023 breach.

The subsequent bankruptcy and asset sale only intensified these concerns. The purchase of 23andMe's assets by its former CEO through the TTAM Research Institute has not resolved the structural governance failures that allowed the breach to escalate. The process exposed significant regulatory gaps – unlike the GDPR or California's CCPA, U.S. law provides limited protections for consumer genetic data during corporate restructuring. Unlike GDPR, U.S. law relies on contractual terms that often fail to reflect meaningful user control (Bradshaw, Millard & Walden, 2011; Gerke, Jacoby & Cohen, 2025). Consequently, millions of users, whose data was collected under earlier privacy policies, are now subject to a new data steward with unclear federal oversight (Kirk, 2025).

Several limitations in our analysis remain. First, users indirectly exposed through genetic linkages received little redress in legal settlements or regulatory action (Calo, 2011; Schwartz & Solove, 2011; Carballo, Schmall & Tumin, 2024). Second, although 23andMe announced new privacy commitments, it did not offer options to retroactively limit profile visibility or reclaim control over already shared data (23andMe, 2023). Third, while TTAM signaled a shift toward public interest governance, its operational model remains unclear, and no re-consent initiative has been launched (Herper, 2025).

This paper argues for systemic reforms in the governance of digital platforms that handle biometric and familial data. Without statutory safeguards, granular user controls, and enforceable transparency, relational architectures will continue to pose privacy risks (Boyd & Crawford, 2012; Narayan, Kohli & Martin, 2025). The 23andMe breach was more than a technical failure; it demonstrated how accumulated design debt, weak governance, and regulatory ambiguity converge to create privacy fragility – a systemic vulnerability where failures in one domain cascade through relational networks, exposing entire populations (Erich & Narayanan, 2014; Hernandez, 2025; Kotlan, Magoon & Yates, 2026).

## 7. REFERENCES

- 23andMe. (2024, Jan. 22). Notice of data breach. *Office of The Attorney General*, Sacramento, CA. <https://oag.ca.gov/system/files/CA%20AG%20-%20CA%20Notification%20Letters.pdf>
- 23andMe. (2023, Dec. 5). Addressing Data Security Concerns – Action Plan. *23andMe Blog*.
- 23andMe. (n.d.). Legal – Privacy Statement. *23andMe*, San Francisco, CA, USA. <https://www.23andme.com/legal/privacy/full-version/>
- Bampoulidis, A., & Lupu, M. (2019). *An abstract view on the de-anonymization process*. arXiv preprint arXiv:1902.09897.
- Barocas, S., & Nissenbaum, H. (2009). On Notice: The Trouble with Notice and Consent. In *Proceedings of the Engaging Data Forum*.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679.
- Bradshaw, S., Millard, C., & Walden, I. (2011). Contracts for clouds: Comparison and analysis of the terms and conditions of cloud computing services. *International Journal of Law and Information Technology*, 19(3), 187–223.
- Carballo, R., Schmall, E., & Tumin, R. (2024, Jan. 26). 23andMe Breach Targeted Jewish and Chinese Customers, Lawsuit Says. *New York Times*. <https://www.nytimes.com/2024/01/26/business/23andme-hack-data.html>
- Calo, R. (2011). The Boundaries of Privacy Harm. *Indiana Law Journal*, 86(3), 1131–1161.
- Erich, Y., & Narayanan, A. (2014). Routes for breaching and protecting genetic privacy. *Nature Reviews Genetics*, 15(6), 409–421.
- Florêncio, D., & Herley, C. (2010). Where do security policies come from? In *Proc. of Symposium on Usable Privacy and Security*.
- Gellman, R. (2025). Fair Information Practices: A Basic History (Version 2.32). *Center for Democracy & Technology*, Washington, DC. <http://dx.doi.org/10.2139/ssrn.5348107>
- Gerke, S., Jacoby, M. B., & Cohen, I. G. (2025). Bankruptcy, genetic information, and privacy – Selling personal information. *New England Journal of Medicine*, 392(10), 937–939.
- Greely, H. T. (2020). The future of DTC genomics and the law. *Journal of Law, Medicine & Ethics*, 48(1), 151–160.
- Greenleaf, G. (2012). Global data privacy laws: Forty years of evolution. *Journal of Law and Information Science*, 23(1), 1–112.

- Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying Personal Genomes by Surname Inference. *Science*, 339(6117), 321–324.
- Hernandez, J. (2025, Mar. 24). 23andMe is filing for bankruptcy. Here's what it means for your genetic data. *NPR*, Washington, DC.
- Herper, M. (2025, Jun. 13). Anne Wojcicki wins back 23andMe, this time as a nonprofit. *STAT*, Boston, MA. <https://www.statnews.com/2025/06/13/23andme-anne-wojcicki-wins-back-from-bankruptcy-will-become-nonprofit-ttam/>
- Holthouse, R., Owens, S., & Bhunia, S. (2025). *The 23andMe Data Breach: Analyzing Credential Stuffing Attacks, Security Vulnerabilities, and Mitigation Strategies*. arXiv preprint arXiv:2502.04303.
- Kawaguchi, K., & Lee, M. H. (2025, Jun. 27). DNA For Sale: The Human Rights Crisis in the 23andMe Bankruptcy. *Health and Human Rights*. Harvard University Press.
- Kirk, R. (2025, Jun. 10). 23andMe Customers Did Not Expect Their DNA Data Would Be Sold, Lawsuit Claims. *New York Times*. <https://www.nytimes.com/2025/06/10/business/23andme-data-lawsuit.html>
- Kotlan, A. M., Magoon, J. A., & Yates, D. J. (2026). Privacy Fragility in Direct-to-Consumer Genetic Testing: Lessons from the 23andMe Journey. *Hawaii International Conference on System Sciences (HICSS)*, Lahaina, Maui, HI.
- Lanzing, M. (2016). The Transparent Self: A Normative Investigation of Changing Selves and Relationships in the Age of the Quantified Self. *Philosophy & Technology*, 29(1), 33–48.
- Lee, J. B. (2025, Jul. 17). *23andMe Bankruptcy: The Privacy Ombudsman's Report*. Loeb & Loeb, New York, NY.
- Macdonald, A. S. (2024). Genetic testing and actuarial science. *Annals of Actuarial Science*, 18(1), 1–4.
- McGuire, A. L., Caulfield, T., & Cho, M. K. (2008). Research ethics and the challenge of whole-genome sequencing. *Nature Reviews Genetics*, 9, 152–156.
- Narayan, S. M., Kohli, N., & Martin, M. M. (2025). Addressing contemporary threats in anonymized healthcare data using privacy engineering. *NPJ Digital Medicine*, 8, 145.
- Nelson, A., Rekhi, S., Souppaya, M., & Scarfone, K. (2025). *Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile* (NIST SP 800-61 Rev. 3). NIST, Washington, DC.
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, 79(1), 119–158.
- Phillips, A. M. (2016). Only a click away – DTC genetics for ancestry, health, love...and more: A view of the business and regulatory landscape. *Applied & Translational Genomics*, 8, 16–22.
- Rothstein, M. A. (2008). Is GINA Worth the Wait? *Journal of Law, Medicine & Ethics*, 36(1), 174–178.
- Rutherford, A. (2025, Mar. 27). As a geneticist, I will not mourn 23andMe and its jumble of useless health information. *The Guardian*.
- Saey, T. H. (2025, Mar. 26). What 23andMe's bankruptcy means for your genetic data. *Science News*, Washington, DC. <https://www.sciencenews.org/article/23andme-bankruptcy-genetic-data-delete>
- Schwartz, P. M., & Solove, D. J. (2011). The PII problem: Privacy and a new concept of personally identifiable information. *NYU Law Review*, 86, 1814–1894.
- Shabani, M., & Borry, P. (2015). Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation. *European Journal of Human Genetics*, 26(2), 149–156.
- Terry, N. P. (2012). Protecting patient privacy in the age of big data. *UMKC Law Review*, 81(2), 385–415.
- Wikipedia. (n.d.). *23andMe data leak*. [https://en.wikipedia.org/wiki/23andMe\\_data\\_leak](https://en.wikipedia.org/wiki/23andMe_data_leak)
- Zettler, P. J., Guerrini, C. J., & Sherkow, J. S. (2019). Regulating genetic biohacking. *Science*, 365(6448), 34–36.

