Good Intentions, Risky Inventions: A Method for Assessing the Risks and Benefits of AI in Mobile and Wearable Uses

MARIOS CONSTANTINIDES, Nokia Bell Labs, United Kingdom

EDYTA BOGUCKA, Nokia Bell Labs, United Kingdom

SANJA SCEPANOVIC, Nokia Bell Labs, United Kingdom
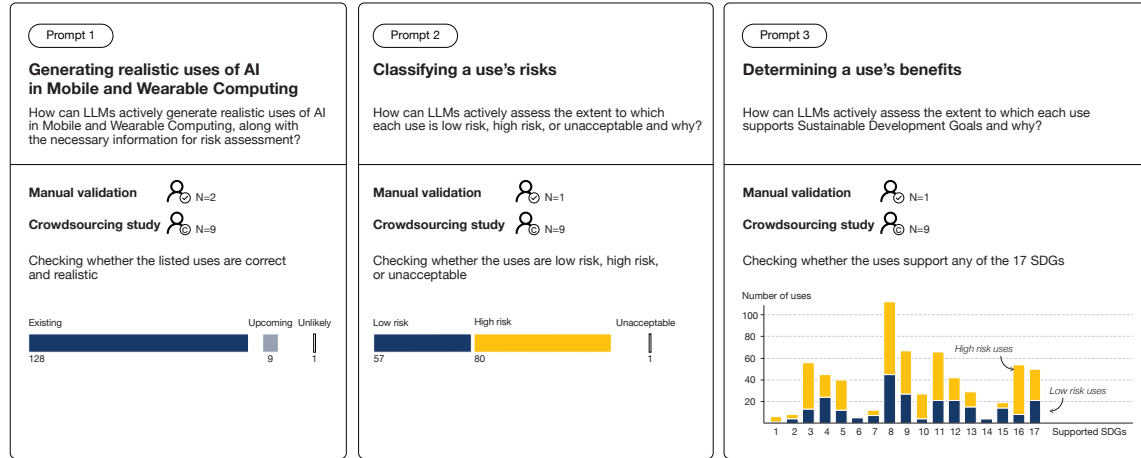
DANIELE QUERCIA, Nokia Bell Labs, United Kingdom

Fig. 1. Overview of our method. A three LLM-prompt pipeline for generating mobile and wearable uses of AI (prompt #1), classifying each use's risks according to the EU AI Act [23] (prompt #2), and determining whether each generated use is beneficial according to the UN's Sustainability Development Goals [114] (prompt #3). Out of 138 generated uses, as many as 80 were considered high risk according to the EU AI Act, primarily aligning with Sustainable Development Goals 3 (good health and well-being), 10 (reduced inequalities), and 16 (peace and justice). Our method was validated by two experts in mobile and wearable technologies, a legal and compliance expert, and a cohort of nine individuals with legal backgrounds who were recruited from Prolific, confirming its accuracy to be over 85%.

Integrating Artificial Intelligence (AI) into mobile and wearables offers numerous benefits at individual, societal, and environmental levels. Yet, it also spotlights concerns over emerging risks. Traditional assessments of risks and benefits have been sporadic, and often require costly expert analysis. We developed a semi-automatic method that leverages Large Language Models (LLMs) to identify AI uses in mobile and wearables, classify their risks based on the EU AI Act, and determine their benefits that align with globally

Authors' addresses: Marios Constantinides, Nokia Bell Labs, Cambridge, United Kingdom, marios.constantinides@nokia-bell-labs.com; Edyta Bogucka, Nokia Bell Labs, Cambridge, United Kingdom, edyta.bogucka@nokia-bell-labs.com; Sanja Scepanovic, Nokia Bell Labs, Cambridge, United Kingdom, sanja.scepanovic@nokia-bell-labs.com; Daniele Quercia, Nokia Bell Labs, Cambridge, United Kingdom, daniele.quercia@nokia-bell-labs.com.

recognized long-term sustainable development goals; a manual validation of our method by two experts in mobile and wearable technologies, a legal and compliance expert, and a cohort of nine individuals with legal backgrounds who were recruited from Prolific, confirmed its accuracy to be over 85%. We uncovered that specific applications of mobile computing hold significant potential in improving well-being, safety, and social equality. However, these promising uses are linked to risks involving sensitive data, vulnerable groups, and automated decision-making. To avoid rejecting these risky yet impactful mobile and wearable uses, we propose a risk assessment checklist for the Mobile HCI community.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing**; • **Computing methodologies** → *Artificial intelligence*.

Additional Key Words and Phrases: mobile, wearables, sustainable development goals, risk assessment, LLM, prompt engineering

1 INTRODUCTION

The integration of Artificial Intelligence (AI) into mobile and wearable devices has unlocked new capabilities, from improving workplace productivity to enhancing healthcare and education. Mobile assistants and wearables, equipped with various sensors such as accelerometers and heart rate monitors, have significantly increased workplace productivity [5, 19, 27, 76]. Smart wearable devices provide low-cost, objective measures of physical activity, mental health, and sleep quality assessments [25, 91, 127], contributing significantly to scalable behavior monitoring in large populations [92]. In sports medicine, wearable technology can detect patterns in physiological variables, assisting athletes in improving performance and preventing injuries [18]. In education, smartwatches can be used to collect real-time data during students' learning activities, improving the analytics and insights into learning processes [22].

Despite the potential advantages, any technology comes with risks. Mobile and wearable technology uses are no exception. Privacy concerns are particularly prominent, as AI integration can lead to data leakage and unauthorized surveillance [26, 32, 51]. However, recent advances in federated learning [69] and differential privacy [36] may enhance the data privacy of AI-based mobile and wearable uses, potentially transforming high-risk uses into low-risk ones. As these uses often involve automated decision-making, it is crucial to provide explanations, especially in high-stakes domains. For example, there is an increasing demand for explainable AI in healthcare to ensure trustworthy AI-based decisions [99]. Security issues are also critical. In addition to well-known challenges such as the potential to compromise sensitive data or use it to maliciously infer private information [30], wearable devices may also cause physically harm to the wearer [75]. For example, there have been reports of skin irritation and allergic reactions caused by the materials used (e.g., nickel in the metal components), or due to the accumulation of moisture under the device, which can lead to skin irritation or even infections [61].

Traditional assessments of risks and benefits have been sporadic, and often require costly expert analysis. In fact, risk identification requires comprehensive domain expertise [52]. According to the OECD, legal advisory costs required to ensure compliance with regulatory systems of different countries is estimated to be nearly 780 billions dollars per year worldwide [57, 81]. That is why traditional assessments of risks and benefits tend to be fragmented, focusing on specific aspects of mobile and wearable use (e.g., efficiency in risk assessment for work-related activities [124], or adoption based on privacy risk versus perceived benefits [64]), and come with technical and ethical challenges. On the technical side, mobile and wearables pose significant privacy risks [15], while, on the ethical side, these devices can potentially

augment (or replace) human intelligence in ways that might not always be desirable [123]. This dichotomy underscores the need for a more systematic approach to evaluate the benefits and risks associated with these technologies. Issues such as data privacy, the algorithms accuracy, and the potential for social and individual opportunities (outlined, for example, by the United Nation's Sustainability Development Goals report [114]) must be considered to ensure that the integration of AI truly benefits users without compromising their safety and privacy [93]. To address these challenges, we implemented a semi-automatic way of conducting a risks and benefits assessment of AI for mobile and wearable uses and, in so doing, we made two main contributions:

(1) We engineered and validated three Large Language Model (LLM) prompts capable of (§4): *(a)* generating a diverse range of realistic uses of AI for mobile and wearable computing; *(b)* classifying each generated use risks (whether it is unacceptable, high risk, or low risk) according to the EU AI Act [23]; and *(c)* determining the benefits associated with each generated use. For classifying risks, we chose the EU AI Act because it is the most advanced legal framework available at the time of writing, which specifically adopts a risk-based approach at a use level [23]. This means that stakeholders involved in the development of an AI technology are required to assess its risks for specific uses, rather than assessing the technology itself for broad risks. For determining the benefits, we chose the Sustainable Development Goals (SDGs). The SDGs are globally recognized and allow for the evaluation of the AI's long-term impacts related to five priority areas that concern people, the planet, prosperity, peace, and partnerships [114]. To achieve these tasks, we used a set of novel prompt elements, including a comprehensive list of *domains*, five *risk concepts* for describing uses, and detailed manual *evaluations* of prompts.

(2) Our method generated 138 uses, which we reviewed and thematically grouped them in two categories (§5): uses promoting SDGs and being low risk; and uses promoting SDGs yet being high risk. We found that specific applications of mobile computing hold significant potential in improving well-being, safety, and social equality. However, these promising uses are linked to risks involving sensitive data, vulnerable groups, and automated decision-making.

In light of these results, we discuss challenges in conducting semi-automatic assessments, and present a Risk Assessment Checklist for Mobile Computing uses as a practical solution for balancing risks and benefits trade-offs in the use of AI for mobile and wearable technologies (§6).

2 RELATED WORK AND BACKGROUND

Next, we surveyed previous literature that our work draws upon, and grouped it into three main areas: *i)* AI regulations and impact assessments (§2.1), *ii)* risk and benefits assessments of AI for mobile and wearable uses (§2.2), and *iii)* the use of LLMs to transition from manual to semi-automatic assessment through prompt engineering (§2.3).

2.1 AI Regulations and Impact Assessments

The increasing deployment of AI has prompted calls for regulatory oversight [10, 41, 110]. The US Office of Science and Technology Policy has released a non-binding AI Bill of Rights Blueprint, highlighting principles such as safety, non-discrimination, data privacy, and AI transparency. Similarly, the European Commission's AI Act, which is legally binding, aims to balance innovation with safeguarding societal values and rights. The Act classifies AI applications based on risk levels, from low to unacceptable, and prohibits AI applications that are harmful or manipulative [23, 24]. It emphasizes a risk-based regulatory framework, requiring comprehensive evaluations for high risk AI systems.

Typically, an impact assessment is conducted and a report is produced that documents the potential impacts of AI. CredoAI's report template includes sections on formal system evaluations and compliance with laws, regulations, and standards [105]. Similarly, the report developed by the National Institute of Standards and Technology (NIST) includes listings of potential biases, evaluations of the impact's magnitude and likelihood, technical specifications, third-party technologies, and legal and compliance issues [83]. More recently, a study has identified the five necessary components (i.e., domain, purpose, capability, AI user, and AI subject) to conduct regulatory risk assessment according to the EU AI Act [47]. The *domain* specifies the industry or sector, such as health. The *purpose* explains the goal, such as improving well-being. The *capability* describes the technology behind it, such as mood and stress level tracking. The *AI user* is the one using the system, such as mental health professionals. The *AI subject* is the one affected by the system, such as patients. Our study operationalizes these five components in a semi-automatic way using LLMs.

2.2 Risk and Benefit Assessments of AI for Mobile and Wearable Uses

With the upcoming enforcement of the EU AI Act [23], AI-based technologies will be subject to legal compliance. This raises particular concerns for the use of AI in mobile and wearables, as the Act highlights factors that could classify certain uses as high risk or even unacceptable. For example, mobile and wearable devices, often placed near or worn on the body, operate in the background unnoticed, and collect personal information about individuals. This characteristic may categorize some uses as high risk due to the sensitive nature of the data handling. However, recent advances in federated learning and differential privacy can ensure data privacy and security for such applications, potentially transforming high-risk uses into low-risk ones. Federated learning, for example, allows AI models to be trained across multiple decentralized devices holding local data samples, and never exchanging them [69]. Additionally, differential privacy introduces mathematical guarantees to ensure that the output of AI models does not reveal sensitive information, and allows data analysis without compromising individual privacy [36]. Similarly, homomorphic encryption methods offer the potential for performing computations on encrypted data, providing robust security assurances [44]. Additionally, mobile and wearable uses are often non-generalizable due to limited training datasets (e.g., skewed towards predominantly white cohorts [107]), introducing biases and limiting their applicability across diverse populations. Therefore, it is important to understand the risks these uses entail. Assessing both risks and benefits allows for a well-informed perspective, puts risks into context, and helps determine which risks are worth taking or which mitigation strategies should be prioritized to preserve the benefits.

The traditional approach to assess risks and benefits in mobile and wearable uses has often been sporadic, required expert knowledge, and primarily confined in research papers. Heidel et al. [53] conducted a scoping review to identify benefits and risks arising from the introduction of health apps and wearables into the German statutory health care system. They noted that health apps and wearables lead to improved treatment quality through enhanced patient monitoring, disease management, personalized therapy, and better health education. However, they also expressed concerns about data privacy and security. The primary concern regarding privacy and human rights with wearables lies in the quantity and type of data that individuals provide. Zuboff, in "The Age of Surveillance Capitalism," [128] wrote that "every time we encounter a digital interface we make our experience available to 'datafication', thus 'rendering unto surveillance capitalism' its continuous tithe of raw-material supplies". Experts typically test mobile and wearable apps for privacy and security vulnerabilities using methods such as man-in-the-middle attacks, eavesdropping, and packet injection [31, 43] to prevent attacks and data misuse.

To sum up, risks and benefits assessments for AI in mobile and wearable uses have been hitherto hindered by two primary limitations. First, these analyses often focus on individual AI use cases or only a selection of few, potentially

overlooking the broader spectrum of uses. Second, assessments are predominantly conducted manually by mobile and wearable experts or researchers, identifying potential benefits and risks through scoping reviews. Overcoming these limitations requires the development of new approaches that automate or semi-automate part of the assessment process. Next, we turn to LLMs, which present a promising ground for exploring this approach.

