Consumer Health Data: Regulation, Governance, and Innovation

Elizabeth Davidson
Shidler College of Business
University of Hawaiʻi at Mānoa
edavidso@hawaii.edu

Jenifer Sunrise Winter
School of Communication and Information
University of Hawaiʻi at Mānoa
jwinter@hawaii.edu

Abstract

Consumers are increasingly turning to mobile health applications (apps) for healthcare needs. Enthusiasm for their transformational potential is widespread, though there are concerns about ensuring ethical and socially beneficial uses of the personal health data they generate. Here we consider, how can consumer health data be governed, so as to preserve individual privacy, choice, and autonomy while also facilitating their potential to contribute to health system innovation and research? Our goals in this paper are to highlight key data governance challenges related to consumer health data, particularly concerning regulatory oversight of these types of data, and then to outline an agenda for IS research, curriculum development, and engagement activities in this domain. We review sources of consumer health data, special governance challenges with this type of data, and emerging regulatory approaches relevant to consumer health data governance, and highlight research and pedagogy opportunities for IS scholars.

Keywords: consumer health data, m-health, data governance, IT regulation, digital innovation, artificial intelligence.

1. Introduction and motivation

Consumers are increasingly turning to mobile health applications (apps) to find relevant health information, work towards their health and wellness goals, manage chronic conditions, communicate with healthcare providers, and assist in care of family members. Entrepreneurial start-up firms as well as tech giants like Apple and Google are flooding the lucrative consumer health marketplace with apps that target a wide variety of consumer needs and populations (Milne-Ives, van Velthoven, and Meinert, 2020; Vesselkov, Hämmäinen, and Töyli, 2019). For instance, mobile apps targeted towards elderly populations include pill reminders and dispensers, personal EKGs (e.g., KardiaMobile), and continuous glucose monitors (e.g., Dexcom).

Enthusiasm for the transformational potential of consumer health apps is widespread, but there is also skepticism among healthcare professionals, funders, and policymakers about their efficacy, safety, and potential to contribute to cost-effective healthcare (Kao and Liebovitz, 2017; Montgomery, Chester, and Kopp, 2018). Policy makers in different countries and regions are developing processes for evaluating and recommending health apps to consumers (Essén et al, 2022), but they are struggling to determine how to effectively regulate the burgeoning marketplace, for instance, whether to treat health apps and related software as medical devices (SaMD) or more simply as consumer products (Yaeger et al. 2019).

In this research we are interested in particular aspects of these broad regulatory concerns: *How can, and should, consumer health data, particularly data generated by mobile health apps, be governed, so as to preserve individual privacy, choice, and autonomy while also facilitating potential contributions to health system innovation and research?* Traditionally, data governance has been viewed as the responsibility of the data-generating organization, which is expected to protect consumer privacy against cybersecurity breaches and unauthorized access to data (Abraham, Schneider, and vom Brocke, 2019; Davidson et al., 2023). Beyond this, consumer firms can govern data primarily to benefit the firm, as long as the firm complies with applicable regulations (Winter and Davidson, 2022). However, it has become increasingly apparent that the innovative potential of data relies on data being shared widely among organizations, governmental agencies, researchers, consumers – that is, stakeholders with potentially conflicting interests in how data are used and variying capacity to govern data effectively (Bari and O'Neill, 2019; Bell and Shimron, 2023; Roski, Bo-Linn, and Andrews, 2014). Thus, data governance today requires a broader set of discussions, practices and regulations about when and how

URI: https://hdl.handle.net/10125/109382
978-0-9981331-8-8

digitalized data should be shared across stakeholder groups in society, what constitutes legitimate uses of personal data within and between organizations, who benefits (or should benefit) from data exploitation, and so on. Public policies developed and enacted through data regulations have a critical role in mediating these issues among stakeholders and in fostering societal benefits, and thus in data governance discussions.

This expanded understanding of data governance and the importance of data regulation are very relevant to consumer health data (Winter and Davidson, 2022). Mobile medical devices prescribed by a healthcare provider, and hence the health data they generate, are typically governed under applicable healthcare regulations. However, many mobile health apps available in the consumer marketplace, as well as the personal health data they generate, are governed under consumer, not healthcare, regulatory authorities. These data reside not only the consumer's mobile device or computer, but also in the cloud infrastructures of app providers (Vesselkov et al., 2019) and therefore are governed by the vendor's consumer data privacy and governance policies. This is concerning, because mobile health apps generate consumer data that are highly personal, sensitive, even intimate.

Consumer health data are among the most valuable sources of data in the "big data" ecosystem (Roski et al., 2014). Depending on their business model and the regulatory environment, app providers may share these data with employers, insurers, healthcare providers, health researchers, among others, and may also monetize data through Internet advertising networks and with data sales to firms outside of the healthcare sector, often without consumers' knowledge or consent (Grundy et al., 2021; Winter and Davidson, 2022). Consumer health data can be combined with other sources of health or consumer activity data. Even when these data are not explicitly linked to health, advanced analytics software enables "digital phenotyping" in which these varied data sources are used to characterize individual consumers, including making predictions about their health conditions, to market to them or potentially to discriminate against them based on health predictions (Montgomery et al., 2018; Perez-Pozuelo et al., 2021; Warzel, 2019). Nonetheless, despite privacy concerns, sharing consumer health data has the potential to generate health system innovations and societal benefits. For instance, machine learning algorithms require large data sets to train AI, and growing stockpiles of consumer health data can provide necessary fuel for AI innovations in healthcare (Favela et al., 2020; Jim et al., 2020; Wolff et al., 2023).

