Lernende Systeme

GERMANY'S PLATFORM FOR ARTIFICIAL INTELLIGENCE

Secure and safe AI Systems in Medicine

White Paper by Jörn Müller-Quade et al.
Working Group IT Security, Privacy, Legal and Ethical Framework

Executive Summary

The use of Artificial Intelligence (AI) promises great improvements in medicine. In the future, self-learning systems can lead to better treatment results in prevention, early diagnosis and patient-oriented therapy, thus improving our health care. The use of AI systems can also help doctors and medical caregivers to improve patient care and reduce the workload on medical staff. The fictitious application scenario „With Artificial Intelligence against Cancer" developed by the working group Health Care, Medical Technology, Care of Plattform Lernende Systeme, provides concrete examples.

At the same time, the use of intelligent and self-learning systems in healthcare demands high requirements on the data management and IT security and IT safety of the systems. Possible risks when using self-learning systems in the health care sector include incorrect or deliberately falsified training data, attacks on AI software or incorrect integration into clinical practice. Along the mentioned fictitious application scenario, members of the working group IT Security, Privacy, Legal and Ethical Framework and members of the working group Health Care, Medical Technology of Plattform Lernende Systeme, have identified such IT security requirements that are necessary for the use of AI systems in medicine.

Requirements for the use of AI in medicine

- Ensuring original, unbiased training data
- Protecting AI software from attacks
- Pooling of training data while respecting privacy
- Secure AI databases
- Providing patient data securely
- Securely integrating AI systems into the clinical process

The experts focus on data management and security and safety aspects – and thus primarily on technical issues. This technical analysis is a fundamental step towards being able to discuss and answer regulatory questions. AI and Machine Learning (ML) also raise socially relevant questions that cannot be answered purely technically. This is particularly true in the health care sector, where data is particularly sensitive and potential risks are serious.

The white paper identifies technical and organisational conditions that are necessary for a medium-term realisation of the fictitious application scenario „With Artificial Intelligence against Cancer" developed by Plattform Lernende Systeme. A technical-organisational analysis is to lay the foundation for a follow-up discussion. Since there are several contingencies regarding the future design of our digitalised health care system that reach beyond the scope of this paper, the following aspects were not included in the analysis, but are yet to be answered in future discussions:

- Access authorization for third parties to the entries of the electronic health record (EHR)
- Voluntary and protected data sharing
- Need to review the legal situation

The aspects addressed are still legally unresolved. Based on the analysis of the application scenario, the experts formulate legal design requirements and possible design options. The focus is thereby on the question of quality assurance of the data used for the training of AI systems, the traceability and explainability of AI systems and their security in terms of safety and IT security.

Legal-regulatory requirements and possible design options

- **Develop common guidelines and test specifications for approval and certification:** A side effect of increasingly dynamic software architectures is the problem that the function and mode of operation of a medical device is less measurable, less verifiable and potentially less certifiable. This also applies to self-learning systems. It is unclear whether a learning medical device should be regarded as a new product with every minor software update. Nevertheless, the approval process should be further developed. Besides the product itself, it is also necessary to consider its operation and certification requirements for updates.
- **Develop common guidelines and test specifications for the accreditation and certification of AI database operators:** Together with the relevant stakeholders, the legislator should also develop guidelines, test specifications and requirements for an accreditation and certification process for certified AI database operators.
- **Obligate manufacturers by law to remedy new types of defects:** New, possibly stricter safety requirements for the applications arise, which must be fulfilled in the context of the approval. For the operation of an AI system, this is regulated and firmly defined in European legislation. Certain product properties can be tested and evaluated before market launch. In addition, malfunctions should also be observed downstream and remedied by the manufacturers in the sense of remedying defects – regardless of whether they are only due to AI functionalities or other system adaptations.

