Universität Zürich[UZH]

Digital Society Initiative

Interdisciplinary Workshop: Ethics of Emotion Computing

Location: Digital Society Initiative (DSI), University of Zurich, Raemistr. 69, 8001 Zurich, Switzerland
Date: September 18-19, 2023

Organizers:
Prof. Dr. Eva-Weber Guskar (philosophy, PI of *INTERACT!* at Ruhr-University of Bochum, member of the *Ethics Community* of the DSI)
Prof. Dr. Anna Tuschling (media science, PI of *INTERACT!* at Ruhr-University of Bochum)

Topic:
This two-day workshop is focusing on emotion/affective computing. The aim of the workshop is to initiate a discussion among computer scientists, media scientists and philosophers about the core concepts of emotion computing as well as about emerging consequences and effects of applications of emotion computing in order to be finally able to properly assess the field from a normative-ethical and media-historical perspective.

Program

Monday, September 18

3 pm	Welcome
3:15 – 4:15	Rafael Calvo (computer science, London; via Zoom): *Affective computing: Value tensions in design engineering*
4:30 – 5:30	Gabriele Gramelsberger (philosophy, Aachen): *Affective Infrastructure. Towards the Naturalisation of the Digital*
6:00 – 7:15	Ana Paiva (computer science, Lisbon/Harvard): *Feelings can't be ignored: the role of emotions in the communication between humans and AI*
7:30	Apéro riche

Tuesday, September 19

9:00 – 10:00	Eva Weber-Guskar (philosophy, Bochum): *Emotion Detection Systems and Emotional Granularity*
10:15 – 11:15	Alexander Campolo (media science, Durham): *Affect as paradigm: The structure of machine learning problems*

11:30 – 12:30	Anna Tuschling (media science, Bochum): *Emotional and Affective States as Quasi-Semantics in the history of interactive Computing*
12:30 – 02:00	Lunch
2:00 – 3:00	Luke Stark (media science, London, Ontario): *Reordering Emotion: Scenes from the History of Affective Computing and Emotion AI*
3:30 – 5:00	General discussion with short inputs by Jonas Blatter (philosophy, Bochum), Daniela Wentz (media science, Bochum) and Piera Filippi (philosophy, Zurich)
Optional	Joint walk to the Zurich lake
7:00	Conference Diner

Abstracts

Rafael A. Calvo

Affective computing: Value tensions in design engineering

As engineers who "make things", we like to focus on the positive impact that our work has on the world, and we occasionally look at its unintended consequences. But what we rarely do is ask "what are the values by which we judge the systems we create?". I discuss an "archaeology of AI" using Weizenbaum's ELIZA chatbot, created in 1966 to reflect on different schools of psychology, philosophy and engineering that today shape the world, and how the inconclusive debates at the dawn of AI still rage through our work on Affective Computing.
In this talk I will be reflecting on the ethical considerations that arise from our affective computing research and how I have personally attempted to address the corresponding tensions within my own work. This is not a lecture on ethics, but rather a personal, and therefore partial, description of the dilemmas faced when looking at the "big picture" impact of our engineering work

Alexander Campolo

Affect as paradigm: The structure of machine learning problems

This paper asks why affect and emotion recognition have emerged as a paradigmatic type of applied scientific problem for machine learning. Rather than entering into contentious debates about the status of emotion science (Leys 2017; Fridlund 2017; Barrett et al. 2019) or further exploring historical connections to ideas like physiognomy (Delaporte 2008; Porter 2005), I analyze similarities between the science of emotional expression and machine learning in order to draw out deeper epistemic implications of the latter as a scientific technique.
There are a number of obvious points of intersection. Most generally, the face as an expressive object has long been the object of physiological and psychological (and, it should be said, criminological) study, and it has also served as a case of *pattern recognition* in machine learning since at least the 1960s, close to the field's inception (Bledsoe 1965). For more recent discriminative forms of modelling, basic affect and emotion theory provide a small, *discrete* set of data labels that have some scientific rationale. This permits the manual creation of training datasets that can be used to infer statistical associations between facial images and labels, presumably linked to interior emotional states.
In addition to these historical and pragmatic connections, I argue that there are a deeper set of epistemic commitments shared between emotion science and machine learning having to do with surfaces and depths, structures and intentionality. The psychological model of affect and emotional expression as an external indication of some interior state resembles machine learning's probabilistic conception of the world as an observable sample of some deeper underlying distribution accessible through dimensionality reduction and interpolation (Chollet 2021, 129). This shared model of deeper structures observable only at some observable surface entails a radical unintelligibility of ultimate causes (e.g. Silvan Tomkins' repeated emphasis on cognitive errors with regard to affects [2008, 64] and frequent characterizations of machine learning systems as black boxes); only associations of different effects can be modelled.
Curiously and perhaps controversially, the very unintelligibility of these causes appears to strengthen our belief in their truth. Machine learning systems are able to achieve "better than human performance" while emotional expressions resist our ability to modulate or dissimulate our inner selves. This sense of an inner

self that is all the more true as it eludes our consciousness and intentions transcends the bounds of affect and emotion. Machine learning is said to be able to use observable surfaces to identify complex interior states, dispositions and identities, from sexuality (Wang and Kosinski 2018) to political ideology (Rasmussen et al. 2023).

