Ethical Risks, Concerns, and Practices of Affective Computing

Document status and date:
Published: 10/09/2023

Document Version:
Publisher's PDF, also known as Version of record

Open Universiteit
www.ou.nl

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

Deniz Iren, Ediz Yildirim, Krist Shingjergji
deniz.iren@ou.nl

Open Universiteit

download

1. Introduction

- AI is progressing fast, raising concerns
- Ethical safeguards are needed
- Rules and regulations are being prepared [1]
- Affective computing is particularly sensitive
- Affective computing community has already taken action to ensure ethical practice [2]
- This study aims at investigating the ethical considerations of our community
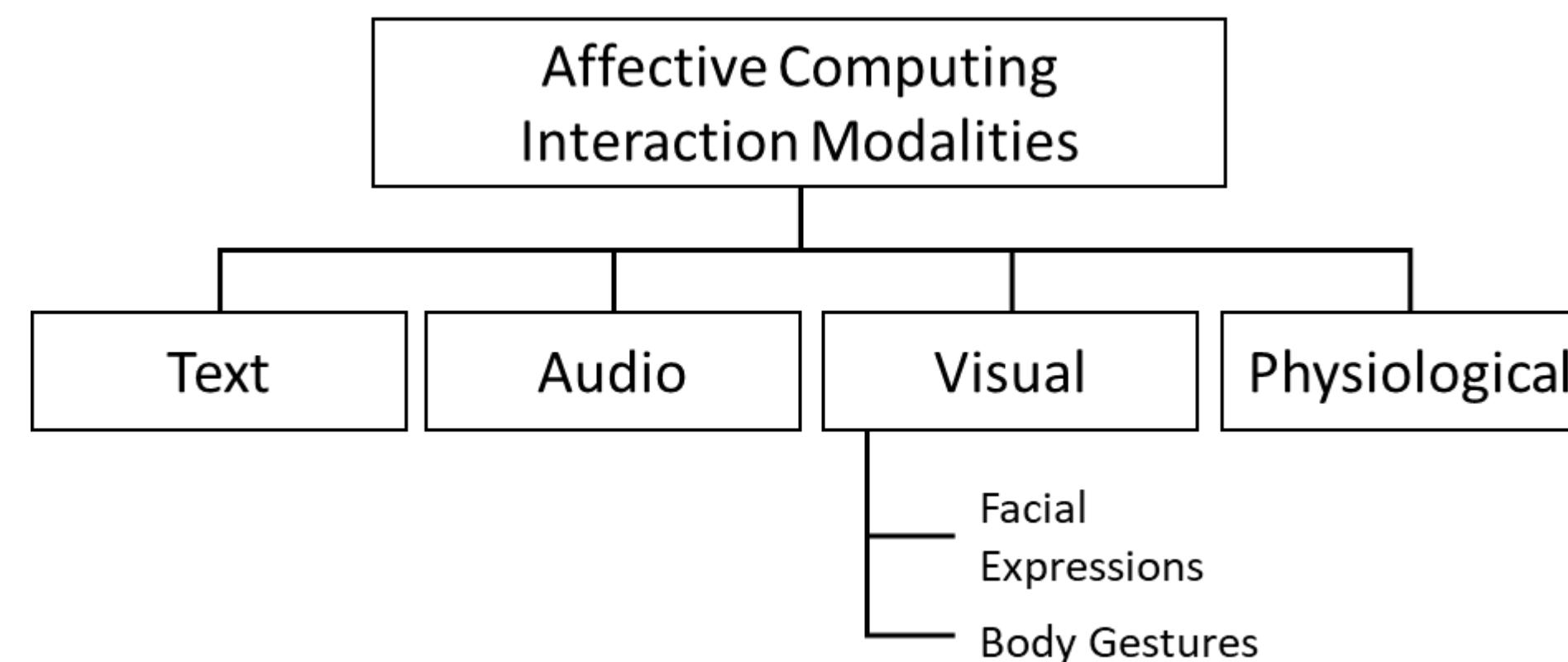
2. Affective Computing Typology

