

The potential impact of the AI Act on affective computing research and development

ACII 2023
10.09.2023 | Boston

MIT Media Lab
Room: E14-244
Time: 13:30 – 16:30

Presenters and Moderators



Andreas Häuselmann

a.n.hauselmann@law.leidenuniv.nl
PhD Candidate
eLaw, Center for Law and Digital
Technologies
Leiden University, The Netherlands

Deniz Iren

deniz.iren@ou.nl
Associate Professor Affective Computing
Department of Information Science
Open Universiteit, Netherlands

Bhoomika Agarwal

bhoomika.agarwal@ou.nl
PhD Candidate
Department of Technology Enhanced
Learning and Innovation
Open Universiteit, Netherlands

Krist Shingjergji

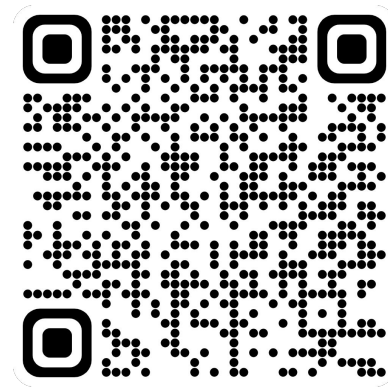
krist.shingjergji@ou.nl
PhD Candidate
Department of Technology Enhanced
Learning and Innovation
Open Universiteit, Netherlands

Goal(s)

- Provide “*digestible*” information regarding **AI Act** (and other AI regulations) specifically **tailored** to the **affective computing community**
- **Discuss** these topics, get **your** **opinions**, and prepare a **report**
- Communicate our community’s response to the policy makers (it is still not too late!)

Participant information package

1. Slides ([url](#))
2. AI Act Index for Affective Computing Community (1-page, [download](#))
3. Paper: Ethical Risks, Concerns, and Practices of Affective Computing: A Thematic Analysis ([download](#))



[slides](#)

Outline

Where: MIT Media Lab | Room: E14-244

When: 10/09/2023 13:30 – 16:30

Part 1: The AI Act proposal (50 mins)

Part 2: LBR (10 mins)

Part 3: Group Discussion (max 90 mins)

Part 4: Presentations of Discussion Highlights (max 30 mins)

Part 1 - The AI Act proposal

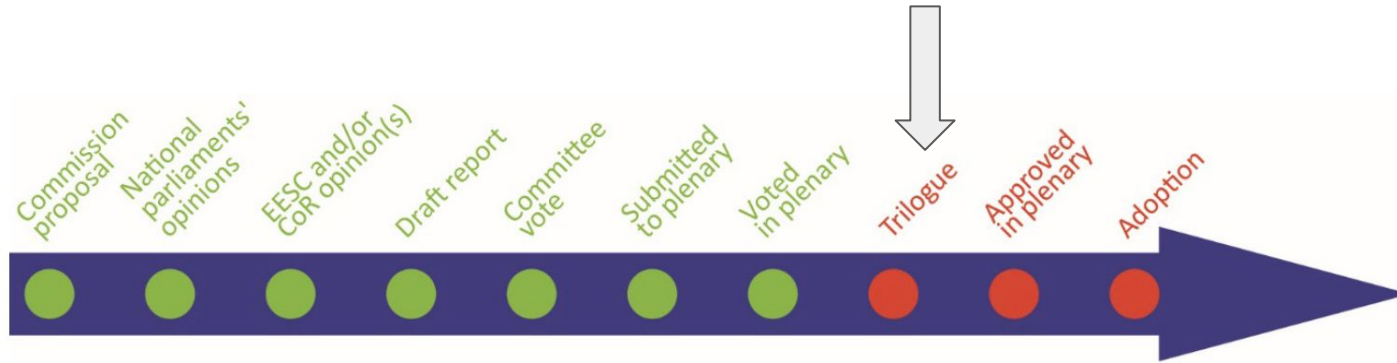
Part 1 - The AI Act proposal

- Relevance and current state
- Subject matter & scope
- Key actors
- Key definitions
- Prohibited AC systems
- Emotion Recognition Systems as high risk systems
- Obligations for high risk systems
- Cooperation & penalties
- Impact on education

Current legislative state



The AI Act is a **moving target** and subject to **changes**!



Our tutorial is based on the latest proposal of the AI Act adopted by the European Parliament

Relevance for AC community



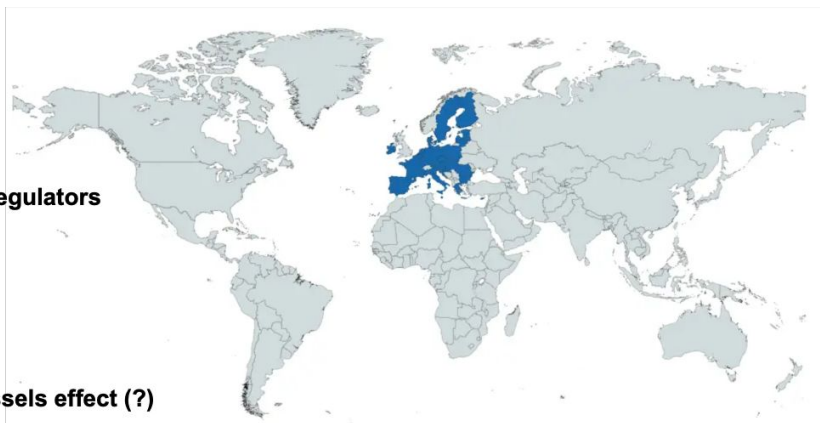
Regulates AC systems



Cooperation with regulators



Brussels effect (?)



