Human-Sensing Platforms and Ethical Considerations Throughout Their Data Life-Cycles

Sung Une Lee, Hye-young Paik, and Salil S. Kanhere

Abstract

Human-sensing platforms, encompassing digitized biometrics, activity sensing, and virtual human modeling, play a pivotal role in leveraging human-related data within the digital realm. With the swift progress of AI techniques, the promising opportunities brought about by these platforms are counterbalanced by the ethical challenges they pose. In considering these challenges, it is necessary to comprehend the characteristics of these platforms from data-centric viewpoints to understand their data management practices throughout their life-cycles. In this study, we first provide a comprehensive overview of various human-sensing platforms, highlighting the range of techniques utilized and the corresponding issues at each stage of the data lifecycle. Based on existing and emerging standards and frameworks on ethics in AI, we show the ethical issues and challenges in each phase in these data life-cycles. By providing this holistic perspective, our research contributes to the ongoing dialogue surrounding responsible data practices in the realm of human-sensing platforms.

Introduction

The decades of advancement in the Internet technology have created today's hyper-connected world. This underpins the current on-going discussion about the systems and infrastructures that enable the integration of the digital and physical worlds in which humans operate.

Human-sensing refers to the process of using various technologies to detect, gather, or analyze information related to human activities, behaviors, or characteristics. This can include data collected from sensors, cameras, social media, wearable devices, and other sources to gain insights into how humans interact with their environment, each other, and technology.

Human-sensing has already applied in many case studies and applications such as healthcare, urban planning, social sciences, and more. They can be used to study patterns of movement in public spaces, monitor patient well-being in medical settings, analyse consumer behaviour in stores, enhance security through surveillance. In Singapore, for instance, the government has implemented a sophisticated human-sensing platform as part of its efforts to create a smart city. The platform combines data from various sensors, cameras, and other sources to make real-time decisions to manage traffic flow (e.g., traffic lights can be dynamically adjusted based on traffic conditions to optimize traffic flow).

However, the data in these platforms are inherently more personal and privacy sensitive. With the increasing emphasis on automated data collection, analysis and decision making at scale (e.g., using AI and deep learning), there are heightened concerns around their impact on individuals and society at large.

Take applying machine learning models to assess or classify a person's risk for certain diseases based on daily activities collected. Besides the legalistic challenge of potential privacy violations, there are still openly debated questions such as who is accountable for the model's decisions, how do you achieve transparency or explain the models' decision making process. These questions are now considered under "ethics of AI" which, according to a global analysis of 36 sets of AI ethics principles done by [1], includes human-oriented values such as privacy, safety and security, fairness, accountability and transparency.

The effectiveness of an AI model hinges upon the quality of the data used for training and evaluation. Notably, the utilization of human-sensed data has introduced significant concerns, encompassing privacy, fairness, bias, and ethical ramifications. Establishing trust in AI systems demands a comprehensive grasp of potential risks and their management across the data lifecycle.

Despite its significance, there remains a noticeable scarcity of studies and research dedicated specifically to AI risks from the perspective of data management. While comprehensive frameworks and guidelines exist for responsible AI, encompassing aspects like algorithmic transparency and accountability, the specific considerations and practices associated with data warrant deeper investigation and attention.

We aim to shed light on the ethical dimensions and AI risks within human-sensing platforms, and provide guidance for practitioners, policymakers, and researchers in effectively managing the ethical challenges and mitigating AI risks within these platforms.

Digital Object Identifier: 10.1109/IOTM.001.2300148

Sung Une Lee is with CSIRO's Data61, Australia; Hye-young Paik and Salil S. Kanhere are with the University of New South Wales, Australia.

