Technical Solutions to Emotion AI's Privacy Harms: A Systematic Literature Review

Shreya Chowdhary
University of Michigan
School of Information
Ann Arbor, MI, USA
schowdha@umich.edu

Alexis Shore Ingber
University of Michigan
School of Information
Ann Arbor, MI, USA
ingber@umich.edu

Nazanin Andalibi
University of Michigan
School of Information
Ann Arbor, MI, USA
andalibi@umich.edu

Abstract

Emotion AI, while contested for its validity, bias, and accuracy, claims the ability to infer individuals' emotions and other affective qualities. While proponents tout its potential (e.g., improving well-being and productivity), critics raise concerns about its impact on data subjects' privacy. We conducted a systematic literature review of scholarship that has explored technical solutions to address emotion AI's privacy concerns, examining the underlying conceptualizations of data subjects (i.e., individuals subjected to emotion AI), data, and privacy that motivated them. Findings reveal patterns of 1) conceptualizations of data subjects as decontextualized and flattened, 2) a heavy focus on the sensitivity of input data for emotion AI systems while neglecting the sensitivity of output data (i.e., emotion inferences), 3) conflating privacy with security and 4) viewing privacy as a burden to the development of large-scale emotion AI systems. We argue these conceptualizations motivate technical solutions which largely fail to address the full range of emotion AI's privacy harms. We discuss what a human-centered and comprehensive conception of privacy would mean for emotion AI development, concluding that while technical approaches can address some privacy concerns, key privacy concerns persist.

CCS Concepts

• **Security and privacy** → **Privacy protections**; • **Computing methodologies** → **Artificial intelligence**; • **Networks** → Security protocols.

Keywords

emotion recognition, affect sensing, emotion detection, passive sensing, affective computing, privacy, responsible AI

ACM Reference Format:

Shreya Chowdhary, Alexis Shore Ingber, and Nazanin Andalibi. 2025. Technical Solutions to Emotion AI's Privacy Harms: A Systematic Literature Review. In *The 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25), June 23–26, 2025, Athens, Greece.* ACM, New York, NY, USA, 26 pages. https://doi.org/10.1145/3715275.3732074

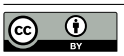
1 Introduction

Emotion artificial intelligence (emotion AI), rooted in affective computing [126], claims to use data—as gathered through inputs such as facial expressions [47, 92], speech [81, 167], language [37], brain activity [69, 93], gait [174], or other multimodal methods [1, 143]—to generate inferences about individuals' emotions. Emotion AI is used across domains [87], including the workplace [18, 31, 80, 104, 139], healthcare [6, 116, 137], cars [49, 94, 122], hiring [105, 138, 152], advertising [95], education [50, 83, 109, 133], and policing [127, 172]. Advocates laud emotion AI's potential to improve societal challenges such as well-being [48, 90], safety [48, 73], and worker productivity [19], fueling a market for emotion AI projected to reach $42.9 billion by 2027 [106]. Although critics raise concerns about emotion AI's ethics [87] including issues with (in)accuracy [34], scientific validity [23, 150], bias [79, 82, 151], and privacy [14, 46, 149], there also exist efforts to mitigate these challenges (e.g., debiasing emotion AI, addressing privacy concerns via technical methods) [35, 71, 72]. In this paper, we examine how current scholarship addresses emotion AI's privacy concerns and implications thereof.

Emotion AI raises concerns around *privacy* given its collection and generation of intimate information [44], often without meaningful consent from individuals subjected to it (i.e., data subjects) [42]. Researchers have called for emotion AI's regulation [22, 51], considering emotion data as sensitive [14], or guidelines for responsible development [71, 110]. In response, computer scientists have sought to address emotion AI's privacy concerns, resembling broader responsible AI development efforts [13, 64]. That being said, the extent to which these efforts adequately address emotion AI's privacy concerns remains unknown. How research conceptualizes "privacy" and associated concepts such as "data" and "data subjects" shape technical efforts to mitigate privacy concerns. As such, our investigation of approaches to addressing emotion AI's privacy concerns seeks to reveal these underlying conceptualizations. This examination will help emotion AI developers and researchers consider what may or may not be realistic goals in their development of emotion AI that preserves privacy.

We conducted a systematic literature review of publications aimed at addressing the privacy concerns surrounding emotion AI (n = 61), revealing the successes and limitations of current state-of-the-art technical approaches. To better understand these technical approaches, we analyzed how this scholarship conceptualizes privacy and key concepts (i.e., data, data subjects). Findings demonstrate that this scholarship seeks to address privacy concerns across a wide range of domains and technical methods (See Table 1).

Our analysis reveals a pattern of superficial engagement with data subjects and their corresponding privacy concerns. While most papers viewed data used as *input* to emotion AI systems as sensitive, surprisingly very few recognized the sensitivity of emotion AI's *output* data (i.e., the inferences made about data subjects' emotions).



FAccT '25, Athens, Greece

ACM ISBN 979-8-4007-1482-5/25/06
https://doi.org/10.1145/3715275.3732074

Further, we find papers to conceptualize privacy as 1) data security, 2) protecting information deemed to be private, and 3) burdensome. The scholarship has leveraged these conceptualizations to address three primary privacy concerns: the risk of input data exposure; the potential for input data to be used to make sensitive inferences (e.g., personally identifiable information or emotion); and the potential for emotion AI systems to enable surveillance.

We applaud researchers and developers for attempting to address the privacy concerns invoked by emotion AI. Still, our analysis supports an argument for future emotion AI development to be more critically grounded by 1) centering the existence of and consequences to data subjects, 2) considering emotion inferences—not just data used to make inferences—as sensitive, and 3) expanding understandings and conceptualizations of privacy beyond security or as restrictive to technological advancement. We suggest, for example, that researchers could design *against* or even *refuse* to design emotion AI. These shifts would orient researchers and developers towards alternative approaches to addressing emotion AI's privacy concerns.

2 Prior Work

While previous work has assessed the literature demonstrating the promise and peril of emotion AI [87] as well as emotion AI's ethical considerations [78], there has not yet been a systematic investigation on the technical solutions to emotion AI's privacy concerns. Our work builds upon the existing critical scholarship investigating emotion AI's privacy concerns [42, 136, 149–151], focusing specifically on how computing scholarship uses technical methods in an attempt to address emotion AI's privacy concerns.

Of note, concerns around privacy of emotions were prevalent long before the development of emotion AI. In their foundational 1890 work establishing an individual right to privacy, legal scholars Warren & Brandeis mention protection of "emotions" 12 times [169]. They contend people should have a right to privacy over their private emotions. Developing AI to make possibly inaccurate emotion inferences [34] and attempting to expose inner feelings represents a violation of privacy which Warren & Brandeis would have urged against. In fact, qualitative scholarship establishes data subjects are wary of significant privacy concerns resulting from emotion AI [14, 42, 46, 139].

2.1 The Promise and Peril of Emotion AI

Emotion AI development has propelled partly due to claims that it can enhance our understanding of people's emotions more accurately than self-report measurements. For example, Latif et al. [88] suggest detecting emotions through speech patterns could mitigate psychological distress. In education, information about students' emotional states has been shown to enhance academic achievement and motivation [33]. Additional purported benefits of emotion AI include addressing automobile safety [177], assisting in criminal investigations [135], or creating more effective consumer advertisements [112]. Even in entertainment, researchers envision affect-aware video games which respond to players' emotions [164] and interactive chatbots that infer emotions through human language [77].

The excitement about emotion AI across industries is not without concerns. Of note, most emotion AI systems rely on the Basic Emotion Theory (BET) which identifies six emotions as encompassing the entire human emotional experience [54]. This theory has been highly contested [23, 29, 89, 150]. As a result, scholars have critiqued emotion AI's accuracy [23] as well as the representativeness and validity of its training data [108]. For example, researchers have empirically shown emotion AI cannot accurately and reliably detect emotions based on facial expressions [34, 79], and can be biased across lines of gender [151], race [107, 145], age [82] and disability [75, 107, 117] among others. Even if emotion AI were to be accurate and free of bias, Ortiz et al. [123] argue its potential to manipulate decision-making threatens civic dignity, a concern echoed by emotion AI data subjects [14]. These concerns call emotion AI proponents' promises into question.

2.2 Privacy Concerns Associated with Emotion AI

Privacy has emerged as a primary concern [123] across emotion AI scholarship attending to its social and ethical implications [14, 108, 109, 113]. A key theme in this discourse is consent. Indeed, emotion AI's inconspicuous nature often results in nonconsensual data collection [110], especially in contexts imbued with power imbalances (e.g., the workplace [2, 42]). Researchers have also identified emotion data as sensitive [14], and suggest data subjects (e.g., workers [43, 46]) experience or anticipate a range of emotion AI privacy concerns.

Despite emotion AI's documented privacy issues, computer scientists and developers continue to push this technology forward. Some motivate emotion AI development given its purported *lack* of invasiveness in data collection, making it protective of privacy [40, 163]. For example, Bethge et al. [26] argue emotion recognition through audio and car speed, among other data points drawn from the driver's smartphone, are unobtrusive ways to predict emotion compared to facial and speech recognition or body-worn physiological sensors. Others similarly suggest that their emotion AI systems' data collection practices are unobtrusive, making their systems protective of privacy [40, 163].

By contrast, some researchers have directly addressed emotion AI's privacy concerns. Latif et al. [87] posit it is possible to mitigate emotion AI's privacy concerns by providing transparency, requiring consent, engaging in ethical data sharing, and promoting data ownership. Researchers have used a variety of technical approaches to preserve privacy while developing emotion AI systems, including federated learning [30], differential privacy [24], encryption [10], and data perturbation [156] (See Table 1 for more information about these methods). Still, the extent to which these approaches can address emotion AI's privacy concerns—as seen through existing scholarship—is unclear. To this end, we ask the following research questions regarding researchers' approach to addressing the privacy concerns of emotion AI:

RQ1: How does this scholarship conceptualize (a) data subjects, (b) data, and (c) privacy?

RQ2: What privacy concerns does this scholarship seek to address?

RQ3: What technical methods does this scholarship use?

3 Methodology

We conducted a systematic literature review of scholarship seeking to address emotion AI's privacy concerns using technical methods. In Table 1, we briefly define relevant technical methods.

Data collection. Following Chancellor and De Choudhury [39]'s process, we first defined our inclusion and exclusion criteria. This study was part of a larger project about understanding technical approaches to addressing ethical issues with emotion AI, including bias, accuracy, and privacy. For inclusion in this larger sample, papers had to 1) be published between 2019 and 2024; 2) indicate a focus on addressing ethical issues with emotion AI within the title or abstract; and 3) implement a technical solution. We excluded work prior to 2019 because of the technology's fast pace of development and the more recent focus on emotion AI's privacy concerns [139, 150].

We predominately used two databases: the Association for Computing Machinery (ACM) and Institute of Electrical and Electronics Engineers (IEEE), representing high impact technical publications, and also searched the proceedings for the Association for Computational Linguistics (ACL), which were not indexed in the ACM and IEEE databases. We searched for papers using keywords informed by prior literature [39, 71, 113, 150].[1] This initial search resulted in 3,736 papers including those addressing a range of ethical issues with emotion AI. Next, we removed papers that did not meet the inclusion criteria stated above. We then manually sifted through references within this sample to identify relevant papers outside the ACM or IEEE databases. This yielded another 79 papers, leading to a total of 3,815 papers.

Given this immense scope, we chose to focus the present analysis solely on privacy rather than all of emotion AI's ethical considerations. We therefore excluded papers that did not address privacy through a technical solution from the 3,815 identified papers. This process excluded 3,753 papers. Our final sample contained 62 papers. Our systematic approach is illustrated in Figure 2 in the Appendix.

Data analysis. The first and last author developed a codebook which guided identification of the themes animating this paper. We generated an initial list of codes based on our research questions, which acted as broad categories and informed our reading of the papers. For instance, **RQ1** explores the conceptualization of data subjects, data, and privacy. Accordingly, we pre-generated the following codes: "characterizing data," "characterizing data subjects," and "conceptualizing privacy." **RQ2** investigates the privacy concerns that papers sought to address. Thus, we pre-generated the following codes: "privacy concerns" and "why address privacy". **RQ3** is concerned with the technical approaches that papers employed. This questeion informed the pre-generation of the following codes: "technical approach used", "why this approach", and "limitations of approach". Additionally, before we began close reading, we generated other codes which reflected particular points of interest in our analysis — such as "domain," which refers to the context papers imagine the application of their approaches.

We followed a process of line-by-line deductive coding [63, 141], applying pre-generated codes while remaining open to any other insights related to our research questions. Line-by-line coding generated subcodes of the pre-generated codes. Once coding was complete, we organized related codes into the themes that constitute our findings. The themes detected were not mutually exclusive, but instead quite overlapping, resulting in both contradictions and connections which we elaborate on in the Results section. All codes are detailed in our codebook in the Appendix.

Dataset characteristics. The papers within our dataset capture efforts to address privacy concerns of emotion AI systems used across contexts such as healthcare, driver monitoring, workplace, and video games. Some scholarship views their approach to be broadly applicable across domains. Publication counts by year are illustrated in Figure 1 in the Appendix. Table 3 in the Appendix offers a detailed breakdown of the dataset.

Limitations and considerations. While we followed a well-established and rigorous process to develop the sample, it is possible that papers aiming to address privacy in emotion AI were overlooked (e.g., in other databases). Additionally, our analysis includes papers published between 2019 and 2024. Although this scope provides a forward-looking view of this space, it excludes possible insights presented prior to 2019. Given the pace of development in the field, we believe focusing our analysis on work published after 2019 is best for assessing technical methods for addressing emotion AI's privacy concerns. This dataset captures contemporary approaches in highly-relevant venues, which is an acceptable limitation for the intended contribution of the paper.

4 Results

We report on patterns in conceptualizations of "who" (the data subject) and "what" (the data)[2] are in need of privacy protection (RQ1a-b), how privacy itself is conceptualized (RQ1c), and the privacy concerns sought to be addressed (RQ2a), and the invoked technical solutions (RQ3).

4.1 Defining "who" and "what" are in need of privacy protection

4.1.1 Conceptualizations of the data subject. Most papers (58/62) broadly gestured to the *existence* of data subjects, mentioning that individuals whose data are used in emotion AI need privacy protections. Only three papers [27, 99, 132] did not mention data subjects, focusing only on the input data or abstractly mentioning a need for privacy.

A few papers (4/62) more substantively characterized data subjects' concerns over data which may be used as inputs for emotion AI [55, 62, 119, 147]. These papers described data subjects' worries, skepticism, or distrust with sharing their input data for emotion AI to varying degrees of specificity. Some papers briefly described data subjects' concerns about particular inferences (e.g., demographic

[1] Like Chancellor et al. [39], we used a pairwise combination of two sets of keywords: emotion-related keywords and ethical mitigation keywords. Emotion-related keywords included: "emotion recognition"; "emotion detection"; "emotion AI"; "mental health" AND "AI"; "affective" AND "prediction"; "wellbeing" AND "ubicomp" OR "wellbeing" AND "sensing"; "mood" AND "inference"; "affective computing". Ethical mitigation keywords included: "emotion prediction"; "bias"; "accountable"; "responsible"; "human-centered"; "fairness"; "participatory"; "ethical"; "trustworthy"; "privacy"; "transparency"; "explainable"; "reliable"; "accuracy".