All told, these developments are both promising but also highly concerning. While much policy and research attention has been devoted to governance of clinical health data (i.e., data generated within hospitals, clinics, and so on) (Rosenbaum, 2010), discussions of consumer health data are just now emerging. Thus, our research goals in this article are (i) to highlight key data governance challenges related to consumer health data, (ii) discuss emerging regulatory oversight to address challenges, and (iii) to outline promising areas for research and pedagogy to help develop and diffuse socially-responsible governance approaches for consumer health data. In the next sections, we outline our research approach, review key sources of consumer health data, identify data governance challenges in this domain, and discuss existing and emerging regulatory approaches. Building on this background and analysis, we outline possible research and pedagogical topics to contribute to effective regulatory governance of consumer health data within and across organizations and stakeholders.

2. Background and perspectives

This agenda-setting essay grew from research we have been conducting on health data governance over the past decade (Winter and Davidson, 2019a, 2019b, 2022). In the course of empirical studies of health data governance organizations, we have examined literature from legal, healthcare, public policy, regulatory, and information systems domains, including academic articles, policy papers, health data privacy legislation and regulations, technology vendor documentation, white papers from advocacy groups, and so on. This broad corpus along with our ongoing engagement with health data developments informs our understanding of data governance issues as health data are increasingly digitized, accumulated, shared, and monetized. In this paper, we foreground developments related to the rapid expansion of digitalized consumer health data, much of it from mobile consumer health apps. This essay presents our summative reading and interpretation of recent regulatory developments as reported by regulators, industry consultants, and academic scholars (and cited herein) and our resulting recommendations for further IS research and practice.

2.1. Sources of consumer health data

Today rapidly expanding sources for personally identifiable health data – termed consumer health data – are generated within the consumer health and wellness industry from mobile health apps and as "digital exhaust" from consumers' Internet activities (e.g., online shopping, social media use, location data). Demographic, social and economic data from non-health related sources (e.g., census data, financial records, social media) are also mined for health-related insights about consumers (Bari and O'Neill, 2019).

Table 1 highlights popular categories of consumer health apps and the types of consumer health data they generate. Some consumer health apps combine biometric monitoring devices and mobile smart phones to collect data from the consumer's body and daily activities, interpret and display the data for the user, and track this biometric data to provide the consumer with customized health-related advuce (Kao and Liebovitz, 2017). Other applications depend on the user to enter data, for instance on diet, activities, moods or emotions, menstrual cycles, and so on. From the perspective that "all data are health data" (Warzel, 2019, para. 4), health conditions such as depression, contagious disease, chronic disease, pregnancy, to name just a few, can be imputed from location data gathered from mobile phones (e.g., via geofencing or other tracking) (Favela et al., 2020), Internet searches and purchases online, memberships in online communities (Tempini, 2017), and many other Internet data sources. Even when consumer data are not explicitly linked to health, advanced analytics software enables "digital phenotyping" in which these varied data sources are combined to characterize individual consumers, including making predictions about their health conditions (Perez-Pozuelo, 2021).

Thus, what counts as "consumer health data", rather than simply as "consumer data", is open to interpretation, contributing to regulatory gaps and confusion in how these data are (or should be) legally and ethically aggregated, applied, shared, or monetized (Fazlioglum 2023; Winter and Davidson, 2022). For instance, if over-the-counter (non-prescription) drug purchases (e.g., sleep aids, medicines for indigestion) are treated merely as consumer data, pharmacies might legally sell identifiable customer health-related data (possibly enhanced via phenotyping) to pharmaceutical firms to directly market their drugs to consumers.

2.2. Regulatory challenges with governance of consumer health data

In the past, personal health data were gathered primarily in healthcare settings where healthcare professionals provide services to their patients, such as in a hospital or clinic. These "clinical" or "medical" data are then governed within the healthcare organization. Given the highly sensitive nature of health data, clinical data have long been closely regulated under comprehensive privacy and/or sectorial data regulations, such as the GDPR in the European Union or HIPAA in the U.S. (de Kok et al., 2023; USHSS, 2022). Sharing data across healthcare settings to benefit patients and for research has been promoted, even mandated, though sharing is highly regulated and monitored to respect patient privacy and consent (European Union, 2024; Lamb, 2023; Marelli, Testa, and van Hoyweghen, 2021; Rosenbaum, 2010).

Table 1. Common sources of consumer health data from mobile-health apps.

Common categories and examples	Consumer health data generated
Health and wellness: Activity trackers (e.g., Fitbit, Apple Watch); fitness studio/exercise programs; nutrition and diet aids; mindfulness tools	Steps, heart rate, O2 saturation; location, activity; calorie and nutrition intake; moods, emotions.
Women's Healthcare: Menstrual cycle and fertility tracker	Details data on bodily functions, family planning goals, emotional responses.
Eldercare: Medication management; pill reminders/ dispensers; location and activity monitoring apps	Age, medical history, prescriptions, activities, diagnoses.
Chronic health condition management: Continuous glucose monitors; blood pressure gauges; heart monitors	Blood glucose readings; blood pressure readings; detailed data on bodily functions; EKGs
Medical record apps: Family health apps; immunization records; data from health record portals transferred from clinical systems (EHR)	Medical histories transferred from clinical electronic health record portals; search history related to portal use
Education and prevention: Symptom trackers; medical information (e.g., WebMD); online health education	Search histories by IP or MAP address linking search activity to other Internet activities
Prescriptions, supplies, service locators: Prescription price comparison apps; pharmacy delivery apps; medical supply apps	Diagnoses, medical coverage, age, address, insurance, financial information