2576-3180/24/$25.00 © 2024 IEEE

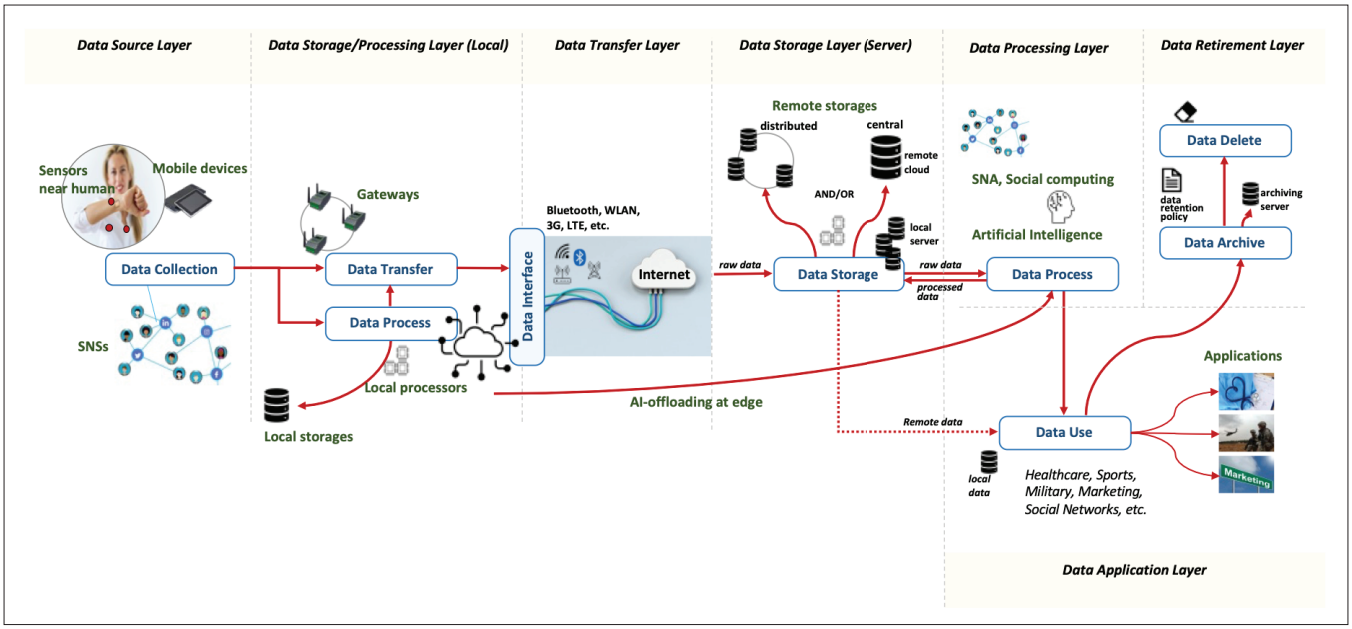

FIGURE 1. A generic architecture of human sensing platforms.

The key contributions of this work are as follows:

- We identify and categorise existing human-sensing platforms into three distinct categories. This effort allows us to delve deeper into their unique features and functionalities, providing a more nuanced comprehension. Our approach hones in on the data dimension, shedding light on the employed techniques as well as the associated challenges at various points in the data lifecycle.
- We present an analysis of key factors and potential ethical risks associated with data management practices of human-sensing platforms, according to five prominent AI risk frameworks and with overarching AI ethics principles [2].

Categorization of Human Sensing Platforms

We first present a categorisation of the systems that we collectively refer to as "human sensing platforms". By using the term "human sensing", we convey the idea of placing the data collected *from humans and about humans* at the centre of concerns for these platforms (Fig. 1). We use the term "sensing" loosely to mean collecting the data originating from the physical objects and digitising it for storage and processing. Sensing may be performed implicitly (e.g., via monitoring) or explicitly (e.g., submitting a form).

There is a newly emerging area commonly refer to as "Internet of Senses" [3]. It is an extension of the Internet of Things (IoT) where the emphasis is on incorporating sensory experiences into digital interactions. It envisions a future where not only data and information are transmitted and exchanged over the internet, but also sensory perceptions such as touch, taste, and smell. This could involve technologies like advanced virtual/augmented reality, haptic feedback devices. The goal is to create a more immersive and multisensory digital experience, enhancing human-computer interactions and expanding the ways we engage with the digital world. Human Sensing Platform focuses on using sensors to monitor and gather *data about human physiological and behavioral parameters*. While there could be some overlap in technologies and applications, the key distinction lies in their primary goals and areas of emphasis. For this reason, we do not consider Internet of Senses.

We categorise the platforms into three types: sensing of bodies, sensing of behaviour and sensing of identity as summarised in Table 1.

Sensing of Bodies

In this category, data is sourced from human bodies equipped with devices that can be small enough to be ingested, implanted, or mounted to the body. These "connected bodies" then exchange data with the platform which can remotely monitor and control the devices. Typical examples of these networks are termed Body Area Network (BAN)} or Body Sensor Network (BSN), and more recently Wireless Body Area Network (WBAN) and Internet of Bodies (IOB). Short-range wireless technologies such as Radio frequency (RF), ZigBee, WiFi and Bluetooth are used to interconnect tiny nodes (small and low power devices) with sensor or actuator capabilities [4]. Typical applications include sports training (e.g., measuring performance), health and wellness monitoring (e.g., heart rate monitors, fall detection).

Sensing of Behaviour

In this category, the data is primarily sourced from interaction activities (e.g., posting, liking) amongst users (e.g., social networks) as well as interactions between users and software systems (e.g., user activity logs of a mobile application). Typical examples are *Internet of People* and *Internet of Behaviour*. In the Internet of People, humans, using their mobile devices as proxy, become active elements of the network nodes. This type of networks can be spontaneous in self-organizing to form social communities based on devices in physical proximity. The Internet of Behaviour (IoB) is a comprehensive framework that leverages sensors and various technologies to observe,

Category	Platform	Definition	Application
Sensing of bodies	Internet of Bodies	A technology to connect the human body to a network through devices for sensing human data.	Healthcare, military
	Body Area Network	A wireless network and technology for sensing human data from wireless/ wearable sensors and personal devices.	Healthcare, sports
Sensing of behavior	Internet of People	An interconnected network composed of "human" nodes (digitized human and the generated data): for example, opportunistic mobile social networks.	Healthcare, social networks
	Internet of Behaviors	The extension of IoT to gather a collection of usage and data derived from human activities.	Healthcare, criminology, marketing, social networks.
Sensing of identity	Virtual/3D	A (autonomous) digital copy of a human based on their historical/current behavior data and persona.	Immersive 3D applications (social networks, education, games)

TABLE 1. Three types of human-sensing platforms.

assess, and anticipate human actions. IoB integrates AI/ML, extensive data analysis with sensors predicting multifaceted perspective on human behavior. For instance, [5] used IoB combined with Explainable AI (XAI) techniques to build a process to influence user behaviour in electrical power consumption.