2.3 LLMs and Prompt Engineering

LLMs such as OpenAI's ChatGPT [90], which are trained on a large corpora of text using cutting-edge deep learning techniques [118], have the capability to produce text that closely resembles human writing in response to various prompts. These models are extremely efficient for tasks such as answering queries and generating content, including text annotation [21, 45, 63] and powering chatbots that assist with mental health concerns [17, 104]. LLMs are known for providing insights that often exceed general public knowledge [45], playing a significant role in collaborative processes between humans and AI [21, 63, 111, 122], and in some instances, their performance can match that of expert opinions [14, 35, 89]. To achieve the desired output from LLMs, a set of best practices for prompt engineering has emerged [100, 106]. Typically, these practices can be grouped into five areas.

The first area concerns the choice of the prompt's elements. This includes the system's role, which is key to initiating interactions with the language model as it primes the model to behave in a manner that aligns with the role's expectations [7]. Whether the model is instructed to act as a mobile developer, a creative assistant, or a project manager, this initial framing is a critical step in guiding the model's responses and ensuring they are contextually relevant and appropriately styled. Additionally, the role can be extended with additional cues to direct the model to the desired output (e.g., initial words for the sentence the model is expected to complete [7, 106]).

The second area is about the prompt's elements arrangement. For example, the order in which elements are placed greatly affects the outputs due to a phenomenon called recency bias. It is recommended to place the most important elements of the prompt at the beginning and end [66].

The third area is about the instructions to the model. Successful instruction requires clear and precise communication of tasks. It involves choosing the right action verbs (e.g., describe, summarize, explain), specifying the preferred format of the output (e.g., JSON or a single sentence), incorporating knowledge that is required for the model to learn. Part of the instructions are also different learning methods that can boost the model's performance. Beyond zero-shot learning, where models depend only on their pre-existing knowledge, few-shot learning provides task-specific examples of input-output pairs, enhancing performance on specialized tasks. Chain-of-thought (CoT) reasoning, which prompts the model to provide step-by-step answers, is especially beneficial for tasks that involve counting and mathematical problems [11, 119, 120].

The fourth area is about adjustment of model parameters. For example, temperature modulates the balance between unpredictability and predictability in the produced text. Higher temperature values encourage more varied responses, while lower values result in outputs that are more deterministic.

The fifth area is about the evaluation of the output. LLMs may generate biased or incorrect content, a phenomenon known as hallucinations [37]. Therefore, for tasks that require certain levels of quality or creativity, the model's output should be evaluated. Evaluations may include human reviewers to guarantee that the content adheres to quality benchmarks and remains authentic [37, 38].

As highlighted, traditional risks and benefits assessments of AI for mobile and wearable uses have been sporadic, fragmented, and resource-intensive. LLMs incorporating these best practices have the potential to uncover benefits and risks in an automated manner.

3 AUTHOR POSITIONALITY STATEMENT

Recognizing the importance of author positionality is essential for transparently examining our perspectives on methodology and analysis [42, 88]. In this paper, we situate ourselves in a Western country (United Kingdom) during the 21st century, writing as authors primarily engaged in academic and industry research. Our team comprises two males and two females from Southern, Eastern, and Central Europe with diverse ethnic and religious backgrounds. Our combined expertise covers a range of areas, including human-computer interaction (HCI), ubiquitous computing, software engineering, artificial intelligence, natural language processing, data visualization, and digital humanities.

It is also important to acknowledge the use of a large language model as part of our methodology. As researchers from a primarily Western institution, we understand the importance of broadening the perspectives on methodology presented in this paper, and retain responsibility for the content and interpretation of the findings. Consequently, our positionality may have influenced the subjectivity inherent in selecting our methodology, designing our study, and interpreting and analyzing our data.

4 METHODS

We developed three LLM prompts (Figure 1). The first prompt generated a wide range of AI uses for mobile and wearables, the second classified each use (whether it is unacceptable, high risk, or low risk) according to the EU AI Act [23], and the third determined whether each use is helpful for reaching Sustainability Development Goals [114]. We chose the EU AI Act for classifying a use's risks because it is the most advanced risk-based legal framework for AI systems, and the SDGs for determining a use's benefits because it is globally recognized and allows for evaluation of long-term impacts related to people, the planet, prosperity, peace, and partnerships. Each prompt was manually validated by two authors to verify the correctness of its outputs. We also consulted a legal and compliance expert to further validate the prompt's outputs as well as conducted a crowdsourcing study with a cohort of nine individuals with legal backgrounds who were recruited from Prolific, confirming the accuracy of our approach to be over 85%. As our LLM model, we chose OpenAI's GPT-4 due to its best performance across benchmarks, and its demonstrated capability in interpreting legal documents [126]. To allow for reproducibility, in line with open-source language models, we made our prompts and code publicly available.

4.1 Prompt #1: Generating Uses

This prompt generates a list of AI uses for mobile and wearable technologies. It achieves this by introducing two innovative prompt elements—a cue for different *domains* where these technologies could be applied, and a cue for five *components* that guide the specific description of each use and enable risk assessment as per the EU AI Act.

Prompt #1 combines three elements: *system role*, *instructions*, and *output format.*

First, the *system role* is a *Senior Mobile and Wearable Systems Specialist, with experience in sensing technologies and their applications in multiple domains;* this choice guides the model to generate content aligned with advanced expertise required of this position [46]. We selected this role because it brings a comprehensive understanding of both the technical aspects and practical deployment of wearable technologies, ensuring our study addresses real-world applicability alongside theoretical advancements. Compared to other potential roles (e.g., a Health Informatics Specialist, Consumer Electronics Product Manager, or IoT Architect), a more broader expertise is particularly suited

---

https://cdn.openai.com/papers/gpt-4.pdf
https://lmsys.org/blog/2023-06-22-leaderboard/
https://github.com/comarios/mhci-risks-benefits

to guide the development and deployment of mobile and wearable technologies. Additionally, we included a cue to further contextualize the role. This cue is a list of mobile and wearable sensors (e.g., accelerometers, gyroscopes, photoplethysmography, cameras, and global positioning systems) along with their possible placement (e.g., placed in various objects or parts of the body, such as the torso, wrist, and pocket) and their technical capabilities (e.g., monitoring human behavior, inferring physical, mental, emotional, and social status).

Second, the *instructions* include: (1) the definitions of five components [47] that are required for risk assessment as per the EU AI Act [23]. The five components are: the *domain*, that specifies the industry or sector (e.g., health); the *purpose*, that explains the goal (e.g., monitoring fitness); the *capability*, that describes the technology behind the use (e.g., mood and stress level tracking); the *AI user*, who operates the system (e.g., doctors); and the *AI subject*, who is affected by the system (e.g., patients); (2) a list of application 46 domains (e.g., health, family, law enforcement), systematically developed based on the EU AI Act by selecting eight high risk domains listed in Annex III and 32 additional domains inferred from the Act's text and amendments [24]. This list guides the LLM to generate specialized and nuanced uses across individual, interpersonal, institutional, community, and public policy contexts; (3) the task of generating three AI uses for each of the 46 application domains.

Third, the *output format* specifies how the LLM's output should be displayed in the five-component format (e.g., for monitoring fitness using wearables [Domain, Purpose, Capability, AI user, AI subject] becomes ["Health", "Monitoring fitness", "Mood and stress level tracking", "Doctors", "Patients"]).

Manual validation. After generating the uses, two authors with experience in mobile and wearable computing reviewed each one of them. The authors independently assessed each generated use to determine whether it is correct and whether it is existing or upcoming. They considered all 138 generated uses to be correct, with 128 (93%) already existing (practically viable and in use), 9 (6%) being upcoming (practically viable but still in early prototype stage), and 1 being unlikely (practically viable but only discussed as possibilities in research papers).

4.2 Prompt #2: Classifying Use Risks

This prompt classifies the risks of each generated use according to the EU AI Act [23]. It combines four elements: *system role*, *input*, *instructions*, and *output format*. First, the *system role* is a *Senior AI Technology Expert, specializing in compliance with the EU AI Act*. Second, the input includes articles from the EU AI Act and its Annex III regarding high risk AI systems. Third, the instructions use a Chain-of-Thought approach (i.e., dividing the task into a series of smaller, intermediate reasoning steps that lead to the final output [120]). This approach involves three steps: (1) drafting a concise description of the system's use through the five component-format; (2) producing a risk classification as per the EU AI Act (whether the system's use is unacceptable, high risk, or low risk); and (3) providing justification for the classification. Fourth, the output format of risk assessment as per the EU AI Act contains three fields: "[Use Risk Classification], [Classification Justification], [Relevant EU AI Act Article]."

Manual validation. Two authors familiarized themselves with the EU AI Act articles, and manually classified each of the 138 generated uses as either unacceptable, high risk, or low risk. The validation criteria aligned with the specific requirements and definitions outlined in the EU AI Act concerning risk levels associated with AI uses. We defined the following criteria: potential impact on privacy and data protection, and the likelihood of adverse outcomes affecting individuals or groups. Additionally, we considered the operational contexts in which a mobile or wearable use is deployed such as the type of data collected and the functionality of the device. 80 uses (58%) were identified as high risk, 57 (41.3%) as low risk, and 1 (0.7%) use as unacceptable risk. The agreement between the two authors and the

LLM's classification was 85%. In the rest 15% of cases in which LLM was wrong, the disagreements were about uses related to vulnerable groups, personalization, government and democracy, and critical infrastructure (e.g., energy), all of which the LLM classified as low risk instead of high risk. The only use categorized as unacceptable risk (where both the two authors and the LLM agreed) was related to enhanced surveillance capabilities using cameras. These disagreements were resolved in two steps. First, the two authors held an initial review meeting in which they discussed their assessments against those done by the LLM, and consulted the rest of the research team. Second, we consulted a compliance expert specialized with the EU AI Act to inspect the 15% of the uses that were wrongly classified by the LLM. The expert is an academic affiliated with TU Delft in Netherlands and works at the intersection of technology, policy and management. The expert manually classified each use's risk (i.e., unacceptable, high risk, or low risk), confirming that indeed these uses were wrongly classified by the LLM. Upon the consultations with the team and the expert, we reached the final decision.
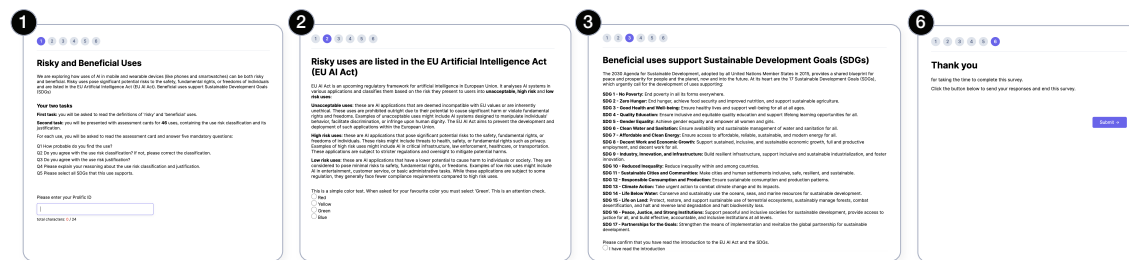
4.3 Prompt #3: Determining Use Benefits

This prompt determines whether each generated use is helpful for reaching Sustainability Development Goals [114]. It combines four elements: *role*, *input*, *instructions*, and *output format*. First, the role is a "Senior Specialist in the field of Mobile and Wearable Technologies with a dedicated focus on understanding, promoting, and implementing the SDGs". Second, the input includes the 17 definitions of the SDGs, covering a broad range of potential areas for beneficial AI applications across the economy, society, and the environment. Third, the instructions (using Chain-of-Thought as in the previous prompt) involve: (1) summarizing the system's use based on the five-component description; (2) listing the benefits, reflecting how the use promotes each SDG. Fourth, the output format contains two fields: "[SDG Goal], [Reasoning for Promoting the Goal]".

Manual validation. Two authors familiarized themselves with the 17 definitions and descriptions of the UN Sustainable Development Goals, and manually reviewed the feasibility of benefits associated with each goals [114]. To ensure a robust manual validation, the authors defined the following criteria: *a)* the applicability of the use to an SDG goal is evident; and *b)* the feasibility of benefits is evaluated based on current technological capabilities. The agreement between the two authors and the LLM's output was 61%. Disagreements were mainly related to goals 5 (gender equality), 16 (peace, justice and strong institutions) and 17 (partnerships for the goals), all of which were manually corrected. Disagreements were resolved in a similar way to classifying uses risks. First, the two authors discussed disagreements among them, and then consulted the rest of the research team to reach a final decision.

4.4 Crowdsourcing Study

Setup and Procedure We developed a web-based survey and administered it on Prolific, which included the LLM classifications for each use (Figure 2). The survey consisted of six pages. The first page outlined the study's description and the two tasks that the crowdworkers had to execute. Initially, they were asked to read the definitions of risky' and beneficial' uses, and then they were presented with assessment cards for 46 uses to assess. The second and third pages included the definitions of risky and beneficial uses according to the EU AI Act and the SDGs, respectively. The fourth and fifth pages presented 46 uses (23 each in page), where trap questions were interspersed to ensure data quality. We added two types of trap questions: the first type included two attention checks: *When asked for your favorite color, you must select 'Green'*; and *When asked for your favorite city, you must select 'Rome'*. Participants had to correctly respond

Fig. 2. Crowdsourcing study survey. Participants were given instructions about the study (1) and were provided the definitions of risk classifications as per the EU AI Act (2) and the definitions of the 17 Sustainable Development Goals (3). Then participants were presented with 46 mobile HCI uses (4,5) and asked to assess how probable is the use (Q1), whether they agree with the LLM-generated risk classification (Q2) and its justification (Q3), explain their reasoning about the classification and justification (Q4), and select the SDGs that they believe the use supports. After annotating all uses, they were redirected to the Prolific confirmation page (6).

to these checks after completing each task. The second type included an unacceptable use (i.e., a social scoring system based on facial recognition in public spaces) according to the EU AI Act, which we deliberately classified as low risk. Participants who failed either of the attention checks or to correct the classification of this use were disregarded and not paid. Additionally, we disabled pasting from external sources and editing previous responses to ensure original and thoughtful answers. The final page included a thank-you note and redirected participants on Prolific. To ensure reproducibility and assist researchers who aim to use our methodology, we have made the code for our survey publicly available.