Consumer health data are generated, aggregated, managed, and shared in very different circumstances. Data generated by a mobile activity tracker such as an Apple Watch or Fitbit resides on the consumer's devices as well as in the technology firm's cloud repositories (Grundy et al., 2021; Vesselkov et al., 2019). Consumers' online purchases of over-the-counter (or even prescription) products are routinely collected by e-commerce sites. Search engines aggregate data on Internet searches for health-related information and services. As these examples illustrate, data relevant to health are intermingled with all manners of consumer data – location, purchases, technology use – gathered by multiple apps and network devices controlled by many different firms, which analyze, package, and sell data into the markets for digitalized consumer data (Bari and O'Neill, 2019; Montgomery et al., 2018; Roski et al., 2014). As a result, the majority of consumer health data are not covered by traditional health data regulatory oversight, particularly in the United States (Winter and Davidson, 2022).

Concerns about effective governance of consumer health data arise from a number of social, technical, and market developments. First, many new technology actors, with little or no historical involvement in healthcare, have entered the consumer health market (Jim et al., 2020). Big Tech companies are avidly entering the healthcare sector as secure cloud services providers, developers of AI-enabled analytics in partnership with health care organizations, creators of consumer health apps, and partners in medical research (Marelli et al., 2021). These firms already accumulate many types of data on consumers that may be used to infer health – shopping habits, biometric data, web searches, and so forth (Perez-Pozuelo, 2021; Warzel, 2019). These data sources also make (re)identifying individuals from putatively anonymized data sources easy. Data gathered from consumers are typically collected, stored, and aggregated on the technology infrastructure ("cloud") of myriad app providers (Vesselkov et al., 2019). Dispersed data sources increase cybersecurity risks for sensitive health data and facilitate monetization of consumer health data.

In the U.S., the Supreme Court ruling in 2021 that overturned women's right to abortion care has raised the specter of women being prosecuted for seeking reproductive services or information, based on their clinical or their consumer health data. The U.S. Department of Health and Human Services limits access to health data generated within clinical settings through the Health Insurance Portability and Accountability Act (HIPAA) (USHSS, 2022). While newly promulgated HIPAA rules limit disclosure of reproductive health data for purposes other than direct healthcare services or legitimate research, consumer health data, including data related to reproductive health such as fertility apps, geo-location, Internet inquiries, and non-prescription drug purchases, are not covered by HIPAA and thus are open to subpoena or even data sales.

Given the sensitive nature of health data, regulatory attention to the governance of consumer health data is growing (Grundy et al., 2021; Kariotis et al., 2020; Lamb, 2023; Sharon and Lucivero, 2019). Two recent cases in the U.S. demonstrate that concerns are well founded. GoodRx (GoodRx.com) is a prescription drug discount vendor and telehealth service app that provides users with coupons for prescription drugs, a beneficial innovation that can help consumers afford high-priced medications. Despite promising consumers that their health data would not be shared, the firm for years did share identifiable health information with advertisers and other third parties, including Facebook and Google. GoodRx monetized customers' personal health data by combining data collected via the prescription app with Facebook data to identify and target market customers with advertisements. GoodRx also failed to maintain control over how third parties utilized their data and misrepresented its privacy policy by claiming the firm was complying with health sector and consumer notification regulations (Federal Trade Commission, 2023). After years of pursuing these data monetization practices, GoodRx was fined by the Federal Trade Commission, and the firm's right to monetize identifiable health data in the future was limited.

Similarly, the U.S. Federal Trade Commission (Federal Trade Commission, 2021) applied consumer data protection regulations (versus drawing on HIPAA health data regulations) against providers of ovulation and period tracking mobile applications Easy Healthcare Corporation (Premom app) and Flow Health, Inc., whose apps have been used by over 100 million consumers. These companies disclosed identifiable health information to third party companies like Facebook and Google for advertising and analytics, in contradiction to the firms' stated data handling policies and without consumers' consent, and they failed to limit how third parties used this sensitive consumer health data. Both firms were fined millions of dollars and their future data monetization practices were restricted by regulatory order.

These instances illustrate the need for effective regulation and pervasive governance practices within and across firms, which acknowledge and preserve consumers' privacy and autonomy regarding access to intimate consumer health data. That said, regulations should also balance the potential for health care research and innovation to promote societal well-being through ethical uses of consumer health data. For instance, activity tracking and other consumer health apps can be helpful in detecting dementia, supporting eldercare givers, and providing services to the elderly (Favela et al., 2020, Wolff et al., 2023). As another example, Apple has shared Apple Watch heart monitoring data with heart disease researchers.

Uncovering such applications of consumer health data and developing innovative products and services are typically private sector activities, which require innovative firms have access to consumer health data resources. However, finding the right balance between privacy protections and innovation through data reuse, analytics and AI is proving difficult to achieve. Stringent omnibus regulations are now being challenged on the grounds they are too restrictive to enable medical research, drug discovery, and improvements in healthcare system efficiencies and access. Medical researchers have expressed concerns that strict data protection laws such as the GDPR or the California Consumer Privacy Act may limit future beneficial uses of the data (Bell and Ahimron, 2023; de Kok et al., 2023). To better understand these governance tensions and challenges, we now consider how consumer health data regulations are developing.