Sensing of Identity

This category is relevant to the user's identity or persona (e.g., Metaverse Avatar) for the digital world. An important distinction of this category from other categories is that the sensing of "persona" could lead to the creation of a real-time counterpart of a user in the virtual world who can eventually learn and act autonomously on behalf of the physical counterpart.

Not dissimilar to *Digital Twins* where physical objects or processes are virtualised for running simulations, applications in this category would build sophisticated models to create a virtual user, based on the real-world counterpart's personal information collected from a number of acquisition routes. The most notable application of this category is in virtual games and virtual spaces (e.g., 3D learning environments). However, the scope and applications of this category will continue expand to as the popularity of the immersive 3D platforms increases.

A Data Lifecycle View of Human Sensing Platforms

This section discusses key techniques and models used by these platforms, organised by their data lifecycle phases. The data lifecycle used in this article is inspired by the international data governance standard (ISO/IEC 38505-1) [6]. The lifecycle comprises of five phases (create, store, process, archive and delete) with a set of activities. We split the "process" into "process" and "use" to detail each phase.

Phase: Create – Data collection Methods

This phase is concerned with data collection.

In human sensing networks, data is gathered from various devices, entities, and activities. This entails diverse management and processes in subsequent phases.

In device-based sensing, IoT sensors in the network measure situational events in physical objects and spaces (machine-sensored/generated data). Prior to reaching the platform, data might undergo pre-processing via personal data processing unit (PDPU). Single/multiple PDPUs filter out redundant data for energy efficiency, network bandwidth, or encrypting data for security and privacy.

In device-free sensing, data is collected via a non-intrusive sensing approach where sensing utilises an existing infrastructure such as camera, WiFi signals, access points or radar [7]. This technology primarily captures human motions and gestures, offering opportunities like cost savings and potential for widespread human-interaction services. However, challenges such as intricate human motions, large-scale technique application, and security/privacy concerns also arise.

Sensing via Social Networking Service (SNS) has emerged as a vital data collection method, particularly for networks like Internet of People or Internet of Behaviour. An individual, acting as an independent sensing node, contributes personal details, media-rich content, preferences, and search keywords. This firsthand data from human subjects enhances the accuracy of predicting their intentions, habits, or emotions. Yet, there can be privacy issues such as personal information exposure without consent and unauthorized data access (e.g., Cambridge-Analytica scandal[1]).

Another data source is *Business processes*, overseeing activities based on predefined workflows (e.g., a new hospital patient's arrival). The network gathers process-mediated data at each step (e.g., patient processing details). This data reveals how "things" and people interact, aiding better integration [8]. The collected data can also facilitate predictive modeling for improved process monitoring by identifying anomalies or critical errors early on [9].

Typically, platforms for sensing of bodies rely on machine-generated data. Yet, others collect data from human or human's activities (Table 2). These human-originated data frequently raise security, privacy, and other data-related concerns, especially personal information's vulnerability to breaches. Constraints, continuous review, and pre-collection steps like categorization and use case definition are essential risk-mitigation strategies.

Phase: Store – Data Storage Models

The choice of suitable storage models can hinge on factors like distributing data control authority among stakeholders and addressing data security and privacy demands.

Data Architecture Models: Conventional data storage can be classified according to the architectural style.

[1] https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge_Analytica_data_scandal

Phase	Category	Technique/model	Type of human sensing platform		
			Sensing of bodies	Sensing of behavior	Sensing of identity
Create	Data source	Machine-generated	O	–	–
		Human-sourced	–	O	O
		Process-mediated	–	O	–
Store	Architecture-based	Centralized model	O	O	O
		Distributed model	O	O	O
		Hybrid model	O	O	O
	Location-based	Cloud model	O	O	O
		Fog model	O	O	–
		Edge model	O	O	–
Process/Use	Analytical solution	Social network analysis	–	O	O
		Social computing	–	O	O
		Smart algorithms	O	O	O
Archive/Delete	Retirement	Data deletion	O	O	O
		Data invisible (e.g., de-indexing)	O	O	O
		Self-sovereign identities	O	O	O

* O *considered*, – *not considered*

TABLE 2. Techniques and models for human-sensing platforms.

Rising privacy concerns include entities avoiding centralized server data storage, prompting exploration of solutions like federated/split learning for AI/ML

Centralized model is designed to collect and transmit data to a remote server. For example, smart sensors near human-body detect blood pressure and transfer the data to the central servers. This model allows for complete data control, maintaining high quality and security (Fig. 2). This, however, comes with inherent risks such as single-point failure (SPOF), communication overheads, bottlenecks, and potential privacy breaches which can undermine business continuity and compromise security. SPOF analysis, load balancing, data flow analysis, and tackling bottlenecks can be mitigation solutions. Rising privacy concerns include entities avoiding centralized server data storage, prompting exploration of solutions like federated/split learning for AI/ML.

