Artificial Intelligence at Trium: Our Position and Practice

Trium Analysis Online GmbH — position paper on the use of Artificial Intelligence

Introduction

Trium develops software that helps midwives and clinicians care for mothers and babies during pregnancy and birth (see Appendix for full Intended Use). We are increasingly asked whether, and how, we use Artificial Intelligence (AI). This paper provides an honest answer: **we use AI in our development process (AI-assisted coding), and we are planning to introduce machine learning into our product.** Our current product, as shipped today, does not rely on statistical learning or large language models – but definitions of AI vary, so we explain our position clearly.

One principle governs all our AI-related activities: **AI supports skilled people; it does not replace their judgement.**

1. What we mean by "AI" – and what the EU regulation says

Discussions of AI are often confused by different definitions. The European Union's AI Act (Regulation (EU) 2024/1689) provides the authoritative legal one. Article 3(1) defines an AI system as:

> "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".

It would be unrealistic to expect a single legal definition to draw a sharp line through every borderline case in one stroke in a field like AI. The definition above is wide: most of its elements are graded to the low end, optional, or open-ended — autonomy is graded, adaptiveness is explicitly optional ("may exhibit"), the listed output types non-exhaustive — so the substantive requirement is that a system *infers* its outputs rather than executing rules defined solely by humans. Read at its broadest, that bar is low: even some of the classical algorithms in our current product — for instance the model-fitting that characterises fetal-heart-rate decelerations — derive their outputs from the data and could, on such a reading, fall within it.

Breadth, however, is not the same as intent — and the Act's purpose is clearer than its width might suggest. Recital 12 sets the floor expressly, excluding systems "based on the rules defined solely by natural persons to automatically execute operations." Trium CTG Online, as shipped today, rests on exactly such human-defined rules — FIGO-based CTG scoring — together with classical signal processing; it does not learn from data and does not adapt after deployment. Read in line with the Act's evident purpose, then, the product as it stands

today is not the kind of system the regulation is aimed at. We therefore interpret the Act by its intent rather than its widest possible literal reach, and we expect notified bodies and courts to do the same.

For the purposes of this document, we use "AI" to refer to two concrete technologies: **Large Language Models (LLMs)** used in software development, and **statistical machine learning** planned for future product versions.

2. AI in our software development

We use Large Language Models (e.g., Claude) as assistants in software engineering. This is a contested area: some developers keep AI out of their craft entirely; others let AI lead. We take a middle path: **AI accelerates, but a qualified human engineer directs and reviews everything.**

Technical background: LLMs are probabilistic language models, not logical reasoners. They just interpolate words and phrases for a given context based on statistical relations between words and phrases from the training data [Moradi et al. 2024]. Since the training data include billions of lines of code, hundreds of thousands of textbooks on computer science, documentation, and technical discussions from the internet, they excel at generating plausible-looking code. But they can produce:

- Logic errors or incomplete case distinctions
- Security vulnerabilities (e.g., injection flaws, unsafe memory handling, LLM poisoning)
- Hallucinations – code that appears correct but is functionally wrong

Asking an LLM to "explain" its own output is **not** explainable AI – the explanation is generated by the same probabilistic process and may be confidently incorrect [Mayne et al. 2025].

Our policy:

- Line by line review of AI generated code is not considered as sufficient. The developer has to understand the code as if it has been written by themselves.
- Automated tests are necessary but not sufficient to catch AI-specific errors.
- The core methods of our signal analysis – e.g. how we compute the baseline and characterise fetal heart rate decelerations – are conceived by our engineers; AI assists only in building out, refining, documenting, and checking that work under human direction.

Our use of AI in development is auditable. As a manufacturer of a medical device, we build software under established verification, validation, and review disciplines.

