Emotion Recognition Systems and the EU AI ACT: a compliance oriented framework

Ignazio Infantino[1], Helga Molbæk-Steensig[2,3], Giovanni Pilato[1], Filippo Vella[1,*] and Gianpaolo Vitale[1]

[1]*Consiglio Nazionale delle Ricerche, ICAR, Via Ugo La Malfa 153, Palermo, Italy*
[2]*Copenhagen University*
[3]*European University Institute*

Abstract

In human relations, emotional awareness and subsequent adjustment is as fundamental a part of communication as the actual information shared. Without it, communication falters and misunderstandings proliferate. Increasingly, emotion recognition is also a foundational component of affective human–machine and human–robot interaction: the emotional state detected in a human partner can trigger a change of task, a switch of interaction strategy, or an adaptation of the working paradigm of a robotic system. The EU AI Act considers emotion inference as a potentially concerning practice and, under Article 5, prohibits its use in workplace and education settings, and defines it as high risk in most other settings. This paper analyzes the relevant provisions of the AI Act and its Recitals , reviews the state of the art of emotion recognition methods used in HRI, and discusses trade-offs between privacy, bias and autonomy versus personalization, trust, and task performance.

As the relevant provisions of the EU AI Act are coming into force following the delay introduced in the AI Omnibus, lack of clarity about the limits and definitions of emotion recognition is becoming a real problem for developers. Therefore, in this paper, we aim to clarify the legal landscape and provide a more detailed framework to effectively distinguish between harmful and beneficial applications of emotion recognition technology, preventing abuse while preserving the benefits of emotion aware models. To support developers and researchers working on affect-aware HRI systems, we propose a compliance-oriented evaluation framework that connects technical performance, ethical risks, and AI Act requirements into a common assessment methodology.

Keywords

Emotion recognition, EU AI Act, Human–robot interaction, Affective computing, AI regulation,

CEUR-WS.org/Vol-4314/ReDH1.pdf

1. Introduction

Emotions are a central channel in human communication, and their absence from an interactive system is quickly perceived by human users as a lack of responsiveness or social competence. In human–robot interaction (HRI), the ability to perceive – even approximately – the emotional state of the user is often what separates a merely functional agent from one perceived as socially intelligent, trustworthy, and cooperative [1, 2]. An emotionally aware robot can modulate the pace, tone, and content of an interaction: it can allow more time for voice input, ask a user if they want to switch languages, warn a sleepy driver to take a break, encourage a frustrated user during a rehabilitation exercise, simplify a task for a confused learner, or back off when a person shows discomfort. This adaptivity is not a cosmetic feature; several studies in HRI and cognitive-robotics research have linked affect-aware behaviour to measurable improvements in task engagement, perceived trust, and long-term acceptance of assistive and collaborative robots [1, 2, 3].

At the same time, emotion recognition is one of the technologies most exposed to misuse: unlike task-oriented perception (e.g., object recognition), it infers a state that is fleeting, internal, and socially sensitive. Emotions are quite possibly the most private data that exists, protected under the right to

ReDH - Rethinking the Future of Digital Humanities through Ethics and Artificial Consciousness
*Corresponding author.
ignazio.infantino@icar.cnr.it (I. Infantino); Helga.Molbaek-Steensig@eui.eu (H. Molbæk-Steensig); giovanni.pilato@icar.cnr.it (G. Pilato); filippo.vella@icar.cnr.it (F. Vella); gianpaolo.vitale@icar.cnr.it (G. Vitale)
0000-0002-0980-7391 (I. Infantino); 0000-0002-7659-8737 (H. Molbæk-Steensig); 0000-0002-6254-2249 (G. Pilato); 0000-0002-2502-0062 (F. Vella); 0000-0002-5439-6835 (G. Vitale)
© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-11-05

privacy and data protection under Article 8 of the European Convention on Human Rights and Articles 7 and 8 of the EU Charter. Emotions are fleeting, they can change rapidly, are only one of many factors that influence individual behaviour. Moreover, emotions are tricky to identify even for human beings, both for persons trying to understand another's emotions [4], and even, on occasion recognize their own [5]. As a result, in European jurisdictions, persons are generally sanctioned for their actions, and, not for their emotions, although a few notable exceptions exist. [6].This double problem, that emotions are fleeting and private, and that they are difficult to ascertain with certainty, has historically left most European jurisdictions with an uneasy relationship with polygraphs or lie detectors, relying on clues such as perspiration rate and blood pressure – factors that human beings are not capable of picking up by sight – for determining emotions. There is an active academic debate on the accuracy of polygraphs as well as on whether reliance on them in high stakes situations such as courts, psychological treatment, and employment is ethical. Some jurisdictions outright outlaw its use by law enforcement as a violation of the right to a fair trial including the right to be silent and not to incriminate oneself, while others have attached little or no evidentiary force to results obtained by polygraph [6].

The perspective of wide proliferation of AI-based emotions and lie detection (potentially but not necessarily more accurate), towards a wide range of private and public applications is therefore not trivial: jurisdictions' experiences with polygraphs might provide an indicator to how constitutional courts and fundamental rights agencies are likely to view the technology.

This contestation, where emotion detection can be both a genuine enabler of better interactions and a potential vector of privacy intrusion, has motivated the regulatory attention that the EU AI Act pays to emotion detection, as well as the present analysis. Existing research on emotion recognition has mainly investigated three complementary research directions, namely *technical performance*, *ethical implications*, and *regulatory compliance*. These perspectives have generally evolved independently, although the design and deployment of emotion-aware Human–Robot Interaction (HRI) systems inherently involve strong interdependencies among them. Indeed, architectural and algorithmic design choices directly influence both the ethical properties of the resulting systems and their compliance with the regulatory obligations introduced by the EU AI Act. Inspired by this motivation, this paper proposes a compliance-oriented evaluation framework that systematically integrates these complementary dimensions by establishing a methodological relationship between the technical characteristics of emotion recognition systems, the ethical principles underlying trustworthy HRI, and the legal requirements defined by the EU AI Act. The resulting framework provides a unified evaluation perspective for the analysis, design, and assessment of compliant emotion-aware HRI systems that may guide compliant development before legal requirements have been tested in courts.

The EU AI Act explicitly aims to bridge the gap between regulation and innovation, prohibiting very few AI applications, and regulates what it defines as high-risk applications with requirements of comprehensive documentation, logging activity, risk management architectures, accuracy and human oversight mechanisms. In principle, this should allow development to go forward, but in practice all of these come with substantial costs to providers and employers. In some fields – including individual human–robot interaction – applications are typically developed for a limited number of users and under significant budget constraint, making compliance with high-risk requirements prohibitively expensive. In response to this, the EU AI Digital Omnibus which entered into force on July 27, 2026, has amended the EU AI Act, delaying the coming into force of high risk requirements to December 2027, and expanding measures to aid innovation from Small-to-Medium enterprises (SMEs); but also including Small Mid-Cap Entities (SMCs) and start-ups both with regard to prioritised access to regulatory sandboxes and opportunities to provide technical documentation for high risk applications as required in Annex IV of the Act in a simplified manner [7]. Nonetheless, the Omnibus does not change the foundational categorisation of AI systems and it categorises: (a) emotion detection in employment and education as unacceptable risk and prohibited under Article 5(f), and (b) emotion recognition technology based on biometric information in other fields as high risk under Article 1(c) of Annex III [7] The definition "emotion recognition systems" in Article 3(39) and Recital 18 has also remained unchanged. These provisions will form the basis for our analysis in section 2.

