Pharmaceutical AI: Transparency and guidelines

Dr Stefan Kahl BearingPoint GmbH



>> Artificial intelligence (AI) is fundamentally changing the pharmaceutical industry. From research and production to the supply chain, enormous potential is opening up. But in the highly regulated environment of the healthcare industry, there is no room for error. Anyone who wants to use AI must take responsibility – with maximum transparency, clear security mechanisms and their own protected LLM framework as a basis. Security and compliance are paramount.

Transparency as a regulatory obligation

Transparency is much more than an ethical ideal – it is a regulatory necessity. Only when it is possible to understand how an AI model arrives at a result can the requirements of authorities, experts and patients be met. Explainable AI (XAI) is therefore becoming a key factor. The EU AI Regulation (AI Act) classifies many pharmaceutical AI applications as high-risk. This means that in future, companies will have to disclose how their models were trained, what data is used and how decisions are made. Black-box algorithms without explainability are a no-go in the pharmaceutical sector. Those who are not prepared for this risk not only fines, but also the trust of their stakeholders. Without transparency, there can be no AI in pharmaceuticals.

AI with a big impact

Despite all the caution, we must not lose sight of the positive effects of AI. Used correctly, AI can be a game changer:

- **Accelerate research and development:** AI models can sift through huge amounts of data in minutes, from scientific literature to genome data. This allows new drug candidates to be identified more quickly, failures to be

weeded out earlier and development times to be significantly shortened.

• **Optimising clinical trials and safety:** AI can improve study designs, for example by learning from previous trials which patient groups are most likely to respond. In pharmacovigilance, NLP systems accelerate the evaluation of adverse reaction reports. Important: With the right controls (guardrails), AI can be prevented from "inventing" symptoms or misinterpreting correlations. When properly integrated, AI significantly accelerates safety monitoring.

Guardrails are a must when using AI

As great as the benefits are, careful precision work is required in the pharmaceutical sector. Using a generic, off-the-shelf AI solution without customisation would be tantamount to a safety risk. LLMs without guardrails? That would be like driving a car without a seatbelt. Guardrails are therefore mandatory. But what does that mean in concrete terms? It refers to technical and organisational measures that ensure that an AI system only operates within defined limits. For example, inputs are filtered (no sensitive patient data without approval) and outputs are checked (no unauthorised or hallucinated content).

Critical facts must be cross-checked, for example by comparing them with validated databases. This prevents an AI assistant from recommending untested off-label use or outdated dosage information. Potentially fatal errors ("never events") must be absolutely ruled out. Equally important: AI should indicate when it is uncertain, rather than filling in the gaps with imagination. Quality assurance principles from the pharmaceutical world are directly applicable.

Keyword: ALCOA

The results of AI must be complete, traceable and correctly documented. In short, AI must be validated and monitored as conscientiously as any laboratory measuring device. And ultimately, a human being always remains in the loop: when it comes to critical decisions, AI can only provide support; the final responsibility lies with an expert.

Strict regulatory requirements

Regulatory authorities on both sides of the Atlantic have begun to set clear guidelines for AI. An overview of important points:

• **EU AI Act:** The EU regulation, which will come into force in 2026, stipulates extensive measures when AI is used in high-risk areas such as medicine and pharmaceuticals. Among other things, it requires a risk management system, high-quality, unbiased training data, detailed documentation of the model, logging of all interactions and human supervision. Violations will be severely punished (up to £27 million or 6% of annual turnover). The message is clear: AI systems in sensitive areas must be transparent, secure and controlled, otherwise they may not be operated.

• **GxP & validation:** Good practice has long been the norm for digital systems in the pharmaceutical industry. Whether in manufacturing (GMP), laboratories (GLP) or clinical trials (GCP) – electronic systems must be qualified and validated everywhere. Data integrity is a top priority: all results must be documented in a complete, traceable and unalterable manner.

Therefore, an AI system must be treated like a computer system: it requires strict validation, regular checks and audit trails that log every AI action. Without this core quality, its use in GxP processes is taboo.

- **FDA:** The US Food and Drug Administration also requires evidence of the reliability of AI models. In early 2025 , the FDApublished draft guidelines that provide for a risk-based "credibility assessment". Companies must demonstrate that an AI model used in regulatory studies, for example, is robust and trustworthy – through validation studies, continuous monitoring and transparent documentation. The FDA recommends early consultation with the authority if AI is to be used in development.
- **EMA:** In 2024, the European Medicines Agency published guidelines that, although aimed at authorities, are also relevant for companies. The EMA emphasises the safe, responsible use of AI and calls for transparency, traceability and human control. Companies should therefore establish robust AI governance, train employees and continuously monitor AI-supported processes. Ultimately, AI that influences drug decisions is subject to the same quality requirements as conventional methods.

Practical example

The BearingPoint tool GenXplore shows how AI can be used safely and profitably. This is an in-house generative AI solution designed to identify hidden efficiency potential in large organisations. GenXplore analyses thousands of activities and processes from job and process descriptions and systematically identifies where automation or AI support can deliver the greatest benefits.

The result is a data-based AI impact score that provides clear priorities for where AI deployment is most worthwhile. In previous projects, the added value quickly became apparent: within a few weeks, measurable efficiency gains were noticeable. This gives companies a precise AI roadmap, tailored to their actual situation rather than assumptions. Applied to the pharmaceutical industry: Pharmaceutical companies also have vast amounts of documented processes (SOPs, laboratory protocols, reports) that could be efficiently evaluated using GenXplore-type analyses. Important: GenXplore runs in its own IT environment, and all data remains internal. This example shows that generative AI can also have a major impact in a protected environment. Transparency, data sovereignty and compliance are maintained, yet tangible results are achieved quickly.

Conclusion

The use of AI in the pharmaceutical market is not a sure-fire success. The industry can benefit enormously from AI – if it combines innovation with responsibility. Transparency, guardrails and proprietary AI frameworks are not bureaucratic hurdles, but the basis for sustainable success. Pharmaceutical companies that invest now in secure, explainable AI solutions not only ensure regulatory compliance, but also gain the trust of doctors, patients and partners and increase their competitiveness. Ultimately, the bottom line is: AI yes, but only with a safety belt. <<

Literature:

1. European Commission. (2025). Guidelines and code of practice on transparent AI systems. https://digital-strategy.ec.europa.eu/en/faqs/guidelines-and-code-practice-transparent-

aisystems [digital-st….europa.eu]

2. European Medicines Agency. (2024). Artificial intelligence: Guidance and workplan 2025–2028. https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-bigdata-other-sources/artificial-intelligence [ema.europa.eu]

3. U.S. Food and Drug Administration. (2025). FDA proposes framework to advance credibility of AI models used for drug and biological product submissions [Press release]. https://www. fda.gov/news-events/press-announcements/fda-proposes-framework-advance-credibilityai-models-used-drug-and-biological-productsubmissions [fda.gov]

4. EU AI Risk Team. (2025). Making AI explainable: A practical guide to transparency and documentation under the EU AI Act. https://euairisk.com/resources/ai-explainability-documentation-guide euairisk.com]

5. International Society for Pharmaceutical Engineering (ISPE). (2025). ISPE GAMP® Guide: Artificial Intelligence. https://ispe.org/publications/guidance-do-cuments/gamp-guide-artificial-intelligence [ispe.org]

Citation

Kahl, S.: "Pharma AI: Transparency and Guidelines," in "Monitor Health Services Research" (01/26), pp. 54–55.
http://doi.org/10.24945/MVF.01.26.1866-0533.2795