[2] Although papers in the corpus use various terms to describe those whose data are fed into emotion AI systems and/or who are impacted by emotion AI systems, we refer to them as *data subjects* for consistency.

Technical method	Definition
Federated learning	Decentralized machine learning technique using data on multiple devices rather than centralizing data on a single server [91]
Differential privacy	An approach to data privacy that prevents identification of whether or not a given individual's data was incorporated into an output [52]
Adversarial learning	A machine learning paradigm that seeks to prevent the recognition or classification of certain features (often sensitive information) by intentionally exposing the machine learning model to slightly modified input data. This then confuses the model, forcing it to unlearn classifications of private information [97]
Data perturbation	A variety of methods that aim to preserve privacy by modifying input data in subtle ways, including removing information, changing information, or adding noise (e.g., [155])
Encryption	A method of data security that scrambles data using a cipher so that only authorized people with access to a key can decipher the information encoded in the data (e.g., [12])

Table 1: Definitions of technical methods

information) [55], input data being permanently stored in third-party servers and abused by hackers [119], or personal data being shared for use as emotion AI inputs emotion AI [62, 147].

A few papers (3/62) described data subjects' concerns within specific domains of use. For example, Pranjal et al. [128] described data subjects' overall skepticism about "sharing their (sensitive) data with entities other than their doctor," referring to evidence demonstrating that data subjects generally support their data being used for biomedical research but perceive a high risk with emotion AI input data due to its sensitivity. Similarly, Ravuri et al. [131] characterized data subjects' concerns about the use of their input data by technology and healthcare companies, referencing other studies which found low comfort with sharing voice data and health data. Testa et al. [155] explained how data subjects are concerned about the exploitation of their private data by advertisers or law enforcement: "Many individuals are concerned by the exploitation of their private data by companies or the government. In the US and the UK, many have significant reservations about sharing smart speaker voice assistant (VA) audio with external stakeholders such as advertisers and law enforcement..." These papers described data subjects' concerns with emotion AI by grounding their arguments in empirical scholarship (e.g., [17, 32, 61, 109]).

In summary, although most papers in our corpus broadly acknowledged data subjects as entities to consider in emotion AI development, only a few papers elaborated on their concerns.

4.1.2 Conceptualizations of data. Our analysis reveals several trends in explicit characterizations of input (i.e., the data used as training data for emotion AI) and output data (i.e., emotion inferences generated by emotion AI) of emotion AI systems. Although many papers characterized input data as sensitive, far fewer papers characterized output data as sensitive.

Input data. Many papers (39/62) characterized input data as sensitive [4, 5, 8–10, 12, 15, 16, 26–28, 30, 36, 41, 57, 59, 66, 70, 101, 119, 124, 125, 128, 131, 147, 159, 161, 163, 168, 176, 178] or containing sensitive information [9, 24, 55, 56, 59, 74, 103, 114, 118, 119, 159]. Several kinds of input data were frequently characterized in this way, such as facial data (12/39) [8, 10, 26, 28, 36, 101, 119, 124, 161, 163, 168, 176], EEG recordings (4/39) [4, 5, 15, 36], or speech data (8/39) [9, 26, 55, 56, 119, 128, 159, 178]. Other kinds of input data characterized as sensitive include electrocardiogram (ECG) data [12, 36], electrodermal (EDA) data [12, 36], blood activity [36], muscle activity [36], blood response [36], respiratory response [36], eye activity [36, 57], body gestures [36], and medical data more generally [66, 168].

In some cases, papers characterized certain types of input data as less sensitive than others. For example, gait [28, 103], facial landmark (e.g., head poses or movements of facial muscles) [124], and environmental data (e.g., odor, sound, illumination, CO_2 concentration) [99] were characterized as less sensitive than facial images. Of note, there were inconsistencies in the corpus in terms of identifying particular types of data as sensitive. For example, though eight papers characterized speech data as sensitive, two papers characterized speech data as not sensitive [40, 98].

Though most papers swiftly characterized input data as sensitive or not sensitive, as described above, two papers offered a more detailed classification [12, 26]. Alshareef et al. [12] categorized physiological data into four groups based on personal and/or identifying characteristics of the data: unique individual identifiers; highly private (very personal but not uniquely identifying); moderately private (some personal aspects); or less private (more general and less personal). Within the context of developing emotion AI to monitor drivers, Bethge et al. [26] drew from an existing framework [175] to categorize input data as sensitive, personal, and/or able to be transferred safely over the internet. For example, they determined that traffic, weather, and road data could be transferred over the internet and are not personal nor sensitive, that personal data about the driver (like the driver's sex or age) are personal but not sensitive, and that facial expressions and speech data are both personal *and* sensitive. From this analysis, the authors decided not to use facial expressions or speech data as input data. Instead, they allowed data subjects to select what personal but non-sensitive data they would consent to being collected for their emotion AI system.

Output data. Only six papers reflected on the nature of output data, or emotion inferences [4, 10, 15, 56, 96, 155]. These papers acknowledged several different characteristics of output data that make it sensitive or intimate. For example, three papers recognized the potential for output data to be used to exploit data subjects [4, 10, 155]. Agarwal et al. [4] and Al-Nuaimi et al. [10] stated more broadly that emotion inferences could be used to manipulate users. Testa et al. [155] more specifically discussed the potential exploitation of emotion inferences by large technology companies to manipulate consumers' decision-making, among other harms.

Only four papers recognized the personal or private nature of emotion AI's output data [10, 15, 56, 96]. This determination was often made contextually, relating either to the setting where emotion AI could be deployed or the types of emotion inferences that could be made. For example, Al-Nuaimi et al. identified that emotion inferences made on people in public without their awareness could be "likened to eavesdropping on personal sentiments" [10]. Similarly, Anwar et al. [15] argued that emotion inferences made on employees in a work-from-home environment need to be preserved as employees may not want to share emotions. Other papers described how certain emotion types are particularly private. For example, Feng et al. [56] identified emotions like anger as sensitive, suggesting data subjects may not want these emotions to be revealed. Low et al. [96] similarly contended inferences of micro-expressions, or involuntary emotional expressions that people intentionally suppress, as fundamentally violating privacy.

In sum, although papers in the corpus frequently characterized input data as sensitive, generally for its potential to reveal data subjects' personally identifiable information (PII), there were inconsistencies in *what* input data was considered sensitive and a general lack of detailed justification for *why* a particular sensitivity classification was made, with the exception of the two papers mentioned above [96, 155]. We found that very few papers characterized output data as sensitive. Those which did characterize highlighted the potential for exploitation or privacy violations, and the inherently private nature of emotions.

4.2 Conceptualizations of privacy

Our analysis identified three conceptualizations of privacy: 1) privacy as data security, 2) privacy as protecting information deemed to be private, and 3) privacy as a burden, or restrictive of technological advancement (See Table 2).

4.2.1 Privacy as data security. About one third of the papers (21/62) conceptualized privacy as data security (i.e., protecting data from leakage) implicitly or explicitly (see Table 2). These papers assumed data subjects' privacy would be protected so long as input data security was ensured. This conceptualization is reflected in papers identifying input data leakage as a central privacy concern [5, 12, 15, 43, 62, 66, 100, 101, 111, 114, 115, 118, 147, 160, 178] or use technical methods such as federated learning or encryption whose primary focus is increasing the security of input data (See Section 4.3). Of note, four papers *explicitly* defined privacy as data security [10, 43, 62, 66, 125].

4.2.2 Privacy as protecting private information. Beyond securing input data deemed to be private, some papers (14/62) conceptualized privacy as protecting private information (See Table 2). As discussed in Section 4.1.1, conceptualizations of what is considered to be "private" data varied across the corpus. Some papers explicitly identified emotion AI input data as containing markers of data subjects' identity [65, 70, 74, 96, 103, 119, 128, 131, 155, 159, 176].

Other papers specifically drew upon data subjects' definition of private information is in their own conceptualization of privacy, including output data, citing this as their rationale for protecting against the transfer of emotion data that data subjects might not want revealed [15, 26, 56].

Finally, some papers [65, 74, 96, 131, 155, 159, 176] conceptualized privacy as protecting private information through their use of technical methods which primarily aimed to prevent inference or collection of private information. These methods included adversarial learning and data perturbation, which are further elaborated on in Section 4.3.

4.2.3 Privacy as restrictive of technological advancement: the small data problem. Some papers (13/62) conceptualized privacy as an *obstacle* to collecting sufficient data for emotion AI (See Table 2). Toto et al. [157] define this as the "small data problem," where an insufficient amount of input data prevents the development of state-of-the-art, large-scale, and robust emotion AI. These papers conceptualized privacy as contributing to the small data problem because it limited available input data.

Papers described several reasons why the small data problem is an issue for developing large-scale, robust emotion AI. Insufficient data can limit the accuracy of emotion AI [30, 157, 158] and can also create what Bethge et al. [27] describe as "domain adaptation" problem, where an emotion AI system can only perform well on one dataset. Finally, the analyzed papers attest insufficient input data can result in less representative emotion AI [8, 30]. For example, Akhyani et al. [8] described how the need to address privacy limits facial data collection that is 1) diverse in terms of gender and ethnicity, and 2) reflects emotions beyond the six basic emotions defined in BET [54].

Ultimately, papers cited privacy for instigating the small data problem. For instance, data subjects may not want to share their data [27, 36, 154, 157, 158]. Additionally, some papers determined the process of collecting input data may compromise data subjects' privacy, which can then limit data collection [8, 86, 102]. Prioritizing privacy—as done through privacy policies— can also limit access to existing datasets [154]. Bondin et al. [30] identified the decentralized methods of storing emotion AI input data can address privacy concerns by creating barriers to accessing sufficient data. Lastly, some papers underscored that privacy concerns might limit the availability of certain types of emotion AI input data such as full facial images [86, 124].

Papers used various methods to address the small data problem while still enabling emotion AI development. Four papers [30, 36, 146, 154] used federated learning as a method to circumvent the limitations created by the need to preserve privacy through using decentralized data. Three other papers [102, 124, 132] used what they deemed to be less sensitive data to avoid compromising privacy; for instance, only using the eye region [132] or facial

Conceptualization of privacy	Associated papers
Privacy as data security	[5, 10, 12, 15, 24, 41, 43, 62, 66, 100, 101, 111, 114, 115, 118, 125, 130, 147, 160, 168, 178] (21 papers)
Protecting private information	[15, 26, 56, 65, 70, 74, 96, 103, 119, 128, 131, 155, 159, 176] (14 papers)
Restrictive of technological advancements	[8, 27, 30, 36, 86, 102, 124, 132, 146, 154, 157, 158] (13 papers)

Table 2: Conceptualizations of privacy and associated papers

landmark data [124] instead of full facial images or using electrodermal signals instead of facial data [102]. Similarly, two papers [8, 86] used synthetic data to avoid a privacy-compromising data collection process. Finally, three papers devised new machine learning frameworks which were specifically designed to work with small input datasets [27, 157, 158].

4.3 Privacy concerns invoked and technical approaches used to address them

Our analysis identified three primary privacy concerns that papers sought to address: the risk of input data exposure; the potential for emotion AI input data to be used to make sensitive non-emotion inferences about data subjects (e.g., identity, demographics); and the potential for emotion AI to enable surveillance. To address these concerns, various technical approaches were employed with goals including 1) preventing exposure of emotion AI input data; 2) preventing the use of emotion AI input data for inference of a) PII or b) emotion; and 3) to develop emotion AI using input data that researchers deemed less sensitive. Table 4 in the Appendix illustrates the technical approaches used to address particular privacy concerns, mapped on to corresponding papers.

4.3.1 A risk of input data exposure. Many papers (33/62) identified the risk of input data exposure as a privacy threat. This reflects the most common conceptualization of privacy — as articulated in the previous section — of privacy as data security. Within this group, some papers (22/33) identified the *centralization of emotion AI input data* (i.e., transfering data from multiple sources to a single server for model training) as the source of this problem. Several of these papers (7/21) further noted specific concerns over the exposure of what they deemed to be particularly sensitive input data [12, 74, 131] like medical records, finance, and personal data [168]; biometric information and fingerprints [15]; EEG recordings [5]; biometric identity, personality traits, location, "emotional state", age, gender, and overall health [114].

Several papers (19/33) identified that input data leakage would create privacy concerns because sensitive input data could be exposed to *malicious actors* who may misuse or exploit the data. Papers referred to malicious actors with various terms, including simply "malicious actors" [9, 12, 24, 65, 74, 115, 131], "adversaries" [9, 74, 114], or "attackers" [10, 15, 59, 85, 115, 119, 176, 178]. One paper, Testa et al. [155], specifically called out technology companies and law enforcement as malicious actors who might use emotion AI input data for targeted advertising or to assess data subjects' mental states, respectively.

Papers mentioning malicious actors' misuse or exploitation of emotion AI input data named several concerns around how malicious actors might compromise privacy. These concerns included a general misuse of emotion AI input data [12, 70, 131], identity recognition [24, 57, 59], targeted advertising [24, 57, 155], exposure of other sensitive information like health conditions [24, 57, 59, 178], or "model poisoning with fake doping of data" [41].

Papers aimed at preventing the exposure of input data leveraged various solutions. 19/33 papers implemented federated learning because it "enables users to locally train the model without compromising his/her privacy" [41] or "allows models to be trained collaboratively on decentralized devices without the need to transfer raw data" [111].

Other papers used alternative technical strategies to secure input data, such as differential privacy, cryptographic methods, or other machine learning approaches. For example, in two studies, Mainsant et al. [100, 101] used continual learning, where only the classification function is persistently stored and the input data are forgotten. In another study, Bethge et al. [27] used a machine learning architecture claiming to only learn generalized emotion-related patterns from the input data rather than PII.

Some papers (7/33) sought to minimize the risk of input data exposure by preventing malicious actors from making sensitive inferences from potentially leaked input data. Most of these papers leveraged adversarial learning (5/7) to unlearn what they identified as private and sensitive information encoded in the input data, like people's age, gender, race, ethnicity, and identity. One paper used synthetic data (i.e, data generated by AI) to prevent the inference of identity by anonymizing facial images used as input data. Finally, two papers used data perturbation, which added noise to the facial image data. One paper was distinctive in addressing concerns surrounding malicious actors by perturbing the data in order to fully prevent emotion inference.

4.3.2 Input data could be used to infer sensitive information. Many papers (26/62) identified the potential for emotion AI input data to be used to *infer sensitive information* as a privacy concern. Papers discussed how various kinds of input data could be used to make "unintended or improper inferences of sensitive information and demographic information" [56]. Several papers described how facial data [3, 40, 70, 99, 119, 161, 163], speech data [9, 24, 56, 114, 159], eye movement data [86], and other contextual information collected for emotion AI [16, 125] could also reveal sensitive information like identity, gender, ethnicity, and age.