1.1. State of the Art in Human Robot Interaction

A first grouping of emotion recognition methods can be performed through the sensing modality they rely on. Table 1 summarizes the main families of approaches, their typical sensors, representative techniques, and their principal strengths and limitations for interactive use. For example, less invasive sensors are preferable when user comfort and natural behaviour must be preserved; large or intrusive devices could restrict movement or cause users to behave unnaturally. On the other hand, advanced sensing hardware becomes appropriate when measurement accuracy is the primary requirement and the resulting increase in device complexity is an acceptable cost. For example, traditional polygraphs would be a bulky and intrusive device, but that is accepted for most current uses.

Facial expression recognition is still the most widely adopted modality of recognition because it requires no physical contact with the user and it can exploit the cameras that are already available on many robotic platforms. Existing approaches combine facial geometry, texture-based descriptors, and convolutional neural networks. The detection reliability, however, decreases when the face is viewed from non-frontal angles, when it is partially occluded, or characterized by weak, ambiguous, or intentionally hidden expressions [8, 1, 9].

Speech emotion recognition is a natural complement in conversational robots, since it requires no additional sensor beyond a microphone, and has moved from hand-crafted acoustic features and classical classifiers toward deep architectures, at the cost of strong sensitivity to language, dialect, and ambient noise [10, 11]. *Physiological and EEG-based recognition* (Electrodermal Activity, Electrocardiography, Heart Rate Variability, Electroencephalography, eye-tracking) are generally considered less susceptible to voluntary masking and therefore more "objective" proxies of internal affective state, but they are intrusive, require wearable instrumentation, and pose non-trivial signal-quality challenges outside laboratory conditions [12, 13, 14]. *Text-based* recognition is applicable whenever the interaction includes a linguistic channel (dialogue systems, chatbots) but is constrained by dataset language coverage and captures expressed rather than embodied affect [10]. Moreover, as evidenced by the emergence of emoticons in human communication, conveying emotions accurately through text is not always easy, doubly so for ascertaining potentially hidden emotions. Finally, *multimodal fusion* approaches, combining two or more of the above channels through early, late, or hybrid fusion (increasingly with transformer-based representations), are consistently reported as more robust than any single

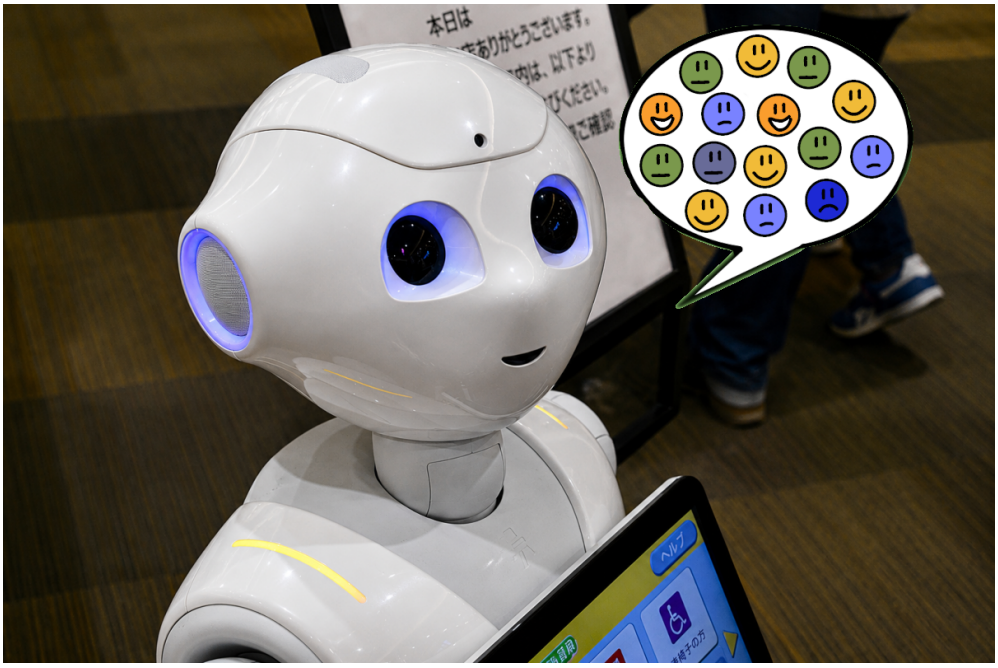


Figure 1: SoftBank Pepper Robot equipped with colored leds into its eyes, ears and shoulders can enhance the verbal communication by displaying colors that represent the emotional content associated with language.

Table 1
Comparison of major emotion recognition methods used in human–machine interaction.

Modality	Typical sensors	Common techniques	Strengths for HRI	Main limitations	Ref
Facial expression	RGB/thermal cameras, Kinect	Landmark geometry, HOG/LBP, CNNs, AU dynamics	Contactless; compatible with existing robot sensors	Sensitive to pose, occlusion, masking	[1][9][8]
Speech emotion	Microphones	Acoustic features, SVM, CNN, LSTM	Natural for conversational robots; camera-free	Language/dialect dependent; noise-sensitive	[10][11]
EEG-based	BCI headsets	SVM, kNN, RF, CNN, RNN/LSTM	Objective, less prone to social masking	Intrusive; hard to acquire outside lab	[12][13]
Peripheral physiological	EDA/GSR, ECG, EMG, PPG, eye-tracking	PCA/LDA, logistic regression, HMM, RF, fusion	Captures internal response; wearables allow mobility	Sensor placement and signal quality challenges	[14][21]
Text-based	Dialogue/chat logs	SVM, Naive Bayes, transformers	Fits language-based interaction	Mostly English-centric; expressed not embodied affect	[10]
Multimodal fusion	Combination of the above	Early/late/hybrid fusion, deep multimodal learning	Compensates single-modality failure modes; most robust	Data scarcity; missing-modality robustness	[15][16] [18]

modality, because different channels compensate for each other's failure modes, although they remain limited by scarce annotated multimodal data and by missing-modality conditions typical of real-world deployments [15, 16, 17, 18]. Recent surveys in robotics show that obtaining strong benchmark results does not mean to get strong outcomes in real-world scenarios. As a matter of fact, the reported performance is often obtained on laboratory datasets containing established expressions and isolated subjects. Real interactions are less controlled: as an example, several people may be present, faces can be partially occluded, users and robots can move, and social context influences how emotions are expressed and interpreted. These factors can substantially reduce recognition reliability [1, 2, 19]. This limitation is directly relevant to the regulatory analysis in Section 2.

