Applied Human Factors and Ergonomics (AHFE2025), Vol. 199, 2025, 105–115
https://doi.org/10.54941/ahfe1006823

Bridging the Privacy Gap: Stakeholder Solutions to Support Transparent Data Management Practices in Digital Health Research

Ramona Pindus[1], Daniela G. Vital[1], Brittany York[1], Woocheol Kim[1], Karandeep Singh[2,3], and Camille Nebeker[1,4,5]

[1]Herbert Wertheim School of Public Health and Human Longevity Science, University of California San Diego (UC San Diego), La Jolla, CA, 92093, United States
[2]Division of Biomedical Informatics, Department of Medicine, University of California San Diego, La Jolla, CA, 92093, United States
[3]Joan and Irwin Jacobs Center for Health Innovation, UC San Diego Health, La Jolla, CA, 92093, United States
[4]The Design Lab, University of California San Diego (UC San Diego), La Jolla, CA, 92093, United States
[5]Qualcomm Institute, University of California San Diego (UC San Diego), La Jolla, CA, 92093, United States

ABSTRACT

Digital health research increasingly incorporates commercial technologies such as wearables, apps, and social media, requiring acceptance of third-party privacy policies that govern participant data. These policies are lengthy, complex, and frequently updated, creating challenges for researchers and institutional review boards (IRBs) in assessing risks. Yet, privacy policy review is not consistently integrated into research oversight, leaving a critical gap. This study identified stakeholder priorities for improving privacy policy communication in digital health research and co-designed solutions using the Double Diamond framework. A four-hour workshop held in March 2025 at UC San Diego engaged 25 participants, all with prior digital health research experience, who represented the perspectives of IRB members, researchers, and research participants. Using the privacy policy of a popular wrist-worn activity monitor as a use case, participants progressed through discovery, definition, development, and delivery activities, including policy analysis, issue prioritization, and prototype creation. Data sources (workbooks, discussion notes, votes, and presentations) were transcribed, labeled by stakeholder group, analyzed with AI-assisted thematic and sentiment analysis and verified by researchers. Participants co-created six prototype solutions: (1) a policy scoring app, (2) a personalized data risk profile app, (3) a gamified learning platform, (4) an interactive consent tool, (5) a multi-format risk/benefit dashboard, and (6) an IRB communication support tool. Five focused on participant communication, while one targeted IRB workflows. Stakeholders prioritized simplification tools, interactive consent interfaces, granular user control, and third-party transparency. Findings highlight how co-design can generate practical, evidence-based strategies to make privacy policies more accessible, support informed consent, and strengthen transparency in digital health research.

Keywords: Privacy policy, Human-centered design, Data management practices, Digital health, Research ethics

INTRODUCTION

Digital health technologies, including wearable devices, mobile health apps and internet-connected medical sensors, are increasingly used in health research. These tools generate rich streams of health data, which may be controlled by third-party privacy policies issued by commercial companies. Although these privacy policies ostensibly inform users about how their data will be collected, used, stored, and shared, they are frequently characterized by legal language, excessive length, and vague descriptions of data practices (McDonald & Cranor, 2009; Nissenbaum, 2004). For prospective research participants, researchers, and Institutional Review Boards (IRBs), such complexity obscures critical information needed to evaluate risks and make informed decisions about data use.

This lack of transparency presents a pressing challenge for digital health research. Researchers are ethically and procedurally obligated to assess potential risks, particularly data risks when partnering with or relying on commercial platforms (McInnis et al., 2024). Ethics and regulatory review boards (i.e., IRBs), acting as gatekeepers of participant protections in health research, must determine whether a proposed study meets ethical standards and federal regulations for data security, privacy, and informed consent. However, when privacy policies are unclear, incomplete, or difficult to interpret, both researchers and IRBs may struggle to identify how third-party companies manage data, particularly with respect to sensitive information, algorithmic profiling, or secondary data use. This raises the potential for underestimating risks, failing to adequately inform participants, and inadvertently compromising participant trust or safety (Vayena & Blasimme, 2017).