Participants. To recruit individuals with a background in legal and compliance, we relied on Prolific's screening criteria. We searched for participants likely involved in revising AI systems as part of their legal roles, using AI at least 1-6 times a week. In total, we recruited 9 participants with "legal" as the "function in organization", suggesting that these individuals had prior legal expertise (Table 1). Six self-identified as males and three as females, with their ages ranging between 21 and 49 years old ($\mu = 37$ and $\sigma = 8.73$). Their current countries of residence were within the European Union, including participants from Italy, Poland, Germany, the Netherlands, Portugal, and Slovenia. On average, participants took 40 minutes to 1 hour to complete a batch of 46 uses, and were compensated €9.29.

https://github.com/comarios/mhci-risks-benefits

Table 1. Demographics of participants in the crowdsourcing study.

Batch	ID	Gender	Age	Country	Weekly AI use
	P1	Male	33	Italy	Every day
B1	P2	Male	41	Poland	2-6 times a week
	P3	Male	38	Germany	2-6 times a week
	P4	Male	40	Germany	2-6 times a week
B2	P5	Female	47	Italy	2-6 times a week
	P6	Male	25	Poland	Multiple times every day
	P7	Female	49	Portugal	Every day
B3	P8	Female	36	Slovenia	Every day
	P9	Male	21	Germany	Multiple times every day

Analysis and Results. Each use was annotated by three crowdworkers, and each of them annotated a total of 46 uses (i.e., all 138 uses were divided into three batches).

When they classified the three risk levels based on the EU AI Act classifications, there was a full agreement in 48% of uses. Full agreement was about uses that have clear benefits and utility (e.g., tracking health status); are widely-adopted (e.g., facilitating mobile payments); or are in already regulated domains (e.g., enhancing food safety). Conversely, disagreements were about uses that present privacy, security, and ethical concerns (e.g., tracking vulnerable groups like children and the elderly); or are deemed emerging (e.g., mobile VR for educational content). We then used majority voting to establish the final classifications by the crowdworkers and compared them against the LLM classification. The agreement was 92.75%. By manually inspecting the disagreements, crowdworkers corrected the LLM classification and cited similar reasons to those reported during the two authors' manual validation (e.g., cases involving uses about democracy and personalization).

When the crowdworkers then determined the 17 SDGs applicable to the uses at hand (one use could have multiple SDGs assigned), there was a full agreement in 11% of uses (i.e., crowdworkers selected the exact same SDG(s) for each use). Full agreement was about uses that had a clear positive impact on daily life (e.g., monitoring health status with wearables); aligned with broader societal goals (e.g., improving election transparency); or improved safety and security (e.g., enhancing surveillance capabilities for public safety). Conversely, disagreements were about uses with varied perceptions of usefulness (e.g., tracking emotions for well-being); or had privacy, security, and ethical concerns (e.g., involving facial recognition). For each use, we then took the union of the three crowdworkers' annotations to establish the final classification and compared this list against the LLM classification. The agreement was 37.89%. Compared to the two authors' validation (61% agreement with the LLM classification), crowdworkers had lower agreement. This greater discrepancy could be attributed to the difficulty in determining the applicability of SDGs, which may require expertise beyond the legal domain. Additionally, it requires complex cross-referencing of the use with 17 different goals and their numerous targets.

5   RESULTS

Having the set of 138 generated uses and their associated risks and benefits (detailed in Table 2 in the Appendix; the notation # indicates the use's ID), we report the results in two steps. First, for each SDG, we computed the proportion of beneficial uses categorized as low risk (§5.1) and high risk (§5.2), and then discuss example uses from the goals with the

highest proportions of each risk category. Second, we conducted a thematic analysis to understand the reasons why a use is risky or not, and why it is beneficial (§5.3). Note that among the 138 uses, only one was classified as unacceptable according to the EU AI Act. This use involves *enhancing surveillance capabilities using mobile cameras in public spaces* (use #64). In our analysis, we did not focus on unacceptable uses because they are banned. Instead, high-risk AI uses, which are allowed to be deployed, must be subject to an impact assessment for compliance purposes. Therefore, our analysis focuses on distinguishing between high-risk and low-risk uses.

5.1 Uses Promoting SDGs and Being Low Risk

Uses that pose low risk primarily promote goals related to environmental protection, such as life below the water (goal 14), life on land (goal 15), and clean water and sanitation (goal 6) (navy bars in Figure 3). These uses are classified as low risk by the EU AI Act. Next, we discuss examples of these uses.

Monitoring crop health and growth, use #49. IoT sensors are used to continuously monitor environmental parameters such as temperature, humidity, and soil moisture, alerting farmers via mobile apps if conditions are unsuitable for their crops [113]. Such uses promote sustainable agricultural practices and environmental conservation (goal 15). Similar uses can involve in vivo plant sensors (e.g., organic electrochemical transistor-based biosensors [116]) that detect nutrient deficiencies or drought stress and allow for the precise application of agrichemicals, fertilizers, and water, thereby maximizing yields [96].

Tracking livestock health and location, use #50. GPS collars have been effectively used to track livestock, including their location, movement patterns, and grazing behavior (e.g., by creating "virtual geofences" which alert farmers if an animal strays from a designated area [56]); uses that enables precise livestock management (goal 15). These collars can also be equipped with more specialized sensors to monitor the behavioral and physiological parameters of livestock, allowing farmers to assess the animals' health and welfare over time. Examples include pedometers and microelectromechanical (MEMS) activity sensors, which are easily attached to animals [54]. Additionally, tracking livestock movements and interactions can help in understanding and managing disease outbreaks, such as those caused by the influenza virus. [85].

Optimizing irrigation systems, use #51. IoT sensors have been used to optimize water usage (goal 6) and predict harvest times (goal 15). Mohamed et al. [80] demonstrated a smart irrigation system equipped with sensors for monitoring water levels, irrigation efficiency, and climate conditions. Wearable soil moisture sensors, strategically placed throughout a field, can now transmit data to a central system, accessible via mobile devices. Farmers can adjust irrigation schedules directly from their phones, optimizing water usage based on real-time soil moisture data. This approach not only preserves (and saves) water but also ensures that crops receive the appropriate amount of water at the right time, which is essential for crop health and yield [87].

Monitoring environmental conditions, use #112. IoT sensors are instrumental in monitoring a wide range of environmental parameters (e.g., heat, sound, pressure, temperature, air quality, water quality), promoting sustainable management of terrestial ecosystems (goal 15). Integrating these sensors into smart environmental monitoring systems allows for large-scale environmental assessments [4]. Battery-free, wireless cameras and sensors can now be deployed in underwater environments to monitor water quality, marine life, and underwater ecosystems [2] (goal 14). These sensors are not only limited to monitoring air and water quality but also extend to waste management [62]. For example, in domestic environments, they can effectively monitor conditions such
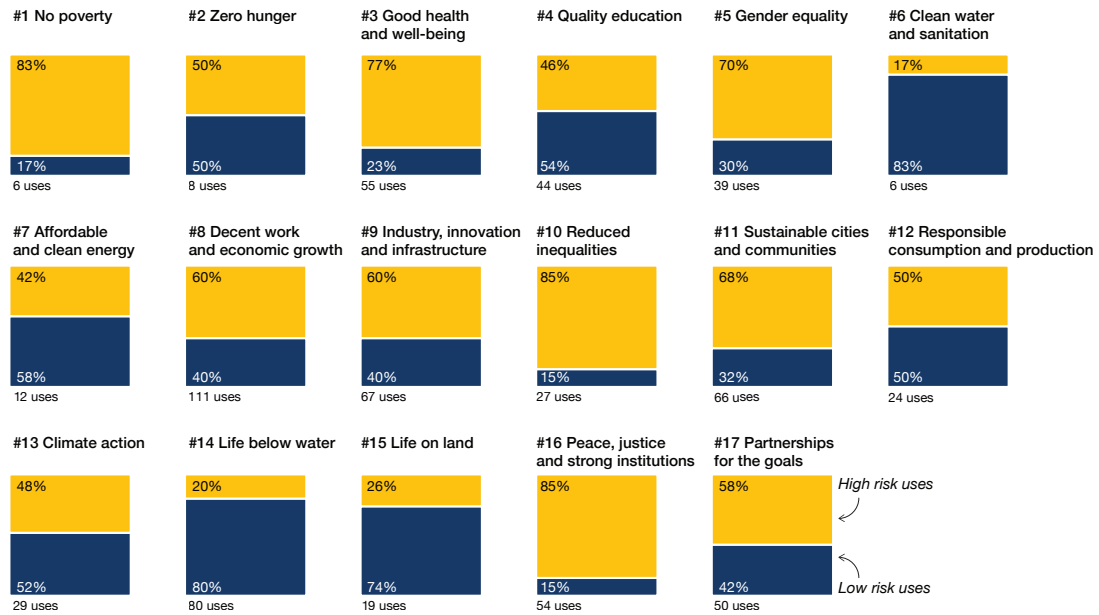
Fig. 3. Percentages of uses categorized as low risk and high risk for each goal. The total number of uses per goal is in parentheses below the goal's name.

as temperature, humidity, and gas levels, with the data being easily accessible through mobile apps or web
    pages [16].

Tracking wildlife movements, use #113. GPS collars are used to study animal behavior, and understand their habitat
    preferences and migration behaviors [3] (goal 15). These technologies also enable real-time monitoring of wildlife
    movements, including proximity, geofencing, movement rate, and immobility [117] and can enable modeling
    epidemics in wildlife hosts, such as outbreaks of dolphin morbillivirus [82]. Similar uses can involve multi-sensory
    wearable devices that employ neural networks for classifying animal behavior in national parks [95] or for
    predicting animal personality [71]. Wireless sensor networks track small turtles, monitoring their micro-climate
    and hibernation periods [59], while GPS-powered systems can effectively track wild animals straying from
    sanctuaries, aiding in poaching prevention [48].

5.2   Uses Promoting SDGs yet Being High Risk

Uses that promote certain goals but pose high risks primarily relate to peace, justice, and strong institution (goal 16),
good health and well-being (goal 3), and reduced inequalities (goal 10) (gold bars in Figure 3). These uses entail high
risks, either due to the use of risky technology (e.g., cameras for facial recognition, or motion sensors paired with
physiological sensors), or because they operate in high risk domains such as migration management, military and
defense, or healthcare involving individuals from vulnerable groups. These uses include:

Enhancing security through facial recognition, use #1. High-performance cameras in smartphones have been typically
    used for various authentication purposes such as phone unlocking, banking, and access control, by analyzing
    face images [12, 29]. Such uses ensure safety and security by preventing unauthorized access (goal 16). However,

as per EU AI Act's Article 6 "biometric identification and categorisation of natural persons" is under the high risk category. However, the very same technology (i.e., cameras) may well be used for crowd surveillance; a use that has unacceptable risks.

Identifying individuals in criminal investigations, use #3. Smartphones and wearable watches can provide crucial information in criminal investigations (e.g., documenting an individual's activities), which can be used to determine the veracity of witness testimony [34]. For example, a smartphone synced with a wearable watch can offer evidence related to the execution of a violent act, aiding in pinpointing the time and nature of the incident [70] and ensuring safety and security (goal 16). Suspicious pattern detection techniques from mobile device data can also help identify individuals involved in activities such as cyberbullying or low-level drug dealing [9]. According to EU AI Act's Article 6, the amendment 711 classifies such a use as risky because it can used by "law enforcement authorities for making individual risk assessments of natural persons in order to assess the risk of a natural person for offending or reoffending or the risk for potential victims of criminal offences".

Monitoring children's location for safety, use #4. Multi-sensor wearable such as wristbands and IoT-enabled devices have been designed to track children's location, body temperature, humidity, and heartbeat, with the capability to alert parents through SMS text messages in case of emergency [20, 109]. Such uses strive for improving children's well-being and safety (goal 3). However, as per the EU AI Act's article 69 and amendment 634, the use is risky because it "may affect vulnerable persons or groups of persons, including children, the elderly, migrants and persons with disabilities".

Monitoring elderly health and activity, use #5. Wearable devices, such as wristbands equipped with sensors like accelerometers, gyroscopes, and heart rate monitors, are used for activity detection and monitoring in elderly care and rehabilitation [86], improving their well-being and safety (goal 3). More broadly, IoT-based smart healthcare systems provide real-time monitoring of the elderly's location, activity patterns, and health status, enabling early detection of health risks and improved response to emergencies [65]. As this use may affect vulnerable groups is classified as risky under the EU AI Act's article 69 and amendment 634. However, the same on-body sensors could be used for surveillance, a concept Yuval Noah Harari termed "under-the-skin surveillance" [50], introducing unacceptable risks. This type of surveillance goes beyond traditional monitoring by collecting intimate physiological data directly from a person's body, potentially invading privacy and autonomy on a deeply personal level.