Literature on benchmarks[20] about the problem with ground truth and annotation of emotions, including in self-reporting [4], gives a picture of an epistemologically unsettled field with foundational problems in how to measure whether current technology is efficient. This is reflected in the EU AI Act [7] in Recital 44 that directly states that *there are serious concerns about the scientific basis of AI systems* as part of the reasoning for its regulation of emotion recognition technology, including its full prohibition in employment and education settings. The inclusion of this concern suggests that regulations could be amended in the future if epistemological concerns were resolved; however, this is not an absolute guarantee since there are also other concerns related to vulnerability and misuse, as explored in section 2.

1.2. Motivations and Applications of Emotion Recognition in HRI

The motivation for equipping robots with emotion recognition is largely application-driven. In *assistive and rehabilitation robotics*, affect-aware feedback loops have been used to adapt exercise difficulty and encouragement to the user's frustration or fatigue level. In *child–robot interaction*, biophysical and affective signals support engagement monitoring in therapeutic and educational contexts, where robots must remain sensitive to signs of distress or disengagement [2]. In *socially assistive robotics for older*

adults, affect recognition contributes to companionship and to early detection of mood changes relevant to well-being monitoring. In *collaborative manufacturing (cobots)*, coarse affect or stress indicators have been proposed as a safety-oriented signal, complementary to physical proximity sensing, to modulate robot behaviour when a human co-worker shows signs of stress or cognitive overload. Finally, in *conversational and social robots*, emotion recognition is often the mechanism through which the system decides how to act during the interaction.

In these applications the inferred emotional state is seen as an input to the robot's control process. Detecting the emotional state allows for adjusting the actions of the robot, its communication style, or the level of assistance provided to the user. This distinction is central to the regulatory analysis in Section 2. As a matter of fact, the AI Act mainly addresses cases in which inferred emotions are used to profile individuals, influence them, or support decisions concerning them. That scope is narrower than the general use of affective information as an input to adaptive robot control.

1.3. Contribution and Structure of the Paper

The paper's main contributions are: *a)* an analysis of the AI Act provisions which are relevant to emotion recognition in HRI, *b)* a critical discussion of the associated ethical trade-offs with risk and benefits of emotion detection systems and *c)* a set of evaluation criteria that researchers and system integrators can use to document compliance and report scientifically comparable results. Finally, another reason for writing this article is the recent publication of the EU Digital Omnibus on AI, officially enacted as Regulation (EU) 2026/1744, which entered into force on July 27, 2026.
Section 1.1 reviews the technical state of the art and its main reliability limits. Section 1.2 examines the principal application domains and the rationale for adopting emotion-recognition systems in HRI. Section 2 analyzes the relevant articles and recitals of the AI Act with particular reference to the role of emotion recognition. Sections 3.1 and 3.2 discuss the main risks, including privacy violations, bias, and manipulation, together with the potential benefits. Section 4 introduces the proposed evaluation criteria and testing protocol. Finally, section 5 concludes the paper.

Beyond reviewing the EU AI Act provisions regulation emotion recognition, this paper introduces a compliance-oriented evaluation framework for Human-Robot Interaction systems equipped with affective computing capabilities. The framework supports the systematic assessment of emotion-aware HRI by integrating three complementary dimensions: *(i)* the technical properties and limitations of emotion-recognition technologies; *(ii)* the rules established by the AI Act; and *(iii)* the evaluation criteria needed to provide structured evidence relevant to compliance assessment without compromising the adaptive behaviour expected from intelligent HRI systems. Accordingly, this work provides a structured methodology for designing emotion-aware HRI systems that satisfy technical, ethical, and regulatory requirements. The proposed framework supports the systematic documentation of compliance, promotes transparent assessment procedures, and facilitates the adoption of adaptive HRI solutions that are aligned with the obligations introduced by the European regulatory framework.

2. The EU AI Act and Emotion Recognition

The EU AI Act famously constructs a risk-based categorization scheme for regulating AI applications that are put on the European market, differentiating between unacceptable risk, high risk, transparency risk and minimal risk. With unacceptable risk applications being prohibited outright, while high- and transparency risk applications are subjected to documentation and risk management requirements. For the purposes of our framework, emotion recognition technologies are regulated directly in Article 5(1)f which prohibits a limited category of emotion recognition technology uses, and in Article 6(2) read together with Annex III, which specifies a larger category of uses as high risk. Guidelines for application and reasoning behind the articles are found in Recitals 18, 44, and 54 with further guidance provided in soft law documents such as the Commission's guidelines on prohibited artificial intelligence practices and forthcoming draft guidelines on high risk classifications.

2.1. Article 5(1)f: Prohibiting emotion recognition systems in a limited number of scenarios

Article 5(1)f prohibits a narrow category of emotion recognition systems, namely,

> [shall be prohibited] ... AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons;

The provision contains three delimiting elements which we will go through in turn: defining inferring emotions, defining workplace and education institutions, defining medical or safety reasons.

2.1.1. Inferring emotions

A relevant detail here is the formulation of the functionality as "infer emotions" – which is different from the terminology "emotion recognition system" which is defined in Article 3(39) of the act as,

> "emotion recognition system" means an AI system for the purpose of identifying or inferring emotions or intentions of natural persons on the basis of their biometric data;

The key difference here is that "emotion recognition system" identifies or infers emotions on the basis of "biometric data". The mere description in Article 5 of "inferring emotions" does not necessarily prescribe that this is done on the basis of biometric data. Given that the AI Act defines biometric data as "physical, physiological or behavioural characteristics of a natural person, such as facial images or dactyloscopic data" (EU AI Act Article 3(34)) and the taxonomy described above, this means that it is unclear whether for example text-based emotion detection is included in the prohibition. If the prohibition in Article 5(1)f uses the definition in Article 3(39), then text-based emotion recognition or sentiment analysis is exempted, potentially opening up for a whole range of applications in text-based parts of hiring procedures or evaluations of university applications, that utilizes emotion recognition based solely on text input, should there be a market for such. If, on the other hand, "infer emotions" is to be understood more broadly than in the definition provided in Article 3(39) then text-based emotion recognition is also covered by the ban and a product relying on it could be subject to a fine of up to 7 percent of its global turnover (AI Act Article 99(3)). This uncertainty may prove highly problematic as LLMs are increasingly displaying something resembling emotion recognition as emergent abilities and LLMs are increasingly applied in a range of applications in both workplaces and education.

The EU Commission Guidelines on prohibited artificial intelligence practices has a solution for this conundrum, reading Article 5(1)f in the light of Article 3(39) and simply stating that "AI systems inferring emotions and sentiments not on the basis of biometric data" are simply "out of scope". While the guidelines are not directly binding, we can expect the European Court of Justice (CJEU) to take them into account should a dispute emerge on the topic.