The majority of these papers discussed sensitive information as PII, with only a few [96, 155] discussing emotion data itself as sensitive. Low et al. [96] claimed emotion AI systems aiming to recognize microexpressions or "involuntary or transient facial expressions, commonly manifested involuntarily when we aim to withhold our emotions" threaten emotional privacy and fundamental human rights. Testa et al. [155] identified speech emotion

recognition systems' generation of emotion data about data subjects can be exploited for surveillance or manipulating data subjects' behavior. This, therefore, makes emotion data sensitive for Testa et al [155].

Other papers identified that even if PII is not explicitly encoded in the input data, or even intentionally hidden, it can still be inferred by models [55, 128, 154]. One paper, Malek et al. [103], explained this phenomenon by invoking the concept of "overlearning" [148], which describes how machine learning models trained to recognize one feature could simultaneously learn to recognize others. Other papers similarly noted that physiological data used as emotion AI input data contains unique identifiers [12, 57, 59]. Several papers connected concerns about PII inference with those about malicious actors (i.e. malicious actors using data subjects' identities for identity theft or targeted advertising) [4, 12, 57, 59, 131].

Papers employed various solutions to address the concern around input data fueling sensitive inferences. Several papers (8/26) used technical methods to prevent input data from being exposed[3] so it could not be accessed for any purpose besides emotion inference. These papers used methods like federated learning (5/8), differential privacy (3/8), and/or a cryptographic approach (1/8).

Other papers aimed to prevent inferences of PII from input data (12/26). Technical approaches here included data perturbation (6/12) through adding noise or methods of anonymization to prevent other sensitive inferences. Some of these papers (6/12) used adversarial learning to unlearn demographic information. As with the papers described in the previous paragraph, these papers' implementations still preserved the ability to make emotion inferences.

Some papers within the corpus addressed the privacy concern around sensitive inferences being made on input data by collecting or using what they deemed to be less sensitive or non-identifiable input data (7/26). Many of these papers (5/7) used "secondary affect data," which [161] contrasts with "primary" affect data as containing less identifiable emotion-related information. For example, Uddin et al. [161] described facial images and speech recordings as primary affect data but facial landmarks and action units as secondary affect data, which allows emotion inferences without revealing identifiable information. In addition to secondary affect data [161, 163], some papers used environmental data (e.g., lighting conditions) [99], or collective conditions rather than individual behavior (e.g., how much people are speaking rather than who is speaking or what they're saying) [16]. Alternatively, Chen et al. [40] and Machanje et al. [98] used speech data to address this privacy concern, arguing that speech data is less sensitive and identifying than facial data. Lan et al. [86], used synthetic data with the argument that collecting eye movement data poses a privacy threat.

Finally, two papers which explicitly and centrally recognized emotion information as sensitive [96, 155] aimed to preserve privacy by designing *against* emotion recognition. Both of these papers implemented methods to perturb input data so that emotion inferences could not be made. For instance, Low et al. [96] implemented an adversarial-learning-based approach that would prevent emotion AI from recognizing micro-expressions in facial images. Similarly, Testa et al. [155] added noise to emotion AI input data so that speech emotion recognition would not be possible.

[3]We acknowledge that the language we are using here ("exposed") is vague. This is a reflection of the language used in the papers.

4.3.3 Emotion AI enables surveillance. A few papers (4/62) argued that emotion AI raises privacy concerns by enabling surveillance. These papers conceived of emotion AI's surveillant role in two ways: through its particular design features, and as a core goal of the technology.

Three papers described how certain *design features* of emotion AI, like its purported "unobtrusive" data collection method or focus on tracking individuals, enable surveillance [10, 16, 26]. For example, Bethge et al. [26] identified that data subjects may perceive certain sensors, like audio and video recorders, as "contribut[ing] to a feeling of surveillance", compared to other sensors (like accelerometers or GPS trackers). Similarly, Augusma et al. [16] suggested emotion AI systems in classrooms could enable surveillance if designed to track the behavior of students or teachers. Finally, Al-Nuaimi et al. [10] noted that emotion AI systems could be used in public spaces, where individuals could be subjected to emotion recognition without their knowledge or consent. In Al-Nuaimi et al.'s words, this could constitute a form of "covert surveillance", which could be seen as an "infringement on personal freedoms and a misuse of personal data" [10].

These three papers offered technical solutions aiming to reduce the potential for emotion AI to enable surveillance. Al-Nuaimi et al. [10] used cryptographic methods to keep input data more secure and data perturbation methods to prevent the inference of sensitive information (e.g., identity or demographics). They implemented this dual-strategy approach to limit the potential for surveillance. Bethge et al. [26] and Augusma et al. [16] addressed the potential for emotion AI to enable surveillance through implementing what they deemed to be less intrusive data collection methods and collecting data they perceived as less sensitive. For instance, Bethge et al. [26] designed an approach to driver emotion recognition that used weather conditions, car temperature, and traffic conditions to "non-intrusively" capture drivers' emotions. Similarly, Augusma et al. [16] implemented a method for inferring emotion in classrooms through "collective monitoring," where the system could perceive the underlying cues of teaching engagement (e.g., student engagement, attention levels) through signals at the classroom-level rather than the behavior of individual students or teachers.

One paper [155] had a different conception of emotion AI's role in surveillance, arguing emotion AI may be designed with the *intent purpose* of enabling surveillance. They elaborated how smart speaker voice assistants are built to support surveillance capitalism [179], where data subjects' private information is datafied for monetization and extraction. Testa et al. [155] further illustrated that technology companies have a vested interest in leveraging data subjects' affective data for purposes like targeted advertising. In other words, Testa et al. [155] fundamentally characterized emotion AI as infrastructure for surveillance capitalism, informing their approach to designing against emotion inferences. Based on this characterization, Testa et al. [155] designed *against* emotion recognition through data perturbation methods aimed at preventing emotion inferences.

5 Discussion

This systematic literature review contributes insights into how emotion AI development scholarship addresses privacy concerns associated with emotion AI. It reveals patterns in conceptualizations of data, data subjects, and privacy, and the privacy concerns associated with emotion AI and technical approaches to address them.

Findings reveal the majority of papers seeking to address privacy concerns implicated by emotion AI 1) take on a flattened and decontextualized perspective of data subjects; 2) focus on securing input data, with much less concern for the potential concerns caused by generating and using output data; 3) conflate privacy and data security, and consider privacy to be a burden; and 4) rely on technical approaches that are limited in addressing the broad range of privacy concerns invoked by emotion AI.

We suggest emotion AI scholars and developers who seek to address emotion AI-related privacy concerns consider privacy as contextual [121], centering data subjects and their corresponding identities. This conceptualization of privacy would broaden the scope of privacy concerns that emotion AI developers engage with. With a broader recognition of emotion AI's privacy concerns, as prior work suggests [14, 43, 139, 150], we conclude that addressing emotion AI privacy concerns through purely technical approaches may not be a realistic goal.

5.1 Conceptualization of Data Subjects: Flattened and Decontextualized

Our analysis revealed a general pattern of a flattened and decontextualized conceptualization of data subjects. Few papers characterized data subjects in terms more specific than "users,"[4], "patients," or "people." Those that characterized data subjects often did not engage with them as situated within specific contexts. This pattern aligns with Chancellor et al.'s [38] examination of conceptualizations of "human" in human-centered machine learning research working to predict mental health, which found this scholarship views humans as data points for machine training and optimization. They argued this risks dehumanizing individuals, contributes to stigma, and violates principles of contextual integrity [121].

We observed similar patterns in our analysis, with most papers lacking an acknowledgment of differences in data subjects' experiences with emotion AI, especially across different contexts of deployment and identity factors. We interpret this flattening of data subjects contributes to the papers' narrow recognition of privacy concerns. Prior work demonstrates data subjects' privacy concerns with emotion AI are shaped by the context of use, such as the workplace [2, 42, 46, 79, 139, 151], healthcare [76, 137], or education [50, 109, 129]. Moreover, research has also shown that different facets of identity, like race [107, 145], gender [151], and disability [75, 107, 117], and age [82], influence data subjects' privacy concerns about emotion AI. Considering this literature and our findings, we urge emotion AI developers and researchers seeking to technically address emotion AI's privacy concerns to develop a contextual [121], identity-sensitive perspective of data subjects.

[4]It is outside the scope of this paper to examine the use of the term "user" in-depth here; however, we note that often *subjects* to emotion AI are not actually those who *use* it, and therefore, the term "user" is misguided in these cases.

5.2 Conceptualization of Data: Under-recognition Emotion Inferences' Sensitivity

We found most papers focused on the need to protect *input* data. This is reasonable, as input data for emotion AI *are* sensitive. However, our analysis revealed an under-recognition of the sensitivity of *output* data—emotion inferences themselves. Most of the papers in the corpus framed emotion AI outputs as innocuous or an unquestioned "good," a characterization we find troubling. First, prior work [14, 140] establishes that emotions have private dimensions, which is violated by emotion AI's very inference of emotions. Several legal scholars' definitions of privacy include emotions as private information. For example, Citron's *intimate privacy* [44] and Richards' *intellectual privacy* [134] identify the need to maintain privacy over intimate aspects of people's internal lives, which might include emotions. In Cohen's reconceptualization of privacy as a social condition [45], she frames "semantic discontinuity" or the right to be somewhat unknown and unknowable, as a fundamental condition of privacy. Emotion AI violates these principles given its express purpose to know aspects of data subjects' internal lives better than they may know themselves [79].

Second, prior work establishes a clear record of ways emotion inferences have or could be used to violate privacy rights. For instance, emotion AI inferences have been used to enforce a regime of border surveillance in the European Union [142, 166]. Similarly, emotion AI inferences have been used in policing, expanding the scope of state surveillance [127, 172]. Other work has established ways that the inferences made by emotion AI on workers could enable increasingly limitless workplace surveillance [7, 42]. Thus, in order to truly recognize and address all of the privacy concerns invoked by emotion AI, researchers and developers need to broaden their conceptualization of data implicated in emotion AI use, and confront the sensitivity of emotion data directly.

5.3 Conceptualization of Privacy and Privacy Concerns

Our analysis identified two notable and problematic conceptualizations of privacy that merit discussion: the conflation of privacy with input data security and viewing privacy as a burden.

5.3.1 Conflation of privacy and security. Many papers conceptualized privacy as data security, with a focus on the privacy of emotion AI input data. This is illustrated not only in explicit conceptualizations of privacy across the corpus, but in the security protection approaches taken to address privacy (e.g., preventing data leakage, exposure of data to malicious actors). This conflation of privacy and security is indeed common [84, 165]. In fact, technology developers tend to use security-related vocabulary to define privacy challenges [68]. Legal scholars similarly conflate privacy and security [21]. This conflation helps explain why many papers in our corpus noted that avoiding malicious actors' intrusion into private data was their motivation to address emotion AI's privacy issues. The conflation of privacy and security also helps explain another contradiction in the corpus: the common recognition of the potential for sensitive inferences to be made from input data co-occurring with a focus

on securing input data. The conflation of privacy and security narrowly orients papers towards using input data security as a solution to all privacy concerns — even those that are more related to the output data.

While maintaining the security of input data *is* important, we argue emotion AI's privacy concerns extend beyond potential misuse or exploitation of input data, including the generation and use of emotion inferences [14, 46]. That is, we do not suggest researchers and developers dismiss input data security. Rather, we encourage them to broaden conceptualizations of privacy beyond input data security.

5.3.2 Conceptualization of privacy as a burden. Many papers conceived of privacy as something they *had to* engage with to avoid regulatory consequences, as opposed to a value they *wanted to* uphold due to its normative or moral importance, demonstrated through recognition of the small data problem [157]. This conceptualization of privacy is not unique to emotion AI, as other technology developers have characterized privacy as a burden that is challenging to implement [11]. Notably, the small data problem may in part stem from regulatory limitations on access to input data. In a recent analysis of large technology companies' response to new privacy regulation (e.g., GDPR, CCPA), Wong et al. [171] found companies see regulation as posing substantial risk to their reputation as well as their business practices. However, the U.S. and much of the world does not actually regulate emotion data nor emotion recognition with the exception of the EU's AI Act which places a ban on emotion AI use in the workplace and education [162]. In the absence of law and policy in the U.S., it is easy for developers of emotion AI to de-prioritize privacy. We argue for further regulation of emotion AI and assert that emotion AI developers and researchers should conceptualize and operationalize privacy as a *responsibility*, not a burden or a necessity for mere legal compliance.

Beyond the threat of regulatory consequences, Ekambaranthan et al. [53] raise that the challenges developers face when building privacy-protective systems are both technical (e.g., design that does not protect privacy by default) and economic (i.e., business models that rely on invasive data collection). Developers should continue working on systems which protect privacy by default, making it more difficult to *not* protect data subjects' privacy. Data subjects *want* to maintain emotional privacy [14, 139]. As such, this desire should be reflected in emotion AI development if its adoption remains a goal.

5.4 The Limitations of Technical Approaches and Potential Paths Forward

To address privacy concerns, papers employed a range of technical approaches. These approaches can be described as "privacy-enhancing technologies" (PETs) [67]. PETs primarily aim to preserve the privacy of data subjects. For example, differential privacy allows emotion inferences to be made while protecting against malicious actors' utilization of the input data (through membership inference or reconstruction attacks) [52]. In this sense, differential privacy limits unauthorized access to and misuse of input and output data. However, research has explored the limitations of PETs in addressing privacy concerns, especially if considering privacy to extend beyond security [20, 67, 129, 144, 153]. These limitations are especially evident when considering the principles of contextual integrity [20, 121], which argues that privacy is achieved when information flows are appropriate to the social context. We draw on contextual integrity to motivate the need for a contextual approach to addressing privacy in emotion AI. As our results demonstrate, the majority of papers have a de-contextualized perspective on emotion AI's privacy concerns, focusing on what loosely-defined malicious actors could do to input data.

We find well-documented limitations of the technical approaches used to address the privacy concerns implicated by emotion AI when we attend to its broader social context. For example, papers frequently used federated learning to secure input data. Although it addresses individual privacy concerns, federated learning is limited because it requires high compliance across the industry and does not attend to broader social impacts caused by technologies such as emotion AI [166]. Similarly, papers often used differential privacy as a mathematically-provable way to protect input data. Seeman & Susser [144], for example, highlight that differential privacy privileges managing the risk of individuals' input data exposure over broader social consequences. However, the use of emotion AI in particular contexts can create broader privacy concerns that may not directly affect individuals whose input data are being used. For instance, Testa et al. [155] and other scholarship [127] highlight the use of emotion AI in the criminal legal context, where individuals' input data can then used as training data to violate others' privacy at a larger scale.

Additionally, several papers used synthetic data to circumvent the privacy issues associated with collecting sensitive input data. However, critical scholars argue synthetic data is not a solution to privacy issues [153]. Susser & Seeman [153] contend synthetic data necessarily retains information about real data subjects, thus risking data exposure. Further, synthetic data, like other PETs, does not address broader privacy concerns caused by inappropriate information flows (as explicated by the theory of contextual integrity [121]). Whitney and Norman [170] argue that using synthetic data circumvents consent for data use. These critical analyses reflect the limitations of many of the common technical approaches used in our corpus to address privacy issues. They over-focus on individual data protection and fail to adequately address privacy concerns.