To address these challenges, interdisciplinary researchers have begun applying *design thinking*, *human-centered design* (HCD), and *value-centered design* (VCD) frameworks to make data management practices more accessible, meaningful, and actionable. Design thinking promotes iterative, problem-solving approaches that involve end users in co-creating solutions (Brown, 2009). HCD emphasizes empathy and usability, ensuring that communication tools, such as consent forms or privacy summaries, reflect the needs and capacities of participants (Norman, 2013). VCD extends these efforts by foregrounding ethical principles, including for example, autonomy, justice, and beneficence, throughout the design process (Friedman et al., 2008). These approaches can support the development of communication strategies about data management practices that are transparent, participant-centered, and aligned with research ethics principles.

This paper describes how design approaches were leveraged to address barriers to the accessibility of privacy policy communication of data management practices in digital health research. In doing so, we explore the responsibilities of researchers and IRBs in evaluating data management practices and related risks. We then present prototypes created during the co-design process intended to increase access to, and understanding of, data management practices of digital health products.

METHODS

This study was guided by the Double Diamond design framework (Ball, 2019) and involved an in-person co-design workshop. The goal was to explore challenges with communicating privacy policy information associated with products used in digital health research and generate potential solutions, informed by key stakeholders. The study involved 25 participants attending a four-hour, in-person, co-design workshop at UC San Diego in March 2025. Prospective participants were recruited via a convenience sampling approach that targeted relevant listservs, newsletters, and direct contact with individuals known for their involvement in digital health research at UC San Diego and nearby areas. The recruitment message included a description of the study objectives, workshop details, participant incentives, and a link to the study screener. Additional recruitment efforts included printed flyers posted at six key locations across the UC San Diego campus including research centers, department bulletin boards, and common areas frequented by faculty, students, and researchers in relevant fields. Those interested in participating in the in-person workshop were asked to complete an online screener to assess eligibility. Eligible participants were 18 years or older with experience in the design, conduct, or review of digital health research. Exclusion criteria included being unable to attend the in-person co-design workshop in its entirety.

This study was reviewed and approved by the UC San Diego IRB [#811681]. Informed consent was obtained electronically from all participants included in the study. Participants agreed to audio and video recordings to assist with data collection and analysis.

Design Thinking Process

The workshop used a privacy policy associated with a popular wrist-worn activity monitor, which was contextualized in a fictional digital health study. Using this policy as a use case, participants were guided through the four phases of the Double Diamond framework: *Discover, Define, Develop, and Deliver* (Ball, 2019). This framework supports an iterative design thinking process to create functional and practical solutions to a problem by following a human-centered approach. In this study, the Double Diamond framework was adapted and implemented to guide participants into exploring the privacy policy communication challenges and co-designing prototypes that increase access to, and understanding of, data management practices of digital health products (Figure 1). The workshop protocol included a structured design thinking process involving four activities, which encouraged engagement and collaboration. Participants at each table were asked to review a specific section of the privacy policy and highlight elements they would label as important and/or confusing (Privacy Deconstruction). Next, they presented what they identified as core privacy issues and communication challenges with other group members with the goal of drafting a problem statement (Problem Identification). Each group member generated and refined creative solutions through an iterative brainstorming process that yielded a list of up to ten potential solutions (Solution

Development). Individuals discussed their lists and narrowed down insights into one solution per group. In the last phase, participants collaboratively designed a low-fidelity prototype using sketching, storyboarding, and other elements of visualized communication strategies (Prototyping).