Additionally, inferring emotions requires the act of categorising a biometric feature as inferring a certain emotion. As such, the mere notation of facial movements (frown, smile etc.), gestures, or voice quality (whisper, raised voice) without using these to identify emotions, is not covered [22](see §256), [7] (Rec.18). This limitation is interesting as it detaches the outward expression (frown) which is objectively detectable as muscle contractions, from the inward cause, sadness, anger, stress etc. which is not objectively detectable[5]. In other words, a frown can mean a lot of different things, and the same anger or sadness may have different expressions in different instances. This is what Recital 44 refers to when it states that "expression of emotions vary considerably across cultures and situations, and even within a single individual".

In practice the delimitation allows for the use of, for example, automatic sign language interpreters in workplaces or education facilities which need to be able to read facial expressions as they form integral parts of sign language signs, but which do not assign emotions of the user on the basis of those facial expressions.

2.1.2. Workplace and education institutions

The Guidelines determine the scope of "workplace" and "education" broadly, including both hiring and ongoing workplace situations as well as admission, teaching and exam situations. In both cases, the intention is to protect two categories of individuals which are expected to be exposed in hierarchical structures, namely employees and students. Other individuals in those contexts are not protected. For example, it differentiates in workplace situations between surveilling call center employees for emotional distress (prohibited) and customers with a view to safeguarding call center employees (not prohibited)[22] (see §254). The Guidelines also state that systems identifying "suspicious customers" in supermarkets are not covered by the prohibition, provided it can be ensured that employees are not being tracked.

2.1.3. Medical and safety reasons

The medical and safety reasons exception from the prohibition of emotion recognition in the workplace or education institution is to be interpreted narrowly and subject to strict proportionality assessments [22] (see §§259-260). General stress-monitoring is for example not allowed, but emotion recognition technologies could be introduced as personal aids for persons with autism, deafness or blindness. Moreover, general assessment of well-being, boredom or concentration would not be allowed unless lack of concentration or elevated stress would pose a specific danger, for example in the management of dangerous machines [22](see §263). Physical feelings like pain or fatigue are entirely exempted from "emotion detection system" (Recital 18) and therefore also from the prohibition and from categorisation as high-risk applications.

2.1.4. Vulnerability and epistemic uncertainty behind prohibition

Recital 44 presents two overarching reasons for the prohibition of inferring emotions in workplaces and education institutions[7]. First, it presents the epistemic uncertainty which we have already touched upon in the introduction and section 2.1.1. raising "serious concerns about the scientific basis" of emotion recognition systems as a whole. This is then coupled with a concern about bias and discrimination,

> Among the key shortcomings of such systems are the limited reliability, the lack of specificity and the limited generalisability. Therefore, AI systems identifying or inferring emotions or intentions of natural persons on the basis of their biometric data may lead to discriminatory outcomes and can be intrusive to the rights and freedoms of the concerned persons.

This epistemic uncertainty coupled with a particular concern for discriminatory outcomes leads the Recital to single out workplaces and education institutions specifically, stating that they are contexts where there is an imbalance of power. This is certainly accurate, but it is not, by far, the only context in which there is a power imbalance. This is also the case in a myriad of interactions between the public and state authorities, which are then not covered by the prohibition, including law enforcement, migration management, and use by the judiciary, all with potentially detrimental consequences to the individuals surveilled. Such uses – and many more – are instead considered high-risk applications under Article 6(2) which are subject to documentation, transparency and risk management requirements but not expressly prohibited.

2.2. Article 6(2) read with Annex III and Article 50: categorising most other emotion systems as high risk with special transparency obligations

Article 6(2) read together with Annex III provide a wide range of AI applications which are to be considered high risk and which are therefore subject to a range of risk management requirements (Article 9), data quality and governance requirements (Article 10), Technical documentation (Article 11), Transparency (Article 13), Record-keeping and human oversight requirements (Articles 12 and 14), robustness and accuracy requirements (Article 15). Annex III Article 1 categorises all emotion

recognition applications (Annex III, Article 1(c)) in so far as they do not fall under Article 5(1)f or are otherwise prohibited – as high risk and therefore subject to compliance with the entirety of Chapter III of the Act (Articles 8-49). Additionally, all emotion recognition applications are subject to special transparency requirements under Article 50(3), including to,

> …inform the natural persons exposed thereto of the operation of the system, and shall process the personal data in accordance with Regulations (EU) 2016/679 and (EU) 2018/1725 and Directive (EU) 2016/680, as applicable.

This too comes with a clawback clause for authorities' use,

> This obligation shall not apply to AI systems used for biometric categorisation and emotion recognition, which are permitted by law to detect, prevent or investigate criminal offences, subject to appropriate safeguards for the rights and freedoms of third parties, and in accordance with Union law.

This exemption is to be understood in line with Recital 132 as applying only in situations where the deployer is a law enforcement authority, and therefore it does not apply to private deployers of emotion recognition software as, for example, in surveillance field.

In practice this means that for example the surveillance of supermarket customers described in section 2.1.2 while not outright prohibited, would be protected by the transparency requirement and therefore have to be informed that they are subject to emotion recognition technology. Moreover, biometric data as used in emotion recognition technology remains sensitive personal information under GDPR subject to strict necessity and proportionality requirements including data minimization practices under GDPR Article 5(1)c.

Moreover, the reliance on the definition of emotion detection system as in line with Article 3(39) for the high risk AI applications, suggests that text-based sentiment analysis is not covered here either, criticized as a loophole at the time[23].

According to Article 6(5) of the EU AI Act, the Commission was supposed to publish guidelines clarifying the exact classification of high risk applications in February 2026, then postponed to August 2026 with the Omnibus (AI Act Article 113), but as of writing in September 2026, only a draft of these guidelines is available.

The draft, however, repeats the concerns raised under Article 5(1)f, namely that emotion recognition systems rest on unsettled science of emotion recognition in general, limiting their reliability and generalizability. Additionally, they are intrusive into the fundamental rights of those subjected to them[22](see §172). This concern about the intrusiveness should be read together with Article 27 which requires deployers of high risk AI systems for public services to conduct a fundamental rights impact assessment which includes proportionality assessments on whether the benefits obtained from the functionality corresponds to the rights intrusions it represents both for the targets and for unrelated third parties.

2.3. General concerns about the logic of the AI Act's treatment of emotion recognition: implications for developers

There has been significant criticism in the literature of the limited reach of Article 5(1)f and the many exemptions for law enforcement including migration management, not only from the ban under Article 5(1)f but also from the transparency requirement under Article 50(3)[24]. Considering the logic as presented by Recital 44, exempting law and border enforcement make little sense. The science is as unsettled outside workplaces and education institutions as inside, the systems may still risk furthering biases leading to discrimination, the interference with the right to privacy also remains intensive and there are clear imbalances of power. In line with this reasoning, the European Parliament adopted at first reading an amendment to Article 5(1)f expanding the ban to law enforcement and migration management and amending certain Union legislative acts [25] (see §52). In the marathon trilogues preceding the final adoption of the Act, the ban was slimmed down to its current form, as part of a

political compromise. Other potential areas with power imbalances that emotion detection technology could exasperate include healthcare – especially in forced institutionalisations in mental health care facilities, and in the administration of justice.