In reflecting on the limitations of common privacy-preserving technical approaches, we do not intend to suggest these approaches have no use, as we believe they could be part of a broader plan for addressing *some* privacy concerns invoked by emotion AI. However, if we construe privacy as primarily a social rather than purely technological problem [45, 120, 129], the technical approaches described above are not sufficient. In other words, recognizing privacy as a social problem necessitates abandoning techno-solutionist approaches to privacy [129, 153]. Abandoning techno-solutionist approaches means developers need to abandon technical approaches as the primary method to address emotion AI's privacy issues.

We suggest that developers can more effectively address emotion AI's privacy issues by taking a more contextually-sensitive approach that contends directly with the sensitivity of both emotion AI's input data and the output data. First, developers should directly engage with data subjects and/or the critical scholarship that has studied data subjects' perspectives towards emotion AI. Through this direct engagement, developers can familiarize themselves with

the contested privacy issues with emotion AI. Within our corpus, Bethge et al. [26] is an exemplar in terms of engagement with data subjects as the only paper that included a user study as part of their approach. This user study surfaced valuable insights on what input data the data subjects— the eventual users—did and did not want collected.

Second, developers should question the inherent positive social good of emotion inferences. Again, engagement with the critical scholarship that has critiqued the very basis of automatic emotion inference [23, 149, 150] could deepen developers' understanding of this. Technical developers can familiarize themselves with the privacy concerns emergent in the domains that emotion recognition is being used in — or consider aggregate information flows [25]. Within our corpus, Testa et al. [155] offer a good example of this practice. They studied Big Tech companies' motivations, referencing existing literature on surveillance capitalism [179] and news reporting which covers the potential harms of these technologies [58], which proliferate beyond initial data collection. This context helped them understand the privacy concerns emotion AI could create for data subjects.

Finally, informed by this deeper contextual exploration, developers could explore novel technical approaches for addressing emotion AI's privacy concerns. They could, as was done by Bethge et al. [26], explore mechanisms that enhance data subjects' ability to choose which types of data they want collected for emotion recognition (if any) and develop models for addressing consent violations. Technical researchers could also leverage their expertise to design *against* emotion AI, as a practice of refusal [60]. Testa et al. [155] and Low et al. [96] are exemplars in this regard from our corpus, as the only two papers that designed *against* emotion inferences.

6 Conclusion

This paper presents a systematic literature review of scholarship which has developed technical approaches to address emotion AI's privacy concerns. Findings demonstrate general trends of decontextualized and flattened conceptualization of data subjects, an under-recognition of the sensitivity of emotion inferences, and a conflation of privacy and data security. These conceptualizations lead to technical solutions which have failed to address the broad range of privacy issues invoked by emotion AI, mainly those experienced by data subjects due to the very act of inferring emotions. We argue that researchers and developers concerned with the privacy issues emotion AI may cause should take on a contextualized understanding of privacy, data, and data subjects which centers the data subjects and contends with the inherent sensitivity of emotions. This conceptualization of privacy can inform using technical expertise to more productively address emotion AI's privacy issues. By maintaining existing practices, emotion AI developers will fail to truly address emotion AI's privacy concerns.

Acknowledgments

We appreciate the reviewers' valuable feedback on this work. This work was sponsored by NSF CAREER award 2236674.

References

[1] Sharmeen M Saleem Abdullah Abdullah, Siddeeq Y Ameen Ameen, Mohammed AM Sadeeq, and Subhi Zeebaree. 2021. Multimodal emotion recognition using deep learning. *Journal of Applied Science and Technology Trends* 2, 01 (2021), 73–79.

[2] Daniel A Adler, Emily Tseng, Khatiya C Moon, John Q Young, John M Kane, Emanuel Moss, David C Mohr, and Tanzeem Choudhury. 2022. Burnout and the quantified workplace: Tensions around personal sensing interventions for stress in resident physicians. *Proceedings of the ACM on Human-computer Interaction* 6, CSCW2 (2022), 1–48.

[3] Ayush Agarwal, Pratik Chattopadhyay, and Lipo Wang. 2021. Privacy preservation through facial de-identification with simultaneous emotion preservation. *Signal, Image and Video Processing* 15, 5 (2021), 951–958.

[4] Anisha Agarwal, Rafael Dowsley, Nicholas D. McKinney, Dongrui Wu, Chin-Teng Lin, Martine De Cock, and Anderson C. A. Nascimento. 2019. Protecting Privacy of Users in Brain-Computer Interface Applications. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27, 8 (2019), 1546–1555. https://doi.org/10.1109/TNSRE.2019.2926965

[5] Manan Agrawal, Mohd Ayaan Anwar, and Rajni Jindal. 2023. FedCER - Emotion Recognition Using 2D-CNN in Decentralized Federated Learning Environment. In *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*. 1–5. https://doi.org/10.1109/ISCON57294.2023.10112028

[6] Kriti Ahuja. 2024. Emotion AI in healthcare: Application, challenges, and future directions. In *Emotional AI and human-AI interactions in social networking*. Elsevier, 131–146.

[7] Ifeoma Ajunwa, Kate Crawford, and Jason Schultz. 2017. Limitless worker surveillance. *Calif. L. Rev.* 105 (2017), 735.

[8] Saba Akhyani, Mehryar Abbasi, Mo Chen, and Angelica Lim. 2022. Towards Inclusive HRI: Using Sim2Real to Address Underrepresentation in Emotion Expression Recognition. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 9132–9139. https://doi.org/10.1109/IROS47612.2022.9982252

[9] Hafiz Shehbaz Ali, Fakhar ul Hassan, Siddique Latif, Habib Ullah Manzoor, and Junaid Qadir. 2021. Privacy Enhanced Speech Emotion Communication using Deep Learning Aided Edge Computing. In *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*. 1–5. https://doi.org/10.1109/ICCWorkshops50388.2021.9473669

[10] Reem AlNuaimi, Fady Alnajjar, and Hassan Abdulmouti. 2023. PEEP with Cloud Encryption: A Dual-Layered Privacy Solution for Emotion Recognition Systems. In *2023 15th International Conference on Innovations in Information Technology (IIT)*. 186–189. https://doi.org/10.1109/IIT59782.2023.10366494

[11] Noura Alomar and Serge Egelman. 2022. Developers say the darnedest things: Privacy compliance processes followed by developers of child-directed apps. *Proceedings on Privacy Enhancing Technologies* (2022).

[12] Moudy Sharaf Alshareef, Mona Jaber, and Ahmed M. Abdelmoniem. 2023. A Differential Privacy Approach for Privacy-Preserving Multi-Modal Stress Detection. In *2023 IEEE 28th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*. 206–212. https://doi.org/10.1109/CAMAD59638.2023.10478412

[13] Marianna Anagnostou, Olga Karvounidou, Chrysovalantou Katritzidaki, Christina Kechagia, Kyriaki Melidou, Eleni Mpeza, Ioannis Konstantinidis, Eleni Kapantai, Christos Berberidis, Ioannis Magnisalis, et al. 2022. Characteristics and challenges in the industries towards responsible AI: a systematic literature review. *Ethics and Information Technology* 24, 3 (2022), 37.

[14] Nazanin Andalibi and Justin Buss. 2020. The human in emotion recognition on social media: Attitudes, outcomes, risks. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–16.

[15] Mohd Ayaan Anwar, Manan Agrawal, Neha Gahlan, Divyashikha Sethia, Gaurav Kumar Singh, and Rishabh Chaurasia. 2023. FedEmo: A Privacy-Preserving Framework for Emotion Recognition using EEG Physiological Data. In *2023 15th International Conference on COMmunication Systems & NETworkS (COMSNETS)*. 119–124. https://doi.org/10.1109/COMSNETS56262.2023.10041308

[16] Anderson Augusma. 2022. Multimodal Perception and Statistical Modeling of Pedagogical Classroom Events Using a Privacy-safe Non-individual Approach. In *2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. 1–5. https://doi.org/10.1109/ACIIW57231.2022.10086029

[17] Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. 2019. Americans and privacy: Concerned, confused and feeling lack of control over their personal information. (2019).

[18] Ezra Awumey, Sauvik Das, and Jodi Forlizzi. 2024. A Systematic Review of Biometric Monitoring in the Workplace: Analyzing Socio-technical Harms in Development, Deployment and Use. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 920–932.

[19] Richard P Bagozzi, Michael K Brady, and Ming-Hui Huang. 2022. AI service and emotion. , 499–504 pages.

[20] Ero Balsa and Yan Shvartzshnaider. 2023. When PETs misbehave: A Contextual Integrity analysis. *arXiv preprint arXiv:2312.02509* (2023).

[21] Derek E Bambauer. 2013. Privacy versus security. *J. Crim. L. & Criminology* 103 (2013), 667.

[22] Jennifer S Bard. 2021. Developing legal framework for regulating emotion AI. *BUJ Sci. & Tech. L.* 27 (2021), 271.

[23] Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella, Aleix M Martinez, and Seth D Pollak. 2019. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest* 20, 1 (2019), 1–68.

[24] Mohamed Benouis, Yekta Said Can, and Elisabeth André. 2023. A Privacy-Preserving Multi-Task Learning Framework For Emotion and Identity Recognition from Multimodal Physiological Signals. In *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. 1–8. https://doi.org/10.1109/ACIIW59127.2023.10388160

[25] Sebastian Benthall and Rachel Cummings. 2024. Integrating differential privacy and contextual integrity. In *Proceedings of the Symposium on Computer Science and Law*. 9–15.

[26] David Bethge, Luis Falconeri Coelho, Thomas Kosch, Satiyabooshan Murugaboopathy, Ulrich von Zadow, Albrecht Schmidt, and Tobias Grosse-Puppendahl. 2023. Technical Design Space Analysis for Unobtrusive Driver Emotion Assessment Using Multi-Domain Context. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 159 (Jan. 2023), 30 pages. https://doi.org/10.1145/3569466

[27] David Bethge, Philipp Hallgarten, Tobias Grosse-Puppendahl, Mohamed Kari, Ralf Mikut, Albrecht Schmidt, and Ozan Özdenizci. 2022. Domain-Invariant Representation Learning from EEG with Private Encoders. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1236–1240. https://doi.org/10.1109/ICASSP43922.2022.9747398

[28] Carmen Bisogni, Lucia Cimmino, Michele Nappi, Toni Pannese, and Chiara Pero. 2024. Walk as you feel: Privacy preserving emotion recognition from gait patterns. *Engineering Applications of Artificial Intelligence* 128 (2024), 107565.

[29] Kirsten Boehner, Rogério DePaula, Paul Dourish, and Phoebe Sengers. 2007. How emotion is made and measured. *International Journal of Human-Computer Studies* 65, 4 (2007), 275–291.

[30] Luca Bondin and Alexiei Dingli. 2021. A Federated Affective Computing Framework To Learn From Small Data. In *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*. 14–19. https://doi.org/10.1109/CSCI54926.2021.00009

[31] Karen L Boyd and Nazanin Andalibi. 2023. Automated emotion recognition in the workplace: How proposed technologies reveal potential futures of work. *Proceedings of the ACM on human-computer interaction* 7, CSCW1 (2023), 1–37.

[32] LaPrincess C Brewer, Karen L Fortuna, Clarence Jones, Robert Walker, Sharonne N Hayes, Christi A Patten, and Lisa A Cooper. 2020. Back to the future: achieving health equity through health informatics and digital health. *JMIR mHealth and uHealth* 8, 1 (2020), e14512.

[33] Müzeyyen Bulut Özek. 2018. The effects of merging student emotion recognition with learning management systems on learners' motivation and academic achievements. *Computer applications in engineering education* 26, 5 (2018), 1862–1872.

[34] Federico Cabitza, Andrea Campagner, and Martina Mattioli. 2022. The unbearable (technical) unreliability of automated facial emotion recognition. *Big data & society* 9, 2 (2022), 20539517221129549.

[35] Erik Cambria, Xulang Zhang, Rui Mao, Melvin Chen, and Kenneth Kwok. 2024. SenticNet 8: Fusing emotion AI and commonsense AI for interpretable, trustworthy, and explainable affective computing. In *International Conference on Human-Computer Interaction*. Springer, 197–216.

[36] Yekta Said Can and Cem Ersoy. 2021. Privacy-preserving Federated Deep Learning for Wearable IoT-based Biomedical Monitoring. *ACM Trans. Internet Technol.* 21, 1, Article 21 (Jan. 2021), 17 pages. https://doi.org/10.1145/3428152

[37] Priya Chakriswaran, Durai Raj Vincent, Kathiravan Srinivasan, Vishal Sharma, Chuan-Yu Chang, and Daniel Gutiérrez Reina. 2019. Emotion AI-driven sentiment analysis: A survey, future research directions, and open issues. *Applied Sciences* 9, 24 (2019), 5462.

[38] Stevie Chancellor, Eric PS Baumer, and Munmun De Choudhury. 2019. Who is the" human" in human-centered machine learning: The case of predicting mental health from social media. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–32.

[39] Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine* 3, 1 (2020), 43.

[40] Shuaiqi Chen, Xiaofen Xing, Guodong Liang, and Xiangmin Xu. 2022. I Feel Stressed Out: A Mandarin Speech Stress Dataset with New Paradigm. In *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 583–589. https://doi.org/10.23919/APSIPAASC55919.2022.9980183

[41] Prateek Chhikara, Prabhjot Singh, Rajkumar Tekchandani, Neeraj Kumar, and Mohsen Guizani. 2021. Federated Learning Meets Human Emotions: A Decentralized Framework for Human–Computer Interaction for IoT Applications. *IEEE Internet of Things Journal* 8, 8 (2021), 6949–6962. https://doi.org/10.1109/JIOT.2020.3037207

[42] Shreya Chowdhary, Anna Kawakami, Mary L Gray, Jina Suh, Alexandra Olteanu, and Koustuv Saha. 2023. Can workers meaningfully consent to workplace wellbeing technologies?. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 569–582.

[43] Md. Reasad Zaman Chowdhury, Mashfurah Afiat, Alvin Rahul Hore, Rabea Akhter, Alex Sarker, Md Humaion Kabir Mehedi, Abid Hossain, and Annajiat Alim Rasel. 2023. Privacy Preserving Federated Learning Approach for Speech Emotion Recognition. In *2023 26th International Conference on Computer and Information Technology (ICCIT)*. 1–6. https://doi.org/10.1109/ICCIT60459.2023.10441577

[44] Danielle Keats Citron and Daniel J Solove. 2022. Privacy harms. *BUL Rev.* 102 (2022), 793.

[45] Julie E Cohen. 2019. Turning privacy inside out. *Theoretical inquiries in law* 20, 1 (2019), 1–31.

[46] Shanley Corvite, Kat Roemmich, Tillie Ilana Rosenberg, and Nazanin Andalibi. 2023. Data Subjects' Perspectives on Emotion Artificial Intelligence Use in the Workplace: A Relational Ethics Lens. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–38.

[47] Chirag Dalvi, Manish Rathod, Shruti Patil, Shilpa Gite, and Ketan Kotecha. 2021. A survey of ai-based facial emotion recognition: Features, ml & dl techniques, age-wise datasets and future directions. *Ieee Access* 9 (2021), 165806–165840.

[48] Saul Davila-Gonzalez and Sergio Martin. 2024. Human digital twin in industry 5.0: A holistic approach to worker safety and well-being through advanced AI and emotional analytics. *Sensors* 24, 2 (2024), 655.

