# Towards Privacy-Aware Research and Development in Wearable Health

Michelle De Mooy
Center for Democracy & Technology
mdemooy@cdt.org

Shelten Yuen
Fitbit, Inc.
syuen@fitbit.com

## Abstract

*Wearable sensor technology has the potential to transform healthcare. The investigation and testing of sensors in the commercial sector offer insight into ways to leverage biometric data, to improve individual health through the better products and to advance the public good through research.*

*However, research with wearable sensor data must be done in a manner that is respectful of ethical considerations and privacy. Not only will the processes that govern this research define the potential public good derived from wearables, they will encourage user trust in wearables and promote participation. The research and development (R&D) teams at companies are not just engines of innovation but also have the potential to be an important part of our social infrastructure. The Center for Democracy & Technology (CDT) embarked on a yearlong partnership with Fitbit. CDT gained rare access to the company's data policies and practices to build recommendations on privacy and ethics.*

## 1. Introduction

The ability to quickly and easily collect detailed biometric information about ourselves—such as how many steps we take each day, how many calories we burn, or how well we sleep—is the result of modern technology, but the desire to quantify is ancient. Great thinkers tracked their behavior and lives throughout the ages, from the Roman philosopher Seneca to Ben Franklin.[1] But big data, mobile computing, the internet of things, the movement to patient-centered care, electronic health records and telemedicine, and augmented reality all provide new ways for us to examine ourselves and scrutinize our behavior for insights. Writer Yang Yesheen calls data "the idiom of the biotechnological age and, increasingly, now the language of the self."[2] Wearable technology, or devices that are placed in clothing or worn on the body in order to record data about the wearer, have been extraordinarily successful in the consumer retail market.[3] Simple wearable technology such as calculator wristwatches, pedometers, and hands-free devices, have been available to consumers for decades, but none of these products have been adopted with the speed and ubiquity that the wearable devices in health, wellness, and fitness have. Approximately one in ten Americans owns a fitness tracker (http://endeavourpartners.net/assets/Endeavour-Partners-Wearables-White-Paper-20141.pdf) and sales of wearables devoted to health wellness and fitness are expected to grow from 29 million units in 2014 to 172 million units in 2018, with a spike in sales in 2015.

Wearables create digital records that track and quantify the physical minutiae of everyday life, including an individual's activity, biometric traits and responses, as well as behavior and habits. Devices that track personal health data (PHD) and wellness metrics are especially popular for people interested in increasing or optimizing their physical activity, improving their diet, identifying sleep patterns, and gaining insight into their overall health. Wearables involved in health and wellness often collect and use sensitive personal health information, but because the data generated by them is created at the direction of the user it is mostly outside of the disclosure restrictions and requirements found in the Health

[1] "Probably since the dawn of humanity, people have been fascinated by even the most minute details of their lives, and kept track of what was going on in their bodies and minds. The Roman philosopher Seneca tracked the food he ate and what he dreamt at night. Benjamin Franklin consistently recorded his performance on 13 measures, such as cleanliness, frugality, and overindulgence, believing it would keep him virtuous. Engineer and architect Buckminster Fuller nicknamed himself "guinea pig b" and kept a diary on his daily life and ideas." http://www.bbc.com/future/story/20130102-self-track-route-to-a-better-life

[2] Yang, Yesheen. Saving the Quantified Self: How We Come To Know Ourselves Now, Winter 2014 issue of Boom: A Journal of California, Available here: http://www.jstor.org/stable/10.1525/boom.2014.4.4.80

[3] Gartner, Inc. forecasts that 4.9 billion connected things will be in use in 2015, reaching 25 billion by 2020.

URI: http://hdl.handle.net/10125/41600
ISBN: 978-0-9981331-0-2

Insurance Portability and Accountability Act (HIPAA). Some wearable users that have expressed uncertainties about how companies will use and share their data, citing the potential for analytics and inferences that negatively affect health benefits or jobs. In response, companies such as Fitbit are increasingly providing clear and comprehensive privacy policies that explain to users the data collected and the limited circumstances under which it may be shared. However, there is a dearth of guidance for companies in this space on appropriate and effective ways to provide privacy and ethical protections for consumers' health data.