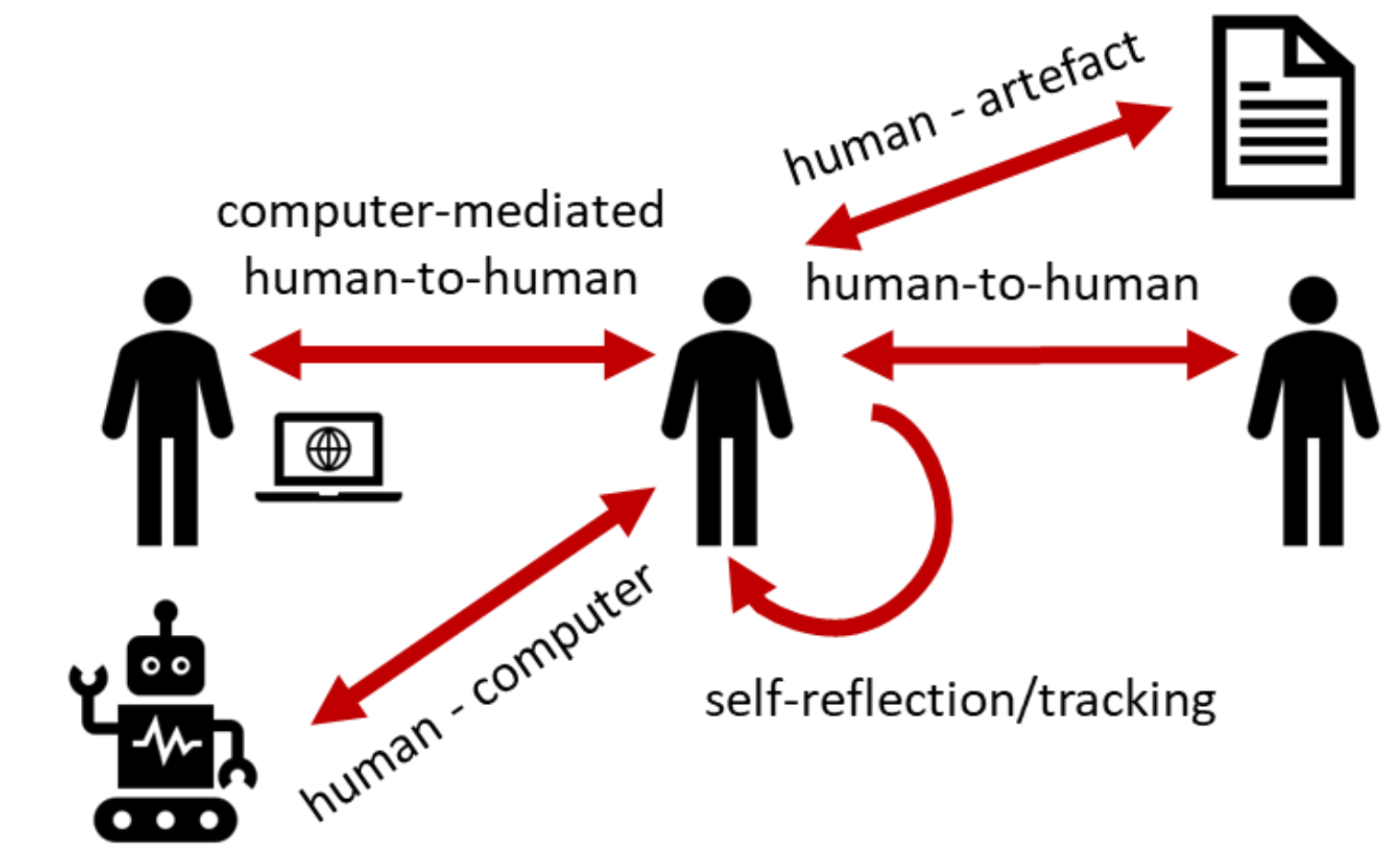**Figure 1.** Typology of affective computing interaction modalities [3]

Figure 2. Typology of communication channels enhanced by affective computing

3. 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?

4. Research Method

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

5. Descriptives

Categories: STUDY | DATA | APPLICATION

Themes:
- Human subjects
- Study design
- Environmental impact
- Data quality
- Nature of data
- Open data
- Application