[49] Luca Davoli, Marco Martalò, Antonio Cilfone, Laura Belli, Gianluigi Ferrari, Roberta Presta, Roberto Montanari, Maura Mengoni, Luca Giraldi, Elvio G Amparore, et al. 2020. On driver behavior recognition for increased safety: a roadmap. *Safety* 6, 4 (2020), 55.

[50] Nathalie DiBerardino and Luke Stark. 2023. (Anti)-Intentional Harms: The Conceptual Pitfalls of Emotion AI in Education. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 1386–1395.

[51] Mateja Durovic and Tommaso Corno. 2025. The Privacy of Emotions: From the GDPR to the AI Act, an Overview of Emotional AI Regulation and the Protection of Privacy and Personal Data. *Privacy, Data Protection and Data-driven Technologies* (2025), 368–404.

[52] Cynthia Dwork. 2006. Differential privacy. In *International colloquium on automata, languages, and programming*. Springer, 1–12.

[53] Anirudh Ekambaranathan, Jun Zhao, and Max Van Kleek. 2023. How Can We Design Privacy-Friendly Apps for Children? Using a Research through Design Process to Understand Developers' Needs and Challenges. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–29.

[54] Paul Ekman. 1992. *Are there basic emotions?* American Psychological Association.

[55] Tiantian Feng, Hanieh Hashemi, Murali Annavaram, and Shrikanth S. Narayanan. 2022. Enhancing Privacy Through Domain Adaptive Noise Injection For Speech Emotion Recognition. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 7702–7706. https://doi.org/10.1109/ICASSP43922.2022.9747265

[56] Tiantian Feng and Shrikanth Narayanan. 2021. Privacy and Utility Preserving Data Transformation for Speech Emotion Recognition. In *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 1–7. https://doi.org/10.1109/ACII52823.2021.9597433

[57] Dario Fenoglio, Daniel Josifovski, Alessandro Gobbetti, Mattias Formo, Hristijan Gjoreski, Martin Gjoreski, and Marc Langheinrich. 2023. Federated Learning for Privacy-aware Cognitive Workload Estimation. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (Vienna, Austria) *(MUM '23)*. Association for Computing Machinery, New York, NY, USA, 25–36. https://doi.org/10.1145/3626705.3627783

[58] Sidney Fussell. [n. d.]. Alexa wants to know how you're feeling today. *The Atlantic* ([n. d.]). https://www.theatlantic.com/technology/archive/2018/10/alexa-emotion-detection-ai-surveillance/572884/

[59] Neha Gahlan and Divyashikha Sethia. 2024. Federated learning inspired privacy sensitive emotion recognition based on multi-modal physiological sensors. *Cluster Computing* 27, 3 (2024), 3179–3201.

[60] Patricia Garcia, Tonia Sutherland, Niloufar Salehi, Marika Cifor, and Anubha Singh. 2022. No! Re-imagining data practices through the lens of critical refusal. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–20.

[61] Rachel L German and K Suzanne Barber. 2018. Consumer attitudes about biometric authentication. *Univ. Texas, Austin, TX, USA, UT CID Rep* (2018), 18–03.

[62] Tapotosh Ghosh, Md Hasan Al Banna, Md Jaber Al Nahian, M Shamim Kaiser, Mufti Mahmud, Shaobao Li, and Nelishia Pillay. 2022. A privacy-preserving federated-mobilenet for facial expression detection from images. In *International Conference on Applied Intelligence and Informatics*. Springer, 277–292.

[63] G Gibbs. 2007. Thematic Coding and Categorizing.

[64] Sabrina Goellner, Marina Tropmann-Frick, and Bostjan Brumen. 2024. Responsible Artificial Intelligence: A Structured Literature Review. *arXiv preprint arXiv:2403.06910* (2024).

[65] Aayush Gupta, Ayush Jaiswal, Yue Wu, Vivek Yadav, and Pradeep Natarajan. 2021. Adversarial Mask Generation for Preserving Visual Privacy. In *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*. 1–5. https://doi.org/10.1109/FG52635.2021.9666933

[66] Chetna Gupta, Vikas Khullar, Nitin Goyal, Kirti Saini, Ritu Baniwal, Sushil Kumar, and Rashi Rastogi. 2024. Cross-Silo, Privacy-Preserving, and Lightweight Federated Multimodal System for the Identification of Major Depressive Disorder Using Audio and Electroencephalogram. *Diagnostics* 14, 1 (2024), 43.

[67] Seda Gürses. 2010. PETs and their users: a critical review of the potentials and limitations of the privacy as confidentiality paradigm. *Identity in the Information Society* 3 (2010), 539–563.

[68] Irit Hadar, Tomer Hasson, Oshrat Ayalon, Eran Toch, Michael Birnhack, Sofia Sherman, and Arod Balissa. 2018. Privacy by designers: software developers' privacy mindset. *Empirical Software Engineering* 23 (2018), 259–289.

[69] Zhipeng He, Zina Li, Fuzhou Yang, Lei Wang, Jingcong Li, Chengju Zhou, and Jiahui Pan. 2020. Advances in multimodal emotion recognition based on brain–computer interfaces. *Brain sciences* 10, 10 (2020), 687.

[70] Fabio Hellmann, Silvan Mertes, Mohamed Benouis, Alexander Hustinx, Tzung-Chien Hsieh, Cristina Conati, Peter Krawitz, and Elisabeth André. 2024. GANonymization: A GAN-based Face Anonymization Framework for Preserving Emotional Expressions. *ACM Trans. Multimedia Comput. Commun. Appl.* (Jan. 2024). https://doi.org/10.1145/3641107 Just Accepted.

[71] Javier Hernandez, Josh Lovejoy, Daniel McDuff, Jina Suh, Tim O'Brien, Arathi Sethumadhavan, Gretchen Greene, Rosalind Picard, and Mary Czerwinski. 2021. Guidelines for assessing and minimizing risks of emotion recognition applications. In *2021 9th International conference on affective computing and intelligent interaction (ACII)*. IEEE, 1–8.

[72] Ayanna Howard, Cha Zhang, and Eric Horvitz. 2017. Addressing bias in machine learning algorithms: A pilot study on emotion recognition for intelligent systems. In *2017 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*. IEEE, 1–7.

[73] Sungjoo Hwang, Houtan Jebelli, Byungjoo Choi, Minji Choi, and SangHyun Lee. 2018. Measuring workers' emotional state during construction tasks using wearable EEG. *Journal of Construction Engineering and Management* 144, 7 (2018), 04018050.

[74] Mimansa Jaiswal and Emily Mower Provost. 2020. Privacy enhanced multimodal neural representations for emotion recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 7985–7993.

[75] Edward B Kang. 2023. On the Praxes and Politics of AI Speech Emotion Recognition. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 455–466.

[76] Nadia Karizat, Alexandra H Vinson, Shobita Parthasarathy, and Nazanin Andalibi. 2024. Patent Applications as Glimpses into the Sociotechnical Imaginary: Ethical Speculation on the Imagined Futures of Emotion AI for Mental Health Monitoring and Detection. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–43.

[77] Mounika Karna, D Sujitha Juliet, and R Catherine Joy. 2020. Deep learning based text emotion recognition for chatbot applications. In *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*. IEEE, 988–993.

[78] Amelia Katirai. 2024. Ethical considerations in emotion recognition technologies: a review of the literature. *AI and Ethics* 4, 4 (2024), 927–948.

[79] Harmanpreet Kaur, Daniel McDuff, Alex C Williams, Jaime Teevan, and Shamsi T Iqbal. 2022. "I didn't know I looked angry": Characterizing observed emotion and reported affect at work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.

[80] Simran Kaur and Richa Sharma. 2021. Emotion AI: integrating emotional intelligence with artificial intelligence in the digital workplace. In *Innovations in Information and Communication Technologies (IICT-2020) Proceedings of International Conference on ICRIHE-2020, Delhi, India: IICT-2020*. Springer, 337–343.

[81] Ruhul Amin Khalil, Edward Jones, Mohammad Inayatullah Babar, Tariqullah Jan, Mohammad Haseeb Zafar, and Thamer Alhussain. 2019. Speech emotion recognition using deep learning techniques: A review. *IEEE access* 7 (2019), 117327–117345.

[82] Eugenia Kim, De'Aira Bryant, Deepak Srikanth, and Ayanna Howard. 2021. Age Bias in Emotion Detection: An Analysis of Facial Emotion Recognition Performance on Young, Middle-Aged, and Older Adults. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) *(AIES '21)*. Association for Computing Machinery, New York, NY, USA, 638–644. https://doi.org/10.1145/3461702.3462609

[83] Yelin Kim, Tolga Soyata, and Reza Feyzi Behnagh. 2018. Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education. *Ieee Access* 6 (2018), 5308–5331.

[84] Konrad Kollnig, Siddhartha Datta, Thomas Serban Von Davier, Max Van Kleek, Reuben Binns, Ulrik Lyngs, and Nigel Shadbolt. 2023. 'We are adults and deserve control of our phones': Examining the risks and opportunities of a right to repair for mobile apps. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 22–34.

[85] Akshi Kumar, Aditi Sharma, Ravi Ranjan, and Liangxiu Han. 2023. FTL-Emo: Federated Transfer Learning for Privacy Preserved Biomarker-Based Automatic Emotion Recognition. In *International Conference on Data Analytics & Management*. Springer, 449–460.

[86] Guohao Lan, Tim Scargill, and Maria Gorlatova. 2022. EyeSyn: Psychology-inspired Eye Movement Synthesis for Gaze-based Activity Recognition. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. 233–246. https://doi.org/10.1109/IPSN54338.2022.00026

[87] Siddique Latif, Hafiz Shehbaz Ali, Muhammad Usama, Rajib Rana, Björn Schuller, and Junaid Qadir. 2022. Ai-based emotion recognition: Promise, peril, and prescriptions for prosocial path. *arXiv preprint arXiv:2211.07290* (2022).

[88] Siddique Latif, Junaid Qadir, Adnan Qayyum, Muhammad Usama, and Shahzad Younis. 2020. Speech technology for healthcare: Opportunities, challenges, and state of the art. *IEEE Reviews in Biomedical Engineering* 14 (2020), 342–356.

[89] Ruth Leys. 2019. *The ascent of affect: Genealogy and critique*. University of Chicago Press.

[90] Han Li, Renwen Zhang, Yi-Chieh Lee, Robert E Kraut, and David C Mohr. 2023. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *NPJ Digital Medicine* 6, 1 (2023), 236.

[91] Li Li, Yuxi Fan, Mike Tse, and Kuo-Yi Lin. 2020. A review of applications in federated learning. *Computers & Industrial Engineering* 149 (2020), 106854.

[92] Shan Li and Weihong Deng. 2020. Deep facial expression recognition: A survey. *IEEE transactions on affective computing* 13, 3 (2020), 1195–1215.

[93] Xiang Li, Yazhou Zhang, Prayag Tiwari, Dawei Song, Bin Hu, Meihong Yang, Zhigang Zhao, Neeraj Kumar, and Pekka Marttinen. 2022. EEG based emotion recognition: A tutorial and review. *Comput. Surveys* 55, 4 (2022), 1–57.

[94] Shu Liu, Kevin Koch, Zimu Zhou, Simon Föll, Xiaoxi He, Tina Menke, Elgar Fleisch, and Felix Wortmann. 2021. The empathetic car: Exploring emotion inference via driver behaviour and traffic context. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–34.

[95] Yuping Liu-Thompkins, Shintaro Okazaki, and Hairong Li. 2022. Artificial empathy in marketing interactions: Bridging the human-AI gap in affective and social customer experience. *Journal of the Academy of Marketing Science* 50, 6 (2022), 1198–1218.

[96] Yin-Yin Low, Angeline Tanvy, Raphaël C.-W. Phan, and Xiaojun Chang. 2022. AdverFacial: Privacy-Preserving Universal Adversarial Perturbation Against Facial Micro-Expression Leakages. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2754–2758. https://doi.org/10.1109/ICASSP43922.2022.9746848

[97] Daniel Lowd and Christopher Meek. 2005. Adversarial learning. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. 641–647.

[98] Daniel Isoso Machanje, Joseph Onderi Orero, and Christophe Marsala. 2019. Distress Recognition from Speech Analysis: A Pairwise Association Rules-Based Approach. In *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. 842–849. https://doi.org/10.1109/SSCI44817.2019.9002972

[99] Koshiro Maeda, Isao Kurebayashi, Nobuyoshi Komuro, and Keita Hirai. 2021. A Study on Time Series Analysis of Environmental Data for Predicting Emotional Conditions. In *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*. 273–274. https://doi.org/10.1109/LifeTech52111.2021.9391865

[100] Marion Mainsant, Martial Mermillod, Christelle Godin, and Marina Reyboz. 2022. A study of the Dream Net model robustness across continual learning scenarios. In *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*. 824–833. https://doi.org/10.1109/ICDMW58026.2022.00111

[101] Marion Mainsant, Miguel Solinas, Marina Reyboz, Christelle Godin, and Martial Mermillod. 2021. Dream Net: a privacy preserving continual leaming model for face emotion recognition. In *2021 9th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. 01–08. https://doi.org/10.1109/ACIIW52867.2021.9666338

[102] Konstantinos Makantasis, David Melhart, Antonios Liapis, and Georgios N. Yannakakis. 2021. Privileged Information for Modeling Affect In The Wild. In *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 1–8. https://doi.org/10.1109/ACII52823.2021.9597417

[103] Matthew Malek–Podjaski and Fani Deligianni. 2021. Towards Explainable, Privacy-Preserved Human-Motion Affect Recognition. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. 01–09. https://doi.org/10.1109/SSCI50451.2021.9660129

[104] Peter Mantello and Manh-Tung Ho. 2024. Emotional AI and the future of wellbeing in the post-pandemic workplace. *AI & society* 39, 4 (2024), 1883–1889.

[105] Peter Mantello, Manh-Tung Ho, Minh-Hoang Nguyen, and Quan-Hoang Vuong. 2023. Bosses without a heart: socio-demographic and cross-cultural determinants of attitude toward Emotional AI in the workplace. *AI & society* 38, 1 (2023), 97–119.

[106] MarketsandMarkets. 2024. Emotion Detection and Recognition Market by Component, Technology, Application, and Region - Global Forecast to 2024. https://www.marketsandmarkets.com/Market-Reports/emotion-detection-recognition-market-23376176.html Accessed: 2024-09-05.

[107] Kerry McInerney and Os Keyes. 2024. The Infopolitics of feeling: How race and disability are configured in Emotion Recognition Technology. *New Media & Society* (2024), 14614448241235914.

[108] Andrew McStay. 2016. Empathic media and advertising: Industry, policy, legal and citizen perspectives (the case for intimacy). *Big data & society* 3, 2 (2016), 2053951716666868.

[109] Andrew McStay. 2020. Emotional AI and EdTech: serving the public good? *Learning, Media and Technology* 45, 3 (2020), 270–283.

[110] Andrew McStay and Pamela Pavliscak. 2019. Emotional artificial intelligence: Guidelines for ethical use. *COMEST/UNESCO* (2019).