Enhancing mental health treatment, use #12. Wearable sensors for electrodermal activity and photoplethysmography provide real-time physiological data crucial for mental health monitoring and treatment [40], ensuring good (mental) health and well-being to individuals (goal 3). A comprehensive study highlighted 46 systems focusing on continuous monitoring, diagnosis, and care of various mental disorders through mobile and wearable technologies [8]. Music streaming services, when repurposed as therapies for affective disorders, exemplify the potential of these technologies in mental health treatment [102]. Additionally, smartphones and wearable devices support mental health research by facilitating the collection of novel naturalistic and longitudinal data relevant to psychiatry [112]. However, similarly to uses affecting children or elderly, this use is again classified as risky the EU AI Act's article 69 and amendment 634 as it may affect vulnerable groups who suffer from mental health problems.

Preventing unauthorized access, use #43. Wearable devices are used in access control systems to protect information and define restrictions on information handling [33] in military facilities, prevent unauthorized data leakage, and safeguarding sensitive information [125] (goal 16). These devices can also be used in user authentication,

improving the security of mobile payment systems and other applications where sensitive data is involved [121]. However, As per the EU AI Act's Article 6 and amendment 711, the use is risky because it can be used for "biometric identification".

Predicting infrastructure failures, use #62. Wearable sensors have been applied to predict health-related issues, such as COVID-19 incidence, by tracking the spread of the virus, demonstrating their potential in predictive health infrastructure [103]. Analyzing mobility patterns in IoT networks (e.g., georeferenced traces that citizens leave when they navigate a city) are used to predict infrastructure failures and optimize road traffic [39] such as road traffic, ensuring inclusive access to society (goal 16). However, the EU AI Act's amendments 713 and 714 classify the use as risky because it is intended to be used as "safety component in the management and operation of road traffic and the supply of water, gas, heating and electricity".

Enhancing surveillance capabilities, use #64. Mobile cameras, accelerometers, and gyroscopes are typically employed for continuous monitoring and activity detection in surveillance applications [28], ensuring safety by maintaining public order and general welfare in a democratic society. The concept of "sousveillance," which refers to surveillance from below as opposed to from above, has been facilitated by wearable devices, generating new types of information in social surveillance situations [67]. Mobile sensors, like those in Android smartphones, are used in real-time surveillance systems for tasks such as human face tracking [98], even in public spaces. High-tech gear like textile-based wearable microphones, miniature cameras, and wireless personal displays can also enhance surveillance capabilities while appearing natural to the user [115]. However, as per the EU AI Act's article 5.1, the "use of real-time remote biometric identification systems in publicly accessible spaces, unless and in as far as such use is strictly necessary".

5.3 Understanding the Reasons Why a Use Is Low Risk, High Risk, and Beneficial

Two authors conducted an inductive thematic analysis (bottom-up) to understand the reasons why a use is risky or not, and why it is beneficial (i.e., as per Prompt 2 and 3 outputs), following established coding methodologies [68, 74, 97]. The authors used sticky notes on the Miro platform [77] to capture the reasons and inductively construct the themes. They held four meetings, totaling 8 hours, to discuss the emerging themes that arose during the analysis process. Next, we discuss the resulting themes and grouped them into three categories, that is, reasons for uses being: *a)* low risk; *b)* high risk; and *c)* beneficial.

Low Risk: Low risk uses were about the *environment and sustainability*, and *logistics*. The reasons why these uses are low risk are because they handle data about non-human subjects and the environment, and operate in domains that are not considered high risk. For environment and sustainability, mobile and wearables enable environmental monitoring (e.g., air and water quality) that allow for timely interventions to mitigate pollution; and monitor energy consumption and provide feedback to users for encouraging energy-saving behaviors and contributing to carbon footprint reduction. For logistics, wearables devices equipped with RFID or GPS technology enable real-time tracking of goods that can significantly reducing the time needed for inventory management and supply chain operations; sensor-equipped wearables (e.g., soil moisture sensors) provide precise information on crop needs to help minimize the use of water and fertilizers.

High Risk: High risk uses were about the *sensitive data handling*, *vulnerable user groups*, and *automated decision making*. The reasons why these uses are high risk are because they generally collect sensitive data, even about vulnerable individuals at time or engage in automatic decision making aimed at behavioural change. For sensitive data handling, the

uses involved collecting data with different levels of sensitivity: biometric data such as facial images or footprints, voice prints or gait data, and other personal data which can reveal information about physical, physiological or behavioural traits of people such as fitness data or financial records. For vulnerable user groups, the uses were affecting vulnerable persons or groups of persons, including children, the elderly, migrants, and persons with disabilities. For automated decision making, the uses were designed in a way to influence habits (e.g., recommender systems), leading to potential negative impacts on mental health and social interactions.

Beneficial: Independent to whether uses are risky or not, they were beneficial mainly because they *foster participation and collaboration* in society, *improve well-being*, and *improve safety and security*. For fostering participation and collaboration in society, mobile and wearables enable seamless communication and connectivity, breaking down geographical and temporal barriers; allow users through apps to share their experiences, achievements, and challenges, fostering a sense of community and mutual support. For improving well-being, mobile and wearables allow for physical activity and sleep patterns monitoring; used to detect early signs of health issues and provide recommendations for stress management (e.g., meditation and psychological support). For improving safety and security, emergency response apps allow users to quickly alert emergency services; GPS tracking features on mobile and wearables are beneficial for monitoring the safety of children, the elderly, or individual in high risk environments.

6 DISCUSSION

Our method offers a semi-automatic way to assess the risks and benefits of AI uses in mobile and wearables using three LLM prompts. We found that low risk uses primarily benefit the environment and logistics. Conversely, high risk uses, while promising improved well-being, safety, and equality, come with significant concerns over sensitive data, vulnerable groups, and automated decision-making. These uses often collect biometric data from cameras or operate in high risk domains such as healthcare, particularly when involving vulnerable groups. Next, we discuss challenges in conducting such an assessment highlighting our lessons learned, and present practical solutions (e.g., a checklist) for balancing the trade-offs between risks and benefits in mobile and wearable technologies.

6.1 Challenges in Conducting Risks and Benefits Assessment Using LLMs

The use of LLMs to assess the risks and benefits of AI (in our case, on mobile and wearable technologies) presented unique challenges, which are likely to be applicable to other domains as well. Although LLMs may be able to provide a comprehensive analysis across various use cases, risks, and benefits, their outcomes should be treated as preliminary steps rather than definitive conclusions. We specifically highlight three areas that require careful attention.

First, a key challenge is ensuring the accuracy and relevance of the LLM' output, as these models heavily rely on the quality and scope of their training datasets. We overcome this challenge by employing a number of best prompt engineering practices [100, 106] such as specifying the model's role (e.g., senior mobile and wearable specialist), detailing its task(s), providing additional knowledge (e.g., excerpts of the EU AI Act), and defining the expected output format.

Second, the complexity and often opaque nature of LLMs make their output's interpretation challenging. To ensure that the language model produced realistic uses, and, in turn, classified each use's risks and determined its benefits, two authors manually inspected the output in every step of the process. A human-in-the-loop approach helps to alleviate model's hallucinations. Additionally, by employing a Chain-of-Though approach [120], we ensured that the model explained its reasoning to arrive at the final output.

Finally, LLMs might not fully capture all ethical and societal dimensions linked to mobile and wearables. More broadly, identifying risks associated with AI systems requires deep organizational, technical, and regulatory expertise [13], which can be achieved by integrating experts' knowledge in the process. The use of AI for evaluating AI systems (and their use) introduces a number of ethical considerations that require careful examination. One prominent concern is the recursive nature of such assessments, wherein AI systems are employed to evaluate outputs generated by similar AI systems, potentially leading to feedback loops that amplify biases or errors without sufficient human oversight [49]. Another concern is about transparency. Documenting the decision-making processes, the criteria used for evaluations, and the limitations of the AI tools employed is essential for allowing independent verification and fostering trust [79]. For example, to allow for reproducibility, we made our code publicly available.

Despite the advancements in LLMs capabilities, there are inherent limitations in their ability to make judgments against complex ethical standards. Current LLMs may struggle to comprehend and apply nuanced ethical principles, leading to potential gaps and inconsistencies in such assessments [58]. Addressing these limitations requires ongoing research aimed at enhancing LLMs ethical reasoning capabilities and ensuring alignment with universally accepted ethical norms and standards. Additionally, we foresee that semi-automated ways for conducting risk and benefits assessments should be complemented with human-in-the-loop validations and/or non-AI approaches (e.g., crowdsourcing studies), which we discuss next.

6.2   Challenges in Conducting AI-Based Risk and Benefit Assessments Versus Human-Based Assessments

Next, we turn into discussing challenges for conducting risks and benefits assessments using AI-based approaches (e.g., LLMs) versus human-based ones (e.g., crowdsourcing). While AI-based assessments offer speed, scalability, and consistency, our findings reveal several challenges that call for a combined approach incorporating both methodologies.

The first challenge is about consistency and contextual judgment. Our AI-based assessment using LLMs demonstrated, to a great extent, high agreement rate with human judgments when assessing the risk classification level of mobile and wearable uses. However, this consistency may overlook nuanced interpretations that human experts can provide. Human assessors, especially those with specialized knowledge, can offer contextual judgments that AI might miss, as evidenced by the adjustments made by crowdworkers based on deeper understandings of democracy and personalization issues. In other words, this illustrates the limitations of AI in areas requiring ethical consideration and complex contextual analysis, suggesting that while AI can guide assessments, it should not replace human judgment.

The second challenge is about the depth of understanding. This was particularly evident in the lower agreement rates on SDGs, where human expertise was crucial for comprehensive evaluation. This difference suggests that AI should operate as an exploratory tool, with final decisions deferred to human experts who can assess the broader ethical and societal implications of a use.

The third challenge is about reliability. The human-based approach introduced diverse judgment perspectives, but at the same time demonstrated potential inconsistencies. This variability, while introducing diverse viewpoints, also necessitates rigorous control measures to ensure reliability (e.g., using trap questions and attention checks). One possibility would be to use AI-based assessments as a baseline for consistency and comparison with human-based ones.

Given these challenges, we believe that a promising approach should combine both AI-based assessment and human-based ones. Such a hybrid approach would use AI for initial assessments due to its efficiency and scalability, and human experts for providing the final judgments. To aid researchers interested in adopting a similar approach in different contexts, we recommend following these steps: *(1)* describe a use case using the five-component format suggested by Golpayegani et al. [47] (i.e., Domain, Purpose, Capability, AI User, AI Subject) as required for risk assessment under the

Risk Assessment Checklist for Mobile Computing

1   Use
1.1   What is the purpose of the use?
1.2   What is the intended domain, industry or sector for the use?
1.3   What are the technical capabilities that enable the use?
1.4   Who are the individuals or groups affected by the use?
1.5   Who are the individuals or entities responsible for overseeing the use?

2   Data
2.1   Does the device/application collect personal data to enable the use? If yes, specify the type
       of collected data (e.g., biometric)
2.2   Does the device/application collect data only after receiving clear consent from users?
2.3   Are users able to opt-out of data collection?
2.4   Are users informed about how their data is used, stored, and shared?
2.5   Are there clear policies on data usage, storage, and sharing?
2.6   Are there encryption methods implemented to safeguard the data?

3   Risks
3.1   Are there aspects of the device/application/use that could present challenges for users
       of different abilities or vulnerable backgrounds?
3.2   Are there aspects of the application/device/use that could be seen as discriminatory
       towards certain users?
3.3   Are there aspects of the application/device/use that could lead to habit-forming behaviors
       with adverse personal or social consequences?

4   Mitigations
4.1   Are there measures in place to prevent unauthorized access to the device/application/data?
4.2   Are there measures in place to prevent the misuse of the device/application/data
       for unethical purposes?
4.3   Are there mechanisms in place that enable users to provide feedback or report problems?
4.4   Are there measures in place to minimize the environmental footprint of the device/application?
4.5   Are there preventive features integrated in the device/application to safeguard personal
       and social well-being?

5   Compliance with Legal and Regulatory Standards
5.1   Does the device/application/use comply with relevant laws and regulations
       (e.g., GDPR, HIPAA, EU AI Act)?
5.2   Are there procedures in place for compliance verification and regular audits?

Fig. 4. Risk Assessment Checklist for Mobile Computing. It helps to systematically consider a mobile or wearable system's: (A) use, (B) data, (C) risks, (D) mitigations and (D) compliance with legal and regulatory standards.

EU AI Act [23]; *(2)* incorporate expert human validation with clear intervention protocols throughout the process; and *(3)* implement continuous feedback loops between the AI-based assessments and the human-based ones.

6.3   Balancing the Trade-Offs Between Risks and Benefits of AI in Mobile and Wearable Uses

The use of mobile and wearables in sensitive or high-stakes domains underscores the need to balance risks and benefits. While these technologies contribute significantly to Sustainable Development Goals such as good health and well-being, reduced inequalities, and peace and justice, they also pose inherent risks due to their sensitive nature. Challenges include privacy concerns, potential misuse for surveillance, and the risk of exacerbating existing inequalities or societal issues. For example, sensitive personal data (e.g., health information) could be misused if not adequately protected, leading to unauthorized access, surveillance, or infringement on individual freedoms and rights. Additionally, it is hard to obtain large representative samples across diverse populations when training models for mobile and wearable uses, making such uses more susceptible to biases [107].

