Artificial Intelligence and Autonomous Systems: 10 open questions

Artificial Intelligence (AI) and Autonomous Systems are inextricably linked in many areas of industry, logistics and transport. Alone and in combination, they hold great economic potential, but also entail risks. The working group "Autonomous Systems" of the VDI-/VDE-Society Measurement & Automatic Control (VDI-/VDE-GMA) has worked out 10 questions, which we have to answer in order to make AI economically successful.

"Autonomous Systems can, if used correctly, take on difficult tasks and bring benefits. But the use of Autonomous Systems also raises questions," says Prof. Alexander Fay, member of the board of the VDI-/VDE-GMA. "These questions should be clarified before an autonomous system is started", adds Prof. Birgit Vogel-Heuser, who together with Prof. Fay heads the working group "Autonomous Systems" of the VDI-/VDE-GMA.

The following 10 questions bundle essential challenges that Autonomous Systems have to overcome:

www.VDI.de/themen/kuenstliche-intelligenz-ki

Question 1: How can we control the autonomous system?

This question is important, because many people fear that Autonomous Systems could elude their control. The user must therefore be able to deactivate the autonomous system at any time according to his wishes or reduce the intensity of the intervention by the autonomous system. This leads to the next question:

Question 2: How autonomous should the autonomous system be for us?

The ideas about the degree of independence of the technical system can be very different. For example, various "degrees of autonomy" are defined for automotive and industrial applications - from assistance and automation to comprehensive autonomy. The levels, such as the "5 steps of automated driving" or the "6 steps to the autonomous factory" help to clarify what the expectations are or what a technical solution offers.

Question 3: How do we make the Autonomous System autonomous?

The "how" is important because there are very different concepts: from behavior designed for all conceivable situations to systems that "train" themselves. Depending on the concept used, you can make different statements about the behavior of the autonomous system.

Question 4: How comprehensible should the behavior of an Autonomous System be?

Traceability strengthens trust in Autonomous Systems and thus their acceptance. Decisions and actions of Autonomous Systems must be understandable and comprehensible for the user of the system. If the autonomous system has developed the decision based on its own learning, then we have to create appropriate additional means to explain and justify the decision. Ideally, we integrate such resources into the autonomous system from the outset. The autonomous system should also communicate on its own initiative if it is uncertain as to the quality of its results.

Question 5: How can we compare Autonomous Systems?

A benchmark - also called "metrics" in metrology - can help in the sound selection of a system and allow an evaluation of the efficiency, reliability and safety of Autonomous Systems.

Question 6: How reliable is the learning autonomous system?

An autonomous system is only accepted and used if it is reliable. When Autonomous Systems learn, they sometimes make decisions that differ from previous decisions in comparable situations. People may perceive this as a lack of reliability if the changed behavior of the autonomous system is not explained.

Question 7: How efficient is the Autonomous System?

On the one hand, Autonomous Systems can take over tasks from humans and bring easement and relief. This can also mean concrete savings in raw materials, energy and money. On the other hand, it is also necessary to consider the effort required for the development and the "training" of Autonomous Systems.

Question 8: How secure is the Autonomous System?

Many people fear that an autonomous system could make wrong choices that put them at risk. It is therefore necessary that mechanisms ensure that even an autonomous system does not violate specified requirements, for example with regard to safety for humans and the environment. This also means that the system cannot be manipulated. Trust in an autonomous system can only exist if its statements cannot be influenced unintentionally, for example by the choice of the training data or the change of the input data. Ideally, the autonomous system detects a manipulation and signals it.

Question 9: Where are the limits of the autonomous system?

How should the autonomous system deal with situations that differ from the standard case for which it is designed and the rules with which it was designed? Should it then pass back the responsibility to the human being? May and can it do so? It is possible that humans have already become accustomed to the autonomous system and have forgotten how to cope with this task themselves. This question is important, since some Autonomous Systems return the action and responsibility to the human being even in the short term, with which he may be overtaxed. Concepts are necessary that react robustly and safely to changing operating conditions.

Question 10: Which values does the autonomous system follow?

This question is particularly important because we expect a technical system to be designed by its developers to comply with applicable laws, standards and regulations. However, the autonomy of an autonomous system allows it to behave differently from its developers. Which rules, which values should it then follow in its behavior? Moreover, how can we ensure that it actually follows these rules and values?

The working group "Autonomous Systems" would like to enter the discussion with the questions raised and develop conclusive and well-founded answers together with all those interested in the topic.

Write to us!

Published on 15 May 2019

Text by:

- Prof. Dr.-Ing. Alexander Fay, Helmut Schmidt University Hamburg
- Prof. Dr.-Ing. Birgit Vogel-Heuser, Technical University of Munich
- Dr. Harald Bosse, PTB
- Dr. Sascha Eichstädt, PTB
- Prof. Dr.-Ing. Klaus-Dieter Sommer, Ilmenau University of Technology
- Other contributors: Board and Advisory Board of the VDI-/VDE-GMA