Together with the limitations in applicability, critics have raised concerns that the provisions include a loophole in the form of the classification of "emotion detection" to not include the biometric elements on which emotion inference is typically based, such as mere facial expressions or voice quality, as clarified under Recital 18. One could easily imagine that a customer service employer could discipline workers by counting smiles and frowns, without expressly having inferred emotions on the basis of such biometric features[26]. In such cases, potential protection for individuals would remain in data protection regulations (GDPR) and the principles of data minimization, and, ultimately, in human rights law.

When academic and NGO criticism of the logic of the regulation matter for developers, it is because it might give an indication of where regulation is heading in the future. At the moment, the AI Omnibus suggests that regulators are leaning towards easing rather than tightening AI regulations for the benefit of innovation, in particular for smaller actors, but the tide may well turn again as cases start making their way through national courts and into the European Court of Human Rights. There can be no doubt that emotion recognition systems applied for crowd control, crime detection and prevention, and in many other potential cases, represent interferences with fundamental rights, and if their use is not necessary and proportional, they will also represent violations of the human right to privacy and personal data (Articles 7 and 8 of the EU Charter of Fundamental Rights, Article 8 of the European Convention on Human Rights). The field is still politically charged, making future regulations uncertain. To that end, there may be an argument for prudence in providers making emotion detection able chat-bots (currently entirely unregulated) or emotion detection systems based on biometric data that are currently allowed but fall among high risk applications.

3. Ethical Issues in Emotion Recognition for HRI

Following the general risk classification coming from The EU AI Act, provided in Section 2, this section describes in detail the risks associated with HRI, goes on to outline the main four potential benefits that can be achieved and proposes mitigation strategies toward a balanced approach.

3.1. Risks

Before introducing the evaluation framework, we make explicit the risks that justify both the caution of the AI Act and the compliance-oriented stance of this paper. These risks are heterogeneous. Some, such as reliability and bias, are mainly technical and can be measured and mitigated through better data, model design, and validation procedures. Other risks, such as manipulation and function creep, depend mainly on the deployment context. They emerge from the purposes for which the system is used, the setting in which it is deployed, and the decisions it supports, even when the underlying model is technically reliable.

The literature identifies a recurring set of risks that can be associated with emotion recognition technologies. Several of these risks are among the reasons behind the AI Act provisions:

- **Reliability and validity.** Emotion models are often trained and validated on predominantly Western datasets and focused on adults; their generalizability to real-world cases, other cultures, age groups, or neurodivergent users is therefore limited, and they could generate incorrect classifications that cause inappropriate robot behavior [1, 19, 4].
- **Bias and discrimination.** Unbalanced training data can result in significantly worse performance for underrepresented groups, with the risk of unequal actions and treatment, or otherwise unfair service delivery for minority groups [13] (see Recital 54).
- **Data privacy and sensitivity.** Facial, vocal, and physiological signals used for emotion inference are biometric- and potentially health-related data. In legal terms, there is no question that it is

personal data under GDPR Article 4(1) but it is unclear whether it also constitutes special category data under Article 9 of the GDPR subject to additional requirements including the granting of explicit consent. The GDPR defines biometric data in connection with the purpose of unique identification, but it is not yet settled law whether that might include biometric data gathered for emotion detection. [27][28][29].

- **Manipulation and power asymmetry.** In contexts with structural power imbalance (workplace, education, healthcare, law enforcement, administration of justice), emotion-aware systems can be used, intentionally or not, to nudge or pressure individuals, which is precisely the concern underlying Article 5(1)(f) and Recital 44, which however prohibit its deployment only in workplace and education contexts.
- **Function creep.** A system introduced for a legitimate purpose (e.g., safety monitoring in a cobot cell) can be repurposed for performance evaluation or surveillance without explicit consent or oversight.

These risks must be considered when evaluating emotion recognition systems. Reliability and bias can be assessed through performance and fairness metrics. Privacy risks require checks on how personal data is collected, stored, and used. Manipulation and function creep depend on where and how the system is used, and for what purpose. These risks are particularly relevant in HRI, where the physical presence of a robot may make emotion recognition more difficult for users to avoid than in on-screen interactions. These aspects are addressed by the evaluation criteria proposed in Section 4.

3.2. Benefits

Despite the risks associated with emotion detection systems, human-robot interaction (HRI) may offer significant advantages when these technologies are implemented responsibly:

- **Personalization and adaptability.** Adjusting task difficulty and communication style to the user's affective state improves performance and perception of support in various contexts, such as care and education. [2]
- **Engagement and trust.** Robots capable of detecting emotions are generally rated as more natural and trustworthy than those without this capability [30, 31].
- **Safety.** Monitoring a driver's fatigue, stress, or discomfort could add to overall road safety. In the current regulatory climate, detection of such physical states is excempted from categorisation as emotion detection, explicitly to encourage such functionality (Recital 18 of the AI Act).
- **Accessibility.** For users with communication limitations, understanding emotions offers an alternative expressive channel, capable of replacing explicit verbal feedback. [32, 33]

These benefits show why evaluating emotion-aware HRI systems requires more than recognition accuracy and regulatory compliance. Such systems should also provide measurable improvements over alternatives that do not use emotion recognition. These improvements may concern engagement, accessibility, or task performance. However, potential benefits do not justify overlooking privacy risks or regulatory restrictions. The proposed framework therefore considers HRI outcomes alongside technical, ethical, and regulatory criteria.

3.3. Toward a Balanced Approach: Mitigation Strategies

Considering one of the central questions of this contribution – whether an accurate norm can discriminate between the negative aspects of this technology and the benefits it can bring – and the uncertainties still connected with the regulatory landscape, we consider that mitigation will in many cases be the most appropriate strategy for developers of emotion detection systems. Some successful mitigation that could be applied to these systems, include: *on-device processing* of biometric signals, that is avoiding transmission or storage of raw facial/voice data, *data minimization*, deriving only the coarse affect category required by the task, not a full emotional profile, *informed consent and opt-out*,

mechanisms integrated into the interaction itself that allow the user decide whether to let the system detect his emotion for an enhanced interaction, *purpose limitation* that technically prevent repurposing of the affect signal beyond its declared use, and *human oversight* over decisions that can have a relevant effect on the user.

These mitigation strategies address the risks identified in Section 3.1 and help protect fundamental rights, particularly privacy (ECHR Article 8) and protection against discrimination (ECHR Article 14). However, adopting these measures does not automatically ensure regulatory compliance or demonstrate that the system provides real benefits to users. A systematic evaluation must therefore assess both the effectiveness of these safeguards and the actual outcomes of the interaction. The compliance-oriented framework introduced in Section 4 addresses this need by combining technical performance, risk mitigation, regulatory requirements, and HRI outcomes.

