Trust and awareness in the context of search and rescue missions

Dr. phil. Martina Philippi

martina.philippi@ruhr-uni-bochum.de

In the BMBF-Project "UAV-Rescue" on UAV-borne remote sensing for AI-assisted support of search and rescue (SAR) missions, an unmanned aerial vehicle (UAV) is developed that is equipped with two instances of AI: one for (semi-)autonomous indoor navigation and another one for the distinguishing of vital signs from another kinds of movements (like blowing curtains) that are both detected by the implemented radar sensor. With these functions, the UAV shall be able to navigate through semi-collapsed buildings, localize survivors and provide a map from the scenario based on lidar data.

There are several reasons why the information given by this system is not fully reliable in an epistemic sense, some coming from the nature of AI that deals with statistics, some coming from sensors and data processing working different in different environments. While some of these errors become visible by non-functional behavior of the UAV – e. g. navigation problems in certain environments with unusual characteristics since the AI for navigation is widely trained by simulations of indoor scenarios; or a lack of data resulting from a sensor breakdown –, many of them occur in an inconspicuous way, especially those related to the output of information. Inconspicuous means here that the result is a map that looks complete, precise, and plausible – but nevertheless does not correctly depict reality.

Whereas such a technology inevitably lacks strong epistemic reliability, it can nevertheless be regarded as trustworthy - on condition that it is embedded in a circumspect practice. In opposite to the usage of AI-supported technology in everyday life where we often weight comfort and celerity over security, the context of SAR missions is characterized by a high awareness, not only towards physical risks, but also concerning the capability of information sources, whether search dogs or technical tools like bio radar and optic or acoustic devices. So, the focus lies on an epistemic issue, namely the gathering of information and the assessment of the quality of this information. In the consequence, the SAR context is a special case where on the one hand a great interest in reliable information is given and on the other hand certain practices help to deal with insecure information – e. g. further context information, cross-checking by other search methods, and knowledge about possible sources of error, based on practical experience and a basic technical understanding.

So, trust is here not replaced by habituation, and this changes the meaning of trust in the usage of the technology: In an everyday application like a navigation system, or in a decision support system that classifies people on the basis of algorithms, such a lack of epistemic reliability might be more problematic that in this special case. In those contexts, information is the basis of an activity to reach an aim, not the aim itself. But in the SAR context, there is an explicit space for awareness; the whole practice is characterized by a highly aware handling of information because the aim of the usage of such an AI-supported technology. This sheds special light on the issue of trustworthiness of technology: The technology is trustworthy *because* it is not trusted blindly but applied with an awareness that is already fundamentally embedded in the usage of search methods. Especially the problem of inconspicuous errors is not as critical as in the other contexts mentioned because there is no need of irritation to verify information – they are cross-checked by default.