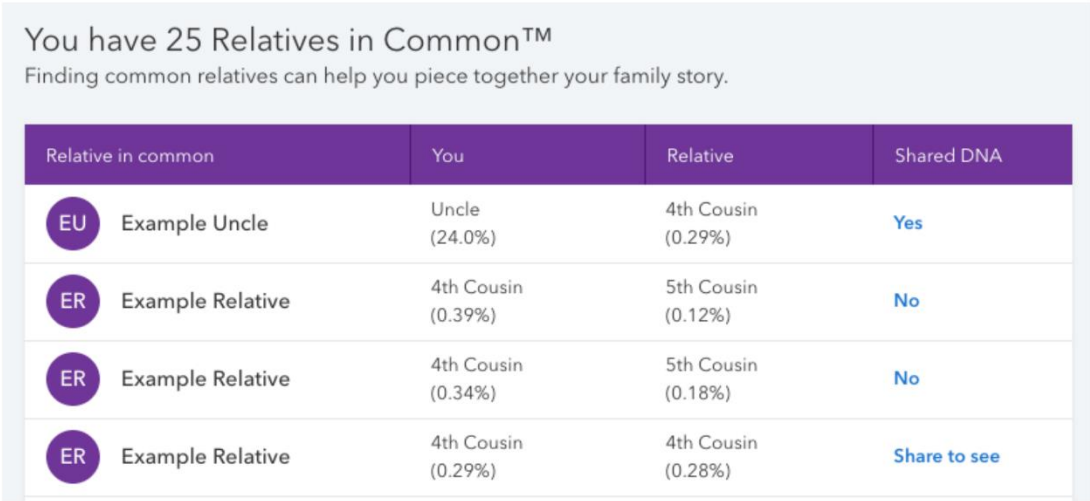
## APPENDIX A





### Sample Output of DNA Relatives Tool

Below are two important paragraphs about the DNA Relatives tool. These were taken directly (verbatim) from the 23andMe web site in June 2025:

"The *DNA Relatives* feature is an interactive 23andMe feature, allowing you to find and connect with your genetic relatives and learn more about your family story. Genetic relatives (also known as DNA Relatives matches) are identified by comparing your DNA with the DNA of other 23andMe customers who are participating in the DNA Relatives feature. When two people are found to have an identical DNA segment, they very likely share a recent common ancestor. The DNA Relatives feature uses the length and number of these identical segments to predict the relationship between genetic relatives.

...  
To see your shared relatives, click on a match in your DNA Relatives list and scroll down to the *Relatives in Common* section. In this section, you can see the list of relatives that you have in common, the predicted relationship between each pair, and in some cases, if you share DNA in the same region of your genome. Keep in mind that only matches with whom you have a sharing connection or those showing ancestry results will display whether or not you share DNA in the same region of your genome."



| Relative in common   | You                   | Relative              | Shared DNA   |
|--|-----------------------|-----------------------|--------------|
|  Example Uncle    | Uncle<br>(24.0%)      | 4th Cousin<br>(0.29%) | Yes          |
|  Example Relative | 4th Cousin<br>(0.39%) | 5th Cousin<br>(0.12%) | No           |
|  Example Relative | 4th Cousin<br>(0.34%) | 5th Cousin<br>(0.18%) | No           |
|  Example Relative | 4th Cousin<br>(0.29%) | 4th Cousin<br>(0.28%) | Share to see |

**Figure 1: Sample Output from DNA Relatives Tool**  
(Source: <https://customercare.23andme.com/hc/en-us/articles/221689668-DNA-Relatives-In-Common-Report-Feature>)

As of September 2025, this tool includes the following disclosure:

"We have temporarily disabled some features within the DNA Relatives tool as an additional precaution to protect your privacy. Read more [here](#)."