To balance risks and benefits requires active research, policy updates, and open dialogue as to what is considered as an acceptable technology use [26]. For example, it is essential to enhance privacy protections beyond those outlined in the EU AI Act as mobile and wearables capabilities continue to evolve. This can be achieved not only by understanding the direct impact of these technologies (e.g., producing impact assessment cards similar to model cards [78]) but also by considering their broader societal implications. It is of prime importance to ensure that the benefits of these technologies do not exacerbate existing vulnerabilities or inequalities. Another solution would be for the mobile computing community to adopt practices for reporting harms and risks associated with the use of specific mobile or wearable sensors and technologies. Conferences such as the Neural Information Processing Systems (NeurIPS) [6] and the International Conference on Machine Learning (ICML) have already started mandating statements that include "risks associated with the proposed methods, methodology, application or data collection, and data usage". More recently, Olteanu et al. [88] argued that even responsible AI research needs impact statements, too; such statements aim at disclosing any possible negative consequences, contributing to more inclusive research. Additionally, researchers have proposed algorithmic impact assessments (AIAs) as a form of accountability for organizations that build and deploy automated decision-support systems [72], and frameworks for bridging the gap between technical and ethical aspects of AI systems [60].

Risk Assessment Checklist for Mobile Computing: Drawing from similar initiatives in AI and Responsible AI, we developed a risk assessment checklist for mobile computing based on: the learnings from developing our semi-automatic method (§4) and the findings about what makes a use risky or not (§5). These two translated into a five-section checklist, which aligns with recently agreed ways of reporting an AI system's use risks (including the indented use, risks, mitigations, compliance) [1, 73, 84, 105]. However, we do not argue that the checklist is exhaustive, but rather aims to serve as a starting point for the Mobile HCI community to facilitate and foster transparency and accountability in the development and use of mobile and wearable uses.

The first section covers the intended use (Figure 4A). This is informed by the EU AI Act's requirements for risk assessment, and the fact that prior work emphasized the importance of documenting a system's use in a structured five-component format [47]. This format is not only suitable for risk assessment but also comprehensive enough to be usable by both technical and non-technical stakeholders (i.e., due to its simple yet compact structure).

The second section is about the data (Figure 4B). From the thematic analysis (§5.3), it was evident that the type of data (e.g., sensitive vs. non-sensitive) that a use handles plays a critical role in determining its risk level. For example, uses that handle data primarily about non-human subjects (e.g., animals, as per use #50) and the environment (e.g., heat, sound, and air quality, as per use #112), they are considered low risk. Conversely, uses that handle primarily sensitive data (e.g., facial images, as per use #1), they are considered high risk.

The third section is about the risks (Figure 4C). From both the inspection of high risk uses (§5.2) and the thematic analysis (§5.3), these uses tend to involve individuals from vulnerable groups or engage in automatic decision making aimed at behavioural change. Therefore, by formulating questions to identify and assess these risks, this section aims to preemptively address and mitigate potential negative impacts on individuals and society.

The fourth section is about the mitigations of risks associated with the specific mobile and wearable uses as well as the data they handle (Figure 4D).

Finally, the fifth section is about the regulatory compliance (Figure 4E). Given that our method is directly influenced by the upcoming enforcement of the EU AI Act, this section aims to prompt for adherence with existing and upcoming

regulations as well as procedures for verification and regular audits. Such procedures will be a legal necessity to monitor a use and determine its risk level.

We foresee that this checklist can be applied across a range of stakeholders and use cases: *AI practitioners* can use it to identify and assess risks at various stages of the product development cycle; *AI teams* can use it to raise awareness and facilitate discussions about potential risks, leading to more comprehensive risk mitigation strategies; *researchers* can use it to navigate the publication process and to help the community avoid rejecting risky technologies that may have positive impact; and *end users* can use it to understand the various aspects (and AI's use) they should consider before using mobile or wearables.

6.4 Limitations and Future Work

We acknowledge three main limitations of our work that call for future research efforts. First, our method incorporates domain experts into the assessment process. The two authors who manually validated each prompt's output have experience in the design, development, and deployment of mobile and wearable technologies. While not achieving full automation, it expedites the experts' tasks. It systematically compiles a list of uses, presenting them in an organized manner for rigorous risk assessment. Consequently, experts can optimize their time, reducing the need for extensive literature reviews to identify potential uses, risks, and benefits. Similarly, our crowdsourcing study experienced relatively low participation because, in reality, only a handful of experts can provide the insights required to validate our method. At the same time, we chose the majority voting to establish the final classification among crowdworkers. However, this may present a situation wherein a majority's decisions overshadow those of a minority (a phenomenon also known as the 'tyranny of the majority" [94, 108]). Future studies could explore ways to form cohorts of experts that are broadly available and accessible to the research community, offering diverse perspectives in evaluating LLM-generated classifications.

Second, the significant level of disagreement among the experts themselves highlights the subjective nature of such assessments. This suggests that even expert evaluations can be heavily biased by individual experiences, domain expertise, and cultural values [55, 101]. The diversity in the cultural and professional backgrounds of our experts, while enriching the study with varied perspectives, also introduced challenges in reaching consensus. This observation highlights the need for developing more robust methods to use expert disagreement constructively.

Third, LLMs may not cover all potential impact areas, particularly emerging ethical issues (e.g., existential risks) or unforeseen societal consequences, leading to an oversimplification of risks. This may lead to overly simplistic approaches in dealing with complex socio-technical AI systems, especially those that deal with applications domains flagged by the EU AI Act as high risk (e.g., concerning vulnerable groups, or law enforcement). The output of such an assessment is highly advisable to be vetted by lawyers and regulatory experts.

7 CONCLUSION

We proposed a method for assessing the risks and benefits of AI in mobile and wearable uses. Using our method, we generated 138 such uses of AI, while also evaluating their associated risks as defined by the EU AI Act and benefits in line with the UN Sustainable Development Goals; upon manual validation of our method, we confirmed its accuracy to be over 85%. We found that a specific set of mobile computing uses (low risk) primarily benefit the environment and logistics. Interestingly, another set of uses (high risk) that hold significant potential in improving well-being, safety, and social equality are linked to risks involving sensitive data, vulnerable groups, and automated decision-making. To not dismiss the potential of these uses, we proposed a checklist for assessing the risks of mobile and wearable uses.

REFERENCES

[1] Ada Lovelace Institute. 2022. *Algorithmic Impact Assessment: AIA Template.* Retrieved January 22, 2024 from https://www.adalovelaceinstitute.org/resource/aia-template/

[2] Sayed Saad Afzal, Waleed Akbar, Osvy Rodriguez, Mario Doumet, Unsoo Ha, Reza Ghaffarivardavagh, and Fadel Adib. 2022. Battery-free wireless imaging of underwater environments. *Nature communications* 13, 1 (2022), 5546.

[3] Blake M Allan, John PY Arnould, Jennifer K Martin, and Euan G Ritchie. 2013. A cost-effective and informative method of GPS tracking wildlife. *Wildlife Research* 40, 5 (2013), 345–348.

[4] Mark Andrachuk, Melissa Marschke, Charlotte Hings, and Derek Armitage. 2019. Smartphone technologies supporting community-based environmental monitoring and implementation: a systematic scoping review. *Biological Conservation* 237 (2019), 430–442.

[5] Bon Adriel Aseniero, Marios Constantinides, Sagar Joglekar, Ke Zhou, and Daniele Quercia. 2020. MeetCues: Supporting online meetings experience. In *2020 IEEE Visualization Conference (VIS)*. IEEE, 236–240.

[6] Carolyn Ashurst, Markus Anderljung, Carina Prunkl, Jan Leike, Yarin Gal, Toby Shevlane, and Allan Dafoe. 2020. A guide to writing the NeurIPS impact statement. *Centre for the Governance of AI. URL: https://perma. cc/B5R8-2B9V* (2020).

[7] Microsoft Azure. 2023. *Introduction to prompt engineering.* Retrieved September 8, 2023 from https://learn.microsoft.com/en-us/azure/ai-services/openai/concepts/prompt-engineering#best-practices

[8] Jakob E Bardram and Aleksandar Matic. 2020. A decade of ubiquitous computing research in mental health. *IEEE Pervasive Computing* 19, 1 (2020), 62–72.

[9] Konstantia Barmpatsalou, Tiago Cruz, Edmundo Monteiro, and Paulo Simoes. 2018. Mobile forensic data analysis: Suspicious pattern detection in mobile evidence. *IEEE Access* 6 (2018), 59705–59727.

[10] Jason Borenstein and Ayanna Howard. 2021. Emerging challenges in AI and the need for AI ethics education. *AI and Ethics* 1 (2021), 61–65.

[11] Tom B. Brown et al. 2020. Language Models Are Few-Shot Learners. In *Proceedings of the 34th International Conference on Neural Information Processing Systems* (Vancouver, BC, Canada) *(NIPS'20)*. Curran Associates Inc., Red Hook, NY, USA, Article 159, 25 pages.

[12] Attaullah Buriro, Bruno Crispo, Filippo Delfrari, and Konrad Wrona. 2016. Hold and sign: A novel behavioral biometrics for smartphone user authentication. In *2016 IEEE security and privacy workshops (SPW)*. IEEE, 276–285.

[13] Zana Buçinca, Chau Minh Pham, Maurice Jakesch, Marco Tulio Ribeiro, Alexandra Olteanu, and Saleema Amershi. 2023. AHA!: Facilitating AI Impact Assessment by Generating Examples of Harms. arXiv:2306.03280 [cs.HC]

[14] Courtni Byun, Piper Vasicek, and Kevin Seppi. 2023. Dispensing with Humans in Human-Computer Interaction Research. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 413, 26 pages. https://doi.org/10.1145/3544549.3582749

[15] Marta E Cecchinato, Anna L Cox, and Jon Bird. 2015. Smartwatches: the good, the bad and the ugly?. In *Proceedings of the 33rd Annual ACM Conference extended abstracts on human factors in computing systems*. 2133–2138.

[16] Apurva Chaudhari, Bhushan Mapari, and Shreenivas Jog. 2018. Effective environmental monitoring & domestic home conditions by implementation of IoT. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*. IEEE, 1–5.

[17] Siyuan Chen, Mengyue Wu, Kenny Q Zhu, Kunyao Lan, Zhiling Zhang, and Lyuchun Cui. 2023. LLM-empowered Chatbots for Psychiatrist and Patient Simulation: Application and Evaluation. *arXiv preprint arXiv:2305.13614* (2023). https://arxiv.org/abs/2305.13614

[18] Swathikan Chidambaram, Yathukulan Maheswaran, Kian Patel, Viknesh Sounderajah, Daniel A Hashimoto, Kenneth Patrick Seastedt, Alison H McGregor, Sheraz R Markar, and Ara Darzi. 2022. Using artificial intelligence-enhanced sensing and wearable technology in sports medicine and performance optimisation. *Sensors* 22, 18 (2022), 6920.

[19] Jun-Ho Choi, Marios Constantinides, Sagar Joglekar, and Daniele Quercia. 2021. KAIROS: Talking heads and moving bodies for successful meetings. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*. 30–36.

[20] Ushashi Chowdhury, Pranjal Chowdhury, Sourav Paul, Anwesha Sen, Partho Protim Sarkar, Shubhankur Basak, and Abari Bhattacharya. 2019. Multi-sensor wearable for child safety. In *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. IEEE, 0968–0972.

[21] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching Stories with Generative Pretrained Language Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 209, 19 pages. https://doi.org/10.1145/3491102.3501819

[22] Monica Ionita Ciolacu, Leon Binder, Paul Svasta, Ioan Tache, and Dan Stoichescu. 2019. Education 4.0–jump to innovation with IoT in higher education. In *2019 IEEE 25th International Symposium for Design and Technology in Electronic Packaging (SIITME)*. IEEE, 135–141.

[23] European Comission. 2021. *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts.* European Comission. Retrieved July 13, 2023 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

[24] European Comission. 2023. *Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts.* European Comission. Retrieved July 13, 2023 from https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html

[25] Marios Constantinides, Jonas Busk, Aleksandar Matic, Maria Faurholt-Jepsen, Lars Vedel Kessing, and Jakob E Bardram. 2018. Personalized versus generic mood prediction models in bipolar disorder. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 1700–1707.

[26] Marios Constantinides and Daniele Quercia. 2022. Good Intentions, Bad Inventions: How Employees Judge Pervasive Technologies in the Workplace. *IEEE Pervasive Computing* 22, 1 (2022), 69–76.

[27] Marios Constantinides, Sanja Šćepanović, Daniele Quercia, Hongwei Li, Ugo Sassi, and Michael Eggleston. 2020. ComFeel: Productivity is a matter of the senses too. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–21.

[28] Maria Cornacchia, Koray Ozcan, Yu Zheng, and Senem Velipasalar. 2016. A survey on activity detection and classification using wearable sensors. *IEEE Sensors Journal* 17, 2 (2016), 386–403.

[29] Barbara Corsetti, Raul Sanchez-Reillo, Richard M Guest, and Marco Santopietro. 2019. Face image analysis in mobile biometric accessibility evaluations. In *2019 International Carnahan Conference on Security Technology (ICCST)*. IEEE, 1–5.

[30] Kirsten Crager and Anindya Maiti. 2017. Information leakage through mobile motion sensors: User awareness and concerns. In *Proceedings of the European Workshop on Usable Security (EuroUSEC)*.

[31] Brian Cusack, Bryce Antony, Gerard Ward, and Shaunak Mody. 2017. Assessment of security vulnerabilities in wearable devices. (2017).

[32] Vedant Das Swain, Lan Gao, William A Wood, Srikruthi C Matli, Gregory D Abowd, and Munmun De Choudhury. 2023. Algorithmic Power or Punishment: Information Worker Perspectives on Passive Sensing Enabled AI Phenotyping of Performance and Wellbeing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

[33] Fidel Paniagua Diez, Diego Suárez Touceda, José María Sierra Cámara, and Sherali Zeadally. 2019. Lightweight access control system for wearable devices. *IT professional* 21, 1 (2019), 50–58.