## 2. How Does the Technology Work?

Sensing is the core function of most wearable devices, but they are also designed to record and analyze data about the person wearing the device to provide personalized motivation and insight. A distinct feature of wearables is their ability to instigate a real-time effect in users by providing information at the exact point of decision-making, such as prompting a person to walk around if he has been sedentary for a long time. Activity trackers, such as those designed by Fitbit, track a range of metrics for the wearer around activity, exercise, sleep, and physiology. These include the number of footsteps taken, stairs climbed, amount of calories burned, the pace and distance of a run or bike ride, when and how much a person exercises, the duration of sleep, and heart rate throughout the day. Underlying most of these tracking abilities are commoditized sensor components that have existed in mobile phones for years: accelerometers.[4]

The modern consumer-grade accelerometer is a micro-electromechanical system (MEMS) packaged into an electronic part that is roughly a couple of millimeters square in size. It is commonly referred to as a motion sensor, although it measures both the static and dynamic accelerations imparted on the sensor. The use of commodity sensors does not diminish the technical feat achieved by wearable devices. Wearables package sensors into form factors that can be worn continuously during exercise and sleep, and are powered by sophisticated algorithms that translate raw sensor data - such as acceleration - into data that people can interpret and use to achieve their goals, such as being more active.

Also, the design of a wearable device has unique technical requirements in that it can have the functions of a mobile phone -- a wireless radio, a bright graphical display, alarms, sensors, and sensor-based applications -- in a smaller form factor and with a commensurately smaller battery, but with battery life that in some cases can exceed the average smartphone by a factor of five or more. There is interest in more physiological measures as the wearables industry grows in adoption. A recent trend is continuous heart rate monitoring with a technique called photoplethysmography (PPG), where light is shone into the skin and the amount of light reflected back modulates with a person's pulse.

## 3. What is the State of the Art?

Next generation wearables are armed with more sensors and smarter algorithms than their predecessors, are pushing the boundaries on smaller fitness wristbands and larger smart watches, and tend to be more focused on biometric monitoring. Some have moved off of the wrist and onto other body parts as conduits for data collection, such as "hearables" (or small devices worn in the ear that stream real-time information about activity or pulse). Companies are working on offering more complex sensing features, such as using environmental context to capture surrounding data (such as smells[5]) or interpreting the emotional state of a user. Epidermal electronics expand the canvas of this technology even further.

## 4. Privacy Concerns

Wearable technologies necessarily collect large amounts of data in order to perform their function for their users. This raises privacy concerns due to the amount of information that can be collected and shared. Advocates and regulators are primarily concerned with questions related to access, sharing, and control of this information: who can access the physical device, how the device is connected to the internet, where personally identifiable information flows beyond the user and the company, and the

[4] Many wearable devices also include barometric pressure sensors, global positioning sensors (via Global Navigation Satellite Systems), gyroscopes, and magnetometers. The accelerometer is the most ubiquitous sensor in wearable devices.

[5] A competition staged at MIT last year brought forth an example of a wearable that uses environmental sensing capability, designed for use by astronauts. https://spaceappsseattle.hackpad.com/Wearable-Environment-Sensor-for-Astronauts-KrxZKiA2Ppy

protocols for companies collecting, using, and storing information on private servers. There is currently no comprehensive set of privacy and security regulations, guidance, standards, or best practices for wearable technology companies.

## 7. Creating the Future

Innovations that both sustain and spur the growth of the industry are developed primarily through internal R&D. The internal R&D teams at technology companies around the globe are the beating heart of future growth and innovation because they have enormous access to and facility with all varieties of data. Consumer-facing entities that collect health data about individuals must consider privacy and security in all aspects of developing and deploying their products. Although users of health and wellness devices purchase and expect insight based on the collection and analysis of their personal information, they also expect companies to protect this data. R&D teams balance innovation and data privacy on a daily basis as they consider what questions to pursue, how to design the technology, and how to test the results. While some companies have a strong data privacy policy and pledge to alleviate user concerns about internal uses, many companies in the wearable space are not as transparent on how this personal data is used outside of the consumer experience. CDT's partnership with Fitbit illuminates the important role that R&D teams play in embracing and embodying privacy principles. Responsible and ethical research using personal health data via wearable devices can produce interesting and valuable insights on wellness, however we believe that the potential for this data to impact people's health will not be realized absent consensus from industry, stakeholders, and the advocacy community on clear and actionable guidelines that protect user interests.