## APPENDIX B

### Timeline of 23andMe Milestones

#### 23andMe Share Price Versus Time

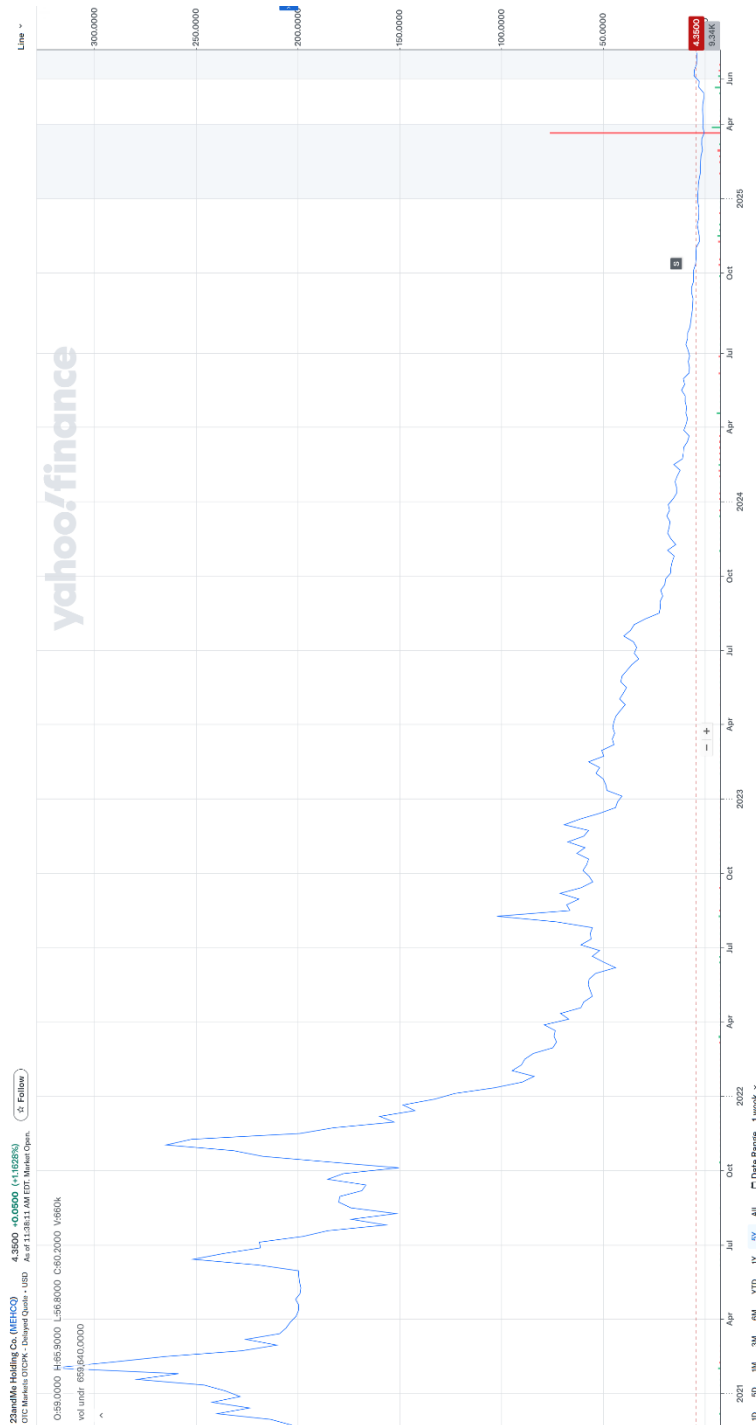


Figure 2: 23andMe Share Price from IPO in 2021 to June 2025

## 23andMe Timeline (2006 – July 2025)

- **2006** – 23andMe is founded by Anne Wojcicki, Linda Avey, and Paul Cusenza with the aim of democratizing access to genetic testing for health and ancestry.
- **2008** – The company gains major public attention; its testing kit is named Time magazine's *Invention of the Year*.
- **2013** – The U.S. FDA orders 23andMe to halt health-related genetic reporting, citing concerns over unvalidated medical risk interpretations. The company temporarily suspends parts of its product.
- **2015–2018** – 23andMe relaunches health reports and enters into a \$300 million partnership with GlaxoSmithKline (GSK) to leverage aggregated genetic data for drug discovery. By Feb 2018, ~3 million customers.
- **2018** – 23andMe solidifies its position as a leader in the direct-to-consumer genetics market.
- **June 2021** – The company goes public via SPAC at ~\$3.5B valuation; ~12M customers at close.
- **2022** – 23andMe faces increased scrutiny over privacy practices and monetization strategies. Growth slows even though company surpasses >12 million genotyped customers.
- **May 2023** – The company reported >14 million customers; lawmakers later referred to more than 15 million during 2025 oversight and bankruptcy proceedings.
- **October 2023** – 23andMe suffers a credential-stuffing attack affecting ~14,000 accounts and indirectly exposing profile data from 6.9 million users through the DNA Relatives feature. The company does not issue full disclosures until months later.
- **Late 2024** – As the stock falls below \$1, the company completes a reverse stock split to avoid NASDAQ delisting. Trust in the platform declines, and user engagement drops. Still, estimates suggest ~15 million cumulative users remain in the database.
- **March 2025** – 23andMe files for Chapter 11 bankruptcy, citing falling revenue, reputational damage, and unresolved legal claims. CEO Anne Wojcicki steps down from leadership.