(extra)territorial scope



Penalties



Burdensome obligations

Non-EU Regulations on AI

- Over **167** different AI ethics guidelines exist at the moment - including governmental initiatives, supra-national efforts, and guidelines by coalitions, institutions and companies
- China has an enforced AI law while Canada, UK, US, Brazil, Australia, Singapore, Japan, Israel, Italy and Germany are working on draft regulations/legislations
- There is an absence of internal enforcement or governance mechanism seen in most of these guidelines and regulations
- The **US** has proposed an '[AI Bill of Rights](#)', which has been designed to combat the pervasive fear of AI misuse and provides recommendations for safely using AI tools in both the public and private sectors, but is not legally binding
- **Canada** released the [Artificial Intelligence and Data Act \(AIDA\)](#), intended as “the first step towards a new regulatory system designed to guide AI innovation in a positive direction and to encourage the responsible adoption of AI technologies by Canadians and Canadian businesses.”

<https://algorithmwatch.org/en/ai-ethics-guidelines-global-inventory/>

<https://www.kwm.com/global/en/insights/latest-thinking/summary-of-ai-regulation-around-the-world.html>

https://en.wikipedia.org/wiki/Regulation_of_artificial_intelligence

<https://www.washingtonpost.com/world/2023/09/03/ai-regulation-law-china-israel-eu/>

Subject matter (Article 1)



AI Act = mix of safety & fundamental rights regulation

Why?	How?
Promote uptake of human-centric and trustworthy AI	Harmonised rules for the placing on the market, putting into service and use of AI systems in the EU
Ensure a high level of protection of health, safety, fundamental rights, rule of law, and environment from harmful effects of AI systems in the EU	<ul style="list-style-type: none">-Risk-based approach (prohibitions, specific requirements for high-risk systems);-Transparency obligations, rules on market monitoring, market surveillance governance and enforcement;-Rules concerning EU's 'AI Office'
....while supporting innovation	Measures to support innovation with focus on SMEs and start-ups, including regulatory sandboxes

Key actors

Provider	actor that develops (or has developed) an AI system with a view to placing it on the market or putting it into service under its own name or trademark	Article 3 (2)
Deployer	actor that uses an AI system under its authority (except personal use)	Article 3 (4)
Importer	actor that places on the market or puts into service an AI system that bears the name or trademark of a natural or legal person established outside the EU	Article 3 (6)
Distributor	actor in the supply chain, other than provider/importer, that makes an AI system available on the EU market	Article 3 (7)
Representative	has received a written mandate from a provider of an AI system to perform and carry out latter's obligations and procedures established by AI Act	Article 3 (5)
Operator	means the provider , the deployer , the authorised representative, the importer and the distributor	Article 3 (8)

Scope (Article 2)



It is **irrelevant** whether provider is located/established in or outside EU



extra-territorial scope

Providers placing AI systems on the market or putting them into service **in the EU**

(para 1 lit a)

Deployers established or located **in the EU**

(para 1 lit b)

Natural persons located **in the EU** that are **adversely affected** by AI system

(para 1 lit cc)



AI Act

Providers located **in the EU** 'exporting' prohibited AI systems to **third countries**

(para 1 lit ca)

Importers, distributors, representatives established/located **in the EU**

(para 1 lit cb)



extra-territorial scope

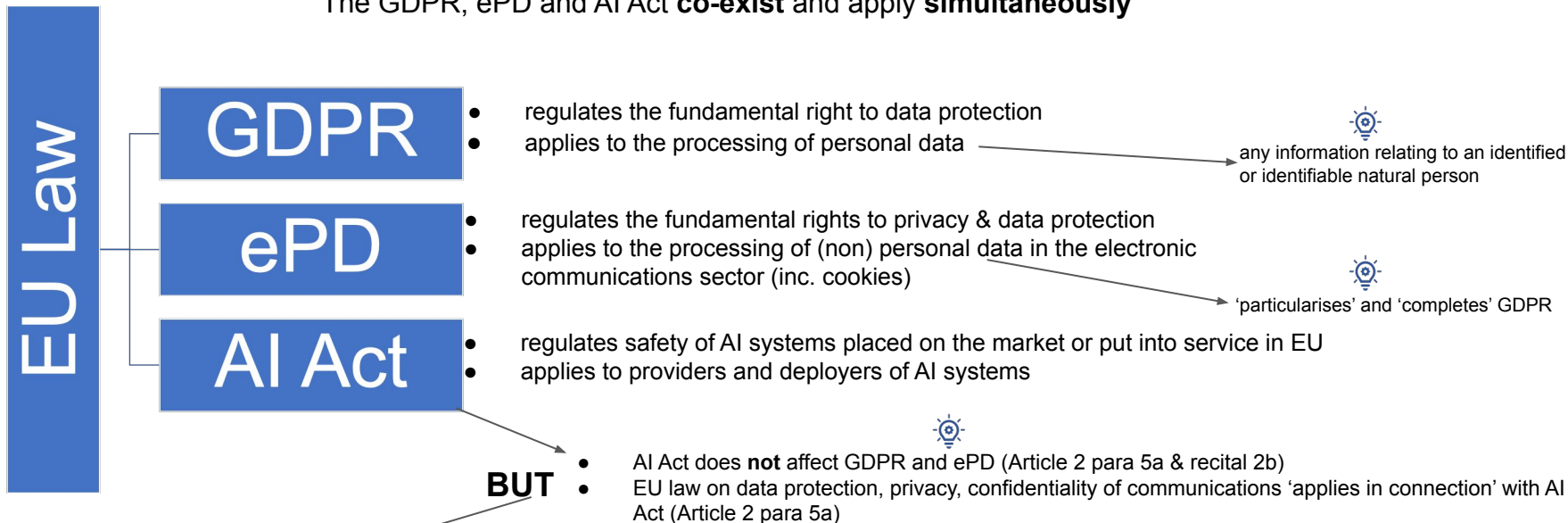
Providers and **deployers** established or located **outside EU** ('third country') if:

- foreseen by public **international law**; or
- **output** produced by AI system is intended to **be used in EU**

(para 1 lit c)

Scope: interplay with EU privacy & data protection law (Article 2 para 5a)



The GDPR, ePD and AI Act **co-exist** and apply **simultaneously**



AI Act **prevails** over GDPR (lex specialis) regarding:

- processing of special data to detect bias in high-risk AI systems (Article 10 para 5)
- requirement to obtain **consent** when processing personal data for **ERS** (Article 52 para 2)
- processing of personal data for regulatory sandbox purposes (Article 54)

Scope: research exception (Article 2 para 5d)

- AI Act does not apply to **research, testing and development activities** regarding an AI system provided that it **respects** fundamental rights and applicable Union law  including fundamental right to data protection
- Recital 2f: exception covers “AI systems specifically developed for the **sole purpose of scientific** research and development”  emphasises focus on scientific research
- Temporary testing of an AI system for its intended purpose in real world conditions **outside** of a laboratory or otherwise **simulated** environment ‘testing in real world conditions’ is **not** covered by this exemption
- AI Act to respect freedom of **scientific** research and not undermine such activities
- Commission and AI Office to further clarify the scope of this exception

Scope: are researchers/institutions developers and/or deployers?

- Are scientific researchers / research institutions:
 - Providers? Arguably **not**, not necessarily an intention to placing AI systems on the market or putting them into service under own name/trademark
 - Deployers? Arguably **not**, as focus lies on development of AI system, not use

Key definitions: AI system

at least some degree of independence of actions from human controls & of capabilities to operate without human intervention

Article 3 (1)
Recital 6

machine-based system that is designed to operate with varying levels of **autonomy** and that can, for **explicit or implicit objectives**, generate **outputs** such as predictions, recommendations, or decisions, that **influence** physical or virtual **environments**

explicit human-defined objectives vs implicit objectives; objectives of system may differ from intended purpose

output generated by the AI system influences environment, even 'by merely introducing new information to it'

contexts in which AI systems operate



- Aligned with [OECD](#) and [NIST](#) definition
- Focus on ML capabilities (see recital 6a)
- Based on 'key characteristics' of AI such as learning, reasoning, modelling capabilities to distinguish it from simpler software/programming approaches

Key definitions: Emotion Recognition System ('ERS')

arguably, notion 'ERS' is misleading as definition covers much more than emotions

Article 3 (34)

AI system for the **purpose** of identifying or inferring **emotions**, **thoughts**, **states of mind** or **intentions** of individuals or groups on the basis of their **biometric** and **biometrics-based data**

no specific recital on ERS, or what emotions are (!)

Article 3 (33)

personal data resulting from specific technical processing relating to the **physical, physiological** or **behavioural characteristics** of a natural person, which **allow** or **confirm** the unique **identification** of that natural person, such as facial images or dactyloscopic data

Article 3 (33a)

data resulting from specific technical processing relating to **physical, physiological** or **behavioural** signals of a natural person



- Definition 'borrowed' from Article 4 (11) GDPR
- Doubtful whether this matches with AC, as definition focuses on identification



- New concept/definition in EU law
- Recital 7 mentions facial expressions, movements, pulse frequency, voice, key strikes or gait as examples

How the AI Act proposal aims to regulate AC systems

High risk

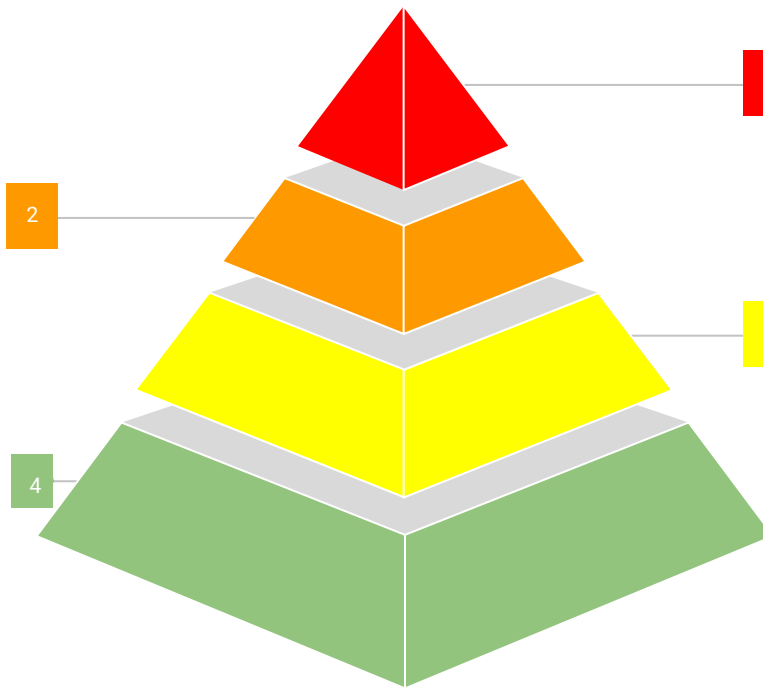
Art. 6 (2); Annex III (1) aa

Emotion Recognition Systems ('ERS')

Subject to specific requirements (Articles 8-15), transparency & consent obligations (Article 52), third party conformity assessment (Article 43 & recital 64)

Minimal risk

Not particularly relevant for AC, e.g. AI-enabled video games, spam filters



Unacceptable risk

Art. 5 (1) dc

Prohibition: AI systems inferring emotions of natural persons in the context of law enforcement, border management, workplace and education institutions

Limited risk

Not particularly relevant for AC, e.g. chatbots

Art. 3 (1) 1a: risk =

probability of an occurrence of harm
+
severity of that harm

Prohibited AC systems

Why is the use of AC systems prohibited in some contexts?