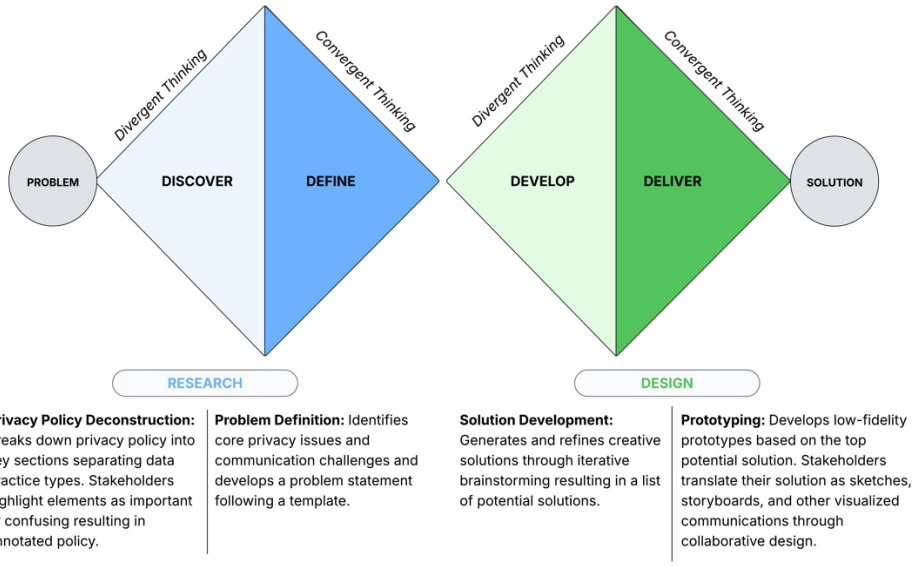

Figure 1: Double diamond framework adaptation to co-designing privacy policy communication solutions.

Data Collection and Analysis

Study data were collected and managed using REDCap electronic data capture tools hosted at UC San Diego (Harris et al., 2009; 2019). Prospective participants received an online eligibility screener survey, which solicited information about their academic and professional affiliations, including their career stage, years of experience in digital health, and a description of their role in digital health research. Based on these criteria, eligible participants were placed in one of three stakeholder groups; *IRB members, Researchers*, and *Research Participants*. Each category was divided into two groups resulting in a total of six distinct sub-groups of 4–5 participants. Participants were prompted to assume their role and views during the workshop activities. Each participant was given a workshop activity guidebook that included a copy of the privacy policy and instructions for completing four activities. This document was collected at the end of the workshop and used as a data source by the research team. Additionally, each group received two poster boards to document their group activities and to draw a prototype of their chosen solution. During the co-design workshop, three research team members and a lead facilitator assisted participants working through the activities. Video recordings of prototype presentations were collected capturing a detailed walkthrough of the functionality and implementation of each group's proposed solution.

Descriptive and frequency statistics were performed to describe participants attending the workshop. Data from individuals and groups were cleaned and organized by group. Recordings and prototype images, including handwritten sticky notes were transcribed using Otter.ai (*Otter Meeting Agent - AI Notetaker, Transcription, Insights*, 2025) and ChatGPT 4o respectively. Initial data analysis was conducted using Anthropic's Claude Sonnet 3.7 (*Meet Claude\Anthropic*, 2025) for artificial intelligence (AI)-assisted thematic and sentiment analyses. AI-assisted analyses and transcription were reviewed and validated systematically by the research team.

AI-generative pre-trained transformers were used throughout various steps during this study. ChatGPT 4o was used to generate an initial workshop plan following the Double Diamond design framework. The plan was refined through a series of team discussions and beta testing of workshop activities. During the systematic validation of the AI-assisted analyses, the research team observed emotion-laden language. Claude Sonnet 4 was prompted to expand the thematic analysis by completing a sentiment analysis of the data. Quality assurance checks of the AI-assisted qualitative data analysis included verification that results accurately reflected the actual data, and that sentiment analysis results were supported by the language used. Examples of steps taken during this quality assurance process include verification of issues prioritized by each stakeholder group, and of differences in sentiment-rich language use between stakeholder groups. One researcher conducted a thorough quality check, followed by a secondary spot-check of 10% of the data by two researchers. Notes documenting workshop activities were used by the researchers during the quality assurance phase as an additional check of the themes identified in the analysis.

RESULTS

Those attending the co-design workshop (N = 25), represented diverse academic roles and varying levels of experience in digital health research. Attendees reported involvement in the design (52%), conduct (72%), and review (68%) of digital health research. Research foci spanned mental health, behavioral interventions, ethics, literacy, and health tracking. Nearly half (48%) reported 1–3 years of research experience in the field suggesting junior-level experience. Current roles show frequent use of digital tools, such as Electronic Health Records (EHR), wearable sensors, eye tracking, ecological momentary assessment, Python, and Zoom. Populations targeted in the digital health research included older adults, individuals with depression, chronic pain, and both underserved and underrepresented communities. Attendees self-identified as female (72%), non-Hispanic (84%), and Asian (40%) or White (28%).

Each of the three stakeholder groups developed two prototypes, for a total of six designs. Figure 2 shows examples of the co-design process and resulting prototype from one group.