Distributed model serves as an alternative approach to enhance data security and accessibility. It involves storing datasets across multiple locations. Blockchain, built on a peer-to-peer network, exemplifies this concept [10]. Self-sovereignty also entails storing personal data on the user's side, granting individuals control and enabling selective data sharing. Yet, handling user behavior and adapting platform goals within this model's decision-making can be challenging, complicated by complex and slow procedures.

The *hybrid model* employs a versatile architecture that combines centralized and distributed storage approaches. For instance, raw data is stored on a remote server accessible solely to authorized users, while metadata about the data is recorded in distributed storage, providing transparent access for all users. This model involves added resources, costs, and energy for maintaining both central and distributed storage, along with the need for extra governance to ensure data consistency, integrity, and handle diverse environments.

Location-Based Data Models. IoT services encompass layers for data collection, processing, and distribution, emphasizing performance, cost reduction, and energy conservation [11].

Cloud, Fog, and Edge are utilized to address these challenges (Fig. 2). *Cloud* is to store data on the remote servers and enable data services. While public cloud is operated in a pay-as-you-go manner, private cloud is an internal data centre of an organisation. There are also hybrid and community cloud. Cloud provides improved usability, accessibility, security, cost-efficiency, and easy data sharing, yet its centralized nature can lead to inefficient data transfer and processing due to exponential growth of data from multiple sources [12].

Fog, introduced by Cisco in 2012, bridges the gap between cloud and edge devices, featuring processing power and temporary storage. It receives abundant data from edge devices, retaining essential information to reduce network traffic, storage usage, and computational resources for the cloud. Despite these benefits, the added layer makes the system more complex, requiring extra resources and costs, and it's less scalable than the cloud.

Edge technology resembles a localized version of *Fog* and *Cloud* [12], operating on a distributed data model with data generated and stored at the network's edge. Connected to nearby sensors, it reduces latency and network load by transmitting data to remote servers when necessary. However, it demands increased storage and energy, while heightened security risks are prominent at the edge. Due to these factors, it's less favored for platforms involving human identity sensing (Table 2).

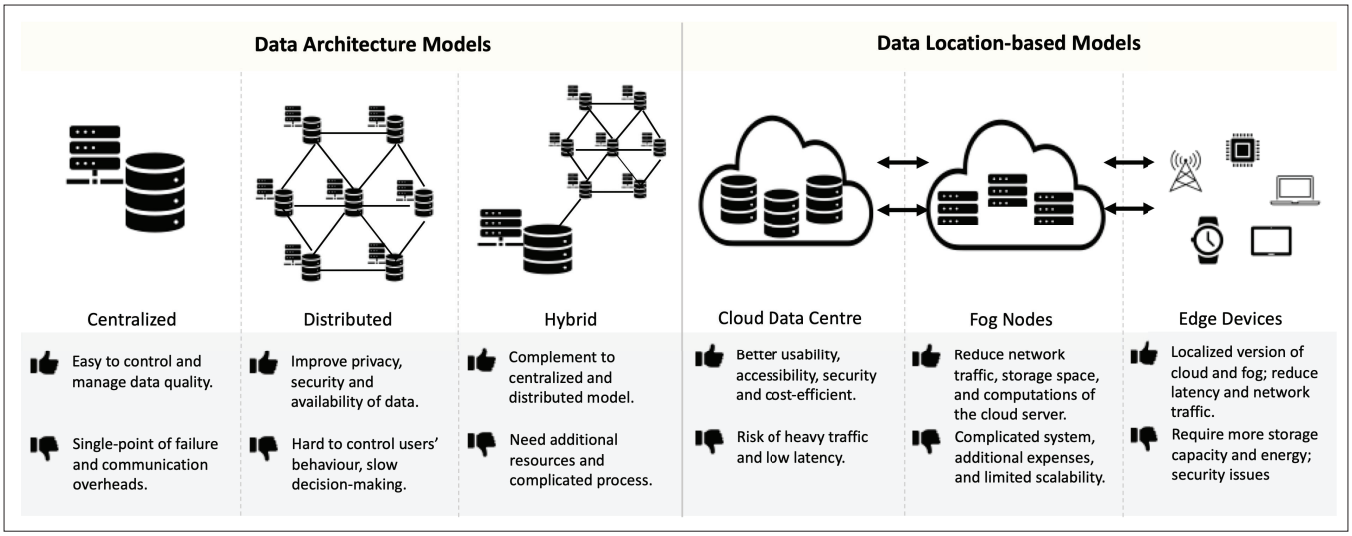

FIGURE 2. Comparison of data storage models.