- limited reliability: emotion categories are neither reliably expressed through, nor unequivocally associated with, a common set of physical or physiological movement (recital 26c)
- major risks for abuse arise when AC systems are deployed in real-life situations related to law enforcement, border management, workplace and education institutions (recital 26c)

Inconsistency: the prohibition in Article 5 (1) dc refers to AI systems *to infer emotions of a natural person*, but not to the *definition of ERS* (?)

Note: AC systems might also be prohibited if they:

- deploy subliminal and manipulative or deceptive techniques (Article 5 (1)a) and/or;
- exploit vulnerabilities of a person or group of persons (Article 5 (1)b)

ERS as high risk systems

Why are ERS classified as 'high risk'?

- Because biometric data are protected as 'special data' under the GDPR (recital 33a)
- Due to "serious concerns" about scientific basis of AC systems (recital 26c)
- Because emotions or expressions/perceptions thereof vary considerably across cultures, situations and even 'within a single individual' (recital 26c)
- Key shortcomings (recital 26c):
 - limited reliability of emotion categories
 - lack of specificity - emotion categories do not 'perfectly match' physical/physiological expressions
 - limited generalisability - effects of context and culture are not 'sufficiently' considered

ERS as high risk systems

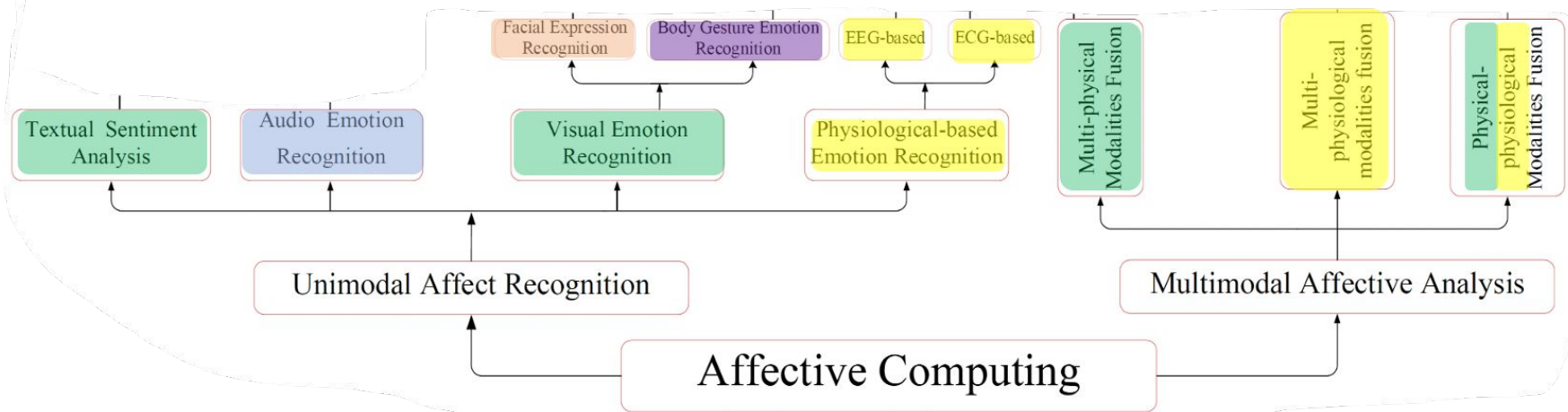
- Definition ERS covers single-modal and multi-modal approaches in AC
- The classification as high risk system under AI Act does not indicate that the use of such a system is necessarily lawful or unlawful under EU law, such as EU data protection law (recital 41)
- Arguably limited relevance of biometric data
 - definition focuses on identification
 - doubtful whether this matches with AC systems
 - legislative flaw (?)
- Biometrics-based data are **highly** relevant



under GDPR, biometric data is **only** protected as 'special' data if processed to **uniquely identify** individual see Article 9 (1) GDPR

Biometrics-based data & AC taxonomy

physical, **physiological** or **behavioural** signals such as **facial expressions**, **movements**, **pulse frequency**, **voice**, **key strikes or gait**



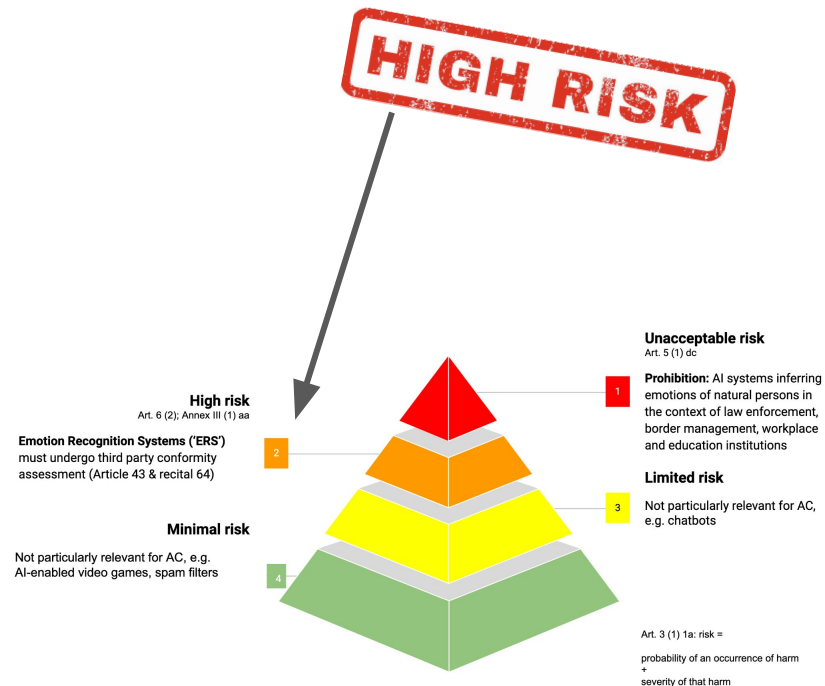
Y. Wang et al (2022) A systematic review on affective computing: emotion models, databases, and recent advances Volumes 83-84 Information Fusion