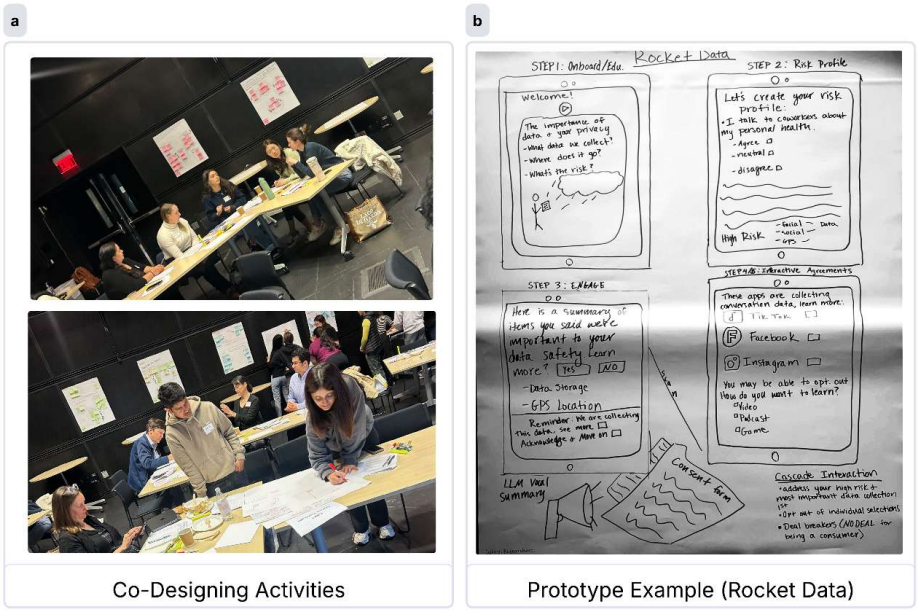

Figure 2: Participant co-design activities and example of one of the low-fidelity prototypes.

Researchers

Prototype 1–A Policy Scoring App: This is a tool to assess and score the privacy policies and data practices of companies and their products. Users can upload privacy policy documents or look up a company and the tool will generate a risk score (e.g., 60/100) based on factors like third-party access and sharing of data. In addition, the tool will provide the company's lawsuit history, data leakage reports, and metrics for the technology's use in research, amounting to a comprehensive reputation report. Users engaging with this tool over time can develop competency and graduate from "Policy Learners" to "Policy Makers."

Prototype 2–The Rocket Data App: Rocket Data is an app that allows users to create their personal risk profile, like Myers Briggs, but for a data risk personality profile. The app takes the privacy policy information from a company and provides options to learn about what they contain in different formats (summary text, video, podcast, games). Once a profile is set up, the app will notify the user if a company has data practices they flagged as sensitive or important to them and prompt the user to choose if they want to take actions to restrict these practices or even stop using the company's product.

Participants

Prototype 3–A Gamified Experience Inspired by Platform Games: This solution is focused on making the experience educational, flexible and accessible. Users can switch at any point between reading the privacy policy

document and playing the game, and also between any languages they want to use. In game, the users start at "Mount Info Collection, moving on to the Risk Railway, diving into the Data Mines, going to Data Mission Control, [...] and then the Local Processes Pathway" where users can learn about differences based on California Consumer Privacy Act (CCPA), General Data Protection Regulation (GDPR), and any other locally relevant variations found based on the location on a map. Finally, users can obtain a summary and explore frequently and not-so-frequently asked questions.

Prototype 4–An Interactive, Closed-loop Consent Tool: This tool can be used in app or website form and is focused on providing flexible and entertaining ways for users to learn about a company's data practices. Users have the option of selecting language and location. The tool provides a summary including data sharing practices, FAQs, and hyperlinks to the source documents. Users can choose between modalities of reviewing the information (e.g., audio, video, chatbot), they can make choices about current and future use of data collected, and they can confirm their understanding of the material by answering short tests.

IRB Members

Prototype 5–An IRB Workflow Tool: This tool is designed to facilitate communication between researchers and IRBs. It walks researchers through steps: first, detail the type and level of information needed (e.g., location data–state, city, acre, room, step); then, asks about the type of device used (e.g., researcher-built vs. Commercial); finally, details requested by the IRB will vary depending on device used (e.g., if the device collects feminine cycle data, the researcher will be asked to justify why this data is needed for the project, or why this device was selected instead of a device that does not collect these data).

