Author(s): Karvonen, Antero; Saariluoma, Pertti

Title: Cognitive Mimetics and Human Digital Twins : Towards Holistic AI Design

Year: 2023

Version: Published version

ercim-news.ercim.eu
Number 132 January 2023

ERCIM NEWS

Special theme:

Cognitive AI & Cobots

Also in this issue

Research and Innovation: The Interactive Classification System

Cognitive Mimetics and Human Digital Twins – Towards Holistic AI Design

by Antero Karvonen (VTT) and Pertti Saariluoma (Jyväskylä University)

AI is replacing and supporting people in many intelligence-requiring tasks. Therefore, it is essential to consider the conceptual grounds of designing future technical artefacts and technologies for practical use. We are developing two new practical design tools: cognitive mimetics and human digital twins for AI designers. Cognitive mimetics analyses human information processing to be mimicked by intelligent technologies. Human digital twins provide a tool for modelling what people do based on the results of cognitive mimetics. Together they provide a new way of designing intelligent technology in individual tasks and industrial contexts.

Today, the designers of industrial processes and technologies must find answers to novel problems of designing intelligent technical systems. The issues emerge at the intersection of major trends. First, AI is predicted to gradually enter further into aspects of industrial processes, in particular replacing and augmenting the work of operators. Second, more industrial work and design is moving into digital formats such as digital twins, augmented reality and the metaverse. Third, Industry 5.0 calls for a human-centric approach to industrial design.

A question facing a designer is how the human aspect should be included in this large-scale digitalisation of the industrial processes (in particular in the cluster of issues arising from AI) and even the design process itself. Human digital twins (HDTs) provide a conceptual design tool that can address some these questions, presenting a concretisation of a methodology called cognitive mimetics. HDTs and cognitive mimetics unify into single design scientific paradigm the basic questions of human work, action, and thought with intelligent technology – and their joint development [1].

The cognitively oriented analysis of human-technology interaction, embedded in cognitive mimetics, provides a set of questions and methods for the gradual development of human digital twins. Thought and action in relation to technical systems can be simultaneously analysed and modelled. As more aspects of the human information processing in a context are modelled in the HDT, implemented using a variety of extant AI or cognitive science, the closer it approaches intelligent technology. The value of such a process is potentially great, since it allows for the mutual development of human operations and technology as a joint cognitive system [2].

Cognitive mimetics is an empirical design scientific method. Like biomimetics mimics solutions of nature, cognitive mimetics constructs technologies by mimicking human intelligent information processing. Mapping between human information processing and computer models need not (and perhaps cannot) be slavishly similar, but similar to the way airplane wings follow the same laws of aerodynamics as bird wings. Intelligent computational processes can provide performance capacities that were previously only possible in people on the grounds of design knowledge collected by means of cognitive mimetics.

Cognitive mimetics does not deal purely with the mapping of abstract theories or models of AI in cognitive science. These are possible implementation methods for cognitive mimetics, which is focused methodologically on the analysis of empirical human information processes and contents. The main idea is to research and “open up” the thought and perception of, for example, human operators in a paper mill, to gain new ideas for the design of intelligent technology. The basic form of this activity is a mapping relation between a source and a target, here a human operator and an intelligent technology target.

Below are the main phases of cognitive mimetic design:
- Phase 1: Understand and define the intelligent technology design problem.
- Phase 2: Collect information about human information processes being able to carry out the task to be designed.
- Phase 3: Generate concepts for intelligent technology.
- Phase 4: Generate a solution structure.
- Phase 5: Finalise and deliver.

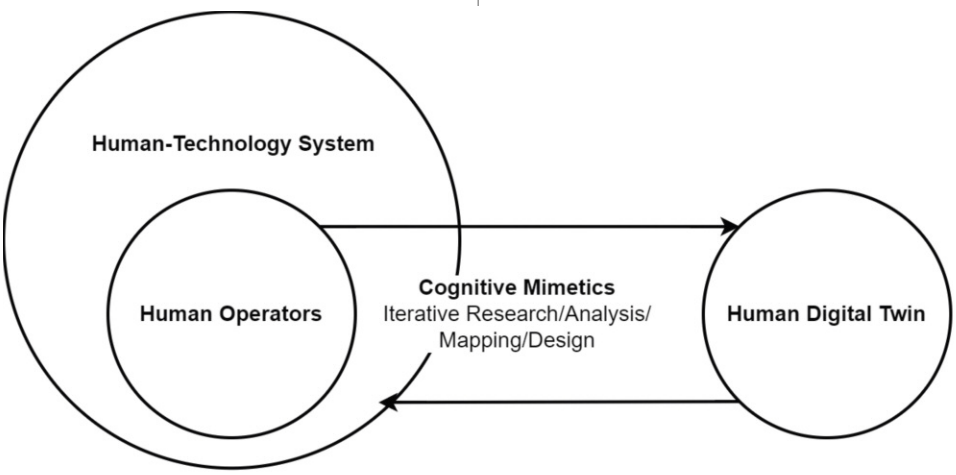

Figure 1: The basic idea of cognitive mimetics is iterative research and design, cycling between a source and a target. Here the source is specified as human operators (in an industrial context) and the target as human digital twins. Design research into human operators thinking progressively opens up the human-technology system as a whole. Ultimately the system is changed via the results of cognitive mimetic design, such as human digital twins.