Compliance high risk systems

- Articles 8-15 focus on requirements of AI system
- Articles 16-23 focus on providers
- Articles 24-26 focus on actors *other* than providers/deployers
- Article 28 focuses on value chain
- Article 29, 29a focus on deployer
- Article 40-51 focus on standards and conformity assessments

Requirements for high risk systems

- Risk management system Article 9
- Data and data governance Article 10
- Technical documentation Article 11
- Record keeping Article 12
- Transparency Article 13 & 52
- Human oversight Article 14
- Accuracy, robustness and cybersecurity Article 15
- Responsibility among AI value chain Article 28
- Fundamental rights assessment Article 29a



Risk management (Article 9)

HIGH RISK

Should run throughout entire lifecycle of AI system -> continuous iterative process



Data & data governance (Article 10)

HIGH RISK

- High quality of training validation & testing data dependent on market segment and scope of application
- Data governance requirements such as
 - design choices
 - data handling
 - assessment of availability, suitability and quantity of data
- Processing of personal data for bias detection and correction allowed under some conditions, i.e.
 - bias detection not possible with processing synthetic or anonymised data;
 - data are pseudonymised,
 - provider takes appropriate technical and organisational measures
 - no disclosure and erasure once bias has been corrected
 - effective measures to ensure availability, security and resilience of processing systems
- Obligations transferable to deployer where provider is not in a position to assess data that is with deployer



including 'special' personal data
(see para 5)

Technical documentation (Article 11 & Annex IV)



- Specified in Annex IV
- **General** description of AI system including, e.g.
 - intended purpose, nature of data, categories of persons or groups likely to be affected by use of system
 - description of hardware, deployer's interface, optimisation goals, expected output and output quality
 - detailed instructions for interpreting system's output
- **Detailed** description of AI system and process for its development including, e.g.
 - architecture, design specifications, key design choices
 - algorithms and data structures including **decomposition** of its components and interfaces, how they **relate** to one another and how they provide for the overall processing or **logic** of the system
 - data requirements, assessment human oversight
 - validation & testing procedures used, including metrics used to measure accuracy, robustness

Record keeping (Article 12)



- State of the art logging capabilities facilitating monitoring of operations according to Article 29(4) and post market monitoring (Article 61)
- Recording of events that may lead to substantial modification of the system
- Logging capabilities enabling the recording of energy consumption, measurement or calculation of resource use and environmental impact of system

Transparency (Article 13)



- Operation of system must enable providers and deployers to **reasonably** understand system's **functioning**
- Transparency = all technical means available are used to ensure that AI system's **output** is **interpretable** by the **provider** and **deployer**
- Instructions for use need to specify (e.g.):
 - characteristics, capabilities and limitations of performance including **accuracy**, robustness, cybersecurity and circumstances that may have an impact thereon
 - possible risks of use
 - degree to which system can provide an explanation for decision it takes
 - performance regarding the **persons**/groups of person on which system is intended to **be used**
 - information about user actions that may influence system performance
 - information about **training, validation** and **testing** data sets used
 - predetermined changes to system
 - human oversight measures
 - maintenance & care measures
- Thus: **extensive** and **formalistic** list of transparency requirements

Transparency of ERS (Article 52)



- Article 52(2a) obliges **deployers** to inform individuals concerned about the operation of ERS
- Recital 70 explains that "natural persons should be notified" when exposed to ERS, but does not further clarify what that precisely entails
- Arguably, it simply means to make natural persons aware that they are exposed to an ERS
- Deployers of ERS are **not** obliged to inform individuals about what specific emotion the system detected

Note: deployers must obtain consent from natural persons exposed to ERS



ERS must be designed/developed in a way which allows to obtain consent **prior** to the processing of personal data!

AI Act = *lex specialis* and prevails over Article 6 GDPR, which contains additional grounds for processing other than consent

Human oversight (Article 14)



- Effective oversight by human, aiming to prevent/minimise risks
- Human oversight must take specific risks, level of automation and context of system into account
- Human must have sufficient level of AI literacy (Art. 4b) and necessary support/authority to exercise function
- Human must be able to
 - understand relevant capacities & limitations of system
 - remain aware of automation bias
 - correctly interpret output generated by system
 - decide to not use, disregard, override or reverse output
 - intervene on the operation of the system (e.g use 'stop button' or similar mechanism)

Accuracy (Article 15)

HIGH RISK

- In light of intended purpose, system must achieve appropriate levels of accuracy, robustness and cybersecurity
- Resilience regarding errors, fault or inconsistencies in particular due to interaction with persons
- AI Office to provide non-binding guidance on how to measure appropriate level of accuracy and robustness
- Levels of accuracy and relevant accuracy metrics must be declared in the accompanying documentation containing inter alia:
 - the overall expected level of accuracy in relation to its intended purpose
 - detailed information about the system's degree of accuracy for *specific* persons or groups of persons on which the system is intended to be used



accuracy under the AI Act more specific than accuracy in data protection law




Fundamental rights impact assessment (Article 29a)



Deployers must perform an assessment **before** using system considering (a.o.):

- lit d: verification that use of system complies with EU and national laws on fundamental rights
- lit e: 'reasonably foreseeable impact' on fundamental rights such as
 - right to human dignity
 - right to privacy and protection of personal data
 - freedom of expression and information
 - non- discrimination
 - right to education and consumer protection
 - workers' rights & rights of persons with disabilities
 - gender equality
 - intellectual property rights
- lit f: specific risk of harm likely to impact marginalised groups and vulnerable groups
- lit h: detailed plan how harms and negative impact on fundamental rights will be mitigated
- lit i: governance system concerning human oversight, complaint handling and redress

 non-exhaustive list, see recital 28a

 should already be taken into account when developing system

Deployers must notify national Supervisory Authority and involve representatives of persons/groups likely to be affected