## 8. Methodology

CDT worked directly with Fitbit to observe Fitbit's researchers in action and understand how they answer the questions posed by their technology. In a series of questions answered by Fitbit's Vice President of Research, CDT was able to get a broad overview of the company's internal process for conducting research. CDT then conducted in-depth interviews with five key Fitbit researchers at their headquarters in San Francisco in April of 2015. These conversations were built around questions designed to elicit a more detailed understanding of individual roles on the research team (Appendix A), as well as internal protections for data and accountability measures. CDT's collection, categorization, and subsequent analysis of this information was guided by a research methodology called grounded theory.[6]

### 8.1. Grounded theory

Grounded theory is a qualitative research method[7] in which the researcher develops her hypothesis after examining the data (rather than the traditional approach of a researcher developing a hypothesis before collecting data). This theory allows the researcher to use both data collection and her own insight about the context of the research question to develop a theory. The project was deployed in five core phases in accordance with grounded theory methodology: (1) Assessment, (2) Mapping, (3) Investigation, (4) Analysis, and (5) Drafting.

### 8.2. Assessment

In the first phase, CDT sought to get an overall understanding of Fitbit's internal research process from the researchers themselves. CDT used self-reported data via emailed surveys, phone calls with company managers, and in-person interviews with five key R&D team members to make this assessment, asking Fitbit R&D staff a series of ten questions [Appendix 1]. Specifically, CDT attempted to learn how R&D projects are scoped and launched, how long projects last, the process for determining which projects to stop and which to pursue, when and how sensitive data is designated, and what privacy and ethical considerations are factored in at each stage. Grounded theory requires researchers to "follow the data" in this stage of a project, rather than place data into categories that confirm or refute a hypothesis.

### 8.3. Mapping

With the information gleaned from the researchers, CDT mapped Fitbit's internal research process, from creating new study ideas to launching new products and services, highlighting particular areas where data privacy and ethics might be implicated [Appendix 2].

---

[6] Glaser, B.G. & Strauss, A.L. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago: Aldine Pub. Co.

[7] This project uses a grounded theory framework developed by researchers at the University of Auckland and Victoria University. It recommends: 1) Focus on theory building as primary goal, rather than theory verification; 2) Use joint data collection and constant comparison (i.e. by adding/enhancing its properties); 3) Use theoretical sampling.

These areas included: when formal policies informed data practices; when management provides oversight on projects; the transition from an informal to a formal project; when a type of data is identified for use in projects; how sensitive data is identified and used; and the deployment of privacy protections.

### 8.4. Investigation

Using the findings gathered, CDT reviewed—or sliced—a selection of the existing information about Fitbit's R&D process in order to uncover commonalities and relationships. For example, CDT looked at the personal and academic backgrounds of Fitbit researchers and found that they shared many commonalities, such as a deep interest in health and wellness, graduate-level education, and an expertise in hardware, software, and sensor systems. The comparison of commonalities teased out compelling areas for further investigation, such as the role individual researchers might play in influencing privacy-protective and ethical data practices. Although experience complying with Institutional Review Boards (IRBs) is not a hiring requirement for a research position at Fitbit, our findings showed that many researchers followed rules on data ethics through IRBs in past projects. This could indicate a heightened awareness of those concerns in their work and inform their thinking about future data usage.

### 8.5. Analysis

At the core of grounded theory is the concept of "coding," or applying categories and/or themes to data based on frameworks and superstructures that relate to the content of the research. In the analysis phase, CDT used the following established information frameworks to "code" our data in order to gain insight into both the unwritten and formally established structures of Fitbit's internal research practices: (1) the Fair Information Practice Principles, (2) the Belmont Report, and (3) the Common Rule.[8]

The Fair Information Practice Principles (FIPPs) inform most modern privacy regimes and CDT believes they offer important guidance when applied to internal research at health wearable companies. The FIPPs were first proposed in 1973 in a report by the U.S. Secretary's Advisory Committee on Automated Personal Data Systems entitled Records, Computers, and the Rights of Citizens. Since then, the FIPPs have become the internationally recognized practices for handling the privacy of information about individuals. A company with practices that are informed by the FIPPs (1) gives individuals control, access, and accountability for their data, (2) is transparent about their data practices, (3) is clear about the provenance or integrity of the data, (4) collects and uses data only within the context that is consistent with the way in which the data was provided, (5) minimizes the amount of data collection and the length of time for which the data is retained, (6) ensures that data is collected for a specific purpose, and (7) secures the data through the use of encryption, de-identification, and other methods.

