FRAUNHOFER INSTITUT FOR INTELLIGENT ANALYSIS AND INFORMATION SYSTEMS IAIS

PRESSINFORMATION

Fraunhofer IAIS publishes new study: AI management systems promote trustworthy Artificial Intelligence

Data protection, risk or compliance management: New AI technologies present companies and developers with new challenges. To enable them to address these systematically and in a structured manner, various institutions are working on guidelines and standards for the management of Artificial Intelligence (AI). The Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS has now published a study entitled "Management System Support for Trustworthy Artificial Intelligence", which compares the draft of the International Organization for Standardization's (ISO) standard for AI management systems and current guidelines. The study, commissioned by Microsoft, shows to what extent AI management systems can support companies in the trustworthy use of AI systems and simultaneously strengthen trust in AI applications.

As a key technology of the future, Artificial Intelligence holds enormous innovation potential for business and society. Particularly powerful AI systems, which are also expected to take on important tasks in the future, e.g., in autonomous driving, are based on the processing of large volumes of data. In order to manage the associated risks, ensure safe use and enable international compatibility, regulatory guidelines and international standards are needed to guide companies and other organizations in the use and development of new AI technologies. Standardized management systems are a common tool in many corporate sectors for successfully dealing with sensitive aspects such as information security. In the context of AI technologies, such management systems are still under development. Currently, an international standard for AI management systems (AIMS) is being developed by the joint working group of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) and currently exists as a draft.

In the newly published study "Management System Support for Trustworthy Artificial Intelligence", Fraunhofer IAIS has examined this AIMS draft to determine the extent to which it can support companies and other organizations in using and developing AI technologies in a trustworthy manner. For this purpose, the AI researchers compared the draft with the current requirements and recommendations for trustworthy Artificial Intelligence that have been formulated so far by the European Commission, its appointed High-Level Expert Group on AI (HLEG), and the German Federal Office for Information Security (BSI).

The Fraunhofer IAIS study shows that the introduction of an AI management system can be an important and appropriate step for companies in the future to define suitable strategies and processes for the trustworthy development and use of AI technologies. "For organizations using AI, the goal of being responsible, trustworthy, and legally compliant should be clearly expressed in their governance, risk, and compliance strategy," recommends Dr. Michael Mock, co-author of the study, who leads the project "KI-Absicherung" for safe AI in autonomous driving at Fraunhofer IAIS, funded by the German Federal Ministry for Economic Affairs and Energy. In this context, AI management systems can also support companies and developers in complying with current and upcoming guidelines and laws in the long term. "Even in the presence of multiple stakeholders and complex supply chains, the use of AI management systems facilitates legal compliance throughout the lifecycle of AI systems," says Mock.

In their study, the researchers emphasize that the implementation of AI management systems in companies can also have a positive influence on the acceptance and trust of AI technologies in society. "In our estimation, AI management systems are an important building block for significantly strengthening the trust of stakeholders – such as customers or employees – in AI applications", says. Dr. Maximilian Poretschkin, co-author of the study and team leader "Trustworthy AI" at Fraunhofer IAIS.

As one of the leading research institutes in the field of Artificial Intelligence in Europe, Fraunhofer IAIS is also working on the trustworthiness and reliability of AI systems and leads, for instance, the project "ZERTIFIZIERTE KI" ("Certified AI") that develops testing criteria and assessment tools for AI-systems. This expertise enables the scientists to evaluate the current draft of the standard on AI management systems. "With the draft regulation of the EU Commission, but also with other documents such as the recommendations of the HLEG, important guard rails for the use of AI applications are emerging. It is therefore very important to compare these with upcoming standards," says Dr. Maximilian Poretschkin. In addition to the evaluation, the AI experts provide recommendations how the draft standard can still be improved.

The 60+ page study first summarizes the draft of the standard as well as the AI guidelines of the European Commission (Proposal for AI Regulation), the HLEG (Assessment List for Trustworthy AI) and the BSI (AIC4 catalog). In the detailed comparison of all documents, both organizational and technical requirements are compared. Here, the structure of the comparison is based on the AI audit catalog recently published by Fraunhofer IAIS, a guideline for the design of trustworthy Artificial Intelligence. In addition, the study addresses a possible certification of AI management systems and concludes with a summary of the findings as well as recommendations for further elaboration.

The study "Management System Support for Trustworthy Artificial Intelligence" was funded by Microsoft.

Further information:

www.iais.fraunhofer.de/ai-management-study

About Fraunhofer IAIS

As part of the largest organization for application-oriented research in Europe, the Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS based in Sankt Augustin near Bonn is one of the leading scientific institutes in the fields of Artificial Intelligence, Machine Learning and Big Data in Germany and Europe. With its more than 300 employees, the institute supports companies in the optimization of products, services, processes and structures as well as in the development of new digital business models. Fraunhofer IAIS thus shapes the digital transformation of our working and living environment.

Focus topics include trustworthiness and reliability in the context of AI systems: Thus, the institute holds the consortium leadership of the project "ZERTIFIZIERTE KI" ("Certified AI") as well as the deputy consortium leadership and scientific project coordination of the project "KI Absicherung". In addition, the institute's director Prof. Dr. Stefan Wrobel is a member of the high-level coordination group on AI standardization and conformity founded by the German government. In the creation of the "German Standardization Roadmap Artificial Intelligence", recently published by the standardization organizations Deutsches Normungsinstitut (DIN) and Deutsche Kommission Elektrotechnik (DKE), Fraunhofer IAIS led the working group "Quality, conformity assessment and certification".

Press Contact

Fraunhofer Institute for Intelligent Analysis
und Information Systems IAIS
Schloss Birlinghoven 1
53757 Sankt Augustin

Silke Loh, Press and Public Relations
pr@iais.fraunhofer.de
Phone + 49 2241 14-2829

Scientific Contacts:

PD Dr. Michael Mock
michael.mock@iais.fraunhofer.de
Phone +49 2241 14-2576

Dr. Maximilian Poretschkin
maximilian.poretschkin@iais.fraunhofer.de
Phone +49 2241 14-2260