[34] Gokila Dorai, Shiva Houshmand, and Sudhir Aggarwal. 2020. Data Extraction and Forensic Analysis for Smartphone Paired Wearables and IoT Devices.. In *HICSS*. 1–10.

[35] Michael Dowling and Brian Lucey. 2023. ChatGPT for (Finance) research: The Bananarama Conjecture. *Finance Research Letters* 53 (2023), 103662. https://doi.org/10.1016/j.frl.2023.103662

[36] Cynthia Dwork. 2008. Differential privacy: A survey of results. In *International conference on theory and applications of models of computation*. Springer, 1–19. https://doi.org/10.1007/978-3-540-79228-4_1

[37] Nouha Dziri, Sivan Milton, Mo Yu, Osmar Zaiane, and Siva Reddy. 2022. On the origin of hallucinations in conversational models: Is it the datasets or the models? *arXiv preprint arXiv:2204.07931* (2022). https://arxiv.org/abs/2204.07931

[38] Ronen Eldan and Yuanzhi Li. 2023. TinyStories: How Small Can Language Models Be and Still Speak Coherent English? *arXiv preprint arXiv:2305.07759* (2023). https://arxiv.org/abs/2305.07759

[39] Peppino Fazio, Miralem Mehic, and Miroslav Voznak. 2021. An innovative dynamic mobility sampling scheme based on multiresolution wavelet analysis in IoT networks. *IEEE Internet of Things Journal* 9, 13 (2021), 11336–11350.

[40] Richard R Fletcher, Ming-Zher Poh, and Hoda Eydgahi. 2010. Wearable sensors: Opportunities and challenges for low-cost health care. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE, 1763–1766.

[41] Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, and Effy Vayena. 2021. *An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*. Springer International Publishing, Cham, 19–39. https://doi.org/10.1007/978-3-030-81907-1_3

[42] Hana Frluckaj, Laura Dabbish, David G. Widder, Huilian Sophie Qiu, and James Herbsleb. 2022. Gender and Participation in Open Source Software Development. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–31. https://doi.org/10.1145/3555190

[43] Jaime Fuster, Sonia Solera-Cotanilla, Jaime Pérez, Mario Vega-Barbas, Rafael Palacios, Manuel Alvarez-Campana, and Gregorio Lopez. 2023. Analysis of security and privacy issues in wearables for minors. *Wireless Networks* (2023), 1–17.

[44] Craig Gentry. 2009. *A fully homomorphic encryption scheme*. Stanford university.

[45] Fabrizio Gilardi, Meysam Alizadeh, and Maël Kubli. 2023. ChatGPT outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences 120 (30)* (2023). https://doi.org/10.1073/pnas.2305016120

[46] Louie Giray. 2023. Prompt Engineering with ChatGPT: A Guide for Academic Writers. *Annals of Biomedical Engineering* (2023), 1–5. https://doi.org/10.1007/s10439-023-03272-4

[47] Delaram Golpayegani, Harshvardhan J. Pandit, and Dave Lewis. 2023. To Be High-Risk, or Not To Be–Semantic Specifications and Implications of the AI Act's High-Risk AI Applications and Harmonised Standards. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (Chicago, IL, USA) *(FAccT '23)*. ACM, New York, NY, USA, 905–915. https://doi.org/10.1145/3593013.3594050

[48] M Gor, Jayneel Vora, Sudeep Tanwar, Sudhanshu Tyagi, Neeraj Kumar, Mohammad S Obaidat, and Balqies Sadoun. 2017. GATA: GPS-Arduino based Tracking and Alarm system for protection of wildlife animals. In *2017 international conference on computer, information and telecommunication systems (CITS)*. IEEE, 166–170.

[49] Thilo Hagendorff. 2020. The ethics of AI ethics: An evaluation of guidelines. *Minds and machines* 30, 1 (2020), 99–120. https://doi.org/10.1007/s11023-020-09517-8

[50] Noah Yuval Harari. 2021. The world after coronavirus.

[51] Muhammad Haris, Hamed Haddadi, and Pan Hui. 2014. Privacy leakage in mobile computing: Tools, methods, and characteristics. *arXiv preprint arXiv:1410.4978* (2014).

[52] Anke Hassel and Didem Özkiziltan. 2023. Governing the work-related risks of AI: implications for the German government and trade unions. *Transfer: European Review of Labour and Research* 29, 1 (2023), 71–86. https://doi.org/10.1177/10242589221147228

[53] Alexandra Heidel, Christian Hagist, et al. 2020. Potential benefits and risks resulting from the introduction of health apps and wearables into the German statutory health care system: scoping review. *JMIR mHealth and uHealth* 8, 9 (2020), e16444.

[54] Amruta Helwatkar, Daniel Riordan, and Joseph Walsh. 2014. Sensor technology for animal health monitoring. *International Journal on Smart Sensing and Intelligent Systems* 7, 5 (2014), 1–6.

[55] Jessica Hoth, Björn Schwarz, Gabriele Kaiser, Andreas Busse, Johannes König, and Sigrid Blömeke. 2016. Uncovering predictors of disagreement: ensuring the quality of expert ratings. *ZDM* 48 (2016), 83–95.

[56] Qazi Mudassar Ilyas and Muneer Ahmad. 2020. Smart farming: An enhanced pursuit of sustainable remote livestock tracking and geofencing using IoT and GPRS. *Wireless communications and mobile computing* 2020 (2020), 1–12.

[57] International Federation of Accountants. 2018. *Regulatory divergence: costs, risks, impacts*. IFAC. Retrieved April 25, 2024 from https://www.ifac.org/_flysystem/azure-private/publications/files/IFAC-OECD-Regulatory-Divergence.pdf

[58] Anna Jobin, Marcello Ienca, and Effy Vayena. 2019. The global landscape of AI ethics guidelines. *Nature machine intelligence* 1, 9 (2019), 389–399. https://doi.org/10.1038/s42256-019-0088-2

[59] Apurva Joshi, I Naga VishnuKanth, Navkar Samdaria, Sumit Bagla, and Prabhat Ranjan. 2008. GPS-less animal tracking system. In *2008 Fourth International Conference on Wireless Communication and Sensor Networks*. IEEE, 120–125.

[60] Atoosa Kasirzadeh. 2021. Reasons, Values, Stakeholders: A Philosophical Framework for Explainable Artificial Intelligence. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery. https://doi.org/10.1145/3442188.3445866

[61] Heather Kelly. 2014. *Fitbit recalls activity tracker due to skin rashes*. CNN. https://edition.cnn.com/2014/02/21/tech/gaming-gadgets/fitbit-force-recall/

[62] Suprava Ranjan Laha, Binod Kumar Pattanayak, and Saumendra Pattnaik. 2022. Advancement of Environmental Monitoring System Using IoT and Sensor: A Comprehensive Analysis. *AIMS Environmental Science* 9, 6 (2022), 771–800.

[63] Mina Lee, Percy Liang, and Qian Yang. 2022. CoAuthor: Designing a Human-AI Collaborative Writing Dataset for Exploring Language Model Capabilities. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 388, 19 pages. https://doi.org/10.1145/3491102.3502030

[64] He Li, Jing Wu, Yiwen Gao, and Yao Shi. 2016. Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. *International journal of medical informatics* 88 (2016), 8–17.

[65] Hsin-Chang Lin, Ming-Jen Chen, and Jung-Tang Huang. 2022. An IoT-based smart healthcare system using location-based mesh network and big data analytics. *Journal of Ambient Intelligence and Smart Environments* Preprint (2022), 1–27.

[66] Nelson F. Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang. 2023. Lost in the middle: How language models use long contexts. *arXiv preprint arXiv:2307.03172* (2023). https://arxiv.org/abs/2307.03172

[67] Steve Mann, Jason Nolan, and Barry Wellman. 2003. Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments. *Surveillance & society* 1, 3 (2003), 331–355.

[68] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-Rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019). https://doi.org/10.1145/3359174

[69] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*. PMLR, 1273–1282. https://proceedings.mlr.press/v54/mcmahan17a/mcmahan17a.pdf

[70] Sarah McNary and Aaron Hunter. 2018. Using Wearable Device Data in a Criminal Investigation: A Preliminary Study. In *2018 European Intelligence and Security Informatics Conference (EISIC)*. IEEE, 85–85.

[71] Lakmal Meegahapola, Marios Constantinides, Zoran Radivojevic, Hongwei Li, Daniele Quercia, and Michael S Eggleston. 2023. Quantified Canine: Inferring Dog Personality From Wearables. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[72] Jacob Metcalf, Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, and Madeleine Clare Elish. 2021. Algorithmic impact assessments and accountability: The co-construction of impacts. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, 735–746. https://doi.org/10.1145/3442188.3445935

[73] Microsoft. 2022. *Microsoft's framework for building AI systems responsibly*. Retrieved January 22, 2024 from https://blogs.microsoft.com/on-the-issues/2022/06/21/microsofts-framework-for-building-ai-systems-responsibly/

[74] Matthew Miles and Michael Huberman. 1994. *Qualitative Data Analysis: A Methods Sourcebook*. Sage.

[75] Adam J Mills, Richard T Watson, Leyland Pitt, and Jan Kietzmann. 2016. Wearing safe: Physical and informational security in the age of the wearable device. *Business Horizons* 59, 6 (2016), 615–622.

[76] Shayan Mirjafari, Kizito Masaba, Ted Grover, Weichen Wang, Pino Audia, Andrew T Campbell, Nitesh V Chawla, Vedant Das Swain, Munmun De Choudhury, Anind K Dey, et al. 2019. Differentiating higher and lower job performers in the workplace using mobile sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 1–24.

[77] Miro. 2022. *Miro | Online Whiteboard for Visual Collaboration*. Retrieved August 2022 from https://miro.com/

[78] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency*. 220–229.

[79] Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society* 3, 2 (2016), 2053951716679679. https://doi.org/10.1177/2053951716679679

[80] Elsayed Said Mohamed, AA Belal, Sameh Kotb Abd-Elmabod, Mohammed A El-Shirbeny, A Gad, and Mohamed B Zahran. 2021. Smart farming for improving agricultural management. *The Egyptian Journal of Remote Sensing and Space Science* 24, 3 (2021), 971–981.

[81] Thiago Guimarães Moraes, Eduarda Costa Almeida, and José Renato Laranjeira de Pereira. 2021. Smile, you are being identified! Risks and measures for the use of facial recognition in (semi-) public spaces. *AI and Ethics* 1, 2 (2021), 159–172. https://doi.org/10.1007/s43681-020-00014-3

[82] Sinead E Morris, Jonathan L Zelner, Deborah A Fauquier, Teresa K Rowles, Patricia E Rosel, Frances Gulland, and Bryan T Grenfell. 2015. Partially observed epidemics in wildlife hosts: modelling an outbreak of dolphin morbillivirus in the northwestern Atlantic, June 2013–2014. *Journal of the Royal Society Interface* 12, 112 (2015), 20150676.

[83] National Institute of Standards and Technology. 2023. . Retrieved May 2023 from https://www.nist.gov/itl/ai-risk-management-framework

[84] National Institute of Standards and Technology. 2023. *The EqualAI Algorithmic Impact Assessment Tool.* Retrieved January 2024 from https://www.equalai.org/aia/

[85] Suresh Neethirajan. 2017. Recent advances in wearable sensors for animal health management. *Sensing and Bio-Sensing Research* 12 (2017), 15–29. https://doi.org/10.1016/j.sbsr.2016.11.004

[86] Henry Friday Nweke, Ying Wah Teh, Ghulam Mujtaba, and Mohammed Ali Al-Garadi. 2019. Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions. *Information Fusion* 46 (2019), 147–170.

[87] Khaled Obaideen, Bashria AA Yousef, Maryam Nooman AlMallahi, Yong Chai Tan, Montaser Mahmoud, Hadi Jaber, and Mohamad Ramadan. 2022. An overview of smart irrigation systems using IoT. *Energy Nexus* (2022), 100124.

[88] Alexandra Olteanu, Michael Ekstrand, Carlos Castillo, and Jina Suh. 2023. Responsible AI Research Needs Impact Statements Too. *arXiv preprint arXiv:2311.11776* (2023).

[89] Ömer Aydın and Enis Karaarslan. 2022. OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare. *SSRN Electronic Journal* (2022), 22–31. https://doi.org/10.2139/ssrn.4308687

[90] OpenAI. 2023. *Research on GPT-4 - latest updates.* Retrieved August 7, 2023 from https://openai.com/research/gpt-4

[91] Sungkyu Park, Assem Zhunis, Marios Constantinides, Luca Maria Aiello, Daniele Quercia, and Meeyoung Cha. 2023. Social dimensions impact individual sleep quantity and quality. *Scientific Reports* 13, 1 (2023), 1–11.

[92] Ignacio Perez-Pozuelo, Dimitris Spathis, Emma AD Clifton, and Cecilia Mascolo. 2021. Wearables, smartphones, and artificial intelligence for digital phenotyping and health. In *Digital Health.* Elsevier, 33–54.

[93] Stefano Puntoni, Rebecca Walker Reczek, Markus Giesler, and Simona Botti. 2021. Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing* 85, 1 (2021), 131–151.

[94] James H Read. 2009. *Majority rule versus consensus: the political thought of John C. Calhoun.* University Press of Kansas.

[95] Antonio Rios-Navarro, Juan Pedro Dominguez-Morales, Ricardo Tapiador-Morales, Manuel Dominguez-Morales, Angel Jimenez-Fernandez, and Alejandro Linares-Barranco. 2016. A sensor fusion horse gait classification by a spiking neural network on SpiNNaker. In *Artificial Neural Networks and Machine Learning–ICANN 2016: 25th International Conference on Artificial Neural Networks, Barcelona, Spain, September 6-9, 2016, Proceedings, Part I 25.* Springer, 36–44.