- **June 2025** – 23andMe had accumulated class-action lawsuits and faced regulatory action in the U.S., U.K., and Canada. The UK ICO fined the company £2.31 million following a joint investigation with Canada's OPC.
- **July 2025** – Wojcicki's new venture, TTAM Research Institute, successfully purchases 23andMe's assets – including its genetic data – for \$305 million in a bankruptcy auction, outbidding firms like Regeneron. This sale raises unprecedented concerns over data transfer ethics, user consent, and platform accountability in the DTC-GT industry.
- **September 2025** – 23andMe seeks approval for a \$50 million class-action settlement; this reflects an increase from the \$30 million preliminarily approved in December 2024.

## APPENDIX C

### Technical Vulnerabilities and Security Failures that Contributed to the 23andMe Breach

| Category                         | Failure or Weakness  | Implication  |
|----------------------------------|--|--|
| <b>Authentication</b>            | No mandatory multi-factor authentication (MFA)                           | Allowed attackers to access accounts using stolen passwords alone                                  |
| <b>Access Control</b>            | Inadequate rate limiting and anomaly detection on DNA Relatives queries  | Enabled lateral exposure of millions of profiles from a small number of compromised accounts       |
| <b>Credential Management</b>     | Susceptible to credential stuffing due to weak password reuse protection | Exploited passwords reused across platforms; lacked protections against bulk login attempts        |
| <b>Logging and Monitoring</b>    | Insufficient real-time monitoring of unusual query behavior              | Delayed detection and containment of attacker activity   |
| <b>Data Minimization</b>         | Broad data exposure via the DNA Relatives feature                        | Enabled visibility of names, ancestry, and relationships beyond the originally compromised account |
| <b>User Consent Architecture</b> | No granular or retroactive consent options for shared data               | Users had no ability to limit relational data exposure post-breach                                 |