Ethical considerations must also be a part of any discussion about research on wearable user data. In the university context, research on human subjects has been regulated since the 1970s, with specific ethical guidelines spelled out in the Belmont Report and administered by Institutional Review Boards. The Belmont Report lists three overarching principles: (1) respect for persons, (2) beneficence, and (3) justice. Respect for persons means that people should be treated as individuals with the right and capability of making informed decisions. This principle thus requires researchers to be truthful, conduct no deception, and to give subjects the chance to consent and withdraw consent. Beneficence means that the research must not harm people and should work to ensure that the benefits of the study are maximized while the risks are minimized. Justice means that researchers must treat people fairly and not unduly influence the decisions of vulnerable individuals or communities to participate.

Federal agencies engaged in research that uses human subjects must comply with the Common Rule,[9] a policy that draws heavily on the findings in the Belmont Report. The Common Rule offers detailed guidance on what constitutes informed consent from research subjects, with special emphasis on protections for vulnerable populations such as women who are pregnant, prisoners, and children. The Common Rule also contains requirements for the creation and functionality of IRBs, which are the formally designated committees that approve and monitor research involving humans.

CDT also used a final framework that we called "Practical/Business Realities" for coding the data. Successful technology companies, in particular, must

[8] Federal Policy for the Protection of Human Subjects, The Common Rule. United States Department of Health and Human Services.

[9] The Common Rule does not apply to federal agencies that have not signed "Human Subjects Protection" agreement.

keep pace with constant demands for higher quality products and increased functionality. Innovating at this speed is no easy task—for example, Fitbit employs approximately 5% of its workforce toward exploratory R&D for new features and insight for their customers. There is tremendous pressure on wearable companies to create devices that offer sophisticated features, are easy to use, and comfortable (as well as fashionable) to wear.

CDT compared and contrasted these buckets of data to highlight areas for analysis and analyzed these areas using two frames: privacy and ethics. Through this lens, CDT made determinations that led to our recommendations around issues such as where privacy and ethics should be considered during the R&D process; what practices should be in place to honor user privacy; constraints that should be placed on the uses of certain types of data; and the real-world factors, such as quick launch times, that companies in this space might experience.

## 9. Findings

Fitbit's corporate mission is to facilitate the improvement of the health and wellness of its users. R&D contributes to this end by developing and building new features and services for users. R&D is not just to create revenue or test boundaries, but also to establish a company's reputation as both an innovator and trusted institution. R&D teams also face the added constraint of time and the need to innovate new products, while ensuring that the privacy of their users is respected. The primary focus of internal R&D is to push an innovative concept into a product within a timeframe of two-to-three years through the creation and testing of new hardware and software. To do this, R&D teams analyze user behavior to figure out how sensors might improve user health or help users meet health goals, though this is not the sole focus of their work. The team's emphasis is centered on achieving these goals by improving sensor functions and creating new technical features. Interviews with core members of the Fitbit R&D team gave CDT an overview of the team structures and different types of projects and studies they undertake. The interviews also provided insight into what motivates researchers, the ways in which they form research questions, and the privacy and ethical considerations that come into play in their work. There are two primary tracks for R&D investigations at Fitbit: hacks and projects. Some projects become larger in scope, requiring more formal research methods and additional data, and are then referred to as studies.

### 9.1. Hacks

"Hack" is the term the R&D team uses for informal and low commitment investigations, often driven from individual interest in a product feature or potential line of insight. Hacks allow researchers the flexibility and creativity to pursue their curiosity. All hacks are shared with the full R&D team every month. An example of a hack is when Fitbit examined the average heart rates of Super Bowl viewers in Seattle and Boston. The company looked at how many steps people lost during football games and discovered biometric patterns unique to certain cities. Researchers investigated these patterns through the measurement of anonymized user data during the Super Bowl. Fitbit researchers released their analysis the day after the game. This investigation required a single researcher to evaluate patterns in user data and correlate some of that data in no more than twelve hours. Because it did not use identifiable user data, this investigation was not reviewed through the same formal process as an R&D project. Short-term hacks have had long-term impact. After a similarly-fast paced hack which evolved into a full R&D project, Fitbit decided that their trackers count minutes as "active" if a user participates in an activity for ten or more contiguous minutes, a definition that echoes recommendations from the Centers for Disease Control.

### 9.2. Projects

"Projects" are more formalized investigations. After a project idea is conceived, it is approved through discussions with the head of R&D. If approved, the R&D team holds a kick-off meeting that defines the team, goals, and timeline of the project. Project updates occur every one to two months and are shared with the entire research team. While one or two researchers perform hacks at most, project work usually occurs in comprehensive teams that contain at least one member from each of the Fitbit core competencies [Appendix 3].