4. Evaluation Criteria and Testing Protocol

The previous sections identified the main technical limitations, ethical risks, and regulatory requirements of emotion recognition systems. Based on this analysis, we propose a compliance-oriented evaluation framework for emotion-aware HRI systems. The framework brings these aspects together through a set of evaluation criteria. Its goal is to help researchers and developers assess system performance, identify potential risks, and document relevant compliance requirements.

To ensure legal compliance, emotion recognition applications, once classified as high risk systems, are subject to strict obligations before they can be put on the market. The list of the requirements includes: adequate risk assessment, mitigation systems, high-quality of the datasets feeding the system to minimise risks of discriminatory outcomes, logging of activity to ensure traceability of results, detailed documentation providing all information necessary on the system and its purpose for authorities to assess its compliance, clear and adequate information to the deployer, appropriate human oversight measures high level of robustness, cybersecurity, and accuracy.

A recurring gap in the emotion-recognition-for-HRI literature, however, is the lack of a commonly accepted protocol connecting technical evaluation, ethical risk assessment, and regulatory compliance. We propose the following evaluation dimensions, intended to be reported jointly whenever an emotion-aware HRI system is described in a scientific publication, so that results are both scientifically comparable and organized to facilitate future compliance reviews or audits. We have identified three main dimensions that group together the criteria for assessing a high-risk application's compliance with the AI Act:*(i)* Performance and Bias Assessment, *(ii)* Data Privacy and Governance, and *(iii)* Results and Objectives. These operational dimensions translate the three complementary perspectives (i.e., technical performance, ethical considerations, and regulatory requirements) introduced in Section 1 into concrete evaluation criteria. Seven evaluation points are derived from the aforementioned dimensions:

1. **Recognition performance.** Standard classification metrics (accuracy, macro-F1, unweighted average recall – the latter preferred over accuracy for imbalanced emotion classes) reported *per demographic subgroup* (e.g., age band, gender, skin tone where facial data is used), not only in aggregate, to expose the bias risks as flagged in Recital 44 and 54 of the EU AI Act.
2. **Fairness metrics.** Analysis of group-wise performance gaps (e.g., difference in recall or F1 across subgroups), demographic parity, and equalized-odds-style metrics, computed on held-out data stratified by subgroup. [34, 35]
3. **Robustness and ecological validity.** Performance degradation from posed/lab data to spontaneous, in-the-wild, or cross-cultural data; performance under occlusion, multi-party scenes, and sensor noise typical of embodied HRI.
4. **Data governance documentation.** Data statement covering provenance that specify demographic composition, and known limitations of the training corpus,such as underrepresented groups, in line with general high risk systems requirements under Articles 10 and 11.[36]

5. **Privacy and data-minimization audit.** Verification of on-device vs. cloud processing, retention period of raw biometric signals, and whether the derived output is limited to the minimum category needed by the task.
6. **Regulatory classification checklist.**
   a) A first discrimination is carried out considering if the AI system detects emotions or a physical state.
   b) If the emotions are detected with biometric data (Article 3(39) scope check) the setting has to be ascertained. According to Article 5(1)(f) the emotion detection systems are prohibited in education institution or workplaces, with the exception of medical or safety reasons.
   c) Outside of the previous cases the systems are allowed falling under Annex III high-risk obligations or under the limited-risk/transparency tier (Article 50).
7. **HRI outcome measures.** Task performance, standardized trust and acceptance questionnaires (e.g., trust-in-automation and robot-acceptance scales commonly used in HRI user studies [30, 31, 37, 38, 39]), and engagement metrics, collected with a comparison condition (affect-aware vs. affect-blind system) to substantiate claims of scientific advancement rather than relying on face validity alone.

Figure 2 shows the decision process used to classify emotion recognition systems under the EU AI Act. The procedure checks whether the system infers emotions. It then identifies possible prohibitions and the applicable regulatory requirements. This decision tree supports the regulatory assessment included in our framework. Elements 1-3 address Performance and Bias Evaluation provide quantitative evidence about system's fairness across different user groups, but they also address core concerns related to accuracy, generalisability and bias expressed by regulators in Recital 44 in the EU AI Act as reasoning for the relatively strict regulations of emotion recognition systems. Elements 4 and 5 address Privacy and Data Governance in line with GDPR requirements which should always be read together with the EU AI Act (Recital 63), verifying that training and session data are properly handled. Finally, Elements 6 and 7 focus on Results and Objectives, evaluating the effectiveness of emotion recognition and the legitimate use of the acquired information. Moreover, element 6 excludes physical states such as pain or fatigue in line with Recital 18 and differentiates between inference of emotions and the mere detection of readily apparent expressions, gestures or movements. The application in educational institutions or workplaces are prohibited, with the only exception of medical or safety reasons. For all the systems that are not prohibited, the obligations listed in Article 6 and Annex III have to be followed. As additional case the deployers and the Providers have different guidelines (Article 26). If deployer acts to operates the Emotion Detection system for law enforcement authorities has no obligation of transparency. In the other cases, as specified in the Article 50(3) has Transparency obligations (e.g., natural persons have to be informed that they are interacting with an AI system). For the Providers that develop an AI system or general-purpose AI model and places it on the market under its own name, the technical documentation requirements depend whether they qualify as a small or medium enterprise(SME), start up or Small Mid-Cap (SMC)(Article 16). In these cases, companies may produce simplified technical documentation. reducing their administrative burden.

The proposed framework can be intended as a structured assessment and reporting methodology. Its purpose is to make explicit the evidence needed to evaluate emotion-aware HRI systems across technical, ethical, and regulatory dimensions. In particular, regulatory admissibility must be established independently of technical performance or demonstrated interaction benefits: a prohibited application cannot become compliant simply because it achieves high accuracy or provides measurable advantages to users.