- **Appoint independent authorized operators of the AI assistance system:** These state-appointed neutral institutions should be commissioned to manage and maintain the analysis procedures and data records. This institution must not be authorised to modify or feed in data, as it may have its own economic interest.
- **Set up an independent audit committee:** An interdisciplinary committee of experts should review the functioning of the certified and deployed AI systems at regular intervals. It would make sense to set up this committee at the Federal Institute for Drugs and Medical Devices. In addition, recall processes should be established at the manufacturers in order to be able to act if a system fails.
- **Health insurance companies should keep blocking lists:** As the releasing bodies of the electronic health cards and the health professional ID cards, health insurance companies should keep blocking lists to prevent unauthorised access to data. These lists must be updated continuously so that in the event of loss, the respective authorization card is worthless. The blocking emergency call 116 116 could be extended to include the electronic health cards and the health professional ID cards. Consideration should also be given to a corresponding voluntary commitment on the part of the health insurance companies to participate in the system.
- **Introduce relapse solution:** A relapse solution could complement the blocking of the electronic health card. It is a mode in which the range of functions is limited, but the most important functions of a system can be maintained.
- **Formulate minimum security requirements for data infrastructures and data centres:** The IT infrastructures required for the implementation of AI in healthcare are already subject to the scope of current legislation. However, this does not exclude the possibility that the legislator may adapt the relevant legal regulations by defining minimum requirements. These should specify that the data may only be stored and processed within the European Union. As a first step, it is important that the AI systems used, and the infrastructures associated with them are also systematically covered by the Regulation on Critical Infrastructure. In a second step, corresponding security requirements for the establishment of the necessary AI systems must also be defined.
- **Introduce a research-compatible electronic patient record:** A research-compatible electronic patient record is needed so that patients can make their data sets available to research after treatment and AI methods can be further developed. This means that the relevant data should be findable, accessible, interoperable and reusable.
- **Developing the electronic patient record into an extended electronic patient record:** Particularly in the field of preventive medicine, more patient data is needed in order to be able to assign patients to statistically reliable possible risk groups.
- **Continue research into IT security and safety issues:** Not all of these described IT security and safety problems that might occur when using AI systems in the healthcare sector can be answered with the technical solutions currently available. Therefore, science is called upon to investigate these problems and develop solutions that are as reliable as possible. Appropriate programmes should be set up for this purpose and the corresponding research funding should be made available.

These possible design options are linked to relevant societal issues, such as the benefits and potential risks of using AI systems in the health care system and the use of anonymised or pseudonymised data. These must be discussed and answered in a broad societal discourse.

Societal relevant questions on the use of AI in medicine

- **Operating, maintaining and caring for the data infrastructure:** Patients can make wise decisions about their data with the help of their electronic health cards. The connected electronic patient record forms an interface between patients, treating physicians and the AI systems. Finally, it is not clear where and how data is (temporarily) stored, transferred and expanded. This concerns both the electronic patient data itself and its meta-data, which the AI software has processed. Distributed cloud infrastructures could be a solution approach, as these are already largely covered by existing regulations. It is unsolved who provides and maintains the necessary infrastructure.
- **Provide and support the AI assistance system:** Questions concerning the operative implementation also need to be clarified. An example is the question, which institutions finance, maintain and continuously train the AI systems and can provide the latest AI software at the request of physicians. These institutions must be independent and may not have their own feed-in or change options.
- **Weigh up the benefits and risks:** Like many other medical methods of diagnosis and therapy, AI assistance systems carry certain risks. In diagnostics, false-positive and false-negative results can lead to incorrect treatment and severe physical, psychological and financial stress. Such risks cannot be completely excluded. With AI, new risk-benefit considerations could become necessary. For example, big-data analyses could help to detect diseases earlier and more frequently. However, this could also be associated with the risk of an increase in false-positive findings. Therefore, a societal discourse should address the question under which circumstances and up to what level are we as a society willing to accept „error rates“ if, on the other hand, high medical benefits can be gained?
- **Use of data:** How patients should and can authorise access to their data and their further anonymised or pseudonymised use (e.g. for research projects) requires further design and specification. It therefore needs to be clarified which data they can share and how narrowly the purpose of voluntary and protected data sharing in exploratory research should be interpreted.
- **Responsibility and liability:** Individuals must remain the final decision-makers both for the course of treatment and for the handling of their data. However, even then, incorrectly processed information could possibly cause serious treatment errors, for example during a medical surgery. It is necessary to discuss who is responsible for errors and whether the use of AI systems should be insurable in terms of liability. This raises the question of how responsibility and liability should be shared between the provider and operator of the AI system and medical staff.

- **Transparency of the results, traceability versus explainability:** The more complex an AI procedure is, the less transparent are the calculation steps used to obtain the results. This bears the risk that users may misinterpret correct results. It is also conceivable that they use distorted or manipulated results unnoticed. For correct treatment, it is therefore important to know why a result was given. As desirable as maximum traceability may seem on the one hand, it could lead to an information overload on the other hand. This area of tension must be balanced in the social discourse. This leads to the question of how much right to information about the calculation of an AI system doctors and patients must have. Furthermore, it should be clarified which rules the legislator should create for the traceability and explainability of AI-based medical devices.

Imprint
Editor: Lernende Systeme – Germany's Platform for Artificial Intelligence | Managing Office | c/o acatech | Karolinenplatz 4 | D-80333 München | kontakt@plattform-lernende-systeme.de | www.plattform-lernende-systeme.de | Follow us on Twitter: @LernendeSysteme | Status: April 2020 | Image credit: Tom Werner/gettyimages

This executive summary is based on the white paper Secure and safe AI systems for medicine – Data management and IT security in the cancer treatment of the future, Munich, 2020. The authors are members of the working group IT Security, Privacy, Legal and Ethical Framework and the working group Health Care, Medical Technology, Care of Plattform Lernende Systeme. The original version of this publication is available at: https://www.plattform-lernende-systeme.de/publikationen.html

SPONSORED BY THE

Federal Ministry of Education and Research

acatech
NATIONAL ACADEMY OF SCIENCE AND ENGINEERING