Phase: Process – Data Processing Techniques

In this phase, data undergoes analysis, transformation, and aggregation to generate new insights and derivations. Artificial intelligence/machine learning (AI/ML) is widely used to gain further insights, predict human behaviour and make automated decisions. Advanced technical methods for building data analytics are being actively developed for various applications. In [13] the following three areas of data analysis techniques on human sensing data are proposed.

Social Network Analysis (SNA). This interdisciplinary analysis technique could utilise classical SNA methods such as analysing node centrality (i.e., the number of directly connected friends), predicting links (i.e., closeness of a relationship). It is often used to understand how humans and objects interact with each other.

Social Computing. This paradigm aimed to use computational methods to understand social systems and human behaviour, even influencing certain behaviours. However, a decade later, the data types and volumes from human sensing platforms surpass its original scope. As human nodes establish relationships within the platform, the shared information becomes a basis for a social computing model, capable of addressing diverse societal challenges, like the spread of misinformation.

"Smart" Algorithms. The advancements in AI and deep learning encompasses a wide range of data processing methods. This spans personalized services, fine-tuned recommendation models, service automation, and the creation of just-in-time services. Smart algorithms are commonly used as a analytical tool by all types of platforms, as presented in Table 2.

Furthermore, *AI-offloading* analyzes and processes incoming data at the edge, deciding whether to use more powerful cloud algorithms or local processing. This optimizes edge computing, addressing privacy concerns, but it's constrained by edge devices' limited computational power and memory.

Phase: Use – Primary/Secondary Use of Data

Ensuring optimal data utilization is a critical goal for human sensing platforms, aligning with the *purpose limitation* principle, which involves considering two data usage models as follows.

Primary use of data is designated for the explicit purposes of the platform's services and operational tasks. For instance, a sports management platform gathers physiological data to monitor well-being, issue alarms, warnings, and perform other functions. A comprehensive roster of these use cases needs to be established and communicated to stakeholders.

Secondary use of data extends beyond initial collection purposes including research and personalization, marketing. Instances of misuse without proper consent are not rare, prompting recent regulations like the General Data Protection Regulation (GDPR) to emphasize ethical secondary use and require mechanisms for data subject protection. To manage human data, including personal and sensitive information, metadata plays a key role in governance and protection, especially in the context of AI.

For example, *Model card* serves as metadata to ensure secure and transparent data usage, covering both primary and secondary purposes. It provides essential details about AI models, their applications, and training data, including datasets used and reasons for their choice.

Datasheets for Datasets also aid AI/ML engineers in identifying ethical issues within training and evaluation data. Gebru *et al.* [14] introduced over 50 questions to extract dataset information, enhancing transparency, accountability, and reducing data biases in AI.

Phase: Archive – Data Retirement

This phase involves storing data at the end of its lifecycle. Archived data is retained in storage for a designated period before permanent deletion. Archived data can be repurposed when users require it or based on data governance decisions, transitioning from inactive to active mode. The term "data retirement" underscores the concept of removing data from active use.

A common practice for retiring data is to do a "logical delete" where the data is flagged as deleted, making them inactive and unavailable for regular operations. It is a reversible method suitable

for situations where data might be required for audit or legal purposes. Data could be marked with expiry dates and be associated with an automated process that handles the retirement process based on predefined policies. Before retiring data, sensitive or personally identifiable information (PII) can be obfuscated or anonymized.

Ethical use of data also concerns if and how retiring data is considered in the platforms when the data when it is no longer actively used. Numerous human-sensing platforms often centralize personal data. In this context, comprehending the regulatory landscape as the primary data custodian stands as a fundamental aspect of data management. Designing the data retirement process should consequently be informed by this understanding.

Efficiently managing storage for archived data is another consideration. Many platforms accumulate extensive data, incurring significant costs over time. Employing a *Tiered Storage Model*, such as Apache Pulsar, can yield cost savings. Archival data can be moved to a more economical archive tier, offering comparable user accessibility with flexible latency requirements.

Platforms utilizing distributed storage, like blockchain-based systems, may encounter challenges with data retirement. Data retirement on blockchains is essentially impractical due to their immutable nature. Attempts have been made to breach blockchain immutability. In the upcoming section, we introduce approaches for archiving/deleting data from blockchain networks.

Personal Online Data Stores (POD) empower individuals to securely store their data in decentralized storage. Data owners can control access (grant/revoke) to any segment of their POD-stored data, eliminating the need to store data elsewhere for sharing with others.

Phase: Delete – Right to Be Forgotten

This phase involves purging unused or unnecessary data from storage. Data disposal is automatically executed following data retention policies or carried out manually upon the data owner's request. GDPR enforces the right to be forgotten regarding personal data.

Data deletion in a centralized model is straightforward. Yet, introducing AI/ML models to a platform poses challenges when attempting to remove specific training data. In some cases, retraining the model from scratch may be necessary to ensure data removal and mitigate potential retention. *Machine unlearning* offers a potential solution. This mechanism erases a training data sample by reversing its effects on extracted features and models. Cao and Yang [15] proposed an unlearning approach wherein learning algorithms rely only on summations, not individual data.

