Translation of the VO_KI_Richtlinie_UKER

AI Governance and Ethics Policy

The text below is a translation of the original German document and is provided for the convenience of the reader. For the binding text, please consult the original at https://www.mik.uk-erlangen.de/aktuelles/nachrichten/detail/regelung-zur-nutzung-von-ki-anwendungen-am-uker/.

Objective

This guideline defines binding requirements for the safe and responsible usage of Artificial Intelligence (AI) systems at UKER.

Scope

This policy is binding for all employees who use AI systems at UKER [05.07.2024].

Notes, terms and information

Preamble

Artificial Intelligence (AI) applications are increasingly becoming part of everyday clinical practice and will significantly change medical care and research in the coming years. Against the background of the dynamic development of AI and the national and international requirements that are still evolving, this guideline provides binding instructions for the introduction and use of such systems at UKER.

Definition of AI

According to EU law (Art. 3 (1) AI Act), an "'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments […]" [5a].

Content

1. **General restrictions on AI usage at the UKER**

   For licensing, copyright, and data protection reasons, AI systems at UKER must be used solely for business purposes. The use of personal accounts is prohibited in the UKER IT environment, for personal or business purposes.

   The use of AI systems to monitor and assess the personality, work performance, physical and mental resilience, cognitive or emotional ability, and to make predictions about persons employed or groups of persons employed by UKER is prohibited without the involvement and approval of the Staff Council.

   For research and development activities on AI systems and their results, which are developed and put into operation for the sole purpose of scientific research and development, the consideration of the individual case applies in principle.
   With the entry into force of the European Union Regulation on Artificial Intelligence (AI Act), there are additional exclusions on the use of AI systems at UKER, in accordance with Art. 5 of the AI Act, taking into account clinical care and the special concerns regarding research and development activities at UKER.

AI systems that threaten fundamental human rights under EU law are generally prohibited at UKER. This applies to the use of AI systems with the following characteristics:

- cognitive behavior manipulation of individuals or vulnerable groups outside of recognized therapeutic purposes;
- social scoring: classification of people on the basis of behavior, socio-economic status, and personal characteristics;
- biometric categorization of natural persons.

The usage of AI systems for illegal activities, the circumvention of disallowed purposes, the violation of label obligations, or the unlawful dissemination of AI-generated results is not permitted.

2. Classification of AI systems at the UKER

AI systems at UKER are differentiated according to application context as follows:

a. **General purpose**
b. **Clinical care (medical devices)**
c. **Research projects**

General-purpose AI systems can be distinguished into:

a.1.1 **Generative AI systems and**
a.1.2 **AI systems to optimize logistical/administrative tasks.**

Furthermore, a second differentiation can be made:

a.2.1 **Freely available and free of charge AI systems (e.g., web applications on the internet, regardless of whether they require registration or not)** and
a.2.2 **Fee-based AI systems that are made available for business purposes.**

3. AI systems

3.a General purpose AI systems

3.a.1.1 **General purpose AI systems: Generative AI systems**	3.a.1.2 **General purpose AI systems: AI systems for optimizing logistical/administrative tasks**
a. Definition of **Generative AI**: Generative artificial intelligence is a	a. Definition: This type of AI systems purpose is to