### 9.3. User studies

Once a project is given the green light by the head of R&D, it often involves studies with volunteer users that require the collection and use of data produced by a person. Data used in internal R&D projects is placed into one of three categories of studies, which are characterized by where the data is from, either from Fitbit employees or Fitbit users, and by how

much of it will be used in the project. The study types are: (1) pilot studies, (2) internal Fitbit employee studies, and (3) Fitbit user studies. Fitbit's privacy controls increase from pilot studies to internal Fitbit employee studies to Fitbit user studies.

#### 9.3.1. Pilot studies

Data is collected on individual employees on Fitbit's R&D team. Data used in these studies are not anonymized unless the data itself is determined to be sensitive. For example, raw optical heart rate sensor data may be collected from researchers to examine how a change in the sensor affects the quality of the data. Depending on the context of the study, this type of information would not be considered sensitive and therefore would not typically be anonymized. In other cases, pilot studies may collect weight and age data on R&D team members, which is considered sensitive data. The data would be anonymized to the extent that only the one researcher who collects the data is able to match it to an individual. R&D team members who participate in pilot studies are informed beforehand that their participation is completely voluntary, that they are free to exit the study at will, and that any data used or created will be destroyed at their request. In pilot studies that involve sensitive data, a privacy policy for the data (explaining, for instance, how the data is anonymized and who has access to it) is also provided.

#### 9.3.2. Internal Fitbit studies

Internal Fitbit Studies Data is collected on Fitbit employees not part of the R&D team who volunteer to participate in studies. Determining data controls and levels of sensitivity for internal Fitbit studies is done in a way similar to pilot studies, which relies on applying context to the data. However, all of the data in this study type is anonymized so that only the researcher collecting the data can match the data to an individual within the company. For smaller internal studies, participation is voluntary and the researcher gives immediate verbal feedback to employee data donors. When they do larger studies internal to Fitbit, there is usually a thank you note and a wrap up of basic findings sent directly to the employees.

#### 9.3.3. Fitbit user studies

Data is collected on users of Fitbit products who are not Fitbit employees. Fitbit views the data in this study type as the most sensitive and therefore anonymizes it, even to the lead researcher. Thus, the lead researcher in Fitbit user studies should not be able to access explicit personally identifiable information for any user. There are some exceptions when necessary; for instance, the researcher may access demographic information such as gender, weight, height, and age in order to perform broader analysis on the data. One hypothetical example of a Fitbit user study would be seeking to understand how many people setting daily step goals using their activity trackers actually meet those goals on a daily basis. To decide how to best protect user privacy in this case, researchers would determine the scope of the project (by asking, for instance, if the target of the research would be all Fitbit users or just one subset of users) and use that determination to decide whether to use the data with or without user IDs.

## 10. Embedded privacy practices

As noted above, Fitbit customer information may be anonymized, or rendered de-identifiable, depending on the context of the research, such as where the data comes from, how large the dataset will be, and/or the sensitivity of data components.[10] Anonymization refers to techniques used to minimize the exposure of personal information to the research team. Techniques such as assigning unique participant codes help minimize exposure of participant information to the bare minimum. In addition to protecting privacy, anonymization can help guard against experimental bias, which can occur when the experimenter is able to tie specific participants to results from an experiment. Anonymization forces the experimenter to "follow the data" that is generated instead of relying on stereotypes or other less scientific heuristics. Before it is anonymized, data must be viewed in the raw, un-anonymized form in order for an initial experimenter to assign unique identifiers. For example, a Fitbit employee might ask five employees to wear a heart monitor while using a treadmill. To anonymize the results, the experimenter could then assign random participant IDs and shuffle the participant order so that the research team does not know which specific participant was tied to a specific set of data. In comparison, de-identification is a much stricter standard, applied when the intention is to share a data set outside the parent organization. For example, a hospital would aim to de-identify user data before sharing it with a university research team, but a university research team performing an experiment might simply strive to anonymize the data internally. De-identified data

[10] All user data is treated as sensitive, with tiered levels of access and application of anonymization and de-identification techniques.

has had mathematical techniques applied in order to make correlating the data with the participant close to impossible.

## 11. Recommendations

Companies are managing several dimensions of trust as they innovate and unveil new products and features. They are working to maintain trust between the company and users, the integrity of internal policies and practices, and the relationship between the company and society. The following recommendations are designed to align with these dimensions to capture a broad view of the underlying question: how can wearable companies perform ethical and privacy-protecting internal R&D? CDT and Fitbit considered existing policy frameworks and the approach of the Fitbit R&D team to form practical recommendations that can be applied to wearable research and development processes. Our research focused on the treatment of individuals by the R&D process and the company's overall culture of stewardship. However, another important consideration for companies and users is the contribution that health-focused technology can make to humanity. To address this, our recommendations set a benchmark for future research on broader social concerns and provide a common language for businesses, media, and advocates to describe the challenges and opportunities for wearables to transform society.