Cooperation (Article 23)



- Upon ‘reasoned request’, providers and deployers obliged to provide all the information and documentation necessary to demonstrate compliance with requirements applicable to high risk systems (including logs) with:
 - National competent competent supervisory authority; or
 - AI Office; or
 - European Commission
 - According to recital 79, this includes access to:
 - the training, validation and testing datasets;
 - the trained and training model of the high-risk AI system, including its relevant model parameters and their execution /run environment;
 - the source code
 - Such information constitute ‘trade secret’ and are subject to confidentiality obligations
- after having exhausted all ‘other reasonable ways’
to assess/verify conformity with requirements
applicable to high risk systems

Responsibilities among AI value chain (Article 28)



- Distributor, importer, deployer or ‘other third party’ become a provider of high risk system if they:
 - put their name or trademark on high risk AI system; or
 - make a substantial modification on high risk system; or
 - make a substantial modification to an AI system so that it becomes a high risk system
- In these cases, the provider that initially placed AI system on the market or put it into service is no longer provider of that system
- However, ‘initial’ provider to disclose technical documentation and all other relevant information to ‘new’ provider



‘substantial modification’ recital 66

- unplanned change, including continuous learning, creating new unacceptable risk and significantly affect compliance of high risk system
- intended purpose of the AI system changes
- for AI systems that continue to learn: changes to the algorithm and its performance that have been pre-determined by the provider and assessed at the moment of conformity assessment do not constitute a substantial modification

Penalties (Article 71)

- Administrative fines of up to € 40'000'000 or 7% of worldwide annual turnover for non-compliance with provisions relating to prohibited AI systems
- Administrative fines of up to € 20'000'000 or 4% of worldwide annual turnover for non-compliance with provisions concerning:
 - data and data governance (Article 10)
 - transparency (Article 13)
- Administrative fines of up to € 5'000'000 or 1% of worldwide annual turnover for the supply of incorrect, incomplete or misleading information to notified bodies and national competent authorities

Regulation mechanisms

- The AI Act specifies a **regulatory sandbox** and **national supervisory authorities** to ensure implementation
- ‘**regulatory sandbox**’ means a controlled environment established by a public authority that facilitates the safe development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan under regulatory supervision; (Article 3, point 44(g))
 - How technologically feasible is this, especially given the fast-paced AI developments and the lack of a current prototype?
- ‘**national supervisory authority**’ means a public authority to which a Member State assigns the responsibility for the implementation and application of this Regulation, for coordinating the activities entrusted to that Member State, for acting as the single contact point for the Commission, and for representing the Member State in the management Board of the AI Office; (Article 3, point 42)
 - A lot depends on how stringent these authorities are and how effective the communication is

EU AI Act and education

Recital 35 acknowledges the importance of AI systems in education

Lists systems that would qualify as high-risk

- since they may determine the educational and professional course of a person's life and therefore affect their ability to secure their livelihood.
- such systems *can be particularly intrusive and may violate the right to education and training as well as the right not to be discriminated against* and perpetuate historical patterns of discrimination

Educational and vocational training systems under high-risk that are prohibited (Annex III, paragraph 1):

- systems that influence the admission decisions
- assessing students
- assessing the level of education that students receive or access
- influencing the level of education that an individual will receive
- monitoring and detecting students' prohibited behaviour during tests in the context of education

Provocative questions - impact on education

- Will these regulations cause the EU to miss out on the latest AI developments?
 - Google Bard as an example
 - Competitive disadvantage for the EU
- Will some educational AC research have to be re-evaluated and revamped before being released as a product?
 - While research is exempted, it can only be evaluated by deploying it in 'real-world scenarios'

<https://www.businessinsider.com/google-bard-chatbot-blocked-in-the-eu-postponed-rollout-2023-6>

Recap

- AI Act **excludes scientific research**. However, it might implicate the real-life applications (e.g., patents, products) of research, potentially hindering grants and sustainability of research
- AI Act is also relevant for actors **outside** the EU (extraterritorial scope, Brussels effect)
- AC systems in **context** of law enforcement, border management, workplace and education institutions are prohibited
- ERS are high risk, **irrespective** of context
- Let's **improve lawmaking** and discuss shortcomings, deficiencies and efficient risk mitigation mechanisms

Questions?



Part 2 - Late Breaking Results Paper

**Ethical Risks, Concerns, and Practices of Affective Computing
A Thematic Analysis**

Ethical Risks, Concerns, and Practices of Affective Computing

A Thematic Analysis

11.09.2023 | @MIT Media Lab, Boston

Deniz Iren | deniz.iren@ou.nl
Associate Professor, Open Universiteit
<https://www.linkedin.com/in/deniziren/>

Ediz Yildirim | ediz.yildirim@ou.nl
PhD Researcher, Open Universiteit

Krist Shingjergji | krist.shingjergji@ou.nl
PhD Researcher, Open Universiteit



[download](#)



Open Universiteit

Introduction

- AI is progressing fast, raising concerns
- Ethical safeguards are needed
- Rules and regulations are being prepared
- Affective computing is particularly sensitive
- Affective Computing community is also taking action to ensure ethical practice
- We aim at investigating the ethical considerations of our community

