As AI get smarter, understand human-computer interactions with the following five premises

Manh-Tung Ho[1,2*] & Quan-Hoang Vuong[1]

1. Centre for Interdisciplinary Social Research, Phenikaa University, Yen Nghia, Hanoi, 100803, Vietnam.

2. Institute of Philosophy, Vietnam Academy of Social Sciences, 59 Lang Ha Str., Hanoi, 100000, Vietnam.

*Correspondence: Manh-Tung Ho (Ph.D.) (tung.homanh@phenikaa-uni.edu.vn)

> "Experienced and quite up-dated on new technologies, Kingfisher says (to the birds about the robot bogeyman): - You guys have met rivals of the 4.0 age. You would sure fail if you didn't study them carefully. You must do thorough research on their behaviors to find a solution."
>
> In *Bogeyman*, The Kingfisher Story Collection, (Vuong, 2022)

The hypergrowth and hyperconnectivity of networks of artificial intelligence (AI) systems and algorithms increasingly cause our interactions with the world, socially and environmentally, more technologically mediated. AI systems start interfering with our choices or making decisions on our behalf: what we see, what we buy, which contents or foods we consume, where we travel to, who we hire, etc. It is imperative to understand the dynamics of human-computer interaction in the age of progressively more competent AI. Below are five premises to serve as starting points for building our understanding the "interactions' in human-computer interactions.

First, is the primacy of the structures. Prior to any individual human and to any AI system, there have always been social, cultural, political, and historical structures that predispose individuals and machines to certain ways of thinking, behaving, and being. An individual is always in the process of being socialized, and he/she/they have to internalize certain socio-cultural norms and values to function as a member of society. Likewise, any algorithms that interact with humans are also trained with datasets that come from the preexisting social worlds. In this fundamental way, as sociologist Massimo Airoldi argues, machine learning systems such as search engines or recommender algorithms are socialized and constantly receive an influx of feedback from millions of users. The networked human-algorithms entanglements, i.e., the machine habitus, Airoldi argues, ultimately reproduce "society, with its arbitrary discourses, invisible boundaries, and structures" (p.112). Airoldi, thus, proposes researchers to systematically follow and investigate: *the machine creators, users, medium, and algorithms* (Airoldi, 2021). The primacy of structures means we need to understand how human values are transmitted, propagated, and reproduced across the machine habitus. This leads to point no.2.

Second, while each human desires differing levels of agency and autonomy, all do have some desire for freedom. Psychologist Paul Bloom explores the philosophy and psychology behind the strange appeal of perverse actions from St. Augustine's The Confessions to contemporary thinkers such as Agnes Callard. Bloom used the term "existential perversity" to refer to "perverse behavior that is deliberate, knowing, and strangely satisfying." He reasons existential perversity comes from the deep desire to show to yourself and others that you are a free, authentic, and autonomous being. This desire, thus, motivates a person to avert from the expected, the reasonable, or even the moral (Bloom, 2019). As humans always seek to assert themselves as autonomous being, even if we have the most moral and superhuman intelligent machines, it is hard to see the future where humans obediently abide to what the machines instruct. This leads to point no.3.

Third, whether human-computer interactions serve the foundational values of Truth-Goodness-Beauty depends on the 'awakening' of the humans. "What does that mean to be enlightened or awakened as a human being?" has been the question that great religious leaders and philosophers investigated since the earliest days of human civilizations. Humans have developed a huge, rich repertoire of philosophy and science of human flourishing. The age of human-computer interactions will give us boundless possibilities, but to choose wisely among these possibilities, humans, individually and collectively, must go back, study, and apply these traditions and practices. This leads to point no.4.

Fourth, from human-computer interactions, there emerge values that, whether being consciously understood and examined or not, will be incorporated into the core values of human society. We might call it *acculturative emergence phenomena*, where values arising from the billions of interactions between humans and algorithms in the hyperweb, once marginal or forgotten, can become mainstream again. David Auerbach gives an example that while big-tech owners such as Mark Zuckerberg or Elon Musk and even each user might personally value democracy, free speech, or inclusiveness, the interactions in the hyperweb and their resulting emergent social dynamics might undermine such values in unexpected ways (Auerbach, 2023). This leads to point No.5.

Fifth, human-computer interactions are already and will change human-to-human and human-nature interactions. This triangle of interactions is intertwined and can influence each other. We believe finding the balance among these three overlapping spheres of interactions would allow for human flourishing and undermining the negative influences of over-interacting with, and over-relying on machines and algorithms. Thus, to comprehensively assess the impacts of human-computer interactions in the age of AI, at both personal and collective levels, researchers need to account for the impacts it has on human-human and human-nature interactions.

Putting together all five premises, we can start mapping human flourishing in a time of ever-increasing human-AI interactions. As a human strives to be an autonomous and authentic self, he or she does so in a world that has norms and values precede his/her existence. In this process, with the growth of AI technologies, this person will interact with networked machines and algorithms that could lead to unexpected emerging norms and values. Frictions and conflicts within this process and how people resolve them largely determine their well-being. As different traditional cultures have different sets of values that put constraints on what are normative human-

to-human and human-to-nature interactions, respecting these values might lead to more normative ways of human-computer interactions. Conversely, constructing new cultures that center around values and norms conducive to human well-being is also an important conclusion of these premises. Finally, the dynamic, complex, and dialectic process of understanding human-AI interactions entails an interdisciplinary exploration of the way the human mind is changing as it interacts with increasingly competent AI systems.

Data availability

There is no data associated with this manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

Airoldi, M. (2021). *Machine habitus: Toward a sociology of algorithms*. John Wiley & Sons.

Auerbach, D. (2023). *Meganets: How Digital Forces Beyond Our Control Commandeer Our Daily Lives and Inner Realities.* Public Affairs.

Bloom, P. (2019). *The Strange Appeal of Perverse Actions*. Retrieved December 26, 2023 from https://www.newyorker.com/culture/annals-of-inquiry/perverse-incentives

Vuong, Q. H. (2022). *The Kingfisher Story Collection.* . https://www.amazon.com/dp/B0BG2NNHY6.