There are naturally iterative loops between the phases. Each phase is completed through an empirical investigation of human action and thought using well-established methods such as observation, interviews, focus groups, think aloud and protocol analysis. Value can be extracted during all phases of the design process. The main point is to start from a general understanding of the context of human action and thought, advance through a concept-generation phase of how the joint cognitive system could be improved upon, and then focus in on a particular concept. At this point, cognitive and expertise research methods are employed to develop a rich understanding of the actual thought processes of operators. These are then mapped computationally to a human digital twin.

Human digital twins are computational models of people processing information (or perceiving, attending, and thinking) to carry out the critical tasks under design. Like digital twins of cyber-physical systems, human digital twins can be used to support design processes. They can be used to create alternative solutions, test the validity of solutions, and refine them.

The picture that emerges now addresses potential for intelligent technology, development of the human tasks and activities, places where human and machine intelligence can support each other, and the gradual emergence of intelligent and useful technologies. Using human digital twins makes the process digital and documented.

Human action can't be analysed separately from a context (and a space of possible action). Thus, human digital twins (constructed via cognitive mimetics) reflect and refer to the technical systems and processes in which industrial operators work. It is a short, logical step then to attach human digital twins to traditional digital twins and metaverse-type technologies. These yield concrete dataflow and control pipelines, as well as (machine and human) learning systems, for bringing together digitally the human and technical sides of industry. Achieving this basic vision is clearly bubbling under many researchers' thinking and visions of future industry. Arriving at this point via the cognitive mimetic route yields a deep and rich picture of human-technology co-agency in industrial processes.

Industry is just one example. AI will gradually change the way people live and work and how society operates. Therefore, it makes sense to develop AI design as a holistic process in which social and technical aspects of intelligent technologies can be simultaneously considered. Such new design practice can be called holistic AI design. Cognitive mimetics and HDTs provide key elements in such AI design.

References:

[1] P. Saariluoma, A. Karvonen, L. Sorsamäki, "Human Digital Twins in Acquiring Information About Human Mental Processes for Cognitive Mimetics", in Frontiers in Artificial Intelligence and Applications, 2022. https://doi.org/10.3233/faia210484.
[2] E. Hollnagel, D. A. Woods, "Joint cognitive systems: Foundations of cognitive systems engineering", 2005.

Please contact:

Antero Karvonen, VTT, Finland
antero.karvonen@vtt.fi

Understanding Partners' Behaviour for Transparent Collaboration

by Maria Dagioglou (NCSR 'Demokritos', Greece) and Vangelis Karkaletsis (NCSR 'Demokritos', Greece)

For this teaser we were asked to describe what is the article about, who is doing what and why. Fittingly enough, this is literally what we discuss in this article: in human-robot collaboration, who is doing what and why?

From carrying a heavy object to supporting a rehabilitation exercise, humans coordinate their bodies and minds to move together and achieve mutually agreed goals. During collaborative tasks, each partner needs to: know what the other perceives (or not), be able to predict his/her actions based on action observation and the requirements of the shared task, decide when and how to act for supporting team performance and efficiency [1]. In addition to all these, the spatio-temporal proximity of actions during collaboration affects the sense of self- and joint-agency. Who bears the authorship of a joint action's outcome? The experience of acting as a team, the nature of the task (competitive or complementary) and the fairness of resource distribution are only some of the factors that shape the perceptual distinctiveness of each partner's actions.

Now, what about human-robot collaboration (HRC)? What if a human were to collaborate with a robot to carry a sofa up the stairs (Figure 1)? In the context of this article, HRC is considered as the case of a human and a collaborative robot (cobot) that work in close proximity (within the intimate space of a human, which, based on Proxemics [L1], includes interactions from physical contact up to approximately 0.5 metres) towards a mutual goal that demands interdependent tasks. Similar to the processes described above for human-human collaboration (HHC), cobots must have the perceptive, cognitive and action capabilities that support joint attention, action observation, shared task representations and spatio-temporal action coordination.

In terms of the sense of agency, the (few) existing studies provide mixed evidence with respect to how this mechanism is manifested during human- (embodied) agent joint action and about the extent to which it follows similar patterns compared to HHC [2]. However, this is a crucial issue in HRC related to the delivery of ethical and trustworthy Artificial Intelligence (AI). Transparency in the authorship of each partner's actions is related to the ethical dimensions [L2] of: a) human agency, including over- or under-relying on the cobot's actions, as well as the social interaction instigated; b) accountability, in relation to the ascription of responsibilities if need be; and c) communicating and explaining the decisions of an AI system.

Thus, it becomes clear that HRC must be observed "in action", that is in real-world and in real-time, in order to not only evaluate the performance of AI methods, but also to study human behaviour towards cobots, and AI agents in general. Luckily, state-of-the-art AI methods now allow us to do so [3]. Towards this, in Roboskel, the robotics activity of SKEL-the AI Lab [L3], we have integrated an HRC testbed [L4] and have initi-