Prototype 6–A Location-Based Risk Evaluation Tool: This tool is focused on evaluating risks and benefits for users of a technology based on their local context. The user initially selects their location and language. Based on this information, the tool provides a version of the privacy policy information curated to highlight risks and benefits relevant to the user's location and cultural context. Users have the option of selecting from multiple options of how information is presented: written, visual, audio, video, game, or AI.

Thematic and Sentiment Analysis Results

Thematic analysis revealed six major themes across all three stakeholders' groups (Table 1). Key insights revealed that stakeholders perceived a lack of clear communication about data practices masked by legal and technical jargon, the use of commercial third-party technologies in research introduces layers of complexity to research participation which stakeholders found difficult to navigate, and which contributes to difficulties when trying to assess risks related to research participation. Moreover, current IRB processes and regulatory frameworks are not designed to handle integration of commercial digital devices in research. Therefore, stakeholders wanted the option to make easy, quick, and granular choices about how their

data is collected, used, and shared; and wanted multiple ways of reviewing information, including interactive methods to learn about data practices.

Table 1: Six themes emerged across three groups.

Theme 1: Disconnect between information needed and information provided *"Seemingly contradictory statements. Too long. No executive summary. Too complex & confusing."* *"Who has access to my data, and why [do] advertisement companies have access to our personal information and activity. - Uses heavy technical words to [...] let us know how data is encrypted, but no explanation."*
Theme 2: Lack of meaningful choice *"Stakeholders needs to know what "certain" info they can request to be deleted and what cannot be because they should be informed and have the autonomy over it. - Participants need to know all the details in the "other information." "All or nothing...Pretends to be consensual Privacy statement is both (misleading) and ultimately leads to (no privacy) [...] Data heist"*
Theme 3: Third-Party data ecosystem complexity *"Researchers need to understand what third-party companies are utilizing participant information"* *"Privacy policy is vague and redirects to other 3rd party privacy policies, making it confusing to the participant. - Stealing data w/consent "*
Theme 4: Risk assessment and harm prevention *"Unclear + vague language about how data is stored. Researchers need clear information on how/what information is kept on [the device] to "prevent harm" and for how long because we need to feel safe in providing participants with this tool." "Researchers need specific information on what identity-linked data is shared with 3rd parties because that could be used to profile them."*
Theme 5: Presenting information in multiple formats *"It has a gamified kind of interface, so you can choose how you want to engage with that. So, it's either a video it creates or a podcast, or, you know, interactive images, or whatever it may be. It has some flexibility there so that you can consume it in the most digestible way for you as a consumer."* *"If you were to click on one of those risks, it'll show why [the device] has a high risk for that item. So, they'll say, like third party access and sharing. And then it will pull up what part of the policy looks at that component and why it's high risk for us as researchers to use that technology in our studies."*
Theme 6: Institutional and regulatory gaps *"The IRB would need more specific information/wording regarding how personal information is being stored, deleted (how & what), and which data is being shared to which 3rd parties. This is because the IRB would find it extremely important to understand the handling of personal data so they can understand and better anticipate any lawsuits or issues that may rise due to the vague wording in the privacy policy." "IRB needs specific information regarding the timing and nature of deletion of data. IRB needs specific information regarding types of data protected for children and why this is not claimed to be the standard of protection for all. IRB needs specific information regarding inclusion of services and corporate affiliates."*

Sentiment analysis identified patterns cutting across all stakeholder groups. Powerlessness was reflected differently but consistently in the language

used by all stakeholder groups. Participants felt the lack of power in making meaningful choices or even withholding consent ("stealing data with consent"), researchers felt a lack of power in taking measures to protect participants due to being "overpowered by commercial defaults," while IRB members felt powerless to evaluate risks due to "vague wording." Language also indicated tension and frustration related to critical information being spread across multiple sources and policy documents, and the need for in-depth technical knowledge to understand the information contained in these policy documents. In addition, stakeholders expressed concern at current consent models failing to respond to the need for dynamic, ongoing, and personalized consent, as these documents are routinely updated.