Type of AI that can create or generate different types of content (e.g., text, images, videos, or audio) based on prompts.	make the work at the UKER easier or more cost-efficient by optimizing logistical processes or supporting other administrative work without using generative AI methods. The use of such AI systems has no direct effect on the patient care process (prevention, diagnostics, therapy, aftercare), so they are not certified medical devices.
b. Examples: DeepL, Chat-GPT, DALL-E2, Microsoft Co- Pilot, Google Gemini, and others.	b. Examples: AI-based optimization of warehousing and/or material procurement; AI-based transport of goods in hospitals.
c. Characteristics: In contrast to AI systems as medical devices or AI systems integrated in certified and approved medical devices (No. 3.b), the manufacturers of generative AI systems do not yet assume the responsibility for AI generated results. Large language models in translation programs do not guarantee the correctness of the translation, text generators do not guarantee the originality, and, in particular, the correctness of the generated results. At best, they provide assistance for the intended purpose of use. Users should be aware that these AI systems have been trained on a variety of data, some of which is protected by copyright (e.g., published research data, other authors' works, and texts), and manufacturers assume no copyright liability when using their software. AI results can also be "wrong" to a certain extent (so-called AI hallucinations) or have AI-specific structures, such as a characteristic sequence of letters in several words as a "digital watermark", which makes the use of AI to create texts, etc., recognizable.	c. Characteristics: As with Generative AI (No. 3.a.1.1, see above), the manufacturers of AI systems for optimizing logistical/administrative purposes do not guarantee the results generated. In contrast to generative AI, these systems from the domains of machine learning and operations research are likely to produce reliable results within a defined margin of error.

d. Consequence/usage rules: In the case of the use of generative AI systems, the review of the "results" generated by AI and the decision on their processing/further use always remains the responsibility of the respective user. In particular, they are responsible for checking non-discrimination, copyright, labeling requirements, and plausibility. AI-generated results may not be used without subsequent user review, and, in case of doubt, should follow the “four eyes” principle (e.g., consulting a supervisor). The respective user is always responsible for the final result. As part of training measures for the introduction of generative AI systems, explicit reference must be made to these issues. **In the case of a direct <u>person-AI system dialog</u> (e.g., through chatbots), there is a mandatory disclosure requirement.**	d. Consequence/usage rules: It is always the responsibility of the respective user to check the plausibility of the results in the context of their application and intended purpose, and to decide on their processing/further use. In particular, they are responsible for checking for non-discrimination; in case of doubt, they should follow the “four eyes” principle (e.g., consulting a supervisor).

3.a.2.1 **AI systems with a general purpose: Freely available and free of charge AI systems**	3.a.2.2 **AI systems for general use: Fee-based AI systems for official use.**
a. Definition: These are IT systems that are not officially provided to UKER employees in the context of their duties, but of which employees become aware via the internet, and whose use is offered free of charge. This includes systems that require registration before use and those that can be used without registration.	a. Definition: These are any kind of general purpose AI systems that do not fall under 3.a.2.1.

b. Examples: DeepL free or Chat-GPT free	b. Examples: DeepL Pro Advanced
c. Characteristics: The AI application scenarios of these systems are often not transparent, for example, when conducting internet searches or using chatbots to obtain information. Applications that are freely available on the internet often use user-entered data for their own purposes, such as training AI systems, as agreed in the terms and conditions of use. In particular, this may raise liability, data protection, and copyright issues for the respective employee at UKER.	c. Characteristics: The procurement and introduction of such AI systems are essentially comparable to prior IT software/hardware procurement and are therefore handled in the same way by the MIK (Medizinisches Zentrum für Informations- und Kommunikationstechnologie) Requirement and Portfolio Management.
d. Consequence/usage rules: UKER employees are subject to a comprehensive duty of care. It is therefore imperative that users ensure that no company or business secrets and no personal or personally identifiable data are disclosed (NOTE: Information such as "the senior physician from ward 3" or "the long-stay patient on neuro 2 with diabetes and additional heart failure" is already personally identifiable). If registration is necessary, only the official UKER account is to be used in accordance with Section 1. Freely available AI systems that are provided by the company are subject to the same procedure as fee-based AI systems (see Section 3.a.2.2). Freely available IT usage must be critically scrutinized. If there is a suspicion of potential damage to the UKER as an employer or to third parties, use must be refrained from and the problematic context of use must be reported to the supervisor and agreed upon the next steps (data protection incident report, IT security incident report, compliance breach, etc.).	d. Consequence/usage rules: The procurement of such AI systems must be reviewed and approved at the UKER in accordance with the IT hardware/software requirements process (see MIK-APM-VA Requirements Management, roXtra document 906) with the involvement of (1) the UKER Data Protection Officer, (2) the UKER IT Security Officer, (3) the MIK Requirements and Portfolio Management, and (4) the UKER IT Steering Committee. In addition to weighing up the cost and benefit effects, all aspects of data protection and IT security are also taken into account in this process - especially when it comes to AI systems that are aimed to be used as an external cloud application either on the internet or as a local AI application with a mandatory internet connection. For AI systems provided for official purposes, the use of personal or personal-related data is permitted only if explicitly stated in the hospital's rules of use for the respective AI system. If the use of the aforementioned data is permitted, it should be used as sparingly as possible. The use of pseudonymized data rather than raw data is always preferable and should be aimed for.

