



# Cooperative Human Artificial Intelligence

Marco Ragni<sup>1</sup>

Accepted: 8 April 2021 / Published online: 22 April 2021  
© The Author(s) 2021



Dear readers,

Why are we so fascinated by Artificial Intelligence? Is it the promise of progress towards a better future? Is it because AI has many practical applications and can make life easier? In an *age of information*, we need systems that can process large amounts of data and draw conclusions from it. In *life-threatening environments* such as radioactive water in Fukushima or on Mars, we need systems that can (mostly) independently pursue their tasks. In (human) *error-prone situations*, we need assistance systems that don't get tired. There are many more situations where a cognitive system not unlike our human one is beneficial. Surveys and polls have identified (at least) three ways in which humans consider the future development of AI: *As a tool*—a highly advanced one, but still a tool that does what we cannot or do not want to do. Humans are *using* the AI, and we—as humans—don't expect or don't want a personal AI with rights—it just shall *do* what we want, like a servant or assistance system. *As a counterpart*—sharing more and more similarities with our cognitive abilities and may possibly develop a personality with rights. In fact, systems can already be built that show specific human/animal characteristics, from communication to even demonstrating emotions and body language. To “raise” a digital entity

with learning capabilities is a consequence and was nicely described by the Hugo Award winner Ted Chiang in his story “The Lifecycle of Software Objects”. This may match the hopes of more than a few AI researchers. *As a superior entity*—superior to us as it has fewer cognitive limitations and access to more knowledge and better reasoning capabilities than humans. This idea may frighten people because they fear that such an AI has no “empathy” for humans. To gain control over AI, we wish to “understand” how it works and change it when we disagree with its operating principles, and this, among other reasons, is why we are interested in explainable and responsible AI. This is an important part in a design-cycle that helps develop systems exactly the way we want them to be. But this may not do full justice to the specifics of human and AI strength if the two are just considered as opposites. In 1972, Michie (pp. 332) wrote: “An interesting possibility which arises from the ‘brute force’ capabilities of contemporary chess programs is the introduction of a new brand of ‘consultation chess’ where the partnership is between man and machine. The human player would use the program to do extensive and tricky forward analyses of variations selected by his own intuition. . .”. To approach more complex and an increasing number of challenges in society and science, we need such a co-operative partnership between mankind and AI. We now need to assess what humans and what AI can do better and focus on that, to not waste precious resources. For instance, in situations that require *ethical considerations and empathy*, most humans prefer humans to make decisions. We expect a human to be able to consider the specifics of a case, feel compassion, and to not just apply “general rules”. In *common sense* reasoning humans still outperform AI systems. On the one hand, our *human intuition* (see above quote) is often regarded as typical human, but on the other might be just formed on processing hundreds of similar examples and making hypotheses based on them. There are many more characteristics necessary to consider, but they all keep returning to the philosophical and psychological question of what defines us as humans? More research at the intersection of AI and psychology is needed to identify and compare the potential of humans and artificial systems—to avoid the “sociopsychological diffusion of responsibility”. *We need* to estimate what we have and what the AI systems have the greatest potential for, in order to cooperatively approach the new challenges of tomorrow.

✉ Marco Ragni  
ragni@sdu.dk

<sup>1</sup> Danish Institute of Advanced Studies, South Denmark University, Odense, Denmark

Marco Ragni

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes

were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.