5. Conclusion

The emotions recognition in the EU Act framework opens up interesting possibilities in both the scientific and strictly legal spheres. The EU AI Act prohibits emotion inference from biometric data in the workplace and education, grounded in well-documented concerns about reliability, bias, and

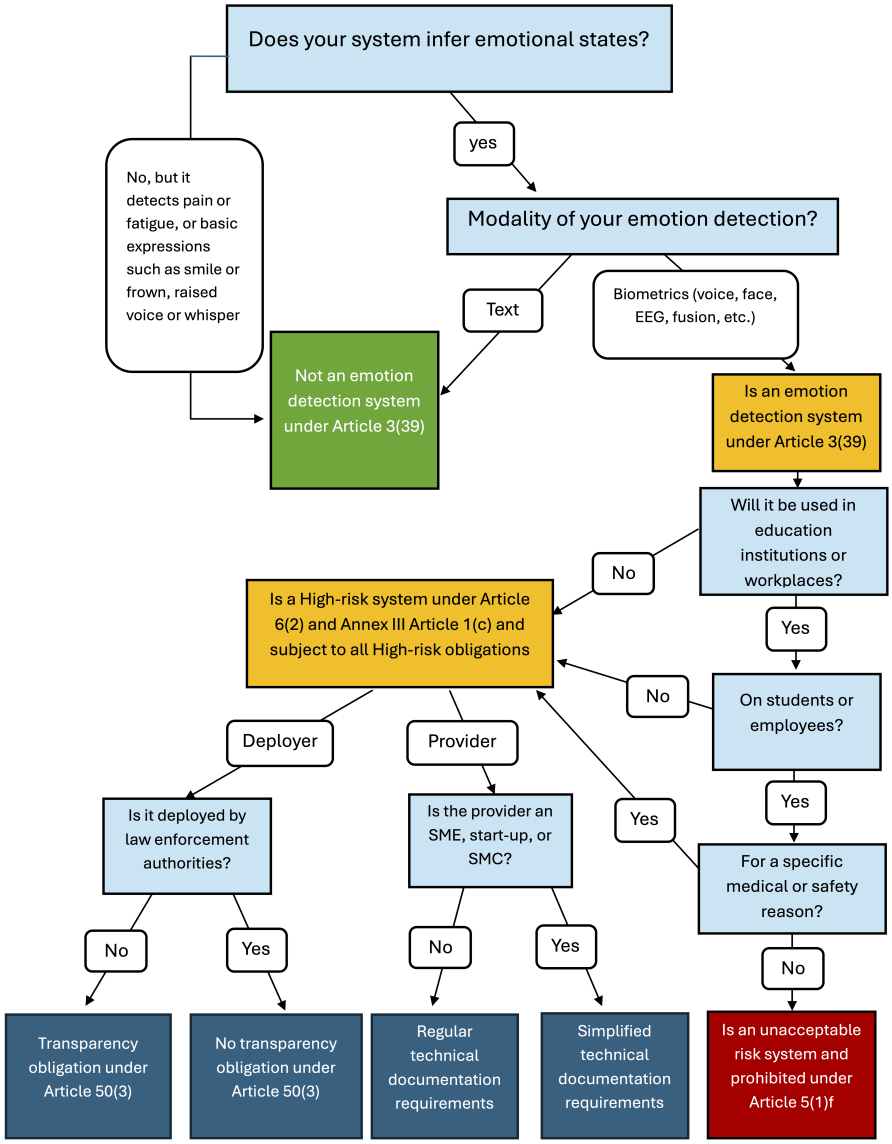


Figure 2: Emotion recognition systems requirements under the EU AI Act: First determine whether your technology is considered an emotion recognition system under Article 3(39) and 5(1)f. If it does not infer emotions or if it does so on non-biometric data, it is not considered an emotion recognition system. It may still be regulated, but not under these provisions. If the emotion recognition system is deployed in education institutions or workplaces to infer the emotions of students or employees, unless there is a specific medical or safety reason, is prohibited. The emotion recognition systems that are not prohibited, are assessed as high risk. For high risk systems, some requirements apply only to some deployers or providers. All deployers except law enforcement and prosecutors are required to disclose that emotion recognition systems are active. All providers must deliver technical documentation, but SMEs, Start-ups and SMCs can deliver simplified technical documentation.

power asymmetry. Outside this specific combination of setting and data modality, however, emotion-aware robotic systems are not banned but fall under the high-risk tiers which impose requirements for documentation, risk-mitigation, and transparency.

The main contribution of this work consists in the definition of a compliance-oriented evaluation

framework for emotion-aware Human–Robot Interaction (HRI) systems. The proposed framework supports the systematic assessment of HRI applications by integrating three complementary evaluation dimensions, namely: (i) technical performance, (ii) ethical considerations, and (iii) the regulatory requirements introduced by the European Union AI Act. To this end, the framework provides a structured reference methodology for the design, evaluation, and documentation of emotion-aware HRI systems with respect to relevant compliance requirements, enabling researchers and practitioners to assess adaptive HRI applications according to a unified technical and regulatory perspective.

Future work should validate this evaluation protocol on concrete HRI deployments, quantify the trade-off between recognition accuracy and privacy-preserving on-device processing, and track how the Commission's forthcoming guidance and enforcement practice further clarify the boundary between legitimate affect-aware interaction and prohibited emotion inference.

6. Acknowledgments

This work has received funding from the European Union's Horizon Europe research and innovation programme under the project ELOQUENCE (Grant Agreement No. 101135916).

Declaration on Generative AI

During the preparation of this work, the author(s) used ChatGPT5.5 in order to: grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] M. Spezialetti, G. Placidi, S. Rossi, Emotion recognition for human-robot interaction: Recent advances and future perspectives, Frontiers in Robotics and AI 7 (2020). doi:10.3389/frobt.2020.532279.

[2] D. R. Faria, A. L. Godkin, P. P. da Silva Ayrosa, Advancing emotionally aware child–robot interaction with biophysical data and insight-driven affective computing, Sensors (Basel, Switzerland) 25 (2025). doi:10.3390/s25041161.

[3] Y. Wang, W. Song, W. Tao, A. Liotta, D.-W. Yang, X. Li, S. Gao, Y. Sun, W. Ge, W. Zhang, W. Zhang, A systematic review on affective computing: Emotion models, databases, and recent advances, ArXiv abs/2203.06935 (2022). doi:10.48550/arxiv.2203.06935.

[4] L. F. Barrett, R. Adolphs, S. Marsella, A. M. Martinez, S. D. Pollak, Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements, Psychological Science in the Public Interest 20 (2019) 1–68. doi:10.1177/1529100619832930.

[5] L. F. Barrett, How Emotions Are Made: The Secret Life of the Brain, Houghton Mifflin Harcourt, Boston, 2017.

[6] E. H. Meijer, P. J. Van Koppen, Lie detectors and the law: The use of the polygraph in europe, Psychology and law: Bridging the gap (2008) 31–50.

[7] European Data Protection Supervisor, AI Act Regulation (EU) 2024/1689 – Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance), Publications Office of the European Union, 2025. doi:10.2804/4225375.

[8] S. C. Leong, Y. Tang, C. H. Lai, C. Lee, Facial expression and body gesture emotion recognition: A systematic review on the use of visual data in affective computing, Comput. Sci. Rev. 48 (2023) 100545. doi:10.1016/j.cosrev.2023.100545.

[9] M. Egger, M. Ley, S. Hanke, Emotion recognition from physiological signal analysis: A review, Electronic Notes in Theoretical Computer Science (2019) 35–55. doi:10.1016/j.entcs.2019.04.009.

[10] A. N. Alslaity, R. Orji, Machine learning techniques for emotion detection and sentiment analysis: current state, challenges, and future directions, Behaviour & Information Technology 43 (2022) 139 – 164. doi:10.1080/0144929x.2022.2156387.