2.3. Emerging consumer health data regulations

With widespread digitization of health data in both clinical and consumer settings, researchers and policy makers are now considering how to regulate consumer health data to protect privacy but also to promote health systems improvements (Bari and O'Neill, 2019; Grundy et al., 2021; Kariotis et al., 2020; Sharon and Lucivero, 2019; Winter and Davidson, 2022). Because the health sector is heavily regulated in many countries, these discourses typically focus on the national regulatory context. Despite common goals for protecting privacy and promoting beneficial data sharing, there are significant differences in national or regional regulatory approaches that affect how governance of consumer health data might be addressed.

Throughout the EU, the GDPR (2018) requires organizations to safeguard and protect personally identifiable information (PII). The GDPR considers "data concerning health" to be a special category with even more stringent protections. Health data under the GDPR are broadly construed and cover any data related to health, including data that infers health (Lamb, 2023). GDPR is supplemented by various national laws that may further protect health data. However, GDPR principles like data minimization and limited data retention and sharing are being re-negotiated as some argue they hinder data-driven innovation enabled by information reuse (Marilli et al., 2021). In early 2024, the EU created a commission proposal for a European Health Data Space (EHDS) to foster development of a single digital health market in the EU while allowing individuals' control of their health data and facilitating the exchange of data for healthcare across the EU (European Union, 2024). However, this proposal will likely address primarily clinical data collected and shared among health services providers, while its utility for protecting consumer health data is less clear. Similarly, the recently passed AI Act (2024) provides omnibus rules for AI models and algorithms derived from data, although it is still unclear what implications this law will have for digital health in the EU and abroad: "Due to the horizontal nature, wide scope, and rapidly changing nature of healthcare AI, there will be many problems created by ambiguities and uncertain intersections with existing laws" (Gilbert, 2024, p. 3).

China enacted the Personal Information Protection Law (PIPL) in 2021, which grants personal health data a higher level of protection and, like the GDPR, requires explicit consent from data subjects and notification of the scope, necessity, processing methods, and retention period for data sharing. To balance privacy protection and innovation, China's National Standard "Information Security Technology Guide for Health Data Security 2020" includes the definition of a "limited data set," allowing de-identified data to be used without consent for medical research and public health purposes (Yao and Yang, 2023). Again, these regulations do not explicitly address consumer health data collected in day-to-day activities by mobile health apps in China's rich mobile-app ecosystem.

In the U.S., data privacy regulation relies on a patchwork of federal sectoral laws and state-level legislation. In the healthcare sector, the federal Health Insurance Portability and Accountability Act (HIPAA) of 1996 mandates protection and confidential handling of some clinical health information but its application is limited to specified covered entities (e.g., physicians or insurers) and their designated business associates, such as technology firms providing health data storage or analysis services. A consumer notification rule in the American Recovery and Reinvestment Act of 2009 (the Health Breach Notification Rule) and the HIPAA Omnibus Rule in 2013 have recently been used in U.S. federal actions against some m-health app providers for their consumer health data sharing practices. However, consumer health data overall are not addressed by HIPAA, leaving significant gaps in consumer health data protection in the U.S. (Bari and O'Neill, 2019; Winter and Davidson, 2022).

The Federal Trade Commission (FTC), a U.S. federal agency that protect consumers from unfair and deceptive business practices, has signaled that it is moving towards a broader definition of consumer health information and will be cracking down on companies that use tracking information to capture or share information about consumer health (Federal Trade commission, 2021, 2023, 2024a, 2024b). In May 2023, the FTC requested public comment on proposed changes to the Health Breach Notification Rule (HBNR) and in April 2024 finalized changes, including revised definitions to clarify the rule's coverage of health apps and technologies not presently covered by HIPAA. The revised rule further clarified that a "breach of security" includes "unauthorized acquisition of identifiable health information that occurs as a result of a data security breach or an unauthorized disclosure" (Federal Trade commission, 2024a, para. 6). Of note, "unauthorized disclosure" includes a firm's intentional use or resale of consumer health data in violation of its stated privacy policies. In recent cases the Federal Trade Commission took action because it was alleged that companies violated the FTC Act with regard to unfair and deceptive practices. In the GoodRx case, the FTC also alleged that the company violated the HBNR, marking the first enforcement action of this rule since it was enacted in 2009. (GoodRx ultimately paid only a modest $1.5 million civil penalty and admitted no wrongdoing.)

To harness the opportunities and mitigate the threats of AI with data reuse, the White House Office of Science and Technology Policy released the Blueprint for an AI Bill of Rights, to "guide the design, use, and deployment of automated systems to protect the American public in the age of artificial intelligence" (White House Office of Science and Technology Policy, 2022, para. 4), and President Biden issued the landmark United States Executive Order no. 14110 on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (United States Executive Order no. 14110, 2023). At the Federal Trade Commission's 2024 Tech Summit on Artificial Intelligence, Chair Lina Khan noted that "we're focused on crafting effective remedies that establish bright-line rules on the development, use, and management of AI inputs. The FTC is making clear that some data is simply off the table for training models" (Federal Trade Commission, 2024b, p. 4), including sensitive health data and location tracking data. Whether this executive order will be effective to guide firms toward authorized, beneficial uses of consumer health data in AI development is unknown.