[111] Kateryna Mishchenko, Samaneh Mohammadi, Mohammadreza Mohammadi, and Sima Sinaei. 2023. Hyperparameters optimization for federated learning system: Speech emotion recognition case study. In *2023 Eighth International Conference on Fog and Mobile Edge Computing (FMEC)*. IEEE, 80–86.

[112] Emmanuel Mogaji, Sunday Olaleye, and Dandison Ukpabi. 2020. Using AI to personalise emotionally appealing advertisement. *Digital and social media marketing: Emerging applications and theoretical development* (2020), 137–150.

[113] Saif M Mohammad. 2022. Ethics sheet for automatic emotion recognition and sentiment analysis. *Computational Linguistics* 48, 2 (2022), 239–278.

[114] Samaneh Mohammadi, Mohammadreza Mohammadi, Sima Sinaei, Ali Balador, Ehsan Nowroozi, Francesco Flammini, and Mauro Conti. 2023. Balancing Privacy and Accuracy in Federated Learning for Speech Emotion Recognition. In *2023 18th Conference on Computer Science and Intelligence Systems (FedCSIS)*. 191–199. https://doi.org/10.15439/2023F444

[115] Samaneh Mohammadi, Sima Sinaei, Ali Balador, and Francesco Flammini. 2023. Optimized Paillier Homomorphic Encryption in Federated Learning for Speech Emotion Recognition. In *2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*. 1021–1022. https://doi.org/10.1109/COMPSAC57700.2023.00156

[116] Scott Monteith, Tasha Glenn, John Geddes, Peter C Whybrow, and Michael Bauer. 2022. Commercial use of emotion artificial intelligence (AI): implications for psychiatry. *Current Psychiatry Reports* 24, 3 (2022), 203–211.

[117] Jeff Nagy. 2024. Autism and the making of emotion AI: Disability as resource for surveillance capitalism. *New media & society* 26, 7 (2024), 3989–4007.

[118] Arijit Nandi and Fatos Xhafa. 2022. A federated learning method for real-time emotion state classification from multi-modal streaming. *Methods* 204 (2022), 340–347.

[119] Vansh Narula, Kexin Feng, and Theodora Chaspari. 2020. Preserving Privacy in Image-based Emotion Recognition through User Anonymization. In *Proceedings of the 2020 International Conference on Multimodal Interaction* (Virtual Event, Netherlands) *(ICMI '20)*. Association for Computing Machinery, New York, NY, USA, 452–460. https://doi.org/10.1145/3382507.3418833

[120] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.

[121] Helen Nissenbaum. 2009. Privacy in Context: Technology, Policy, and the Integrity of Social Life.

[122] Jonathan Shi Khai Ooi, Siti Anom Ahmad, Yu Zheng Chong, Sawal Hamid Md Ali, Guangyi Ai, and Hiroaki Wagatsuma. 2016. Driver emotion recognition framework based on electrodermal activity measurements during simulated driving conditions. In *2016 IEEE EMBS conference on biomedical engineering and sciences (IECBES)*. IEEE, 365–369.

[123] Luis Felipe Ortiz-Clavijo, Carlos Julián Gallego-Duque, Juan Camilo David-Diaz, and Andrés Felipe Ortiz-Zamora. 2023. Implications of Emotion Recognition Technologies: Balancing Privacy and Public Safety. *IEEE Technology and Society Magazine* 42, 3 (2023), 69–75.

[124] Yuchen Pan, Yuanyuan Shang, Zhuhong Shao, Tie Liu, Guodong Guo, and Hui Ding. 2024. Integrating Deep Facial Priors Into Landmarks for Privacy Preserving Multimodal Depression Recognition. *IEEE Transactions on Affective Computing* 15, 3 (2024), 828–836. https://doi.org/10.1109/TAFFC.2023.3296318

[125] A. Parkavi, Tejas B N Shetty, Shane George Shibu, Roshan Ismail, and Ryan Anil Muthirakalayil. 2023. Customer Feedback Analysis Based on Emotion Detection Using Machine Learning Techniques with Privacy Preservation. In *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*. 1617–1624. https://doi.org/10.1109/ICCPCT58313.2023.10244876

[126] Rosalind W Picard. 2000. *Affective computing*. MIT press.

[127] Lena Podoletz. 2023. We have to talk about emotional AI and crime. *AI & SOCIETY* 38, 3 (2023), 1067–1082.

[128] Ravi Pranjal, Ranjana Seshadri, Rakesh Kumar Sanath Kumar Kadaba, Tiantian Feng, Shrikanth S. Narayanan, and Theodora Chaspari. 2023. Toward Privacy-Enhancing Ambulatory-Based Well-Being Monitoring: Investigating User Re-Identification Risk in Multimodal Data. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1–5. https://doi.org/10.1109/ICASSP49357.2023.10096235

[129] Paul Prinsloo, Sharon Slade, and Mohammad Khalil. 2022. The answer is (not only) technological: Considering student data privacy in learning analytics. *British Journal of Educational Technology* 53, 4 (2022), 876–893.

[130] Fan Qi, Zixin Zhang, Xianshan Yang, Huaiwen Zhang, and Changsheng Xu. 2022. Feeling Without Sharing: A Federated Video Emotion Recognition Framework Via Privacy-Agnostic Hybrid Aggregation. In *Proceedings of the 30th ACM International Conference on Multimedia* (Lisboa, Portugal) *(MM '22)*. Association for Computing Machinery, New York, NY, USA, 151–160. https://doi.org/10.1145/3503161.3548278

[131] Vinesh Ravuri, Ricardo Gutierrez-Osuna, and Theodora Chaspari. 2022. Preserving Mental Health Information in Speech Anonymization. In *2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. 1–8. https://doi.org/10.1109/ACIIW57231.2022.10086012

[132] Narsi Reddy and Reza Derakhshani. 2020. Emotion Detection using Periocular Region: A Cross-Dataset Study. In *2020 International Joint Conference on Neural Networks (IJCNN)*. 1–6. https://doi.org/10.1109/IJCNN48605.2020.9207542

[133] Stefan Reindl. 2021. Emotion AI in education: a literature review. *International Journal of Learning Technology* 16, 4 (2021), 288–302.

[134] Neil Richards. 2015. *Intellectual privacy: Rethinking civil liberties in the digital age*. Oxford University Press, USA.

[135] Lisa S Roberts. 2012. *A forensic phonetic study of the vocal responses of individuals in distress*. Ph. D. Dissertation. University of York.

[136] Kat Roemmich and Nazanin Andalibi. 2021. Data subjects' conceptualizations of and attitudes toward automatic emotion recognition-enabled wellbeing interventions on social media. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–34.

[137] Kat Roemmich, Shanley Corvite, Cassidy Pyle, Nadia Karizat, and Nazanin Andalibi. 2024. Emotion AI Use in US Mental Healthcare: Potentially Unjust and Techno-Solutionist. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–46.

[138] Kat Roemmich, Tillie Rosenberg, Serena Fan, and Nazanin Andalibi. 2023. Values in emotion artificial intelligence hiring services: Technosolutions to organizational problems. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–28.

[139] Kat Roemmich, Florian Schaub, and Nazanin Andalibi. 2023. Emotion AI at work: Implications for workplace surveillance, emotional labor, and emotional privacy. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–20.

[140] James A Russell, José-Miguel Fernández-Dols, Anthony SR Manstead, and Jane C Wellenkamp. 2013. *Everyday conceptions of emotion: An introduction to the psychology, anthropology and linguistics of emotion*. Vol. 81. Springer Science & Business Media.

[141] Johnny Saldana. 2014. *Thinking qualitatively: Methods of mind*. SAGE publications.

[142] Javier Sánchez-Monedero and Lina Dencik. 2022. The politics of deceptive borders:'biomarkers of deceit'and the case of iBorderCtrl. *Information, Communication & Society* 25, 3 (2022), 413–430.

[143] SKB Sangeetha, Rajeswari Rajesh Immanuel, Sandeep Kumar Mathivanan, Jaehyuk Cho, and Sathishkumar Veerappampalayam Easwaramoorthy. 2024. An Empirical Analysis of Multimodal Affective Computing Approaches for Advancing Emotional Intelligence in Artificial Intelligence for Healthcare. *IEEE Access* (2024).

[144] Jeremy Seeman and Daniel Susser. 2024. Between privacy and utility: On differential privacy in theory and practice. *ACM Journal on Responsible Computing* 1, 1 (2024), 1–18.

[145] Abdallah Hussein Sham, Kadir Aktas, Davit Rizhinashvili, Danila Kuklianov, Fatih Alisinanoglu, Ikechukwu Ofodile, Cagri Ozcinar, and Gholamreza Anbarjafari. 2023. Ethical AI in facial expression analysis: racial bias. *Signal, Image and Video Processing* 17, 2 (2023), 399–406.

[146] Apoorva Singh, Siddarth Chandrasekar, Sriparna Saha, and Tanmay Sen. 2023. Federated meta-learning for emotion and sentiment aware multi-modal complaint identification. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. 16091–16103.

[147] Ankit Kumar Singh, Ajit Kumar, and Bong Jun Choi. 2022. Privacy-Preserving Digital Intervention for Mental Health Using Federated Learning. In *International Conference on Intelligent Human Computer Interaction*. Springer, 213–224.

[148] Congzheng Song and Vitaly Shmatikov. 2019. Overlearning reveals sensitive attributes. *arXiv preprint arXiv:1905.11742* (2019).

[149] Luke Stark. 2016. The emotional context of information privacy. *The Information Society* 32, 1 (2016), 14–27.

[150] Luke Stark and Jesse Hoey. 2021. The ethics of emotion in artificial intelligence systems. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 782–793.

[151] Luke Stark, Amanda Stanhaus, and Denise L Anthony. 2020. "i don't want someone to watch me while i'm working": Gendered views of facial recognition technology in workplace surveillance. *Journal of the Association for Information Science and Technology* 71, 9 (2020), 1074–1088.

[152] Hung-Yue Suen, Kuo-En Hung, Che-Wei Liu, Yu-Sheng Su, and Han-Chih Fan. 2024. Artificial Intelligence Can Recognize Whether a Job Applicant Is Selling and/or Lying According to Facial Expressions and Head Movements Much More Correctly Than Human Interviewers. *IEEE Transactions on Computational Social Systems* (2024).

[153] Daniel Susser and Jeremy Seeman. 2024. Critical Provocations for Synthetic Data. *Surveillance and Society* (2024).
[154] Chao Tan, Yang Cao, Sheng Li, and Masatoshi Yoshikawa. 2023. General or Specific? Investigating Effective Privacy Protection in Federated Learning for Speech Emotion Recognition. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1–5. https://doi.org/10.1109/ICASSP49357.2023.10096844
[155] Brian Testa, Yi Xiao, Harshit Sharma, Avery Gump, and Asif Salekin. 2023. Privacy against Real-Time Speech Emotion Detection via Acoustic Adversarial Evasion of Machine Learning. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 3, Article 126 (Sept. 2023), 30 pages. https://doi.org/10.1145/3610887
[156] M Thenmozhi and K Narmadha. 2020. Privacy-enhanced emotion recognition approach for remote health advisory system. In *Artificial Intelligence and Evolutionary Computations in Engineering Systems*. Springer, 133–142.
[157] Ermal Toto, ML Tlachac, and Elke A. Rundensteiner. 2021. AudiBERT: A Deep Transfer Learning Multimodal Classification Framework for Depression Screening. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (Virtual Event, Queensland, Australia) *(CIKM '21)*. Association for Computing Machinery, New York, NY, USA, 4145–4154. https://doi.org/10.1145/3459637.3481895
[158] Ermal Toto, ML Tlachac, Francis Lee Stevens, and Elke A. Rundensteiner. 2020. Audio-based Depression Screening using Sliding Window Sub-clip Pooling. In *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 791–796. https://doi.org/10.1109/ICMLA51294.2020.00129
[159] Minh Tran and Mohammad Soleymani. 2023. A Speech Representation Anonymization Framework via Selective Noise Perturbation. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1–5. https://doi.org/10.1109/ICASSP49357.2023.10095173
[160] Vasileios Tsouvalas, Tanir Ozcelebi, and Nirvana Meratnia. 2022. Privacy-preserving Speech Emotion Recognition through Semi-Supervised Federated Learning. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. 359–364. https://doi.org/10.1109/PerComWorkshops53856.2022.9767445
[161] Md Taufeeq Uddin, Lijun Yin, and Shaun Canavan. 2024. Spatio-Temporal Graph Analytics on Secondary Affect Data for Improving Trustworthy Emotional AI. *IEEE Transactions on Affective Computing* 15, 1 (2024), 30–49. https://doi.org/10.1109/TAFFC.2023.3296695
[162] European Union. 2024. Artificial Intelligence Act. https://artificialintelligenceact.eu/recital/44/ Accessed: 2024-09-05.
[163] Dhruv Verma, Sejal Bhalla, Dhruv Sahnan, Jainendra Shukla, and Aman Parnami. 2021. ExpressEar: Sensing Fine-Grained Facial Expressions with Earables. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 129 (Sept. 2021), 28 pages. https://doi.org/10.1145/3478085
[164] Panagiotis Vrettis, Andreas Mallas, and Michalis Xenos. 2024. An Affect-Aware Game Adapting to Human Emotion. In *International Conference on Human-Computer Interaction*. Springer, 307–322.
[165] Ari Ezra Waldman. 2018. Designing without privacy. *Houston Law Review* 55, 659 (2018).
[166] Xiaowei Wang and Shazeda Ahmed. 2023. Bodily Harms: How AI and Biometrics Curtail Human Rights. https://www.accessnow.org/bodily-harms-how-ai-and-biometrics-curtail-human-rights/
[167] Taiba Majid Wani, Teddy Surya Gunawan, Syed Asif Ahmad Qadri, Mira Kartiwi, and Eliathamby Ambikairajah. 2021. A comprehensive review of speech emotion recognition systems. *IEEE access* 9 (2021), 47795–47814.
[168] Dinah Waref and Mohammed Salem. 2022. Split Federated Learning for Emotion Detection. In *2022 4th Novel Intelligent and Leading Emerging Sciences Conference (NILES)*. 112–115. https://doi.org/10.1109/NILES56402.2022.9942417
[169] S Warren and L Brandeis. 1890. The Right to Privacy, Harvard Law Abstract. 1890. *No IV* (1890).
[170] Cedric Deslandes Whitney and Justin Norman. 2024. Real Risks of Fake Data: Synthetic Data, Diversity-Washing and Consent Circumvention. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (Rio de Janeiro, Brazil) *(FAccT '24)*. Association for Computing Machinery, New York, NY, USA, 1733–1744. https://doi.org/10.1145/3630106.3659002
[171] Richmond Y. Wong, Andrew Chong, and R. Cooper Aspegren. 2023. Privacy Legislation as Business Risks: How GDPR and CCPA are Represented in Technology Companies' Investment Risk Disclosures. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 82 (April 2023), 26 pages. https://doi.org/10.1145/3579515
[172] James Wright. 2023. Suspect AI: Vibraimage, emotion recognition technology and algorithmic opacity. *Science, Technology and Society* 28, 3 (2023), 468–487.
[173] Claire Xu. 2023. HarmonyVisage: Ethical Facial Emotion Dataset Using Advanced Generative AI. In *2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. 0558–0563. https://doi.org/10.1109/UEMCON59035.2023.10316128
[174] Shihao Xu, Jing Fang, Xiping Hu, Edith Ngai, Wei Wang, Yi Guo, and Victor CM Leung. 2022. Emotion recognition from gait analyses: Current research and future directions. *IEEE Transactions on Computational Social Systems* 11, 1 (2022), 363–377.
[175] Syeda Sana e Zainab and Tahar Kechadi. 2019. Sensitive and private data analysis: A systematic review. In *Proceedings of the 3rd International Conference on Future Networks and Distributed Systems*. 1–11.
[176] Yong Zeng, Zhenyu Zhang, Jiale Liu, Jianfeng Ma, and Zhihong Liu. 2023. PriEMO: A universal perturbation method for privacy preserving facial emotion recognition. *Journal of Information and Intelligence* 1, 4 (2023), 330–340.
[177] Sebastian Zepf, Javier Hernandez, Alexander Schmitt, Wolfgang Minker, and Rosalind W Picard. 2020. Driver emotion recognition for intelligent vehicles: A survey. *ACM Computing Surveys (CSUR)* 53, 3 (2020), 1–30.
[178] Huan Zhao, Haijiao Chen, Yufeng Xiao, and Zixing Zhang. 2023. Privacy-Enhanced Federated Learning Against Attribute Inference Attack for Speech Emotion Recognition. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1–5. https://doi.org/10.1109/ICASSP49357.2023.10095737
[179] Shoshana Zuboff. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs, New York.