There was a noticeable shift in stakeholders' sentiment during the workshop. In the problem discovery phase, the dominant sentiments expressed were of betrayal, powerlessness, and frustration. By the end of the workshop, when solutions were shared and discussed, there was a shift towards collaborative energy and enthusiasm for the innovative solutions proposed. These sentiments and the themes identified are reflected in the solutions prioritized by the workshop participants who focused on interactive learning and consent options, simplification and personalization of language and processes, enabling granular control over data management decisions, and developing transparent ways of understanding, selecting, and using third-party technologies.

DISCUSSION

Our study found that people want choices that go beyond clear, accessible, and customizable ways of learning what a privacy policy contains. Stakeholders' understanding of these privacy policies is fragmented, but powerlessness is a common theme. Participants expressed a desire for options that allow them to understand the implications of these policies, how data practices affect them in their private and professional capacities, and they want granular and personalized control options in response to data practices described by the privacy policy. If implemented, the ideas surfaced in our prototyping could address myriad issues experienced by different stakeholders. Researchers and IRBs need to achieve a mutual understanding of the data types needed for the project to be successful and of the data-associated risks introduced by using third-party technologies. This will contribute to a better understanding of how third-party technologies are used in research. We recommend that policies and regulations prioritize enabling technology-users' understanding, choice, and continuous control over their data. We also recommend bringing clarity to processes and responsibilities involving the interplay between technology companies, researchers, research institutions, and IRBs. To demonstrate commitment to ethical technology development and data use, commercial entities should build transparency enabling, dynamic data management, and culturally responsive features as core elements of their designs. Future research should explore to what extent the views and sentiments expressed by the participants in this study are representative of the wider community of digital health technology users.

Perhaps the most important takeaway is that new solutions must address the critical erosion of trust technology users seem to experience.

CONCLUSION

Design thinking is a strategy applied to co-create alternatives to the current privacy policy communications found in products used in digital health research. Data management transcends all data touchpoints including the collection, storage, sharing and use. Our study found that for consent to be informed, prospective participants, researchers, and IRBs, need to know what a 3rd party vendor is doing with respect to data management. The six prototypes developed during the co-design process offer creative solutions for making data management practices accessible to those using the product, and for technology companies to demonstrate the ethical principle of respect when considering user experience.

ACKNOWLEDGMENT

The research reported in this publication was funded through a Patient-Centered Outcomes Research Institute® (PCORI®) Award (ME-2020C3-21310) Supplemental Funding. The statements presented in this work are solely the responsibility of the author(s) and do not necessarily represent the views of the PCORI®.

REFERENCES

Ball, J. (2019, October 1). *The Double Diamond: A universally accepted depiction of the design process—Design Council*. Design Council. https://www.designcouncil.org.uk/our-resources/archive/articles/double-diamond-universally-accepted-depiction-design-process/.

Brown, T. (2009). *Change by Design: How Design Thinking Creates New Alternatives for Business and Society*. Harper Collins.

Meet Claude\Anthropic. (2025, July). https://www.anthropic.com/claude.

Otter Meeting Agent—AI Notetaker, Transcription, Insights. (2025, March). https://otter.ai/.

Friedman, B., Kahn Jr., P. H., & Borning, A. (2008). Value Sensitive Design and Information Systems. In *The Handbook of Information and Computer Ethics* (pp. 69–101). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470281819.ch4.

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208. https://doi.org/10.1016/j.jbi.2019.103208.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010.

McDonald, A. M., & Cranor, L. (2009). The Cost of Reading Privacy Policies. *2008 Privacy Year in Review Issue*, *4*, 543–568. https://www.semanticscholar.org/paper/The-Cost-of-Reading-Privacy-Policies-McDonald-Cranor/4b512e2f5ff42ef00ccceca200d888676e6c506f.

McInnis, B. J., Pindus, R., Kareem, D., & Nebeker, C. (2024). Considerations for the Design of Informed Consent in Digital Health Research: Participant Perspectives. *Journal of Empirical Research on Human Research Ethics*, *19*(4–5), 175–185. https://doi.org/10.1177/15562646241290078.

Nissenbaum, H. (2004). Privacy as Contextual Integrity. *Washington Law Review*, *79*(1), 119. https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10.

Norman, D. A. (2013). *The Design of Everyday Things*. MIT Press.

Vayena, E., & Blasimme, A. (2017). Biomedical Big Data: New Models of Control Over Access, Use and Governance. *Journal of Bioethical Inquiry*, *14*(4), 501–513. https://doi.org/10.1007/s11673–017-9809–6.