A decentralized model also presents potential data deletion issues. Blockchains, by design, are immutable, making natural data removal unfeasible. Several potential solutions exist. For instance, *selective data deletion* extends the blockchain with a summary block, summarizing prior chain data in a new block while omitting undesirable information. An alternative is the *de-indexing* approach, akin to hiding a web page from search results or introducing *new transactions* on blockchain networks [10]. Another strategy involves storing sensitive data externally from blockchain networks, ensuring data invisibility and facilitating easy deletion, while upholding blockchain integrity and consistency.

Data Management Gaps and Ethical Concerns

This section presents the results of our analysis of ethical concerns associated with these platforms. Taking the well-known AI ethics frameworks, and the lifecycle phases presented, we examine what ethical questions apply in managing human data in the these platforms.

It is noted that the finding revolves around a shared framework of ethical considerations applicable to all human-sensing applications. Further to these common concerns, other application specific issues can be discussed along with application-specific mitigation.

Besides the challenges arise from the fact that human-sensing platforms encompass a wide array of technologies, the deployment of human-sensing platforms raises significant ethical considerations, particularly in the context of AI integration. We prioritize ethical data practices, identifying risks in platforms, with a focus on AI ethics in data management, given AI's widespread integration in the data lifecycle.

AI ethics principles are designed to ensure AI is safe, secure and reliable. For example, Australia AI Ethics Principles suggests eight principles such as *Human, social and environmental wellbeing*, *Human-centred values*, *Fairness*, *Privacy and security*}, *Reliability and safety*, *Transparency and explainability*, *Contestability*, and *Accountability* [2]. While all the principles pertain to the human-sensing platform, we focus more on three principles: *Human, social and environmental wellbeing*, *Human-centred values* and *Privacy and security* as they bear a more direct relevance.

Human wellbeing highlights that AI systems should benefit individuals, society, environment, and respect human rights. This means that human-sensing platforms should ensure secure use of data and prevent abuse/misuse of data which negatively impact on human.

Human-centred values means that it is crucial to consider comprehensive rights for data owners at all times, while also providing users with transparent information regarding the utilization of their data within the systems.

Privacy and security emphasizes that data governance should be designed and implemented complying relevant laws and regulations. It also requires that data processes should be monitored and assessed throughout the data lifecycle. There are varying data requirements depending on the context and architecture in use, such as local (centralized) learning and federated learning. Federated learning is particularly crafted to protect the privacy of sensitive data by keeping it localized, enabling access to data spread across multiple locations while sharing solely encrypted model updates. However, this approach frequently faces transparency challenges, complicating the understanding of its decision-making processes.

To gain deeper insights into ethical concerns and risks, we conducted an investigation of several globally renowned or commonly used AI risk frameworks, ultimately selecting five frameworks: EU Assessment List for Trustworthy Artificial Intelli-

Ethical use of data also concerns if and how retiring data is considered in the platforms when the data when it is no longer actively used.

Encompassing all stages of the lifecycle, conducting continuous audits to ensure the ethical use of data within human-sensing platforms is essential.

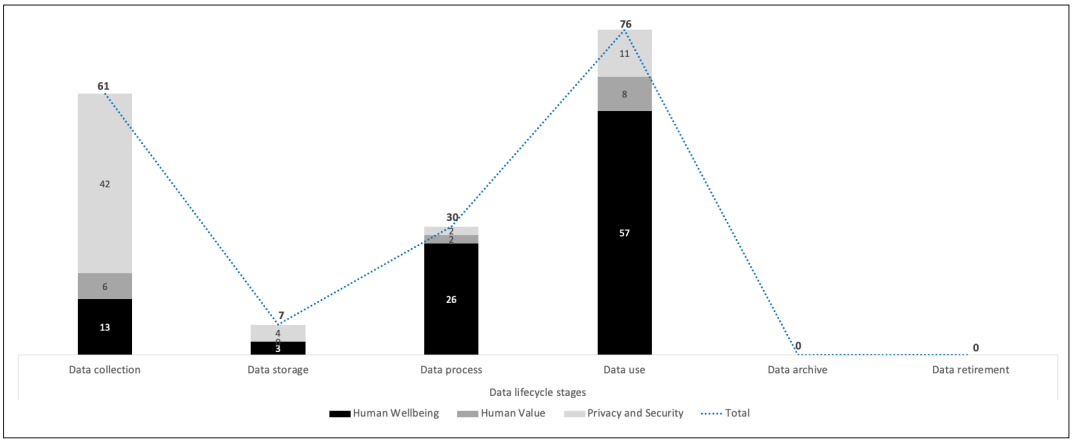

FIGURE 3. Result of the investigation on the five AI risk frameworks: distribution of the data risk questions; *there are no risk questions considered for the "data archive" and "data delete" stages in the existing AI frameworks.*

gence,[2] Canada Algorithmic Impact Assessment,[3] Australia NSW AI Assurance Framework,[4] Microsoft Responsible AI Impact Assessment,[5] and NIST AI Risk Management Framework.[6] We collected a total of 380 risk questions from the frameworks, out of which 174 questions were finally selected for the three principles. We identified critical data risks based on the questions, followed by the analysis and categorization of the data according to the data lifecycle.