3. AI in our product – current and planned

3.1 Current product (no statistical AI)

Trium CTG Online is a CE-marked medical device (Class IIb decision-support system for central fetal monitoring). Its analysis module extracts well-defined features from the fetal heart rate (FHR) signal and scores them according to FIGO rules. It uses classical signal processing – e.g., fitting a parametric model to decelerations to characterise onset, depth, timing, and recovery, and detecting sinusoidal patterns. These are classical algorithms, not statistical learning.

We do not try to sell this as AI, because doing so would mislead. However, we are actively planning to introduce machine learning.

3.2 Planned: Feature-based machine learning for CTG classification

We have established a collaboration with clinics to build a large database of CTG recordings together with clinical outcomes. Together with our prior experience with SVM-based classification from another field (walking speed estimation from accelerometry [Schimpl et al. 2011]), this will enable us to apply machine learning techniques like Support Vector Machines (SVM) on extracted features – including but not limited to FIGO features – to produce an **evidence-based scoring** beyond current FIGO guidelines. The SVM will learn associations between feature patterns and clinical outcomes directly from the database.

Our safeguards:

- **Strict validation policy:** No machine-learning based scoring will replace or augment the current system without a validation on independent data from different sources.
- **Human accountability preserved:** The system remains decision support, not autonomous diagnosis. The clinician always makes the final clinical decision.
- **Explainability:** We will use post-hoc explanation techniques (e.g. LIME [Ribeiro et al. 2016]) to show which features drove each classification – for example, "The SVM classified this CTG as pathological because the feature 'late deceleration area' contributed +0.7, while 'baseline variability' contributed -0.2." This explains the **model's internal logic**, not the clinical truth, but allows a clinician to reject the output if the explanation reveals reliance on artifact.

4. Looking ahead

Across both what we build and how we build it, one principle holds: **AI assists; accountable people decide.** We treat the human's place in this not as a limitation but as a commitment – to safety, to quality, and to the trust on which medical care depends.

We will continue to adopt AI responsibly: with meaningful human oversight, careful evaluation of every capability we introduce, and an unwavering place for the clinician at the centre of care.

References

Schimpl, M., Lederer, C., & Daumer, M. (2011). Development and validation of a new method to measure walking speed in free-living environments using the actibelt® platform. *PloS one*, *6*(8), e23080.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should i trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).

Moradi, M., Yan, K., Colwell, D., Samwald, M., & Asgari, R. (2024). A critical review of methods and challenges in large language models. *arXiv preprint arXiv:2404.11973*.

Mayne, H., Kearns, R. O., Yang, Y., Bean, A. M., Delaney, E. D., Russell, C., & Mahdi, A. (2025, November). LLMs Don't Know Their Own Decision Boundaries: The Unreliability of Self-Generated Counterfactual Explanations. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing* (pp. 24172-24197).

Appendix

Excerpt from "Intended Use Trium CTG Online" 30.06.2022

Main Function

Trium CTG Online is designed to acquire, transfer, store, process and display digital CTG signals in real-time via Intra- or Internet for central surveillance.

Specification of the Main Function

The software is installed on a central server that is connected to the Intra- and/or Internet. No other software installation on other PCs is necessary. The CTG signals are acquired from the CTG devices' digital interface and transmitted to the server, where they are processed by the software. Acquiring CTG signals is carried out with the builtin data acquisition software component "Trium CTG Online – DAQ". In addition the software allows to receive data from remote systems.

Every authorized user can see and supervise all current CTG recordings with a common internet browser location-independently and online over the Intra- and/or Internet. Built-in classification algorithms, which were designed according to common international guidelines, provide decision support. The summary of data allows the professional user to make essential decisions and initiate further actions for the patient, though Trium CTG Online is not intended to be used as an automated diagnosis system. Trium CTG Online is, therefore, an active diagnostic medical product intended to be used as central surveillance system.

Acknowledgements

We would like to thank Stephen Gilbert and Andrew Gibson for their comments. The views and opinions expressed in the paper are those of Martin Daumer, Christian Lederer and Marcin Mider