Beyond the actions of the U.S. federal government, several states recently enacted legislation enhancing consumer privacy generally or targeting consumer health data protection specifically (Reilly and MacDonald, 2023a, 2023b). For instance, in the state of Washington, the "My Health My Data Act" (2023) is the first privacy law in the U.S. to focus on health data falling outside of HIPAA (Office of the Attorney General, Washington State, 2023). Connecticut has followed suit with its own consumer health data regulations. Other states, such as California, Colorado, and Virginia have enacted comprehensive data privacy legislation since 2020. Such legislation extends the definition of health data as "sensitive data" under omnibus privacy legislation to include any data that might be used to identify a consumer's past, present, or future physical or mental health condition or diagnosis, including gender-affirming health data and reproductive sexual health data as well as data derived from non-specific health information such as location data, internet searches, purchases, and so on (Fazlioglu, 2023; Office of the Attorney General, Washington State, 2023; Reilly and MacDonald, 2023a, 2023b). Broad definitions of consumer health data will potentially impact the collection, use, and sharing of consumer data across a wide variety of businesses and markets, e.g., pharmacies, health clubs, online shopping and beyond. However, these laws rely on firms' self-regulation, and their efficacy has yet to be tested.

To summarize, despite some progress in recognizing the need for more actionable regulation of consumer health data, national and state-level health data regulations primarily address clinical health data. Thorny governance questions about whether and how to regulate consumer health data are still being addressed. To the degree some regulations do apply to consumer health data, they rely primarily on industry self-regulation and have limited enforcement options. As we see in the cases described above, firms may state their own terms for consumer health data governance, and they may violate even those terms for extended periods without regulators' notice or action.

3. Towards an agenda for consumer health data governance research and pedagogy

Establishing or extending regulatory guardrails for consumer health data governance are essential steps in balancing competing societal goals for consumer privacy and autonomy with innovation and research. However, regulation alone will not ensure effective governance. With the burgeoning market for consumer health apps, regulators simply are not able to keep up with monitoring and enforcement, and the penalties for violations are dwarfed by the financial gains from monetizing these data. With the jumble of overlapping, confusing, and conflicting consumer health data regulations across sectors and national contexts, even well-intentioned firms will find it difficult and expensive to be compliant. Regulatory limits and ethics guidance for data use and reuse must be explicated in organizational practices, structures and technologies. Along with a growing cadre of IS researchers (e.g., Abraham et al., 2019; Davidson et al., 2023; Kariotis et al., 2020; Link et al., 2017), we argue that new research into data governance as well as pedagogical innovations focused on data governance are desperately needed. Here we focus on such efforts for consumer health data.

We adopt a broad definition of data governance that considers not only how organizations comply with regulations through data governance practices but also how and why regulatory oversight develops with regards to societal goals and priorities. We also consider that effective data governance structures and practices necessarily extend beyond organizational boundaries to platforms and data markets (Davidson et al., 2023; Kariotis et al., 2020). In doing so, we seek to problematize which social actors decide data governance questions and how trade-offs between regulatory control and promoting innovation might be overcome or mitigated. This broad view of data governance highlights myriad opportunities for IS scholars to contribute to enhanced governance and regulation of consumer health data.

3.1. Research opportunities in consumer health data governance

With big data analytics and AI/machine learning developments, there is growing interest among IS researchers on the organizational, societal, and ethical implications of data governance (Abraham et al., 2019; Davidson et al., 2023; Link et al., 2017). We point to some notable research opportunities that illustrate this rich potential; many more are possible.

IS scholars interested in design science research are well situated to examine how data governance policies are embedded (or ignored) in the technical infrastructures for mobile health apps. For instance, Grundy et al. (2021) highlights studies that contrast firms' stated policies and technical app designs that prioritize data sharing. Vesselkov et al. (2019) develops analytic algorithms and methods to examine data sharing and governance protocols embedded in the code of health apps in mobile app platforms. Analytic tools such as these could be leveraged at scale to assist regulators overwhelmed by the rapid growth in the mobile health apps marketplace. Design science principles can also inform app design approaches and IS curriculum design to build data governance into consumer health apps and data management software.

Organizational approaches to consumer health data governance research could produce in-depth case studies to surface strategies and practices for, as well as barriers to, regulatory compliant governance of these data to address privacy and innovation goals in specific circumstances (Anderson, Baskerville, and Kaul, 2023). Cross sectional surveys or cases could assess how aware firm leaders are of data governance issues and whether governance concerns influence strategic data (re)use plans (Black et al., 2023). Scholars studying data ecosystems could explore and propose new approaches to bridge the discursive and methodological gaps between IS and legal / regulatory research (Burmeister et al., 2022) so as to inform mobile health design and organizational practices.

Individual-level studies of consumers' preference for "opt in" and "opt out" of consumer health data sharing, such as voluntary contributions to research studies (Hillebrand et al., 2023) and their preferences for government vs. private governance of consumer health apps (Binzer et al., 2024), can inform app design as well as regulatory policy, particularly as research approaches such as data analytics using mobile app data are incorporated to better understand users' actions to protect privacy and control of their data.