Figure 3 demonstrates how the risk questions are classified according to the three AI principles and the stages of the data lifecycle. It provides an overview for evaluating the potential data risks associated with AI-powered human-sensing systems. It also highlights that the majority of data risks are concentrated in the data collection and data use stages, while the data archive and delete phases are noticeably overlooked, lacking proper consideration.

As our analysis has revealed instances where existing AI risk frameworks lack coverage, we have supplemented these frameworks by identifying and incorporating pertinent data practices, effectively addressing the gaps (Table 3).

It becomes evident that identifying the intended use of data and continuously evaluating and monitoring its usage throughout the lifecycle play a pivotal role in promoting human wellbeing. These measures ensure that human-sensing platforms prioritize and align with ethical standards for data usage. During the data archive and delete stages, the primary considerations revolve around the ethical utilization of archived data and confirming the permanent erasure of user data from AI models. Adhering to ethical standards and guidelines facilitates these processes, enhancing the trustworthiness of AI systems and fostering user confidence.

There is comparatively less emphasis on the inclusion or prioritization of human values. The five frameworks primarily focus on assessing the effects of data on individuals, including those engaged in decision-making processes, with a primary goal of safeguarding human rights and mitigating potential data risks such as overconfidence or excessive reliance on the data.

The gaps in coverage (storage, archive, and delete) have been supplemented by referring to the GDPR, which extensively covers sensitive human data and user rights; the standard is closely aligned with the human-centred value principle.

The GDPR introduces a range of data subject rights (article 13-22) that encompass various data lifecycle stages. These rights encompass access, rectification, and consent for effective data storage management, informed consent for proper data retirement procedures, as well as user autonomy and control, including the right to erasure for lawful data deletion.

In terms of privacy and security, the utilization of personal information or sensitive data (e.g., details about children, criminal records) emerges as a paramount concern. This underscores the significance of proficiently handling and protecting human data to uphold privacy rights and ensure data confidentiality. Furthermore, data retention policies play a crucial role as a mechanism for managing archived data in a lawful manner. This should be followed by a thorough compliance check with internal and external policies, regulations, and standards.

Encompassing all stages of the lifecycle, conducting continuous audits to ensure the ethical use of data within human-sensing platforms is essential. Auditing AI-powered human-sensing systems should involve well-documented audit trails and comprehensive logging to enhance system transparency and provide insights into the collection, processing, storage, modification, and removal of human data. The audit process should vary based on the potential risks associated with the AI system; organizations with lower risks may conduct self-audits, while external independent parties can be involved in high-risk cases.

In the context of the EU AI Act,[7] a "high-risk AI system" is defined as one that poses significant risks to individuals' health, safety, or fundamental rights. For instance, Healthcare Diagnostics systems that gather data from blood pressure monitor sensors and utilize the data for medical diagnoses fall into this category. Given their potential to yield inaccurate or misleading recommendations that impact patient care, such systems are categorized as high-risk. Consequently, they are subject to stringent requirements for responsible and ethical operation to safeguard both patient well-being and fundamental rights.

Conclusions

Through our investigation, we have gained valuable insights into the ethical considerations and AI risks inherent in human-sensing platforms. Our

[2] https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

[3] https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html

[4] https://www.digital.nsw.gov.au/policy/artificial-intelligence/nsw-artificial-intelligence-assurance-framework

[5] https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2022/06/Microsoft-RAI-Impact-Assessment-Guide.pdf

[6] https://www.nist.gov/itl/ai-risk-management-framework.

[7] https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Principle	Collection	Storage	Process	Use	Archive	Delete
HW	Identification of intended and beneficial uses of data	Prevention of data misuse and single point of failure	Benefits of stakeholders from the data	Evaluation and monitoring of the use of data	***Ethical use of archived data avoiding potential harm human**	***Confirmation of permanent erasure of user data from the AI models**
HV	Assessment of impacts on human values in collecting data	***User right to access and rectification and consent for data storage**	Human control or involvement	Identification of data risks (e.g., overconfidence or overreliance on the data)	***User right to be informed and consent for data retirement**	***User autonomy and control: right to erase**
PS	Inclusion of sensitive data and regulations	Consideration of privacy/security by design	Measurement of data quality (privacy/security)	Verification/control of the use of sensitive data	***Implementation of clear and well-defined data retention policies**	***Compliance check with policies, regulations and standards**
***Implementation of continuous (self-/independent-) audits throughout all lifecycle stages**						
HW: Human wellbeing, HV: Human-centered values, PS: Privacy and security						

TABLE 3. Key data concerns and risks at different data lifecycle stages; *the highlighted portions (*) represent the gaps filled by this study as they have not been adequately covered by the existing AI frameworks.*

analysis has revealed the significance of data governance, underscoring the need for heightened ethical considerations within this research field. Furthermore, we have identified notable gaps in existing AI risk frameworks, particularly concerning the ethical treatment of data archiving, deletion, and the incorporation of human values. These findings emphasize the importance of adopting a holistic approach to address the ethical challenges and AI risks associated with human-sensing platforms. Moving forward, we plan to develop an appropriate assessment framework for evaluating the ethical use of human-sensing platforms. Additionally, we aim to conduct empirical evaluations through case studies to further validate and expand upon our research findings.