### 11.1. The individual: Digital dignity

Individual data subjects are often employees of the wearable technology company. This is a natural outcome of the research process, especially in a small, start-up environment. That said, the inevitability of this behavior does not release researchers from ethical and other obligations. Research conducted on employees raises unique questions. We recommend that wearable technology companies consider the following guidelines to preserve the dignity both of employees when they offer their personal data for experiments and for users whose data is involved throughout the R&D process. Individuals should be given a choice to determine how their data is used for internal research whenever possible. Wearable companies should have privacy policies that clearly state that user data generated by the wearable device is used for R&D, and individuals should be entitled to share as much data as they want with the company (as long as they are sufficiently informed), as well as stop the collection and use of their data if they so choose. Wearable users should have the means to delete identifiable data from their personal account (de-identify the data) or alternatively, to delete the account itself.

| |
|---|
| 1) *Use individual expectations to guide consents*. Device users expect that some of their data will be used for routine internal research and development, and thus it is not necessary to offer users an explicit opt-in consent for this purpose. However, researchers should require all users, including their colleagues, to opt-into participation in internal research when that research uses their identifiable data and falls outside a user's reasonable expectations. Companies should consider the purpose of research and whether it would have a negative impact on the user when determining whether opt-in consent is necessary. |
| 2) *Honor individual participation in research by offering rewards judiciously, not coercively*. Volunteers in human research studies may be remunerated in a way that is small but meaningful. This can include small benefits, like gift cards or a free month of a subscription product, but should not be big enough that they become a proxy for a penalty, or would constitute an excessive reward. |
| 3) *Innovations should serve the best interest of the individual*. Innovative technical strategies should be applied to augment privacy protections and offset ethical considerations. For example, concerns about employees feeling pressure to participate in research may be mitigated by technological solutions. Volunteers should have a mechanism for anonymity when participating in large studies[11] as well as the ability to withdraw their involvement at any time without fear of identification or reprisal. This is particularly important, and complicated, for participants who are also employees. For device users, companies should avoid incentivizing consent by unnecessarily removing functionality for certain features, or offering service upgrades conditional on consent. |
| 4) *Respect an individual's identity by applying appropriate protections*. User identity protection must be embedded in all research design through pseudonymous IDs and anonymized data. In particular, any data gathered from or about company employees should be considered sensitive and be stored separately from other employee-related data sets. Data aggregation should be the default research method, as it provides a broader view of sensor function, user behavior, and user trends without posing substantial privacy risks. Projects that utilize a |

[11] Anonymization is impractical for very small data sets (such as when the data is from the researcher herself and one volunteer) and thus should not be required.

larger data set and require more time and effort from volunteers should have strict anonymization standards. In addition, appropriate privacy protections and human subjects research training should be in place for studies whose results provoke the need for identification, such as when researchers need verbal or written feedback from a specific data volunteer. Researchers can use techniques to identify users if there are outliers in data without compromising the identity of the user. For example, researchers may create a "map" of pseudonymous identifiers to real identities, but use it only when a need to identify arises, destroying the map when this analysis is complete. The research quality may depend on determining the contributing factors for an extreme data point, and this investigation may even expose a lower quality result for underrepresented populations and prompt further investigation.[12] Another option for obscuring the data is "data permutation," which involves randomly selecting and changing data cells.[13]

5) *Address the special needs of vulnerable populations thoughtfully*. If the marketing or design of a product creates the expectation that its users might be considered a "vulnerable subject," such as the mentally challenged, a guardian capable of reviewing the material must give consent in a manner that accommodates the individual's disability. In the wearable context, where the health and wellness of users compels a more thoughtful approach to users with special needs, companies might build in a prompt to designate an authorized caregiver or they might design consents that allow various kinds of accessibility.

*6) Uphold individual trust through an exchange of straightforward information about data practices for internal research*. Companies should provide clear and detailed information about internal research on user and employee data in the company's privacy policies and related consent notices. It should be clear to an employee and an individual using a wearable device when data is being collected for internal research; what types of data is being collected for internal research; what that data is used for; what partners it is shared with (and how they use it); how long the data is retained; and what security measures are in place to protect it. Notices to all users on internal research practices should be clear, timely, and concise, but they do not need to be solely written documents (like privacy policies or real time messages)—they could be relayed through audio or visual methods that may be more accessible on small screens. The ideal moment to present data disclosure and sharing choices is when users first connect the device to the Internet. Information about internal research must be comprehensive, truthful, and easy to understand.