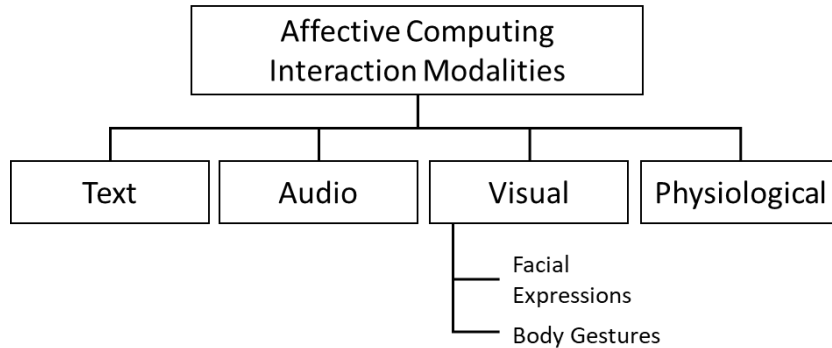




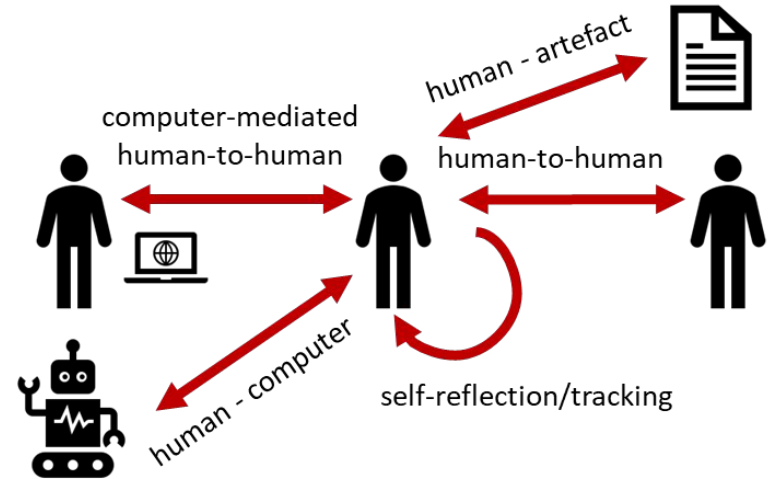
Research Questions

- RQ1: What are the ethical risks and concerns reported by affective computing researchers?
- RQ2: What are approaches proposed by affective computing researchers to mitigate these risks?
- RQ3: What is the potential impact of the regulations (e.g., The AI Act) on different types and applications of affective computing?

Background



Typology of affective computing **interaction modalities**



Typology of **communication channels** enhanced by affective computing

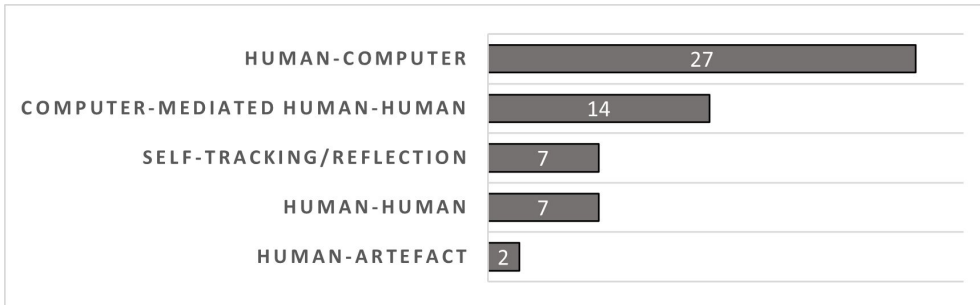
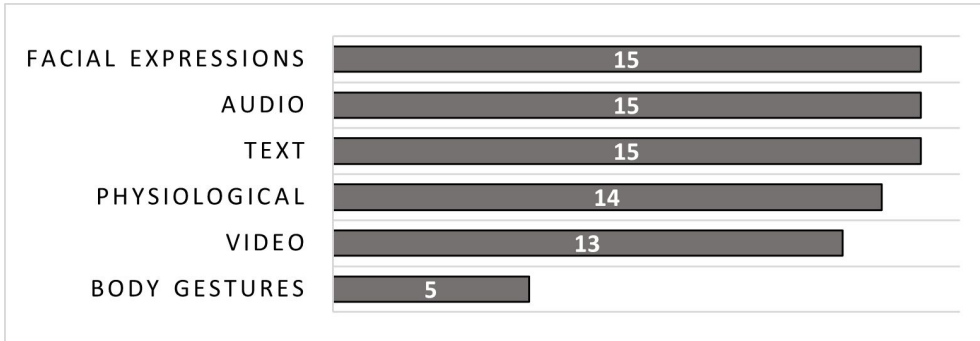
Methodology

- Data: Ethical impact statements, N=70
- Method: Thematic analysis
- Goal: To identify reported **limitations, risks/concerns, and mitigation strategies**
- Code groups: **study-related, data-related, application-related**



Findings

Descriptive analysis



Codes:

- 40 x limitations
- 42 x mitigation strategies

Categories

STUDY

DATA

APPLICATION

Themes

- Human subjects
- Study design
- Environmental impact
- Data quality
- Nature of data
- Data accessibility
- Application

Findings: Study-related

THEMES		CODES		
		LIMITATIONS	RISKS	MITIGATION
STUDY	HUMAN SUBJECTS	<ul style="list-style-type: none"> ⇨ Participant selection and compensation (3) 	<ul style="list-style-type: none"> ⇨ Limited oversight (2) ⇨ Harm to participants (2) 	<ul style="list-style-type: none"> ⇨ Involve IRB(26) ← ⇨ Apply informed consent (22) ← ⇨ Participants can drop-out at will (4) ⇨ Transparent reporting (2)
	STUDY DESIGN	<ul style="list-style-type: none"> ⇨ Context-specific (2) ← 	<ul style="list-style-type: none"> ⇨ Results are not generalizable (6) ← ⇨ Reduced construct validity (2) 	<ul style="list-style-type: none"> ⇨ Improve the study (5) ← → Conduct more research (4) → Improve the performance (3)
	ENVIRONMENTAL IMPACT		<ul style="list-style-type: none"> ⇨ Environmental Impact (5) ← 	<ul style="list-style-type: none"> ⇨ Examine and report environmental impact (2) ⇨ Train small models (1) ⇨ Use pretrained models (1) ⇨ Avoid over-personalization of models (1)