References

[1] J. Fjeld *et al.*, "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI," Research Publication No. 2020-1, Berkman Klein Center for Internet & Society at Harvard University, 2020.

[2] DISER, 2019, Australia's Ai Ethics Principles; https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework, (visited on 2023-08-15).

[3] D. Panagiotakopoulos *et al.*, "Digital Scent Technology: Toward the internet of Senses and the Metaverse," *IT Professional*, vol. 24, no. 3, 2022, pp. 52–59.

[4] A. Celik, K. N. Salama, and A. M. Eltawil, "The Internet of Bodies: A Systematic Survey on Propagation Characterization and Channel Modeling," *IEEE Internet of Things J.*, vol. 9, no. 1, 2021, pp. 321–45.

[5] H. Elayan *et al.*, "Internet of Behavior and Explainable AI Systems for Influencing IoT Behavior," *IEEE Network*, vol. 37, no. 1, 2023, pp. 62–68.

[6] ISO/IEC, "Iso/iec 38505-1:2017 Information Technology — Governance of It — Governance of Data," ISO/IEC, 2017.

[7] J. Xiao *et al.*, "A Survey on Wireless Device-Free Human Sensing: Application Scenarios, Current Solutions, and Open Issues," *ACM Comput. Surv.*, vol. 55, no. 5, Dec. 2022.

[8] C. Janiesch *et al.*, "The Internet of Things Meets Business Process Management: A Manifesto," *IEEE Systems, Man, and Cybernetics Mag.*, vol. 6, no. 4, 2020, pp. 34–44.

[9] A. Gal, A. Senderovich, and M. Weidlich, "Online Temporal Analysis of Complex Systems Using IoT Data Sensing," *2018 IEEE 34th Int'l. Conf. Data Engineering (ICDE)*, 2018, pp. 1727–30.

[10] H.-Y. Paik *et al.*, "Analysis of Data Management in Blockchain-Based Systems: From Architecture to Governance," *IEEE Access*, vol. 7, 2019, pp. 186,091–186,107.

[11] M. Yaghoubi, K. Ahmed, and Y. Miao, "Wireless Body Area Network (WBAN): A Survey on Architecture, Technologies, energy consumption, and security challenges," *J. Sensor and Actuator Networks*, vol. 11, no. 4, p. 67, 2022.

[12] A. Al-Qamash *et al.*, "Cloud, Fog, and Edge Computing: A Software Engineering Perspective," *2018 Int'l. Conf. Computer and Applications (ICCA)*, 2018, pp. 276–84.

[13] F. Shi *et al.*, "The Internet of People: A Survey and Tutorial," ArXiv, vol. abs/2104.04079, 2021.

[14] T. Gebru *et al.*, "Datasheets for Datasets," *Communications of the ACM*, vol. 64, no. 12, 2021, pp. 86–92.

[15] Y. Cao and J. Yang, "Towards Making Systems Forget with Machine Unlearning," *2015 IEEE Symp. Security and Privacy*, 2015, pp. 463–80.

Biographies

Sung Une Lee [SM] (sunny.lee@data61.csiro.au) is a Research Scientist at CSIRO's Data61, Australia. Previously, she held positions as an Associate Lecturer/Sessional Academic and a Research Fellow at Queensland University of Technology and Deakin University. She obtained her Ph.D. from the University of New South Wales in 2019. Her research interests span Responsible AI, Software Engineering, and Data Governance. With over 10 years of industry experience in Software Engineering and Project Management, she brings a strong practical background to her academic work.

Hye-young (Helen) Paik [SM] (h.paik@unsw.edu.au) is an Associate Professor at School of Computer Science and Engineering, University of New South Wales (UNSW). She is also research staff in the Cybersecurity Cooperative Research Centre. Her broad research background and expertise come from areas such as middleware, distributed processes/application integration and Service-Oriented software architectures. Recently, Distributed Ledger Technology have become her main focus of research projects, looking into various security and privacy enabling technologies such as Decentralised Identity Management and privacy-preserving data analytics and platforms, privacy and security in cyber physical systems.

Salil S. Kanhere [SM] (salil.kanhere@unsw.edu.au) is a Professor with the School of Computer Science and Engineering, UNSW Sydney, Australia. He has held visiting appointments with I2R Singapore, TU Darmstadt, TU Graz, RWTH Aachen, and the University of Zurich. His research interests include Internet of Things, cybersecurity, distributed systems, pervasive computing, and applied machine learning. He received the Friedrich Wilhelm Bessel Research Award in 2020 and the Humboldt Research Fellowship in 2014 from the Alexander von Humboldt Foundation in Germany. He is the Editor-in-Chief of the *Ad Hoc Networks Journal* and an Associate Editor of the *IEEE Transactions on Network and Service Management*, *Pervasive and Mobile Computing*, and *Computer Communications*.