## 11.2. The Company: Operational stewardship

Privacy and ethics are not only a concern for the data procedures and practices of one corporate team. The overall culture of an institution, from its written policies to how it interacts with employees, echoes throughout the R&D process and will be reflected in the product design. The formal processes and decisions created by the institution deserve scrutiny, as they will chart the course for the evolution of a product and the ultimate success of a company. Building a culture of data stewardship is foundational for the implementation and sustainability of privacy-aware and ethical internal research practices. These recommendations illustrate ways a wearable technology company can institutionalize operational stewardship, either in the R&D process or throughout the company structure.

1) *Invest in employees with a background in privacy and ethics*. Companies in wearable health should hire individuals with a background or experience in health, health care, sociology, ethics, and/or human subject research.[14] Data anthropologists with experience in the health and wellness arena, for example, offer a broad perspective on design interface, product usability, and user behavior.

2) *Mitigate power asymmetries that result from employer access to employee data*. Companies that do research using their employees as data subjects should have formal, written policies that place limitations on sharing of and access to data and analysis. In particular, restrictions on access by management or human resources staff, insurance companies, and third parties are paramount. While a

[12] "How big data is unfair: Understanding sources of unfairness in data driven decision making," https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de#.s3qlex7ia "...less data leads to worse predictions. Unfortunately, it's true by definition that there is always proportionately less data available about minorities. This means that our models about minorities generally tend to be worse than those about the general population."

[13] Using data permutation researchers can still perform statistical analysis on aggregate data but it becomes harder in general for the data to be re-identified.

[14] At Fitbit, many individual researchers have experience applying ethical considerations to use of human subject data in research, either through experiences in graduate school or in prior employment. They also frequently expressed a deep interest in health and wellness overall. Thus, it is the researchers themselves that seed a culture of data stewardship by embedding privacy and ethics values into research practices.

record of employee participation in studies may be kept (e.g., for the purposes of study coordination), declining to participate should not be penalized or adversely affect performance evaluations.

3) *Empower researchers with flexible, embedded tools for data stewardship*. Provide researchers with a rubric for evaluating the harms and benefits to users for any project that analyzes user data. This rubric should allow researchers to assess the privacy risks for each project, including the purpose of the research, the sensitivity of the data in context, and the reasonable expectation of privacy by the user. It should provide company rules for escalating data protections, consent, and increasing ethical considerations, depending on sensitivity.

4) *Security protocols and practices must guide all interactions with data*. Researchers should be aware of both established[15] and emerging[16] security protocols for protecting data in a health and wellness context. Formal protocols should: a) Combine de-identification techniques with contractual obligations that restrict third parties from attempting to re-identify data and maintain data security standards that minimize the chance of data breach b) Retain and share data that has been de-identified for internal research as long as the wearable company and individual researchers that access and use the data commit to not re-associate it with an individual or device without the individual's consent c) Periodically assess whether to delete large datasets of anonymized or de-identified historical user data when this data is no longer necessary for ongoing internal research projects in order to mitigate any risk to user privacy and security. Additionally, companies should secure data compiled from wearable devices for research purposes while the data is in transit (such as being wirelessly sent to a base station, phone, or computer) or at rest on a company's servers. If data cannot be protected in transit from the device to the base station, it is important to offer an option that allows a device to be only associated with an identified base station, phone, or computer through mutual paired authentication. Companies should also establish well-founded technical, administrative, and personnel security measures, and include regular auditing and frequent updating of security systems. [17]

5) *Sensitive personal information should trigger limitations on data collection and use*. To adequately minimize and protect data involved in R&D, companies should securely store user data once researchers destroy any correlations between data that are no longer relevant or a part of an active project. It is important for researchers to be able to understand instructive correlations or patterns by combining data points but not necessary for them to use identifiable data. Research teams should consider the potential benefit and risk to the user of a research project, the users' expectations for how and why their data is used, the sensitivity of the data involved, and any material negative impact on user experience when deciding to initiate a research project, particularly if the project will involve the correlation of sensitive data points.

6) *Establish formal accountability measures to create sustainability and opportunity*. Wearable companies must create and enforce formal accountability measures that address the privacy, ethics, and security of user data for internal research practices, including dynamic checks and balances during research process.

