

Summary

During the workshop, eight topics were suggested and discussed, some of them directly related to each other, like *cognitive offload* and *shared workload*, while *implementation methods for industry* could be seen as rather orthogonal to most of the other topics.

This document is an attempt to extract and summarise the main points of the discussion, that would lead into central research questions and challenges to be integrated into a form of roadmap for **Human AI Teaming** research in Sweden.

Trust, reliability, failures

The high level question *How can we deal with silent failures?* was marked the most as one that should be taken into consideration. Together with topics in the area of *Trust and Transparency*, which also received quite many importance markers, a line of research in Human AI Teaming could thus be (should be?) the establishment of reliable, transparent and trustworthy mechanisms for teamwork between humans and AI systems. In both areas, sub topics touched upon the problem of “understanding what went wrong and why”, hence, the above mentioned mechanisms would have to rely on the ability of AI systems to perform introspection and to handle the possibility of a situation offering several ways of interpretation. At the same time, humans teaming up with AI systems must have the opportunity to observe and interpret the actions of these AI systems, which is where transparency comes into play.

From:

<https://hait.cs.lth.se/> - **Human AI Teaming**

Permanent link:

https://hait.cs.lth.se/topics:roadmap_stubb?rev=1671800754

Last update: **2022-12-23 13:05**