[11] G. Udahemuka, K. Djouani, A. Kurien, Multimodal emotion recognition using visual, vocal and physiological signals: A review, Applied Sciences (2024). doi:10.3390/app14178071.

[12] E. H. Houssein, A. Hammad, A. A. Ali, Human emotion recognition from eeg-based brain–computer interface using machine learning: a comprehensive review, Neural Computing and Applications 34 (2022) 12527 – 12557. doi:10.1007/s00521-022-07292-4.

[13] S. Khare, V. Blanes-Vidal, E. Nadimi, U. Acharya, Emotion recognition and artificial intelligence: A systematic review (2014-2023) and research recommendations, Inf. Fusion 102 (2023) 102019. doi:10.1016/j.inffus.2023.102019.

[14] A. Dzedzickis, A. Kaklauskas, V. Bučinskas, Human emotion recognition: Review of sensors and methods, Sensors (Basel, Switzerland) 20 (2020). doi:10.3390/s20030592.

[15] B. Pan, K. Hirota, Z. Jia, Y. Dai, A review of multimodal emotion recognition from datasets, preprocessing, features, and fusion methods, Neurocomputing 561 (2023) 126866. doi:10.1016/j.neucom.2023.126866.

[16] S. Kalateh, L. A. Estrada-Jimenez, S. Nikghadam-Hojjati, J. Barata, A systematic review on multimodal emotion recognition: Building blocks, current state, applications, and challenges, IEEE Access 12 (2024) 103976–104019. doi:10.1109/access.2024.3430850.

[17] S. Abdullah, S. Ameen, M. Sadeeq, S. R. M. Zeebaree, Multimodal emotion recognition using deep learning, Journal of Applied Science and Technology Trends (2021). doi:10.38094/jastt20291.

[18] A. Geetha, T. Mala, D. Priyanka, E. Uma, Multimodal emotion recognition with deep learning: Advancements, challenges, and future directions, Information Fusion 105 (2024) 102218. URL: https://www.sciencedirect.com/science/article/pii/S1566253523005341. doi:https://doi.org/10.1016/j.inffus.2023.102218.

[19] Y. Cai, X. Li, J. Li, Emotion recognition using different sensors, emotion models, methods and datasets: A comprehensive review, Sensors (Basel, Switzerland) 23 (2023). doi:10.3390/s23052455.

[20] M. D. D. Kumar, M. S. Rao, K. Narendra, Multimodal emotion recognition: A comprehensive survey of datasets, methods, and applications, IEEE Access 13 (2025) 201067–201097. doi:10.1109/access.2025.3636186.

[21] R. Gasz, Z. Bougriche, J. Osuchowski, M. Tomaszewski, A review of current capabilities and future directions in machine-based emotion recognition., Cognitive, affective & behavioral neuroscience (2026). doi:10.3758/s13415-025-01398-7.

[22] Commission guidelines on prohibited artificial intelligence practices established by regulation (eu) 2024/1689 (ai act), 2026. URL: https://ai-act-service-desk.ec.europa.eu/sites/default/files/2025-08/guidelines_on_prohibited_artificial_intelligence_practices_established_by_regulation_eu_20241689_ai_act_english_ied3r5nwo50xggpcfmwckm3nuc_112367-1.PDF.

[23] O. Uguz, These are the major loopholes on emotion-recognition in eu artificial intelligence act, 2025. URL: https://euobserver.com/4547/these-are-the-major-loopholes-on-emotion-recognition-in-eu-artificial-intelligence-act/, visited on 1/9/2026.

[24] G. Bouchagiar, Leave us alone, but also let us be opaque: The right to opacity as a conceptual response to law enforcement's emotional surveillance under eu law, 2026. URL: https://hdl.handle.net/20.500.14017/7ad2423d-2d80-435c-b71d-95827662d60a.

[25] Amendments adopted by the european parliament on 14 june 2023 on the proposal for a regulation of the european parliament and of the council on laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts (com(2021)0206 – c9-0146/2021 – 2021/0106(cod)), 2023. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/

?uri=ep:P9_TA(2023)0236.
[26] A. Prégent, Is there not an obvious loophole in the ai act's ban on emotion recognition technologies?, AI & SOCIETY 40 (2025) 5581–5582. URL: https://doi.org/10.1007/s00146-025-02289-8. doi:10.1007/s00146-025-02289-8.
[27] M. Veale, F. Zuiderveen Borgesius, Demystifying the draft EU artificial intelligence act, Computer Law Review International 22 (2021) 97–112. doi:10.9785/cri-2021-220402.
[28] D. C. Ong, J. Zaki, N. D. Goodman, Computational models of emotion inference in theory of mind: A review and roadmap, Topics in cognitive science 11 (2019) 338–357.
[29] R. Smith, T. Parr, K. J. Friston, Simulating emotions: An active inference model of emotional state inference and emotion concept learning, Frontiers in psychology 10 (2019) 2844.
[30] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. de Visser, R. Parasuraman, A meta-analysis of factors affecting trust in human-robot interaction, Human Factors 53 (2011) 517–527. doi:10.1177/0018720811417254.
[31] K. E. Schaefer, J. Y. C. Chen, J. L. Szalma, P. A. Hancock, A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems, Human Factors 58 (2016) 377–400. doi:10.1177/0018720816634228.
[32] A. Bussolan, S. Baraldo, L. M. Gambardella, A. Valente, Multimodal fusion stress detector for enhanced human-robot collaboration in industrial assembly tasks, in: 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN), IEEE, 2024, pp. 978–984.
[33] E. Lutin, S. A. Elprama, J. Cornelis, P. Leconte, B. Van Doninck, M. Witters, W. De Raedt, A. Jacobs, Pilot study on the relationship between acceptance of collaborative robots and stress, International Journal of Social Robotics 16 (2024) 1475–1488.
[34] P. Garg, J. Villasenor, V. Foggo, Fairness metrics: A comparative analysis, in: 2020 IEEE international conference on big data (Big Data), IEEE, 2020, pp. 3662–3666.
[35] A. Beutel, J. Chen, T. Doshi, H. Qian, A. Woodruff, C. Luu, P. Kreitmann, J. Bischof, E. H. Chi, Putting fairness principles into practice: Challenges, metrics, and improvements, in: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, 2019, pp. 453–459.
[36] M. Al-Ruithe, E. Benkhelifa, K. Hameed, A systematic literature review of data governance and cloud data governance, Personal and ubiquitous computing 23 (2019) 839–859.
[37] B. C. Kok, H. Soh, Trust in robots: Challenges and opportunities, Current robotics reports 1 (2020) 297–309.
[38] G. Campagna, M. Rehm, A systematic review of trust assessments in human–robot interaction, ACM Transactions on Human-Robot Interaction 14 (2025) 1–35.
[39] D. M. Brock, I. G. Sarason, H. Sanghvi, R. A. Gurung, The perceived acceptance scale: Development and validation, Journal of Social and Personal Relationships 15 (1998) 5–21.