[96] Jenna M Roper, Jose F Garcia, and Hideaki Tsutsui. 2021. Emerging technologies for monitoring plant health in vivo. *ACS omega* 6, 8 (2021), 5101–5107.

[97] Johnny Saldaña. 2015. *The Coding Manual for Qualitative Researchers.* Sage.

[98] Simone Saraceni, Andrea Claudi, and Aldo Franco Dragoni. 2012. An active monitoring system for real-time face-tracking based on mobile sensors. In *Proceedings ELMAR-2012.* IEEE, 53–56.

[99] Deepti Saraswat, Pronaya Bhattacharya, Ashwin Verma, Vivek Kumar Prasad, Sudeep Tanwar, Gulshan Sharma, Pitshou N Bokoro, and Ravi Sharma. 2022. Explainable AI for healthcare 5.0: opportunities and challenges. *IEEE Access* (2022).

[100] Elvis Saravia. 2022. *Prompt Engineering Guide.* DAIR.AI. Retrieved August 7, 2023 from https://github.com/dair-ai/Prompt-Engineering-Guide

[101] Mike Schaekermann, Graeme Beaton, Minahz Habib, Andrew Lim, Kate Larson, and Edith Law. 2019. Understanding expert disagreement in medical data analysis through structured adjudication. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–23.

[102] Karl Schriewer and Grzegorz Bulaj. 2016. Music streaming services as adjunct therapies for depression, anxiety, and bipolar symptoms: convergence of digital technologies, mobile apps, emotions, and global mental health. *Frontiers in public health* 4 (2016), 217.

[103] Dhruv R Seshadri, Evan V Davies, Ethan R Harlow, Jeffrey J Hsu, Shanina C Knighton, Timothy A Walker, James E Voos, and Colin K Drummond. 2020. Wearable sensors for COVID-19: a call to action to harness our digital infrastructure for remote patient monitoring and virtual assessments. *Frontiers in Digital Health* (2020), 8.

[104] Ashish Sharma, Kevin Rushton, Inna Wanyin Lin, David Wadden, Khendra G. Lucas, Adam S Miner, Theresa Nguyen, and Tim Althoff. 2023. Cognitive Reframing of Negative Thoughts through Human-Language Model Interaction. *arXiv preprint arXiv:2305.02466* (2023). https://arxiv.org/abs/2305.02466

[105] Eli Sherman and Ian W. Eisenberg. 2023. AI Risk Profiles: A Standards Proposal for Pre-Deployment AI Risk Disclosures. arXiv:2309.13176 [cs.AI]

[106] Jessica Shieh. 2023. *Best practices for prompt engineering with OpenAI API.* OpenAI. Retrieved August 7, 2023 from https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-openai-api

[107] Michael W Sjoding, Robert P Dickson, Theodore J Iwashyna, Steven E Gay, and Thomas S Valley. 2020. Racial bias in pulse oximetry measurement. *New England Journal of Medicine* 383, 25 (2020), 2477–2478.

[108] Colin D Sumrall, Christopher A Brochu, and John W Merck. 2001. Global lability, regional resolution, and majority-rule consensus bias. *Paleobiology* 27, 2 (2001), 254–261.

[109] Dhiraj Sunehra, V Sai Sreshta, V Shashank, and B Uday Kumar Goud. 2020. Raspberry pi based smart wearable device for women safety using gps and gsm technology. In *2020 IEEE International Conference for Innovation in Technology (INOCON)*. IEEE, 1–5.

[110] Mohammad Tahaei, Marios Constantinides, Daniele Quercia, Sean Kennedy, Michael Muller, Simone Stumpf, Q Vera Liao, Ricardo Baeza-Yates, Lora Aroyo, Jess Holbrook, et al. 2023. Human-Centered Responsible Artificial Intelligence: Current & Future Trends. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–4.

[111] Suzanne Tolmeijer, Markus Christen, Serhiy Kandul, Markus Kneer, and Abraham Bernstein. 2022. Capable but Amoral? Comparing AI and Human Expert Collaboration in Ethical Decision Making. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 160, 17 pages. https://doi.org/10.1145/3491102.3517732

[112] John Torous, JP Onnela, and Matcheri Keshavan. 2017. New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. *Translational psychiatry* 7, 3 (2017), e1053–e1053.

[113] Kirti Tyagi, Aabha Karmarkar, Simran Kaur, Sukanya Kulkarni, and Rita Das. 2020. Crop health monitoring system. In *2020 International Conference for Emerging Technology (INCET)*. IEEE, 1–5.

[114] United Nations: Department of Economic and Social Affairs Sustainable Development. 2023. *The 17 Goals*. Retrieved November 8, 2023 from https://sdgs.un.org/goals

[115] Ronda Venkateswarlu, Choo Hui Wei, Ngiam Li Lian, ET Lim, Zijian Zhu, and Mingjiang Yang. 2006. Wearable high-tech gear for homeland security personnel. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense V*, Vol. 6201. SPIE, 382–391.

[116] Filippo Vurro, Riccardo Manfredi, Manuele Bettelli, Gionata Bocci, Alberto Luigi Cologni, Sandro Cornali, Roberto Reggiani, Edoardo Marchetti, Nicola Coppedè, Stefano Caselli, et al. 2023. In vivo sensing to monitor tomato plants in field conditions and optimize crop water management. *Precision Agriculture* (2023), 1–21.

[117] Jake Wall, George Wittemyer, Brian Klinkenberg, and Iain Douglas-Hamilton. 2014. Novel opportunities for wildlife conservation and research with real-time monitoring. *Ecological Applications* 24, 4 (2014), 593–601.

[118] Chen Wei, Yun Cheng Wang, Bin Wang, and C.-C. Jay Kuo. 2023. An Overview on Language Models: Recent Developments and Outlook. *ArXiv* abs/2303.05759 (2023). https://api.semanticscholar.org/CorpusID:257482688

[119] Jason Wei, Maarten Bosma, Vincent Y. Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V. Le. 2022. Finetuned language models are zero-shot learners. *The International Conference on Learning Representations (ICLR)* (2022).

[120] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc V. Le, and Denny Zhou. 2022. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. In *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Eds.), Vol. 35. Curran Associates, Inc., 24824–24837. https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf

[121] Kok-Seng Wong and Myung Ho Kim. 2016. An enhanced user authentication solution for mobile payment systems using wearables. *Security and Communication Networks* 9, 17 (2016), 4639–4649.

[122] Tongshuang Wu, Michael Terry, and Carrie Jun Cai. 2022. AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 385, 22 pages. https://doi.org/10.1145/3491102.3517582

[123] Xiangyun Xu and Wei Ma. 2023. The regulation of the body by smart wearable devices and their social risk progression. *Wearable Technology* 2, 1 (2023), 60–72.

[124] Liyun Yang, Ke Lu, Jose A Diaz-Olivares, Fernando Seoane, Kaj Lindecrantz, Mikael Forsman, Farhad Abtahi, and JöRgen AE Eklund. 2018. Towards smart work clothing for automatic risk assessment of physical workload. *Ieee Access* 6 (2018), 40059–40072.

[125] Zhice Yang, Qianyi Huang, and Qian Zhang. 2018. Backscatter as a covert channel in mobile devices. *GetMobile: Mobile Computing and Communications* 22, 1 (2018), 31–34.

[126] Zhe Zheng, Ke-Yin Chen, Xin-Yu Cao, Xin-Zheng Lu, and Jia-Rui Lin. 2023. LLM-FuncMapper: Function Identification for Interpreting Complex Clauses in Building Codes via LLM. *arXiv preprint arXiv:2308.08728* (2023).

[127] Ke Zhou, Marios Constantinides, Daniele Quercia, and Sanja Šćepanović. 2023. How Circadian Rhythms Extracted From Social Media Relate to Physical Activity and Sleep. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 948–959.

[128] Shoshana Zuboff. 2023. The age of surveillance capitalism. In *Social Theory Re-Wired*. Routledge, 203–213.

APPENDIX

Table 2. List of LLM-generated AI uses for mobile and wearable technologies. Each entry includes an ID, description, an indication of whether it is high risk, the goals it supports, and the specific clauses from the EU AI Act that categorize the use as high risk or not. Use #64 is classified as unacceptable, and was marked with a (*).

ID	Description	High Risk	SDG Goals	AIActText
1	Enhancing security through facial recognition using high-performance cameras in smartphones	Yes	16, 5, 8	Article 6, Amendment 711 - Biometric identification and categorization of natural persons
2	Monitoring and analyzing crowd behavior using mobile cameras	Yes	16, 11, 5, 8, 3	Article 6, Amendment 711 - Biometric identification and categorization of natural persons
3	Identifying individuals in criminal investigations using smartphones and wearable watches	Yes	16, 10, 5, 8, 3, 11	Article 6, Amendment 711 - Biometric identification and categorization of natural persons
4	Monitoring children's location for safety using wristbands	Yes	3, 11, 16, 5, 8	Amendment 38 - AI Systems for perceiving or exploiting vulnerabilities of children and people
5	Monitoring elderly health and activity using wristbands	Yes	3, 11	Amendment 38 - AI Systems for perceiving or exploiting vulnerabilities of children and people
6	Coordinating family calendar and tasks on mobile phones	No	5, 8, 12	N/A
7	Sharing calendars and reminders on mobile phones	No	5, 8, 12	N/A
8	Sharing location for coordination through smartwatches	No	11, 5, 3	N/A
9	Tracking emotions using mobile app for improving emotional well-being	Yes	5, 3	Article 6, Amendment 712 Biometric identification and categorization of natural persons
10	Tracking health status using smart rings for chronic conditions	Yes	8, 3	Article 6 - Access to essential private services
11	Tracking movements using accelerometers for improving physical therapy outcomes	Yes	8, 3	Article 6 - Access to essential private services
12	Enhancing mental health treatment using smartwatches	Yes	8, 3	Article 6 - Biometric identification and categorization of natural persons
13	Promoting physical fitness through activity tracking on smartwatches	No	5, 8, 3	N/A
14	Tracking sleep quality	No	8, 3	N/A
15	Providing meditation guidance on a mobile app	No	8, 3	N/A
16	Performing gesture recognition using accelerometers for UI accessibility	No	4, 8, 3, 9	N/A
17	Tracking motion and environmental interaction using VR headsets	No	8, 4, 3	N/A
18	Enabling hands-free device control through voice recognition and command execution	No	8, 4, 3	N/A
19	Tracking expenses and monitoring personal finances using a mobile app	No	12, 5, 8, 4, 1	N/A
20	Enabling secure mobile payments and transfers with a smartwatch	Yes	17, 10, 8, 1	Article 6, Amendment 64 - AI systems as safety component
21	Providing investment and financial advice using a mobile app	Yes	10, 5, 8, 4, 1	Article 6, Amendment 64 - AI systems as safety component
22	Delivering interactive educational content using a mobile app	Yes	4, 8	Article 6 - Education and vocational training
23	Providing personalized learning recommendations using a tablet	Yes	4, 8	Amendment 715, 716 - AI systems for assigning natural persons to educational and vocational training
24	Enabling online exams using a 360 degree camera	Yes	4, 17, 10, 8	Article 6 - Education and vocational training
25	Monitoring employee productivity using smart badges	Yes	8, 4, 5, 16	Article 6 - AI systems as safety component
26	Monitoring physical movements for workplace safety using smartwatches	Yes	3, 8, 5, 16, 11	Article 6 - AI systems as safety component and for monitoring and evaluating performance
27	Tracking work hours and tasks during remote work using a mobile app	Yes	12, 5, 8, 4, 3	Article 6 - AI systems as safety component and for monitoring and evaluating performance
28	Tracking travel patterns for public transportation efficiency using smart cards	Yes	17, 13, 12, 11, 5, 9, 8	AI Article 6 - AI systems as safety component
29	Monitoring health status and symptoms using headbands	Yes	10, 9, 8, 3	AI Amendment 721 - evaluate the eligibility of natural persons for public assistance benefits and services
30	Tracking eligibility criteria and application status for social benefits using a mobile app	Yes	10, 8, 5, 3, 1	AI Amendment 721 - evaluate the eligibility of natural persons for public assistance benefits and services
31	Tracking user preferences for personalized online shopping experience using a mobile app	No	12, 8	N/A
32	Customizing music and video recommendations using earbuds	Yes	12, 8	Amendment 740 - AI systems for recommending user-generated content available on the platform
33	Tailoring news and information feeds using a mobile app	Yes	5, 4, 8, 12	Amendment 740 - AI systems for recommending user-generated content available on the platform

Table 2. List of LLM-generated AI uses for mobile and wearable technologies. Each entry includes an ID, description, an indication of whether it is high risk, the goals it supports, and the specific clauses from the EU AI Act that categorize the use as high risk or not. Use #64 is classified as unacceptable, and was marked with a (*).