IS researchers have been interested for some time in how individuals have turned to social media and online communities for support for their healthcare needs (cf., Barrett, Oborn and Orlikowski, 2016). Building on earlier research that adopted a virtual community lens, IS researchers could examine the data governance principles and goals of these communities, some of which developed to facilitate health data sharing and reuse (Tempini, 2017) or data monetization (Danatzis, Chandler, Akaka, and Ng, 2024). Novel organizational forms such as data trusts and data collaboratives have developed to shift governance decision rights towards the individuals who contribute data to these collectives (Kariotis et al., 2020); the viability and effectiveness of such collective data governance of consumer health data can be examined and possible exemplars of governance developed (Winter and Davidson, 2022).

3.2. Developing pedagogical resources for health data governance education

Educational programs are critical leverage points to influence information systems specialists and general business students – future IS designers and business mangers – if data governance and privacy protection are to be "baked into" information systems and infrastructures, business models, and systems design practices. Model curricula for IS undergraduate and Masters of IS programs include information management and ethical uses of IS as core competency areas (Lyytinen, Topi and Tang, 2021). However, do such classes address data governance in depth or detail, or do they focus more on technical data infrastructure concerns? For instance, are the principles of "privacy by design" (Cavoukian, 2021; Gürses, Troncoso, and Diaz, 2011) integrated into core IS courses? Principles such as limiting data collection to essential elements, defaulting to "opt in" choices for data collection, making opt-in granular rather than global, providing a mechanism for consumers to withdraw consent and then ensuring data lineage tracking is possible to execute on the consumer's choice, communicating policies and privacy options to consumers in transparent ways via interface design, and so on, are all relevant to data governance values. These approaches could be infused into the pedagogy of courses such as user experience and interface design, as well as systems analysis and database design, where training on how requirements about which data to collect and how data will be used and encoded into information systems is conducted.

A comprehensive approach for "governance by design" could infuse data governance principles throughout business curricula, extending discussion and guidelines to general business students at undergraduate and graduate levels. Doing so could balance powerful messages about exploiting data for the firm's economic self-interest so evident in business and analytics classes. Courses on data ethics and governance are included in some analytics programs (For an example, see https://www.pathways.prov.vt.edu/students-and-advisors/courses/bit-4604.html). However, we suggest that considering the application of data governance and data regulations in requirements determination processes and critically evaluating information system projects that are justified based on data reuse or sales practices, or alternatively, that intentionally restrict societally beneficial data sharing (e.g., information blocking) for a firm's sole advantage, could be more fully explored in core IS-major and general business classes, to counterbalance typical justifications of IS investments through data exploitation. Data management technologies such as metadata management, data dictionary tools, and master data management applications could be introduced through hands-on exercises to bring data governance practices to life, for instance, demonstrating how the principles of data use agreements, opt-out choices for consumers on data sharing, and segregation of highly sensitive health-related data from other consumer data can be implemented through data lineage analysis.

3.3. Collaborating with practitioners on consumer health data governance

Business practitioners have developed a substantial body of knowledge about organizational-level data governance that prescribes governance functions within organizations through frameworks to guide best practices (e.g., DAMA International, Data Governance Institute). These frameworks outline practices, policies, procedures, and technologies for data storage and access, cybersecurity, regulatory compliance, data architecture, technical infrastructure, and data stewardship. We advocate a somewhat different approach to data governance -- an interorganizational and societally focused perspective with a broad understanding of the implications for data governance – but we do think academics have much to learn from practitioners about designing and implementing data governance practices and technologies within organizations, which could inform IS research and curriculum in ways noted above (Vial, 2023).

In particular, medical informatics researchers and health system practitioners have engaged for some time in health data governance research to promote acceptable data sharing and reuse for research and innovation in health services delivery (Rosenbaum, 2010). Technical approaches for compliant data sharing have been developed, for instance, for deidentification of protected health data (PHI) through synthetic data algorithms that can allow for sharing anonymized, individual-level data linked across multiple sources while also preserving privacy (Gonzales, Guruswamy, and Smith, 2023; Murtaza et al., 2023). In healthcare, organizational practices and structures for engaging data stakeholders, authorizing health data reuse, and articulating data use agreements with third parties are well recognized, if not always well implemented.

As definitions of "consumer health data" are broadened under emerging regulations to cover broad swathes of businesses and their data collection and use practices, data governance approaches developed within healthcare settings could be introduced to business firms that are accustomed to firm or industry self-regulation, to help managers adjust to heightened consumer health data regulation. IS scholars are well positioned to work across these industry boundaries through their educational programs and research projects.

4. Concluding remarks

Consumer health data can be valuable resources for health system innovations that could benefit society as well as individuals. These data also present significant challenges to individuals' privacy, autonomy, and control over health data and the possible consequences of data use (or misuse). In this paper, we highlight the importance and ubiquity of consumer health data. We bring to the foreground substantive data governance challenges these data represent. These challenges occupy the intersections of health IT, mobile health, data analytics, and digital transformation, which are of interest to the IS field. We explain how current and developing approaches to regulatory oversight of these data are developing, as well as the shortfalls in regulatory oversight, topics that are seldom highlighted per se in IS research. Building on this understanding, we identify research areas where information systems scholars might help promote societally beneficial uses and reuses of consumer health data resources by delving into data governance and associated regulatory approaches. Consumer health data governance must be designed into information system technologies and infrastructures and embedded into individual and organizational practices to be effective. In the discussion above, we highlight ways in which IS scholars could contribute to these goals through pedagogy and engagement with practitioners. Finally, while our focus in this paper has been on consumer health data governance, we suggest similar approaches are applicable to address data governance and regulatory concerns related to digital transformation projects more generally, and importantly, to the rapid infusing of AI technologies into healthcare specifically and business practices overall.