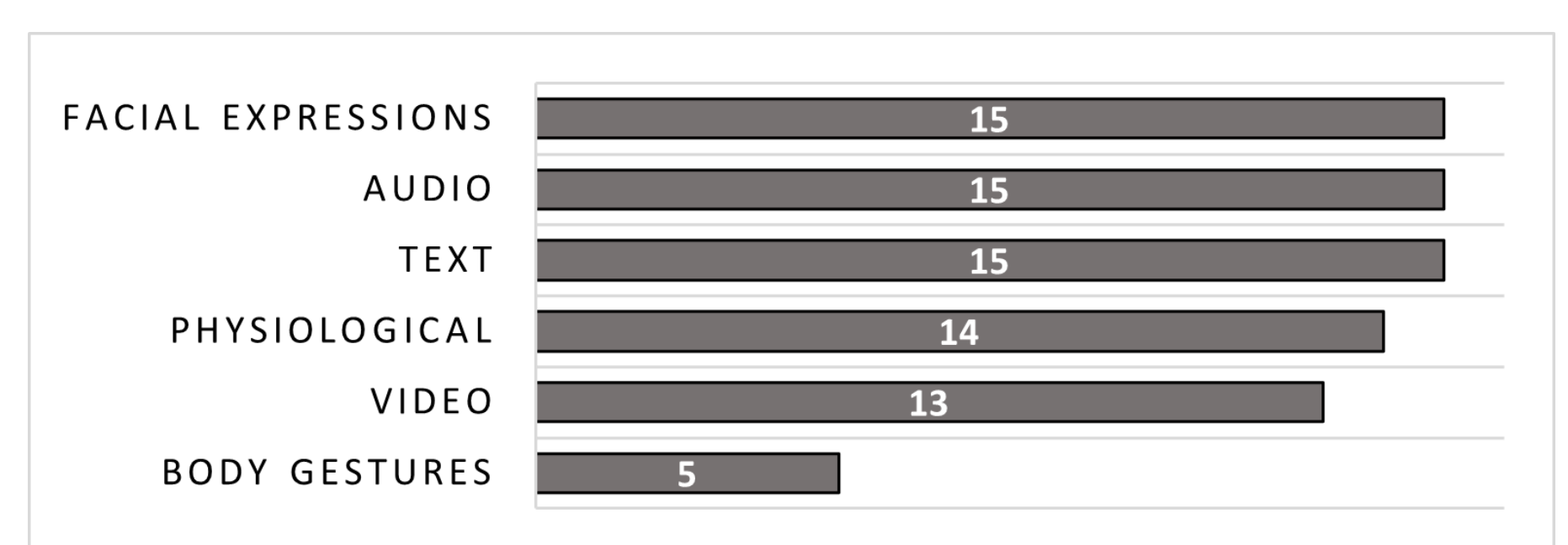**Figure 5.** Categories and themes identified in thematic analysis