Findings: Data-related

THEMES		CODES		
		LIMITATIONS	RISKS	MITIGATION
DATA	DATA QUALITY	<ul style="list-style-type: none"> ⇨ Small sample size (10) ← ⇨ Sample is not representative (4) <ul style="list-style-type: none"> → Demographics (4) → Limited set of emotions (1) ⇨ Data imbalance (2) 	<ul style="list-style-type: none"> ⇨ Results are not generalizable (6) ⇨ Discrimination (3) ⇨ Biases (24) ← ⇨ Reduced accuracy (3) 	<ul style="list-style-type: none"> ⇨ Improve the data (10) ← <ul style="list-style-type: none"> → Collect more data (7) → Collect more diverse data (4) → Apply sampling strategies (2) → Balance data (3) → Examine the biases (4) → Use multiple datasets (2)
	NATURE OF DATA		<ul style="list-style-type: none"> ⇨ Sensitive data (5) ← <ul style="list-style-type: none"> → Healthcare/mental → Offensive content ⇨ Private data (14) ← ⇨ Personally identifiable data (1) ⇨ Unauthorized access to the data (2) ⇨ Unclear IP rights and licensing (2) 	<ul style="list-style-type: none"> ⇨ Anonymization/De-identification (22) ← ⇨ Setup data protection policy (2) ⇨ Establish data protection measures (2)
	OPEN DATA	<ul style="list-style-type: none"> ⇨ Private/unavailable research data (2) 	<ul style="list-style-type: none"> ⇨ Reproducibility is hindered ⇨ Misuse of data 	<ul style="list-style-type: none"> ⇨ Make research data available (5) ⇨ License the published datasets (2) ⇨ Establish EULA for published datasets (2)

Findings: Application-related

THEMES		CODES		
		LIMITATIONS	RISKS	MITIGATION
APPLICATION	APPLICATION	<ul style="list-style-type: none"> ⇨ Limited stakeholder involvement (2) ⇨ Critical domains and application fields ← → Healthcare (20) → Education (4) → Social services (9) → Law enforcement and border control (0) → Workplace (2) 	<ul style="list-style-type: none"> ⇨ Harmful applications (18) ← → Surveillance → Deception → Manipulation → Restrict autonomy ⇨ Societal adverse impact (2) → Limit fundamental rights → Controversial subjects ⇨ Failure consequences (1) 	<ul style="list-style-type: none"> ⇨ Identify and address failure consequences (1) ⇨ Provide transparent information to user (2)

Conclusion

- Please, see the paper for more details
- Let's meet
 - LBR Flash Talks on **Monday 3:00-4:30**
 - LBR Poster Session on **Tuesday 3:00-4:30**
- Tune in, for the journal extension
- Reach out: deniz.iren@ou.nl



[download](#)

Part 3 - Group Discussion

Part 3: Group Discussion (max 90 mins, with coffee :))

1. Goals:

- Discuss AI Act (and AI Regulations)
- Share concerns, risks, improvement suggestions, clarifications...
- Then, we will prepare a report

2. Consent

Please let us know if you do not give **consent** to data collection with the purpose of creating a report:

- Moderators notes during group discussions.
- Participant notes on the provided templates.
- No names, personal info, only group numbers

Part 3: Group Discussion

3. Expert availability

- Andreas will be available to answer your legal questions and provide clarifications if needed.

4. Form groups

- 3-5 people | according to similarity (domain > modality > risk-level)

		DOMAIN				
		HEALTH	EDU	INDUSTRY	DEFENSE	OTHERS
MODALITY	FACIAL EXPRESSION					
	BODY LANGUAGE					
	SPEECH					
	WEARABLES					
	OTHERS					

Part 3: Group Discussion

5. Guiding questions

- Use the guiding questions to lead the discussion (if needed)

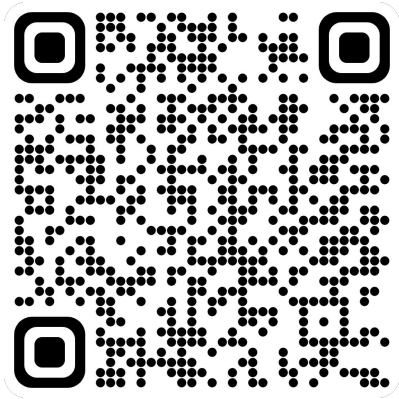
6. Take notes

- Preferably on the templates

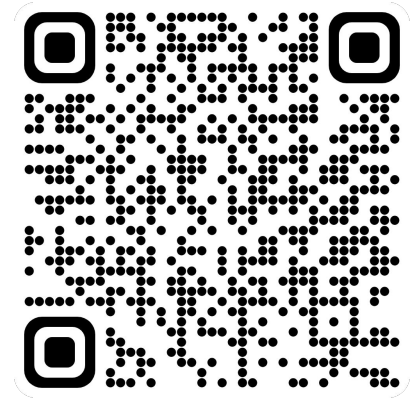
7. Group highlights and discussion

- At the end, groups will shortly present their highlights and discuss with other groups.

Guiding Questions



[questions](#)



[slides](#)

Part 4 - Presentations of Discussion Highlights

Part 4: Group Presentations (max 30 mins)

- Presentation of the highlights
- Open discussion

The End

Thank you for your participation

If you want to be notified regarding the results

- Give us your email
- Contact us