7. References

Anderson, C., Baskerville, R., & Kaul, M. (2023). Managing compliance with privacy regulations through translation guardrails: A health information exchange case study. Information and Organization, 33(1), 100455.

Abraham, R., Schneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. International journal of information management, 49, 424-438.

Bari, L., and O'Neill, D. (2019). Rethinking patient data privacy in the era of digital health. Health Affairs Blog: https://www.healthaffairs.org/do/10.1377/hblog20191210.216658/full/

Barrett, M., Oborn, E., & Orlikowski, W. (2016). Creating value in online communities: The sociomaterial configuring of strategy, platform, and stakeholder engagement. Information Systems Research, 27(4), 704-723.

Bell, L. C., & Shimron, E. (2023). Sharing data is essential for the future of AI in medical imaging. Radiology: Artificial Intelligence, 6(1), e230337.

Binzer, B., Kendziorra, J., Witte, A. K., & Winkler, T. J. (2024). Trust in public and private providers of health apps and usage intentions: a sectoral privacy calculus and control perspective. Business & Information Systems Engineering, 63(2), pp. 273 - 297.

Black, S., Davern, M., Maynard, S. B., & Nasser, H. (2023). Data governance and the secondary use of data: The board influence. Information and Organization, 33(2), 100447.

Burmeister, F., Zar, M., Böhmann, T., Elkin-Koren, N., Kurtz, C., & Schulz, W. (2022, November). Toward Architecture-Driven Interdisciplinary Research: Learnings from a Case Study of COVID-19 Contact Tracing Apps. In Proceedings of the 2022 Symposium on Computer Science and Law, pp. 143-154.

Cavoukian, A. (2021). Privacy by design: The seven foundational principles. IAPP Resource Center, https://iapp. org/resources/article/privacy-by-design-the-7-foundational-principles.

Danatzis, I., Chandler, J., Akaka, M., & Ng, I. (2024, in press). Designing Digital Platforms for Social Justice: Empowering End-Users Through the Dataswyft Platform. MIS Quarterly.

Davidson, E., Wessel, L., Winter, JS., and Winter, S. (2023). Future directions for scholarship on data governance, digital innovation and grand challenges. Information and Organization, 33(1), 100454.

de Kok et al. (2023). A guide to sharing open healthcare data under the General Data Protection Regulation. Scientific data, 10(1), 404.

Essén, A., Stern, A., Haase, C., Car, J., Greaves, F., Paparova, D., et al. (2022). Health app policy: international comparison of nine countries' approaches. NPJ Digital Medicine, 5(1), 31.

European Union. (2024). European Health Data Space. https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

Favela, J., Cruz-Sandoval, D., Morales-Tellez, A., and Lopez-Nava, I.H., 2020. Monitoring behavioral symptoms of dementia using activity trackers. Journal of Biomedical Informatics, 109, p.103520.

Fazlioglu, M. (2023). Filling the void? The 2023 state privacy laws and consumer health data. Retrieved from: https://iapp.org/about/person/0011a00000DlDcFAAV/

Federal Trade Commission (2021). Developer of Popular Women's Fertility-Tracking App Settles FTC Allegations that It Misled Consumers About the Disclosure of their Health Data, January 13, 2021. Retrieved from: https://www.ftc.gov/news-events/news/press-releases/2021/01/developer-popular-womens-fertility-tracking-app-settles-ftc-allegations-it-misled-consumers-about

Federal Trade Commission (2023). FTC Enforcement Action to Bar GoodRx from Sharing Consumers' Sensitive Health Info for Advertising. Retrieved from: https://www.ftc.gov/news-events/news/press-releases/2023/02/ftc-enforcement-action-bar-goodrx-sharing-consumers-sensitive-health-info-advertising

Federal Trade Commission. (2024a). FTC Finalizes Changes to the Health Breach Notification Rule. https://www.ftc.gov/news-events/news/press-releases/2024/04/ftc-finalizes-changes-health-breach-notification-rule

Federal Trade Commission. (2024b). Remarks of Chair Lina M. Khan - FTC Tech Summit. February 2024. https://www.ftc.gov/system/files/ftc_gov/pdf/2024.01.25-chair-khan-remarks-at-ot-tech-summit.pdf

Gilbert, S. (2024). The EU passes the AI Act and its implications for digital medicine are unclear. Npj Digital Medicine, 7, 1-3. https://www.nature.com/articles/s41746-024-01116-6.

Gonzales, A., Guruswamy, G., & Smith, S. R. (2023). Synthetic data in health care: A narrative review. PLOS Digital Health, 2(1), e0000082.

Grundy, Q., Jibb, L., Amoako, E., and Fang, G. (2021). Health apps are designed to track and share. BMJ, 373:n1429.

Gürses, S., Troncoso, C., & Diaz, C. (2011). Engineering privacy by design. Computers, Privacy & Data Protection, 14(3), 25.

Hillebrand, K., Hornuf, L., Müller, B., & Vrankar, D. (2023). The social dilemma of big data: Donating personal data to promote social welfare. Information and Organization, 33(1), 100452.

Jim, H. S., Hoogland, A. I., Brownstein, N. C., Barata, A., Dicker, et al. (2020). Innovations in research and clinical care using patient- generated health data. CA: A Cancer Journal for Clinicians, 70(3), 182-199.