Figure 3. Number of papers addressing different interaction modalities

FACIAL EXPRESSIONS	15
AUDIO	15
TEXT	15
PHYSIOLOGICAL	14
VIDEO	13
BODY GESTURES	5

Figure 4. Number of studies addressing different types of communication channels

HUMAN-COMPUTER	27
COMPUTER-MEDIATED HUMAN-HUMAN	14
SELF-TRACKING/REFLECTION	7
HUMAN-HUMAN	7
HUMAN-ARTEFACT	2

6. Main Findings

| THEMES		LIMITATIONS	RISKS	MITIGATION
STUDY	HUMAN SUBJECTS	⇨ Participant selection and compensation (3)	⇨ Limited oversight (2) ⇨ Harm to participants (2)	⇨ Involve IRB(26) ⇨ Apply informed consent (22) ⇨ Participants can drop-out at will (4) ⇨ Transparent reporting (4)
	STUDY DESIGN	⇨ Context-specific (2)	⇨ Results are not generalizable (6) ⇨ Reduced construct validity (2)	⇨ Improve the study (5) → Conduct more research (4) → Improve the performance (3)
	ENVIRONMENTAL IMPACT		⇨ Environmental Impact (5)	⇨ Examine and report environmental impact (2) ⇨ Train small models (1) ⇨ Use pretrained models (1) ⇨ Avoid over-personalization of models (1)
DATA	DATA QUALITY	⇨ Small sample size (10) ⇨ Sample is not representative (4) → Demographics (4) → Limited set of emotions (1) ⇨ Data imbalance (2)	⇨ Results are not generalizable (6) ⇨ Discrimination (3) ⇨ Biases (24) **[4]** ⇨ Reduced accuracy (3)	⇨ Improve the data (10) → 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		⇨ Sensitive data (5) → Healthcare/mental → Offensive content ⇨ Private data (14) ⇨ Personally identifiable data (1) ⇨ Unauthorized access to the data (2) ⇨ Unclear IP rights and licensing (2)	⇨ Anonymization/De-identification (22) ⇨ Setup data protection policy (2) ⇨ Establish data protection measures (2)
	OPEN DATA	⇨ Private/unavailable research data (2)	⇨ Reproducibility is hindered (2) ⇨ Misuse of data	⇨ Make research data available (5) ⇨ License the published datasets (2) ⇨ Establish EULA for published datasets (2)
APPLICATION	APPLICATION	⇨ Limited stakeholder involvement (2) ⇨ Critical domains and application fields → Healthcare (20) → Education (4) → Social services (9) → Law enforcement and border control (0) → Workplace (1)	⇨ Harmful applications (18) → Surveillance → Deception → Manipulation → Restrict autonomy ⇨ Societal adverse impact (2) → Limit fundamental rights → Controversial subjects → Failure consequences (1)	⇨ Identify and address failure consequences (1) ⇨ Provide transparent information to user (2)
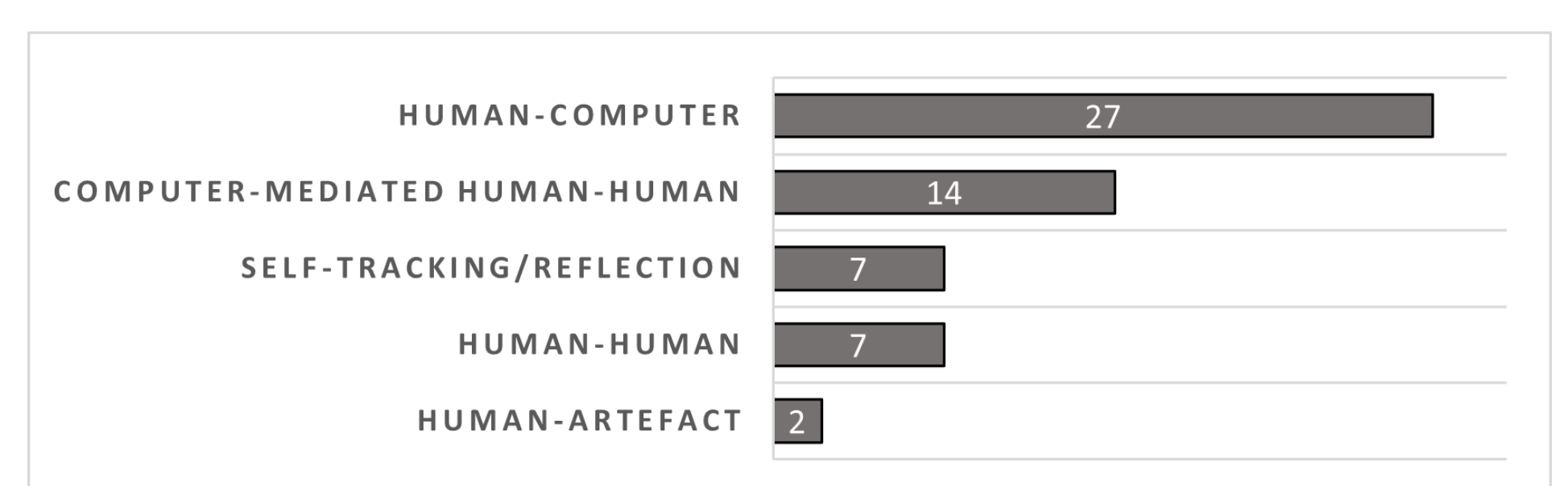
7. Conclusion

- Affective computing community has taken important steps to highlight ethical research.
- However, our findings indicate several gaps and non-standard ethical practices.
- We could benefit from more systematic guidelines for ethical research practice and reporting.

Limitations
- This thematic analysis focuses only on the ethical impact statement sections, and not on the other sections of the paper.

Future work
- We plan to prepare reports and open a communication dialogue between affective computing community and policymakers.
- We will extend our work to cover the leading journals of affective computing (e.g., IEEE Transactions of Affective Computing)

Authors

Deniz IREN | Ediz Yildirim | Krist Shingjergji
deniz.iren@ou.nl
Open Univesiteit, Netherlands

References

[1] Artificial intelligence act, https://eur-lex.europa.eu/legalcontent/EN/TXT/HTML/?uri=CELEX:52021PC0206, accessed 18-June-2023, 2021.

[2] D. Ong, J. Hernandez, R. Picard, et al., Writing an ethical impact statement for acii2023, https://acii-conf.net/2023/wp-content/uploads/2023/03/instructions-ethicalstatement.pdf, accessed 18-June-2023, 2023.

[3] Y. Wang, W. Song, W. Tao, et al., "A systematic review on affective computing: Emotion models, databases, and recent advances," Information Fusion, 2022.

[4] B. Aysolmaz, D. Iren, and N. Dau, "Preventing algorithmic bias in the development of algorithmic decision-making systems: A delphi study," 2020.