Until further notice, persons employed at UKER are not permitted to process (e.g., storage, transmission) within the policy of the use of freely available, free of charge AI systems (a) Trade secrets (Geschäftsgeheimnis) (in accordance with §2 (1) of the German Trade Secrets Protection Act (GeschgehG)), (b) security-relevant information, (c) trade secrets (Betriebsgeheimnis) and (d) copyrighted information to which the UKER does not have the corresponding rights of use is prohibited. UKER employees are also prohibited from processing (a) personal or personally identifiable data within the meaning of Art. 4 (1) GDPR, or (b) Information that may not be published due to corresponding regulations (e.g., specific legislation), in the context of the use of freely available, free AI systems (see also information on data secrecy in accordance with Art. 11 BayDSG).	

3.b **AI systems for use in clinical care (medical devices)**

a. Definition:
These are AI applications whose use has an impact on control and decisions in direct patient care, generate diagnostic results, or influence therapeutic procedures by means of generated patient classification/prognosis. With the exception of research systems (see 3.c), these are medical devices in accordance with the Medical Device Act or the EU Medical Device Regulation with corresponding training, maintenance, and validation requirements. However, there are currently (still) no separate requirements for software medical devices with AI components (as of January 2024).

b. Examples:
Transpara Breast Care; (Transpara artificial intelligence (AI) for radiologists; https://screenpoint-medical.com; Screenpoint Medical)

This AI system is based on an algorithm trained to detect "suspicious" signs in mammograms. It has learned the characteristics of various abnormalities and can tell within minutes whether there are signs of possible cancer. This "decision aid" is intended to serve as a second-opinion system for radiologists. Radiologists will be able to evaluate 3D mammograms faster than before.

TREWS (Targeted Real-Time Early Warning System); (AI Platform for Healthcare | Bayesian Health);
AI early warning system for sepsis, which analyzes the patient's medical history together with current symptoms and laboratory results to determine whether there is a risk of sepsis. The AI system also suggests treatment protocols, such as the use of antibiotics.

c. Characteristics:
In contrast to software medical devices, which are implemented based on, for example, a set of medical rules or decision trees, AI medical devices are based on data that may change over time. In the latter case, we also speak of self-learning AI systems. The data used to train an AI system often comes only from a specific environment and may not be representative of the patient population at the UKER. There is also a risk that the data used to train the AI system's decisions may introduce bias that could lead to discrimination against specific patient groups.
This means, among other things, that an AI developer cannot guarantee that the system works correctly in all situations.

Furthermore, with a "conventional" software medical product, it is possible to clearly and comprehensibly explain to a user the logic on which the result achieved is based. This is not possible with today's AI products (the results of research into so-called explainable AI (XAI) to date are not yet sufficient for this). For users, these therefore remain a "black box" in which they have to "blindly" trust.

Even if the majority of AI system medical devices do not yet operate autonomously, and the final decision to act and therefore the responsibility still lies with the user, this will certainly change at some point in the future, and AI medical devices will be offered whose operation will gradually show a higher degree of autonomous operation.

d. Consequence:
In line with a risk-based approach, **autonomously operating AI systems** at the UKER are classified as high-risk AI systems with a fundamentally unacceptable risk to the health and safety of patients in clinical care. If the intended purpose and the "instruction manual" of an AI medical device even suggest the possibility of a low degree of autonomous operation, this must be clearly stated. During procurement, the use of AI systems should be explicitly checked, evaluated, and documented by the procuring party, in particular, the extent of their autonomy.