Kao, C. K., and Liebovitz, D. M. (2017). Consumer mobile health apps: current state, barriers, and future directions. PM&R, 9(5), S106-S115.

Kariotis, T., Ball, M., Tzovaras, B., Dennis, S., Sahama, T., Johnston, C., et al. (2020). Emerging health data platforms: From individual control to collective data governance. Data & Policy, 2, e13.

Lamb, S. (2023). Health data in the EU AND UK – regulatory trends and developments. https://www.mwe.com/insights/health-data-in-the-eu-and-uk-regulatory-trends-and-developments/

Link, G.J., Lumbard, K., Conboy, K., Feldman, M., Feller, J., et al., 2017. Contemporary issues of open data in information systems research: Considerations and recommendations. Communications of the Association for Information Systems, 41, pp. 587-610.

Lyytinen, K., Topi, H., and Tang, J. (2021). Information Systems Curriculum Analysis for the MaCuDE Project. Communications of the Association for Information Systems, 49. https://doi.org/10.17705/1CAIS.04939

Marelli, L., Testa, G., and van Hoyweghen, I. (2021). Big Tech platforms in health research: Re- purposing big data governance in light of the General Data Protection Regulation's research exemption. Big Data & Society, 8(1): 20539517211018783.

Milne-Ives, M., van Velthoven, M. H., and Meinert, E. (2020). Mobile apps for real-world evidence in health care. Journal of the American Medical Informatics Association, 27(6), 976-980.

Montgomery, K., Chester, J., and Kopp, K. (2018). Health wearables: Ensuring fairness, preventing discrimination, and promoting equity in an emerging Internet-of-Things environment. Journal of Information Policy, 8, 34–77.

Murtaza, H., Ahmed, M., Khan, N. F., Murtaza, G., Zafar, S., & Bano, A. (2023). Synthetic data generation: State of the art in health care domain. Computer Science Review, 48, 100546.

Office of the Attorney General, Washington State. (2023). Protecting Washingtonians' personal health data and privacy.https://www.atg.wa.gov/protecting-washingtonians-personal-health-data-and-privacy

Perez-Pozuelo, I., Spathis, D., Gifford-Moore, J., Morley, J., and Cowls, J. (2021). Digital phenotyping and sensitive health data: Implications for data governance. Journal of the American Medical Informatics Association, 28(9), 2002-2008.

Reilly, B. and MacDonald, A. (2023a). Washington's My Health, My Data Act: What to know (and what may surprise you). Retrieved from: https://www.manatt.com/insights/newsletters/client-alert/washingtons-my-health-my-data-act-what-to-know

Reilly, B. and MacDonald, A. (2023b). Connecticut's new law on consumer health data is now in effect. Retrieved from: https://www.manatt.com/insights/newsletters/privacy-and-data-security/connecticuts-new-law-on-consumer-health-data-is-n

Rosenbaum, S. (2010), Data governance and stewardship: Designing data stewardship entities and advancing data access. Health Services Research, 45(5p2), 1442-1455.

Roski, J., Bo-Linn, G. W., and Andrews, T. A. (2014). Creating value in health care through big data: opportunities and policy implications. Health affairs, 33(7), 1115-1122.

Sharon, T. and Lucivero, F. (2019). Introduction to the special theme: The expansion of the health data ecosystem–Rethinking data ethics and governance. Big Data & Society, 6(2), 2053951719852969.

Tempini, N. (2017). Till data do us part: Understanding data-based value creation in data-intensive infrastructures. Information and Organization, 27(4), 191-210.

United States Department of Health and Human Services (USHSS). (2022). Summary of the HIPAA Privacy Rule. Retrieved from https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html

United States Executive Order no. 14110 (30 October, 2023). Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

United States, White House Office of Science and Technology Policy. (2022, October). Blueprint for an AI Bill of Rights. https://www.whitehouse.gov/ostp/ai-bill-of-rights/

Vesselkov, A., Hämmäinen, H., and Töyli, J. (2019). Design and governance of mHealth data sharing. Communications of the Association for Information Systems, 45(1), 18.

Vial, G. (2023). Data governance and digital innovation: a translational account of practitioner issues for IS research. Information and Organization, 33(1), 100450.

Warzel, C. (2019, Aug. 13). All your data is health data: And Big Tech has it all. The New York Times: https://www.nytimes.com/2019/08/13/opinion/health-data.html

Winter, J. and Davidson, E. (2019a). Big data governance of personal health information and challenges to contextual integrity. The Information Society. 35(1): 36-51.

Winter, J. and Davidson, E. (2019b). Governance of artificial intelligence and personal health information. Digital Policy, Regulation and Governance, 21(3), 280-290.

Winter, J. and Davidson, E. (2022). Harmonizing regulatory regimes for the governance of patient-generated health data. Telecommunications Policy, 46 (1), 102285.

Wolff, J.L., DesRoches, C.M., Amjad, H., Burgdorf, J.G., Caffrey, M., et al. (2023). Catalyzing dementia care through the learning health system and consumer health information technology. Alzheimer and Dementia, 19(5), pp.2197-2207.

Yaeger, K. A., Martini, M., Yaniv, G., Oermann, E. K., and Costa, A. B. (2019). United States regulatory approval of medical devices and software applications enhanced by artificial intelligence. Health Policy and Technology, 8(2), 192-197.

Yao, Y. and Yang, F., 2023. Overcoming personal information protection challenges involving real-world data to support public health efforts in China. Frontiers in Public Health, 11: 1265050.