7 Appendix

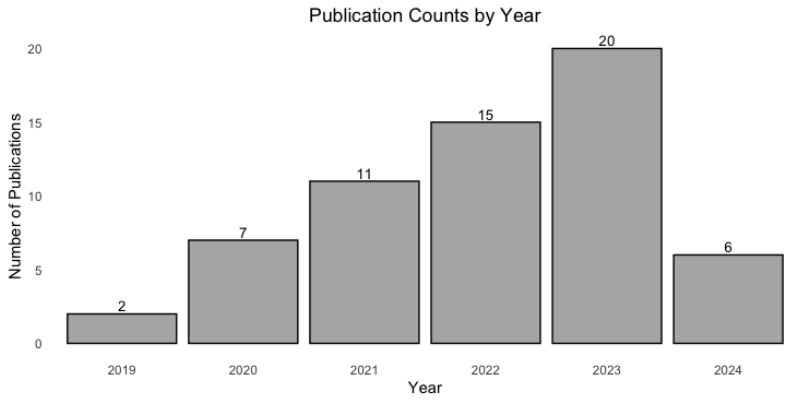


Figure 1: Publication counts in our corpus by year.

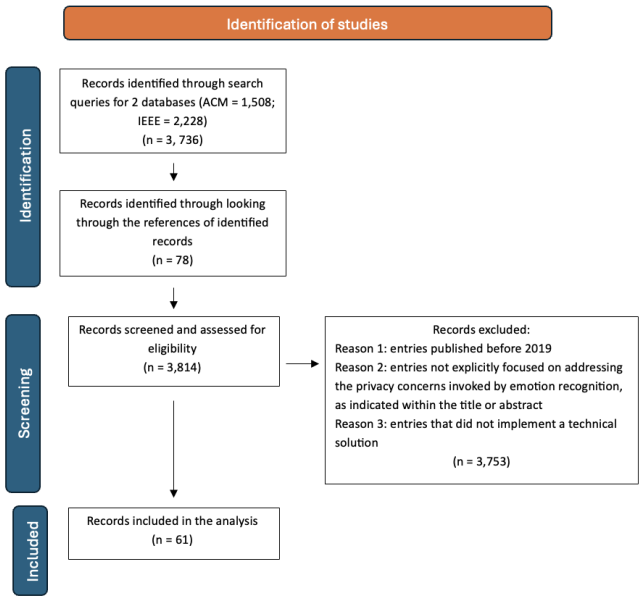


Figure 2: Flowchart detailing article selection procedure

Table 3: Characteristics of the reviewed papers

Author names and year	Domain	Input data	Technical approach
Agarwal et al. (2019) [4]	Driving	EEG	Cryptographic methods
Agarwal et al. (2021) [3]	Unspecified	Facial data	Anonymization; use of synthetic data
Agrawal et al. (2023) [5]	Unspecified	EEG	Federated learning
Akhyani et al. (2022) [8]	Robotics	Facial data	Use of synthetic data
Ali et al. (2021) [9]	Unspecified	Speech data	Adversarial learning
Al-Nuaimi et al. (2023) [10]	Unspecified	Facial data	Data perturbation; encryption
Alshareef et al. (2023) [12]	Healthcare	Multimodal	Differential privacy
Anwar et al. (2023) [15]	Unspecified	EEG	Federated learning
Augusma (2022) [16]	Education	Multimodal	Collecting less sensitive data
Benouis et al. (2023) [24]	Unspecified	Multimodal	Federated learning; differential privacy
Bethge et al. (2023) [26]	Driving	Multimodal	Collecting less sensitive data
Bethge et al. (2022) [27]	Unspecified	EEG	Alternative machine-learning architecture
Bisogni et al. (2024) [28]	Unspecified	Gait data	Federated learning
Bondin and Dingli (2021) [30]	Healthcare	ECG	Federated learning
Can and Ersoy (2021) [36]	Healthcare	Heart activity	Federated learning
Chen et al. (2022) [40]	Unspecified	Speech	Collecting less sensitive data
Chhikara et al. (2021) [41]	Workplace	Facial data	Federated learning
Chowdhury et al. (2023) [43]	Unspecified	Speech, audio-visual	Federated learning
Feng et al. (2022) [55]	Unspecified	Speech	Adversarial learning; data perturbation
Feng and Narayanan (2021) [56]	Unspecified	Speech	Adversarial learning
Fenoglio et al. (2023) [57]	Healthcare	Multimodal	Federated learning
Gahlan and Sethia (2024) [59]	Unspecified	Multimodal	Federated learning
Ghosh et al. (2022) [62]	Unspecified	Facial data	Federated learning
Gupta et al. (2021) [65]	Unspecified	Facial data	Adversarial learning
Gupta et al. (2024) [66]	Healthcare	Speech; EEG	Federated learning
Hellmann et al. (2024) [70]	Healthcare	Facial data	Anonymization; use of synthetic data
Jaiswal and Provost (2020) [74]	Unspecified	Multimodal	Adversarial learning
Kumar et al. (2023) [85]	Healthcare	EEG	Federated learning
Lan et al. (2022) [86]	Unspecified	Eye gaze	Data perturbation; use of synthetic data
Low et al. (2022) [96]	Unspecified	Facial data	Data perturbation; use of synthetic data
Mainsant et al. (2022) [100]	Unspecified	Facial data	Alternative machine learning architecture (continual learning)
Machanje et al. (2019) [98]	Healthcare	Speech	Collection of less sensitive data
Maeda et al (2021) [99]	Unspecified	Environmental data	Collection of less sensitive data
Mainsant et al. (2021) [101]	Unspecified	Facial data	Alternative machine learning architecture (continual learning)
Makantasis et al. (2021) [102]	Video games	Multimodal	Collecting less sensitive data
Malek-Podjaski and Deligianni (2021) [103]	Unspecified	Gait data	Adversarial learning
Mishchenko et al. (2023) [111]	Unspecified	Speech	Federated learning
Mohammadi et al. (2023) [114]	Unspecified	Speech	Differential privacy; federated learning
Mohammadi et al. (2023) [115]	Unspecified	Speech	Cryptographic methods; federated learning
Nandi and Xhafa (2022) [118]	Unspecified	Multimodal	Federated learning
Narula et al. (2020) [119]	Healthcare	Facial data; multi-modal	Anonymization

Author names and year	Domain	Input data	Technical approach
Pan et al. (2023) [124]	Mental health	Facial data	Collecting less sensitive data
Parkavi et al. (2023) [125]	Customer experience	Facial data	Federated learning; data perturbation
Pranjal et al. (2023) [128]	Workplace	Ambulatory	Anonymization
Qi et al. (2022) [130]	Unspecified	Facial data	Federated learning
Ravuri et al. (2022) [131]	Healthcare	Speech	Anonymization
Reddy and Derakhshani (2020) [132]	Unspecified	Facial data	Collecting less sensitive data
Singh et al. (2022) [147]	Mental health	Multimodal	Federated learning
Tan et al. (2023) [154]	Unspecified	Speech	Differential privacy; federated learning
Testa et al. (2023) [155]	Home voice assistants	Speech	Data perturbation
Thenmozhi and Narmadha (2020) [156]	Healthcare	Speech	Data perturbation; cryptographic methods
Toto et al. (2020) [158]	Healthcare	Speech	Alternative machine learning architecture
Toto et al. (2021) [157]	Healthcare	Speech	Alternative machine learning architecture
Tran and Soleymani (2023) [159]	Unspecified	Speech	Federated learning
Tsouvalas et al. (2022) [160]	Unspecified	Speech	Federated learning
Uddin et al. (2023) [161]	Unspecified	Facial data	Collecting less sensitive data
Verma et al. (2021) [163]	Unspecified	Ear movements	Collecting less sensitive data
Waref and Salem (2022) [168]	Unspecified	Facial data	Federated learning
Xu (2023) [173]	Unspecified	Facial data	Use of synthetic data
Zeng et al. (2023) [176]	Unspecified	Facial data	Data perturbation
Zhao et al. (2023) [178]	Unspecified	Speech data	Federated learning

Table 4: Characteristics of the reviewed papers (cont.)

Concern and associated papers	Strategy	Technical approach
Risk of input data exposure — concerns that the centralization of emotion AI input data or other architectures of emotion AI systems will lead to input data being exposed (to, for example, malicious actors who could misuse the data) [5, 9, 10, 12, 15, 24, 27, 41, 43, 57, 59, 62, 65, 66, 70, 74, 85, 100, 101, 111, 114, 115, 118, 119, 130, 131, 147, 155, 156, 160, 168, 176, 178] (33/62 papers)	Protect the input data [5, 9, 10, 12, 15, 24, 27, 41, 43, 57, 59, 62, 65, 66, 70, 74, 85, 100, 101, 111, 114, 115, 118, 119, 130, 131, 147, 155, 156, 160, 168, 176, 178] (22 papers)	Federated learning [5, 15, 24, 41, 43, 57, 59, 62, 66, 85, 111, 114, 115, 118, 130, 147, 160, 168, 178] (19 papers)
		Differential privacy [12, 24, 114] (3 papers)
		Cryptographic methods [10] (1 paper)
		Other machine learning approaches [27, 100, 101] (3 papers)
	Prevent sensitive inferences from being made on the data [9, 65, 70, 74, 119, 131, 155, 156] (8 papers)	Adversarial learning [9, 65, 74, 119, 131] (5 papers)
		Use of synthetic data [70] (1 paper)
		Data perturbation to anonymize data [10, 156] (2 papers)
		Data perturbation to prevent emotion inferences altogether [155] (1 paper)
Risk of inference of sensitive information — concerns that emotion AI input data could be used to make inferences of sensitive information — including PII and emotion inferences [3, 4, 9, 12, 16, 24, 40, 55–57, 59, 70, 86, 96, 99, 103, 114, 119, 125, 128, 131, 154, 155, 159, 161, 163] (26/62 papers)	Prevent input data from being exposed (to malicious actors) [4, 12, 24, 57, 59, 114, 125, 154] (8 papers)	Federated learning [24, 57, 59, 114, 154] (5 papers)
		Differential privacy [12, 24, 125] (3 papers)
		Cryptographic approaches [4] (1 paper)
	Prevent inference of PII from input data [3, 9, 55, 70, 103, 119, 125, 128, 131, 155, 159] (11 papers)	Data perturbation [3, 55, 70, 96, 125, 159] (6 papers)
		Adversarial learning [9, 103, 119, 128, 131, 155] (6 papers)
	Collect less sensitive data [16, 40, 98, 99, 161, 163] (6 papers)	Secondary affect data [16, 99, 161, 163] (4 papers)
		Data they deemed to be less sensitive (i.e., speech instead of facial) [40, 98] (2 papers)
	Prevent inference of emotions [96, 155] (2 papers)	Adversarial learning [96] (1 paper)
		Data perturbation [155] (1 paper)
Emotion AI enables surveillance, as a core goal of the technology and through design features that enable invasive data collection. [10, 16, 26, 155] (4/62 papers)	Keep input data secure [10] (1 paper)	Cryptographic approaches [10] (1 paper)
	Less intrusive data collection methods [16, 26, 155] (3 papers)	Collecting less sensitive data [16, 26] (2 papers)
	Prevent emotion recognition altogether (1 paper) [155]	Data perturbation [155] (1 paper)

Table 5: Technical strategies to address different concerns about emotion AI

Table 6: Codebook

Code	Subcode	Definition	Example from Corpus
Characterizing data		Authors' descriptions and characterizations of the nature of emotion AI input or output data	
	Intimate exploited data	Explicit description of input data as intimate and commodified or exploited	"the most intimate exploited data sources" [4]
	Proprietary	Explicit description of input or output data as proprietary	"Privacy issues are usually a concern when storing proprietary data" [101]) "This work aims to bring data privacy and emotion recognition together into a synonymous and simultaneous system that predicts human emotions using technology without the expense of handing out proprietary information on any individual." [15]
	Sensitive data	Explicit description of input or output data as sensitive	"However, speech recordings are a rich source of sensitive personal information" [131]
	Some data are more dangerous	Characterizations of the collection, analysis or storage of some types of data as more risky or dangerous for data subjects than others	"Compared with visual data, speech signals have proven more difficult to anonymize" [40]
Characterizing data subjects		Broad characterizations of their data subjects, taking the form of sentences where data subjects are the subject of the sentence. These sentences describe data subjects' perspectives, fears, preferences, actions, etc.	"Moreover, many people are not comfortable sharing their personal images for emotion-aware applications." [62]
	Data subjects' agency	Specific characterizations related to the agency that data subjects do or do not have, or what actions they can or cannot take. Was also used to code data subjects' preferences about emotion recognition.	"The purpose of this work is to empower smart speaker VA users to protect their private emotion information." [155]
Domain		The context for deployment that the papers designed their approach for.	
	Cars	Purpose of technology was for the use case of detecting whether people driving are drowsy or fatigued.	"...estimating the drowsiness of drivers, which is the cause of 1000s of fatal crashes each year." [4]
	Education	Purpose of technology was for use in education for predicting or assessing students' and/or teachers' emotional states.	"Notably, continuously worn wearable sensors enable researchers to collect egocentric speech data to study and assess real-life expressed emotions, offering unprecedented opportunities for applications in the fields of assessive agents, medical diagnoses, and personalized education." [56]
	Healthcare	Purpose of technology was for the use case of healthcare, or within hospitals, specifically for patients' wellbeing or mental healthcare.	"Speech emotion recognition (SER) has attracted growing attention due to its beneficial role in building human-centered context-aware intelligent systems in many fields, such as customer support call review and analysis, mental health surveillance, multimdea retrieval, and smart vehicles." [160]
	Workplace	Purpose of technology was to improve workplace wellbeing.	"...we can enhance the work environment in offices post-pandemic..." [41]

Table 7: Codebook (cont.)