Gabriele Gramelsberger

Affective Infrastructure. Towards the Naturalisation of the Digital

Following Mark Weiser's vision from 1991 that the digital of the 21st century weaves itself "into the fabric of everyday life" until it is "indistinguishable from it," we are today facing an evolution of the digital toward its naturalization. In this context, the miniaturization of technology plays an important role, but even more so the increasingly natural access to the digital, which brings the digital closer to emotions, affects and intuitions. In English, the term naturalization ambiguously refers on the one hand to the digital as a technology that becomes (second) nature, and on the other hand to the naturalization (citizenship) of the digital into the human sphere. The digital as an increasingly affective infrastructure that enables a more natural (intuitive) access to the digital plays a crucial role in this evolutionary process. The paper analyzes from a philosophical perspective the emergence of the digital as an affective infrastructure as well as the phenomenological consequences of the naturalization of the digital in the human domain.

Ana Paiva

Feelings can't be ignored: the role of emotions in the communication between humans and AI

Social agents, chatbots, or social robots have the potential to revolutionize the way we interact with technology. As their affordability increases, they will play a greater role in our daily activities, performing various tasks and communicating with us naturally. Ultimately, as AI researchers, our goal is to create machines that can socially and collaboratively partner with humans. And, as we build these social systems to interact with people, emotions cannot be ignored. In recent years, researchers have been investigating the significance of emotions in human-machine communication. The aim is not only to find ways to adapt AI to human emotional responses, but also to react in a manner that promotes collaboration, leading to more engaging and trustworthy interactions. However, the ethical implications of constructing such artificial social systems need to be addressed. Should machines be designed to detect, adapt, and respond to human emotions? Should they attempt to mimic human emotional responses? In this talk, I will provide an overview of the numerous challenges we face as we strive to create increasingly social AI. I will also explore whether, when, and how emotions should come into play.

Luke Stark

Reordering Emotion: Scenes from the History of Affective Computing and Emotion AI

In fields ranging from health and wellness and advertising and marketing to public safety and security, and political campaigning, data analytics tools united with techniques from the psychological and behavioral sciences are being deployed in real-world application defining, tracking, measuring, and modulating our moods, feelings, and physical affective responses.
Scandals like the 2014 Facebook "emotional contagion" study and the "psychographic profiling" of Trump campaign contractor Cambridge Analytica have brought these technologies into public consciousness. Yet despite their increasing centrality to digitally-mediated life, the historical genealogy of these technologies, and the broader role of psychological models of emotion and personality in computing, remains largely unexplored in the history of computing.
In this talk, I connect the historical treatment of human emotion by computer science to parallel wider debates within psychology, psychology, and cognitive science from late 1940s up to the present, examining how the genealogy of cybernetic technologies for emotion measurement and tracking shaped the slow growth of affective computing as an organized field in the 2000s and its speedy growth today. In doing so, I connect elite debates in the early days of computing around the best way to manage social disorder with today's reliance on digital technologies to manage and modulate individuals and populations using AI systems.

Anna Tuschling

Emotional and Affective States as Quasi-Semantics in the history of interactive Computing

This contribution is focusing on the functionality of the applications provided by the field of affective computing in the recent history of interactive computing. The problems and possible applications of computing that processes, simulates and influences emotions and affects (Picard) have gained more and more attention in recent years. Less understood are the role and the functions of affective computing in the history of computing and interface culture. To close this gap the contribution will make first suggestions with an emphasis on the concept of quasi-semantics.
I will make four arguments and develop them into more detail in regard to two examples, fields of research and experiments: a) Emotion Research in/with Neurocinematics (Tikka et al.) and b) The Embodiment Virtual Reality System (EVR) provided by the interdisciplinary project "The machine to Be Another" (https://beanotherlab.org/):

1. All forms of affective computing have to formulate emotional and affective states (or use concepts that allow the processed data to be described in terms of emotional and affective states). In reference to the classical debate about digital states (Goodman et al.) I will describe what can be called emotional and affective states in this context.
2. The formulation of emotional and affective states is twofold: It reduces the holistic experiences of emotions to discrete entities and can therefore be evaluated as reductionist (Bösel et al.). It can also be described as the key to link individual experiences with sign systems and technical systems (especially digital environments, provided by modern, networked computers).
3. I will point out the importance of a media-historical perspective to fully understand the development and functionality of affective computing in this context and propose the concept of quasi-semantics.

4. In a critical review I will highlight the problematic forms of discourses in affective computing that conceal important gaps (see the idea of "body swaps" in Virtuality) in perspectives and perceptivity.

Eva Weber-Guskar

Emotion Detection Systems and Emotional Granularity

An interesting but less explored ethical criticism of the use of affective computing applications in the literature to date is that widespread use of emotion recognition systems would lead to a loss of emotional granularity. In other words, it is suggested that our ability to recognise, and perhaps even experience, finely differentiated emotions would diminish. In this talk I will explore this criticism.
To do so, I will first elaborate on the criticism by referring to the basic theory of emotions and its limitations. I will then focus on the assumption that the systems themselves lack emotional granularity. This focus will bring into play suggestions for rather simple technical solutions. However, I will then argue that emotional granularity is lacking in a second, more fundamental sense, which eludes technical solutions. Finally, I will discuss possible ways to promote emotional granularity in the context of affective computing, including the interactional approach to emotions (as opposed to the informational approach).

Relevant links:

https://www.dsi.uzh.ch/de.html; https://ethics.dsi.uzh.ch/
https://www.pe.ruhr-uni-bochum.de/philosophie/i/phil-ethik-emotion/index.html.en
http://ifm.rub.de/institut/personen/tuschling/
https://project-interact.jimdosite.com/