ID	Description	High Risk	SDG Goals	AIActText
34	Tracking user interactions for enhancing user engagement using a mobile app	No	8, 12	N/A
35	Facilitating social connections using a mobile app	Yes	3, 8	Amendment 740 - AI systems for recommending user-generated content available on the platform
36	Monitoring harmful content using a mobile app	Yes	10, 12, 5, 4, 3	Amendment 740 - AI systems for recommending user-generated content available on the platform
37	Monitoring physical activity and vital signs for improving athletic performance using armbands	No	5, 3, 4	N/A
38	Tracking environmental conditions in outdoor spaces using a mobile app	No	5, 3, 4, 8, 11, 16	N/A
39	Monitoring player health for team management using armbands	No	5, 3, 4	N/A
40	Tracking user motion for enhancing user experience using VR headsets	No	5, 8, 3, 4	N/A
41	Monitoring audience reactions for facilitating interactive performances using mobile cameras	Yes	5, 8, 4	Amendment 52, 191 - AI for inferring emotions, thoughts, states of mind individuals or groups
42	Analyzing user preferences for personalized content recommendations using a mobile app	No	5, 8, 4	N/A
43	Preventing unauthorized access through biometric authentication using mobile scanners	Yes	16, 5, 8	Article 6, Amendment 711 Biometric identification and categorization of natural persons
44	Analyzing network traffic for security threats using mobile dashboards	Yes	16, 5, 8, 9	AI Article 6, Amendment 713 - AI systems as safety component
45	Tracking user location and providing emergency alerts for personal safety using smartwatches	Yes	5, 8, 3, 16, 11	N/A
46	Analyzing user behavior for personalized advertising using a mobile app	No	8	N/A
47	Tracking user interactions with ads for customer engagement using a mobile app	No	8	N/A
48	Analyzing marketing campaign performance data using a mobile app	No	8, 12	N/A
49	Monitoring crop health and growth using IoT sensors	No	17, 11, 13, 9, 12, 2, 15, 8	N/A
50	Tracking livestock health and location using GPS collars	No	13, 12, 17, 11, 9, 2, 15, 8	N/A
51	Optimizing irrigation through soil moisture detection using IoT sensors	No	17, 13, 9, 12, 2, 11, 15, 8, 6	N/A
52	Improving customer engagement through personalized recommendations using a mobile app	No	12, 8	N/A
53	Automating task management in business operations using a tablet	No	12, 17, 9, 8	N/A
54	Enhancing product development using customer feedback from mobile apps	No	12, 8	N/A
55	Facilitating autonomous navigation using mobile cameras	No	11, 9, 8, 15	N/A
56	Enabling human-robot interaction using mobile cameras	Yes	9, 3, 8	Article 6 - AI as safety component
57	Improving the quality of multi-sensor data fusion in fall detection smartwatches	Yes	9, 3, 8	AI Article 6 - AI systems as safety component
58	Accelerating scientific discovery through analyzing data from sensor-equipped backpack	No	17, 12, 9, 8, 4	N/A
59	Facilitating remote collaboration using mobile cameras	No	17, 13, 9, 8, 4	N/A
60	Improving research data management using a mobile app	No	17, 9, 8, 4	N/A
61	Monitoring infrastructure health using pressure sensors	Yes	17, 11, 3, 15, 13, 8, 16, 9, 12, 7	Article 6 - AI systems for management and operation of critical infrastructure
62	Predicting infrastructure failures using wearable sensors	Yes	17, 11, 3, 15, 13, 8, 16, 9, 12, 7	Article 6 - AI systems for management and operation of critical infrastructure
63	Optimizing infrastructure operations using tablets	Yes	17, 11, 3, 15, 13, 8, 16, 9, 6, 12, 7	Article 6 - AI systems for management and operation of critical infrastructure
64	Enhancing surveillance capabilities using mobile cameras	*	16, 5, 8, 11	Article 5, 6 - Remote biometric identification systems in publicly accessible spaces used by law enforcement
65	Improving crime prediction using mobile cameras	Yes	16, 5, 8	Article 6 - AI Systems for predicting criminal offence based on profiling of natural persons
66	Facilitating evidence collection using a mobile app	Yes	16, 5, 8	Amendment 728 - AI Systems for investigation or prosecution of criminal offences
67	Enhancing border security using mobile fingerprint scanners	Yes	16, 5, 8, 17, 10, 11	Amendment 736 - AI Systems for border control management
68	Improving asylum application processing using mobile document scanners	Yes	16, 1, 3, 5, 8, 17, 10, 11	Amendment 735 - AI Systems for assessment in relation to applications for asylum
69	Tracking migrant location using GPS	Yes	16, 5, 17, 10	Amendment 736 - AI Systems for border control management

Table 2. List of LLM-generated AI uses for mobile and wearable technologies. Each entry includes an ID, description, an indication of whether it is high risk, the goals it supports, and the specific clauses from the EU AI Act that categorize the use as high risk or not. Use #64 is classified as unacceptable, and was marked with a (*).

ID	Description	High Risk	SDG Goals	AIActText
70	Enhancing voter participation using a mobile app	Yes	16, 4	Amendment 739 - AI Systems for influencing the outcome of an election or referendum or the voting behaviour
71	Improving election transparency using a mobile app	Yes	16, 10, 8, 9	Amendment 739 - AI Systems for influencing the outcome of an election or referendum or the voting behaviour
72	Facilitating citizen engagement using digital kiosks	No	16, 11, 5, 17, 8, 4, 9	N/A
73	Enhancing immersive storytelling experiences using smart glasses	No	4	N/A
74	Facilitating real-time language translation using earbuds	No	4, 8, 10	N/A
75	Improving content accessibility using mobile screen readers	No	16, 8, 4, 9	N/A
76	Assisting visually impaired individuals using smart canes	Yes	11, 16, 4, 10, 8, 3	Amendment 38 - AI Systems for perceiving or exploiting vulnerabilities of children and people
77	Supporting individuals with hearing impairment using earbuds	Yes	11, 16, 4, 10, 8, 3	Amendment 38 - AI Systems for perceiving or exploiting vulnerabilities of children and people
78	Tracking movement for assisting individuals with mobility issues using accelerometers	Yes	11, 16, 4, 10, 8, 3	Article 6 - AI system as a safety component
79	Monitoring energy consumption using infrared sensors	No	13, 12, 17, 11, 15, 9, 7	N/A
80	Optimizing renewable energy production using heat sensors	No	13, 12, 17, 11, 15, 9, 7	N/A
81	Improving energy efficiency using infrared sensors	Yes	13, 12, 11, 15, 9, 7	Article 6 - AI system as a safety component
82	Enhancing soldier training using VR headsets	Yes	16, 17, 9, 8, 4	Article 6 - AI system as a safety component
83	Improving battlefield situational awareness using smart glasses	Yes	16, 17, 9, 8, 4	Article 6 - AI system as a safety component
84	Enhancing soldier health monitoring using smart belts	Yes	16, 17, 9, 8, 4, 3	Article 6 - AI system as a safety component
85	Facilitating remote court proceedings using a mobile app	Yes	17, 10, 16, 11, 8, 9	Amendment 738 – AI systems for assisting in judicial authority in researching and interpreting facts
86	Improving voter registration and verification using a mobile app	Yes	17, 16, 10, 11, 9	Article 6 - AI Systems for administration of justice and democratic processes
87	Enhancing public participation in democratic processes using smart kiosks	Yes	17, 16, 10, 11, 5, 8, 4, 9	Amendment 738 – AI systems for assisting in judicial authority in researching and interpreting facts
88	Improving public service delivery using a mobile app	Yes	16, 11, 5, 8, 4, 3, 9	Amendment 67 - AI for applying for or receiving public assistance benefits and services
89	Enhancing disaster management using accelerometers	Yes	17, 13, 9, 8, 3, 11, 16, 14, 15	Amendment 724 - AI systems for dispatching of emergency first response services
90	Improving traffic management through cameras	Yes	13, 12, 11, 16, 9	Article 6 - AI systems as safety component
91	Facilitating diplomatic communication using earbuds	No	16, 17, 10, 8, 4	N/A
92	Monitoring global protest actions using a mobile app	No	16, 17, 5, 13, 9, 11	N/A
93	Improving international negotiation processes using a mobile app	No	16, 17, 10, 4	N/A
94	Enhancing food quality inspection using mobile cameras	Yes	2, 12, 3, 17, 11, 9, 8	Amendment 207 - AI Systems for critical infrastructure
95	Improving food traceability using RFID tags	Yes	2, 12, 3, 17, 11, 9, 8	Amendment 207 - AI Systems for critical infrastructure
96	Enhancing food safety compliance using mobile cameras	Yes	2, 12, 3, 17, 11, 9, 8	Amendment 207 - AI Systems for critical infrastructure
97	Enhancing disaster response using GPS	Yes	17, 13, 8, 3, 16, 9, 11	Amendment 724 - AI systems for dispatching of emergency first response services
98	Monitoring health status of victims using wristbands	Yes	3, 9, 8, 11	Amendment 724 - AI systems for dispatching of emergency first response services
99	Coordinating rescue operations using a mobile app	Yes	17, 13, 8, 9, 3, 11, 16	Amendment 724 - AI systems for dispatching of emergency first response services.
100	Improving aid distribution efficiency using motion sensors	Yes	12, 17, 16, 11, 1, 2, 3, 13, 8, 9	Amendment 721 - AI systems for evaluating the eligibility of natural persons for public assistance
101	Monitoring health status of aid recipients using wristbands	Yes	17, 3, 9, 8, 11	Amendment 721 - AI systems for evaluating the eligibility of natural persons for public assistance
102	Facilitating communication in disaster zones using a mobile app	No	16, 17, 13, 8, 9, 11	N/A
103	Improving fleet management efficiency using GPS	No	12, 11, 17, 9, 8	N/A
104	Enhancing driver safety using motion sensors	Yes	3, 12, 11, 8, 9, 16	Amendment 713 - AI Systems for management and operation of road traffic

Table 2. List of LLM-generated AI uses for mobile and wearable technologies. Each entry includes an ID, description, an indication of whether it is high risk, the goals it supports, and the specific clauses from the EU AI Act that categorize the use as high risk or not. Use #64 is classified as unacceptable, and was marked with a (*).

ID	Description	High Risk	SDG Goals	AIActText
105	Optimizing fuel consumption using pressure sensors	No	13, 12, 11, 9, 7, 15	N/A
106	Monitoring urban mobility patterns using GPS	Yes	13, 12, 11, 16, 9	Article 6 - AI Systems as safety component
107	Assessing public space usage using mobile cameras	Yes	16, 11, 12, 9	Article 6 - Biometric identification and categorisation of natural persons
108	Improving public transportation using mobile cameras	Yes	12, 11, 9	Article 6 - AI Systems as safety component
109	Detecting suspicious behavior using motion sensors	Yes	16, 10, 17, 8, 3, 11	Article 6 - Biometric identification and categorisation of natural persons
110	Enhancing surveillance capabilities for counter-terrorism using GPS	Yes	16, 10, 17, 8, 3, 11, 9	Article 6 - Biometric identification and categorisation of natural persons
111	Improving communication in field operations using earbuds	Yes	17, 16, 8, 9, 4	Article 6 - AI Systems as safety component
112	Monitoring environmental conditions using heat sensors	No	13, 17, 12, 6, 3, 14, 9, 15, 8, 11, 7	N/A
113	Tracking wildlife movements using GPS collars	No	17, 14, 15, 13, 9, 8	N/A
114	Assessing resource usage efficiency using motion sensors	No	13, 12, 17, 11, 9, 6, 7, 15	N/A
115	Enhancing cross-border surveillance using mobile fingerprint scanners	Yes	16, 17, 10, 8, 11, 9	Amendment 736 - AI Systems for border control management
116	Improving communication in joint operations using smart glasses	Yes	16, 17, 8, 9, 4	Article 6 - AI Systems as safety component
117	Detecting illegal activities using motion sensors	Yes	16, 10, 11, 8, 5	Amendment 209 - AI Systems used for determining social behavior
118	Monitoring climate change using heat sensors	No	13, 17, 11, 6, 9, 15, 10, 7	N/A
119	Tracking impact of climate change on wildlife using GPS collars and heat sensors	No	17, 11, 15, 14, 13, 9, 8	N/A
120	Assessing effectiveness of energy-saving interventions using smart electricity meters	No	13, 12, 17, 6, 11, 7, 9, 8, 14, 15	N/A
121	Enhancing immersive gaming experience using VR headsets	No	9	N/A
122	Tracking physical activity for improving physical fitness using wristbands	No	8, 3, 4	N/A
123	Facilitating virtual social interaction using VR headsets	No	4, 8	N/A
124	Tracking and improving musical performance using a mobile app	No	4, 8	N/A
125	Monitoring gardening activities using motion sensors	No	11, 12, 2, 15, 8	N/A
126	Enhancing photography skills using mobile cameras	No	4, 8	N/A
127	Improving home security using mobile cameras	Yes	5, 11	Article 6 - AI Systems as safety component
128	Enhancing energy efficiency using heat sensors	Yes	13, 11, 12, 9, 7	Article 6 - AI system as a safety component
129	Facilitating home automation using smart thermostats	Yes	13, 12, 9	Article 6 - AI system as a safety component
130	Improving public safety using mobile cameras	Yes	11, 5, 8, 3, 16	Article 6 - AI Systems as safety component
131	Enhancing accessibility for disabled individuals using smart glasses	Yes	11, 10, 8, 3	Article 6 - AI Systems for access to essential private services and public services and benefits
132	Facilitating disaster response using a mobile app	Yes	17, 13, 9, 8, 3, 11, 16, 10	Amendment 724 - AI systems for dispatching of emergency first response services
133	Improving road safety using mobile cameras	Yes	3, 11, 9, 8, 16	Article 6 - AI Systems as safety component
134	Enhancing navigation and route planning using GPS	No	12, 11, 8, 9	Article 6 - AI systems as safety component
135	Facilitating autonomous driving using motion sensors	Yes	11, 9, 8	Article 6 - AI systems as safety component
136	Enhancing language learning using a mobile app	No	4, 8	N/A
137	Improving remote communication using mobile cameras	No	4, 8, 9	N/A
138	Facilitating social networking using a mobile app	No	4, 8, 9	N/A