Code	Subcode	Definition	Example from Corpus
Limitations of approach		Explicit discussion of the limitations of the approach that papers developed, whether related to privacy protection or not.	
	Could be more personalized	References to a limitation of the approach being a lack of personalization for the data subject's privacy preferences.	"In the future, we plan to discuss personalized privacy with an adaptive noise scale of LDP mechanisms that are tailored to each client's privacy preference." [114]
	Data source agnostic	Explicit references to a limitation of the approach being agnostic of data sources	"The proposed model requires that the input data have spatial and temporal components. However, the model is data source agnostic. Users of the proposed model may need to be aware of the inheritance of noise from data sources as it might have some negative influence on the network, and subsequently on its properties."
	Doesn't account for noise	References that approach does not account for noise in the data. Also includes description of the sources of noise	"We recognize that this methodology is prone to noise and renders diminished nuance, but it is practically viable for in-the-wild driving contexts." [27]
	Hardware requirements	References to a limitation of the approach being that it requires certain hardware – often expensive or difficult to acquire hardware.	"However, our running times were obtained using powerful machines and much work is needed to make these protocols practical in constrained computing devices." [4]
	Issues with annotation methods	References to a limitation of the approach being that there may have been issues in the annotation of training data.	"The main limitations of this research stem from the use of a screening survey as the ground truth for depression..." [158]
	Issues with training data	References to a limitation related to the training data.	"The model does not work hundred percent of the time as there are certain issues with the training data that we have used. The FER2013 dataset is overly biased towards happy and sad images as it has those images in maximum." [125]
	Issues with validation approach	References to a limitation of the approach being, broadly, issues with the validation	"Our experimental studies do not have control groups. These groups can help us validate the functionality of the wearable devices and provide information based on the condition of the volunteers." [36]
	Lower performance	References lower performance or accuracy of emotion recognition compared to a non-privacy-preserving model.	"We further show that the improved privacy comes at a cost of a minor utility loss for the target application." [56]
	Not enough data	References approach's limitations being linked to a lack of enough input data.	"The main limitations of this research stem from the use of a screening survey as the ground truth for depression and the limited quantity of voice clips in the EMU dataset." [158]
	Not generalizable	Approach cannot be generalized beyond the training dataset used to broader datasets or even contexts.	"Our dataset contains a preliminary study of in-the-wild driver emotions. Our dataset contains multiple caveats that affect the model's generalizability: imbalanced emotion class labels, not all registered participants drove multiple sessions, and heterogenous in-the-wild data acquisition setting. To not overfit specific participants, we decided to report the results on a leave-one-participant-out cross-validation and were able to show that the model outperforms baselines." [27]

Table 8: Codebook (cont.)

Code	Subcode	Definition	Example from Corpus
Limitations of approach		Explicit discussion of the limitations of the approach that papers developed, whether related to privacy protection or not.	
	Not robust	Reference that accuracy decreases when environmental conditions change or when noise is introduced, etc.	"We also found that the accuracy of time series emotion prediction depends on the specific time period in a day." [99] 'The challenges discussed so far, especially user variability and sensitivity to emotion artifacts, constrain the practical applicability of our system. While the user dependence can only be addressed through an extensive evaluation to accommodate substantial user variations or a calibration step in the adaptive model, Expressear shows success, albeit limited in modeling facial AUs in mobile scenarios. However, the inconsistency in the noise introduced by different mobile settings eliminates the possibility of building a 'one-fits-all' model." [163]
	Not tested for diverse data subjects	Approach not tested for accuracy with a broad set of data subjects ranging in ethnicity, neurodiversity, etc.	"Although EyeSyn embodies several psychology findings in the literature, its current design cannot fully replicate the complex mechanisms of human visual processing to synthesize eye movements for all subject groups. For instance, people with neurodevelopmental or mental disorders, such as autism spectrum disorder, schizophrenia, or social anxiety disorder, may exhibit atypical eye movement patterns in social interactions..." [86]
Privacy concerns		Explicit or implicit descriptions of the privacy concerns related to emotion AI that the approach is seeking to address.	
	Centralizing distributed data	Explicit descriptions of the centralization of distributed input data as creating other privacy concerns — like the potential for data to be leaked.	"Existing SER approaches are largely centralized, without considering users' privacy." [160]
	Cyber-attacks	References to a concern of cyber-attacks or other 'attacks' on input data.	"In fact, recent reports highlight how the increasing use of ambulatory devices is creating new opportunities for cyberattacks, such as identity theft by using another person's voice to fool voice authentication systems and impersonation attacks via utilizing speech synthesis or voice conversion to fake another person." [131]
	Inherent privacy concerns	Descriptions of some privacy concerns as "inherent" to emotion AI – attributed to the overall goal and nature of emotion AI, not specific architectures (like the centralization of distributed data).	"In contrast to conventional facial expressions that are visually obvious to humans, micro-expressions are involuntary and transient facial expressions, commonly manifested involuntarily when we aim to withhold our emotions. Advanced micro-expression recognition techniques exist that can reveal the genuine emotions that people attempt to conceal, thus threatening individual emotional privacy, as fundamental human rights would dictate that one should have a choice what emotion is being shown or not shown." [96]

Table 9: Codebook (cont.)

Code	Subcode	Definition	Example from Corpus
Privacy concerns		Explicit or implicit descriptions of the privacy concerns related to emotion AI that the approach is seeking to address.	
	Intrusive mode of data collection	Descriptions of certain modes of input data collection as intrusive and therefore creating privacy concerns.	"Body-worn physiological sensors are intrusive, while facial and speech recognition only capture overt emotions."
	Risk of inferring sensitive information from data	Descriptions of a specific privacy concern being the risk of input data being used to infer information papers explicitly characterize as sensitive.	"Since the physiological data from the mentioned sensors contain private data, they can also lead to privacy threats by exposing highly sensitive information. To address this issue, we combine differential privacy and federated learning approaches with multi-task learning to efficiently recognize the user's mental stress while perturbing private user identity information." [24]
	Threat of exploitation	Input data could be exploited for malicious purposes.	"Thus, malicious users who gain unauthorized access to non-anonymized speech can potentially misuse PII from this signal, facilitating attacks on other systems and causing safety risks." [131]
	Threat of exploitation for advertising	Emotion AI input data will be exploited for targeted advertising.	"As for companies' motivation to collect this information, the work of Lerner et al. and many others, emotion has a profound impact on decision-making. Authors at The Atlantic noted these trends in speech emotion recognition patents amongst technology companies and came to a reasonable conclusion: these technologies could be applied for targeted advertising based upon a user's emotions." [155]
	Threat of identity theft	Emotion AI input data could be used for stealing the identity of a data subject.	"In fact, recent reports highlight how the increasing use of ambulatory devices is creating new opportunities for cyberattacks, such as identity theft by using another person's voice to fool voice authentication systems and impersonation attacks via utilizing speech synthesis or voice conversion to fake another person." [131]
	Threat of malicious actor	General or specific references to malicious actors who may want to use emotion AI input or output data for malicious purposes.	"However, speech data contain vulnerable information that can be used maliciously without the user's consent by an eavesdropping adversary." [9] "Unauthorized, unaccountable exposure of emotion information is a real concern beyond just commercial use cases. In government, affective computing is beginning to influence law enforcement activities. In recent years, United States law enforcement has dramatically increased subpoenas for user interaction recordings from smart speaker companies to assess individuals' mental states." [155]

Table 10: Codebook (cont.)

Code	Subcode	Definition	Example from Corpus
Conceptualizing privacy		Statements that implicitly or explicitly reflect the conceptualization of privacy being operationalized in the paper.	
	Data ownership	Explicit statements that associate privacy with individuals' ownership over data.	"The idea is that training data can remain with the data producers, which improves privacy and ownership, while the model is shared between multiple users" [41]
	Ability to withhold information	Privacy is being able to withhold information from others	"In this paper, we focus on safeguarding individual emotional privacy against automated micro-expression spotting and recognition tasks. Micro-expressions are emotions that humans intentionally suppress. Thus, humans will feel their privacy is violated in automatic algorithms, such as those increasingly deployed in social media platforms, can recognize their micro-expressions through the shared videos. Essentially, this notion of privacy is in terms of ensuring micro-expression mis-classifications, and thus withholding the true emotion currently felt by the human." [96]
	Not being able to see others' data	Implicit or explicit statements that define privacy as a data subject not being able to see other data subjects' data.	"For privatizing the user's identity while preserving stress recognition accuracy, we adopted a multi-task FL approach that can effectively improve the performance of stress recognition while limiting the risk of inferring sensitive information from the training model since the client does not want to be exposed to the cloud service provider." [24]
	Preserving individuals' data sovereignty	Privacy involves preserving individuals' ability to decide what does and does not happen with their data.	"Often, the identity of individuals gets revealed through images or videos without their consent, which affects their privacy." (Agarwal et al. 2020)
	Prevent others' use of data	Privacy is preventing use of data.	"Intuitively, our goal is to protect the privacy of users' facial expressions. The approach used is to modify the facial expression image datasets in a small and imperceptible way before they are released so that the facial expression recognition models trained on these facial expression images learn the wrong features about the user's facial expressions such as happy, which look like other expressions. The model considers it to be successful because it correctly identifies the modified emoji image sample (happy) as a happy label. By this method, we successfully block the recognition of user's facial expressions by unknown facial expression recognition models and protect the privacy of users." [176]

Table 11: Codebook (cont.)

Code	Subcode	Definition	Example from Corpus
Why address privacy		Statements that offer motivations for addressing privacy concerns.	
	Ethical affective computing	Statements addressing privacy concerns as part of a larger goal for ethical affective computing or that frame privacy concerns as ethical concerns.	"Privacy issues are usually a concern when storing raw proprietary data. In this paper, we want to propose a way to overcome this important ethical problem." [101]
	Increase trust in systems	Statements addressing privacy as part of increasing trust in emotion AI or to address data subjects' skepticism or hesitation around emotion AI.	"The findings from this study can lay a foundation towards trustworthy speech-based technologies or accessible MH diagnosis, monitoring, and prevention. MH-preserving speech anonymization can reduce public resistance towards ambulatory technologies and can potentially decrease user skepticism in data sharing, since users can consent in sharing only the anonymized speech." [131]
	Infringing on personal freedoms	Statements that characterize emotion AI as creating privacy issues through infringing on personal freedoms	"However, concerns about privacy loss and consent issues arise due to the use of facial emotion recognition in such public applications. For example, in public spaces like shopping malls, airports, or city streets, individuals may be subject to emotion recognition without their explicit knowledge or consent. This covert surveillance can be seen as an infringement on personal freedoms and a potential misuse of personal data. The gathering and analysis of emotional data without consent can be likened to eavesdropping on personal sentiments, potentially leading to situations where information could be used manipulatively or commercially without the individual's awareness." [155]
	Invoking legislation	Statements that invoke legislation as part of motivation to address privacy issues with emotion AI.	"It abides by the data protection laws such as EU General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA)." [24]
	Lack of consent	Statements that highlight nonconsensual collection of input data as a major privacy concern that needs to be addressed.	"The widespread use of devices, including cameras, smartphones, short video applications, and facial scanning systems, has enabled the collection of people's facial emotion data without their awareness or consent." [176]
	Risks versus benefits to user	Statements that motivate addressing privacy in order to balance the risks and benefits to users.	"While these models show benefits for detecting various medical conditions from the time-series data (e.g., stress detection), the risk of undermining the privacy of users' data outweighs the sheer benefits attained from these models." [12]
	To enable large-scale applications	Statements that motivate addressing privacy in order to be able to develop larger-scale emotion AI — or references to the small data problem.	"Also, due to privacy concerns, depression labeled voice datasets often have a limited number of participants." [158]
	To show privacy-preserving and accuracy as mutually compatible	Statements that motivate addressing privacy to demonstrate that it's possible to build a privacy-preserving and accurate emotion AI system.	"Instead, our aim is to show that the computations needed to train and use such regression models can be performed in a fully PP way, i.e. so that none of the parties involved has to disclose its data to anyone else in an unencrypted way." [4]

Table 12: Codebook (cont.)

Code	Subcode	Definition	Example from Corpus
Technical approach rationale		Statements that motivate the specific approach(es) used in the paper and justify their selection.	
	Allows better analysis of data	Approach is successful and good because it enables better analysis of input data.	"More precisely, modeling AUs and affect reports as a whole network system will allow measuring network-centric properties which will augment and complement traditional measurement techniques such as statistical analysis. It will also allow us to incorporate well-known findings from different fields in affective computing." (Uddin et al. 2024)
	Approach compatible with type of data	Approach is successful is compatible with the type of input data being analyzed.	"Network science has the potential to model this type of mixed spatio-temporal secondary affective computing data. One of the major advantages of this type of modeling is it allows us to model the data as a whole system." [161]
	Can handle different data streams	Approach can handle many different data streams of different types.	"More importantly, DP offers the required flexibility in applying adequate obscuring methods to different data streams." [12]
	Balancing privacy and accuracy	Approach enables simultaneous high privacy protection and high accuracy.	"We aim to anonymize speech signals while preserving MH information, specifically for the task of estimating depression severity from speech." [131]
	Performs as well as non-privacy-preserving	Their approach is successful because it performs as well (accuracy-wise) as a non-privacy-preserving model.	"Our results show that by adding the adversarial training in the Cloak framework, the injected noise can effectively prevent demographic attributes, such as gender, from being inferred. The prediction results also show that the injected noise on original input data does not decrease the emotion recognition performance." [55]
	Does not transfer sensitive data	Approach does not transfer data deemed to be sensitive (often input data) to other servers, where it could be exposed. "Federated Learning (FL) is particularly suitable for this purpose thanks to its unique characteristic of collaboratively training machine learning models without sharing local data and compromising users' privacy." [160]	
	Enhances data security	Approach enhances the security of input data.	"...it not only mitigates privacy concerns but also enhances data security by keeping data localized." [43]
	No third party	Approach does not require third party services, which makes it more secure and privacy-preserving.	"Hence we recommend a feature set without sensitive data and where all features are preferred to be locally computable and third-party API independent." [27]
	Fewer resource constraints	Approach does not require as many resources (like powerful computers).	"Using a hybrid Split Federated Learning model that overcomes the resource constraints of Federated Learning and reduces the computational time of Split Learning, we can decentralize the training process and keep the sensitive data safe." [168]

Table 13: Codebook (cont.)

Code	Subcode	Definition	Example from Corpus
Technical approach rationale		Statements that motivate the specific approach(es) used in the paper and justify their selection.	
	Practical and applied	Approach is practical and easy to apply in real-world situations.	"The use of federated learning helps us preserve the customer's privacy and it provides us with effective analysis of the data. It is also a practical solution, allowing the central model to deal with time drifts in the data distribution of the target domain." [125]
	Real-time	Approach works with a continuous stream of data.	"DARE-GP provides: a) real-time protection of previously unheard utterances, b) against previously unseen black-box SER classifiers, c) while protecting speech transcription, and d) does so in a realistic, acoustic environment. Further, this evasion is robust against defenses employed by a knowledgeable adversary." [155]