The procurement of an AI medical device must be reviewed and approved at the UKER in accordance with the IT hardware/software requirements process (see MIK-APM- VA Requirements Management, roXtra document 906) with the involvement of (1) the UKER Data Protection Officer, (2) the UKER IT Security Officer, (3) the MIK Requirements and Portfolio Management, and (4) the UKER IT Steering Committee. For an AI medical device, the customer must always complete a "project application" (document FO), supplemented by the AI Governance and Ethics Policy - Checklist. Using this checklist, the customer should analyze the system being applied for with respect to the aspects listed under "Characteristics" as examples, and attach the results of this analysis to the project application.

3.c **AI systems for use in research projects that result in a notable impact on patient care**

a. Definition:
AI systems that have been developed at university hospitals (among others) and have not yet been certified as medical devices, but are used for evaluation in the context of a research project, the results of which are used by UKER physicians in the patient context and may have an impact on the medical care of UKER patients.

b. Examples:
AI systems developed in collaboration between UKER clinics/departments/institutes/central facilities with the Department of Artificial Intelligence in Biomedical Informatics.
AI systems developed in cooperation between UKER clinics/departments/institutes/central facilities and industry partners.
AI systems developed by UKER clinics/departments/institutes/central facilities or other university hospitals (e.g., in BMFTR projects such as NUM-RACOON, MII-OMI, or MII-Interpolar).

c. Characteristics:
These AI systems may entail risks for patients when used in research due to their application in the healthcare context. In contrast to the AI systems considered under 3b, these systems are not yet certified as medical devices, so formal documentation from their certification process is not yet available.

d. Consequence:
Since the AI system is planned as a research project involving humans, the corresponding research project must be submitted to the Ethics Committee of Friedrich-Alexander-Universität Erlangen-Nürnberg for review, and undergo an ethics voting process.

The use of this AI system in the context of a research project may have an impact on the management and decisions in research-related care in a similar way to the AI systems considered under 3.b, as it is used to generate diagnostic results or to influence therapeutic procedures by means of generated patient classification/prognosis. For this reason, the IT hardware/software requirements process described in QM-roXtra document 906, "MIK-APM-VA requirements management," must be conducted through MIK requirements and portfolio management using a "project application form" (document FO), and the associated checklist for an AI medical device must be submitted before the introduction of such an AI system, addressing the same points as in the case of AI systems considered under 3.b.

4. **Sanctions**
A breach of this guideline may constitute a breach of duty under the employment contract and be sanctioned accordingly. Claims for damages (under civil law) may also arise from culpable violations, particularly in the event of breaches of duty.

5. **References / Sources**

This guideline was created with the help of the following sources (All sources are in German; here we provide an English translation of the references):

a. REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL laying down harmonized rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Regulation) - KI-VO - , online at eur- lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CONSIL:PE_24_2024_REV_1, accessed on 18.06.2024.
b. Bavarian State Ministry of Finance and Home Affairs (BayStMFH)(Hg)(2024): Artificial intelligence. Guide for employees. Status March 2024.
c. BayStMFH (ed.)(2024): Artificial intelligence. Use of artificial intelligence in public administration. Guidelines for departments. Status March 2024.
d. Bavarian State Office for Data Supervision (BayLDA)(Hg)(2024): Data protection- compliant artificial intelligence. Checklist with test criteria according to GDPR. Version 24.01.2024.
e. The Hamburg Commissioner for Data Protection and Information Security (ed.) (2023): Checklist for the use of LLM-based chatbots. As of 13.11.2023.
f. State Chancellery of Schleswig Holstein (ed.)(2023): Permissibility of the use of text- based dialog systems using machine learning in the state administration of Schleswig Holstein here: Use of ChatGPT from the company OpenAI. Status 02.05.2023.
g. Artificial intelligence (AI) in medical devices - MDR guidelines (quickbirdmedical.com)

Applicable documents (in German)

a. MIK-APM-VA-Anforderungsmanagement
b. UKER Projektantrag (Dokument FO)
c. UKER KI-Checkliste