### 11.3. The community: Social good

Wearable companies such as Fitbit are devoted to increasing individual health and wellness. By design, their business models work to augment social good, one person at a time. As an important and growing part of the health care ecosystem, the wearable industry has an ethical obligation to acknowledge this role and dedicate resources toward broader research that benefits humankind.

1) *Commit to improving humanity through research*. Companies should adopt policies that encourage socially conscious research projects and direct resources to internal research that focuses on improving the lives of users and society as a whole.

2) *Ensure diverse communities are represented*. Advocates have raised awareness of the pitfalls of big data as a tool to design broadly applied algorithmic rules. Wearable companies should ensure that data used in research is inclusive of traditionally underrepresented groups and of a range of demographic and geographic populations.

3) *Incorporate cultural sensitivity in internal policies*. Companies should also establish institutional policies of awareness and sensitivity to the many ways a product, service, or feature of a device can impact different communities. For example, the analysis of

[15] Garfinkel, Simpson. National Institute of Standards and Technology Internal Report 8053 vi, October 2015. Available at: http://dx.doi.org/10.6028/NIST.IR.8053

[16] Ann Cavoukian and Khaled El Emam, De-identification Protocols: Essential for Protecting Privacy (June 25, 2014). http://www.privacybydesign.ca/content/uploads/2014/06/pbd-de-identifcation_essential.pdf

[17] While the HIPAA Security Rule does not cover much of the data that flows from individuals to wearable devices, the law's focus on encryption is a helpful standard for developers and device manufacturers to consider when designing their security programs.

health-based data can inadvertently reveal sensitive information, such as ethnicity or sexual orientation.

*4) Share broad insights on health and wellness publicly*. Wearable companies should consider communicating, via notices that are separate from consent notices, the results of studies that use customer data, particularly if that research is geared toward understanding larger societal health or wellness issues. Researchers should also periodically provide users with examples of what the company achieved or learned through research using their data.

## 12. Conclusion

Success in today's competitive global technology market depends in large part on how companies balance corporate citizenship with innovation. To achieve loyalty and trust from users while constantly evolving and offering new products and services, companies must do more than implement good data practices—they must build a culture of privacy and security that embeds and formalizes values of digital dignity and data stewardship, and contributes to the social good. As the wearable industry grows, and as products and services become more intimately connected to our personal lives, questions about the role of individual dignity, data stewardship, and corporate citizenship will increase. Committing to privacy-aware and ethical guidelines for R&D is an important step toward building a sustainable and socially conscious industry that offers the public a trusted voice in wellness and the quantified self.

### APPENDIX 1: INTERVIEW QUESTIONS

Survey questions included the following: 1. Please take me through a typical day for you. 2. What sort of research projects do you work on? 3. What are examples of project goals? 4. What kind of user data is most valuable in terms of research potential and/or achieving a research goal? 5. How do you form a research question? 6. How do you decide which research questions require further exploration? 7. How do you determine when a research project is complete? 8. How long do projects typically take? 9. Describe a typical process for embarking on a research project. 10. Who is involved in each stage of the process? 11. Do you report outcomes to users at the end of a project? 12. When do privacy and ethical considerations about data typically come up?

### APPENDIX 2: FITBIT PRIVACY POLICY

The sections of the Fitbit privacy policy that allows for the company to use and study consumer data are: "Fitbit uses your data to provide you with the best experience possible, to help you make the most of your fitness, and to improve and protect the Fitbit Service…data and logs are used in research to understand and improve the Fitbit Device and Fitbit Service...de-identified data that does not identify you may be used to inform the health community about trends; for marketing or promotional use; or for sale to interested audiences." https://www.fitbit.com/legal/privacy-policy

### APPENDIX 3: RESEARCH TEAMS AT FITBIT

While one or two researchers perform hacks at most, project work usually occurs in comprehensive teams that contain at least one member from each of the Fitbit core competencies: 1) Hardware engineers (e.g., electrical and mechanical engineers) who focus on designing sensors and other hardware components and/or finding new uses for existing sensor technologies. 2) Software engineers (e.g., firmware and algorithm engineers) who develop on-device signal processing software and interactive experiences. 3) Data scientists who analyze data and develop algorithms to spot interesting health trends. 4) Human subjects researchers who work on validating theories generated by data scientists and evaluating the usability of hardware and software via human subject experiments.

### APPENDIX 4: MAP OF FITBIT R&D