Synthetic Data Q&A

Webinar in collaboration with LROI, TNO and BlueGen.ai - March 26, 2026

Questions sourced from the application form

FAQ: The LROI Case

Core Message: LROI is developing a synthetic data infrastructure with BlueGen.ai. BlueGen has no access to the environment where the software runs: the software is delivered and runs entirely locally at LROI, in an encrypted environment to which BlueGen's team has no access.

↓ *Visual support — see appendix*

What specific plans does LROI have for synthetic data, and what access does BlueGen.ai have?

LROI is working on a synthetic data product for research exploration and international collaboration. We provide the software for generation and validation, but have no access ourselves. The software runs entirely locally in an encrypted environment at LROI. BlueGen receives only meta-data logs (averages, standard deviations, min/max, etc.), never the source data or the synthetic output itself.

How does research continuity and dependency on LROI proceed?

Synthetic data makes it possible to prepare and validate analyses without repeatedly submitting formal data requests. This increases the speed of the research cycle. Dependency on LROI remains for confirmatory steps; synthetic data is a supplementary tool, not a replacement for the regular data request process.

What developments are there to compare a hospital's own data with comparable clinics?

LROI is working on case-mix adjusted benchmark reports, where a hospital can compare its outcomes with a peer group based on patient characteristics. Synthetic data can play a role by decoupling the peer-group comparison from direct data sharing between clinics.

Explanation of complication registration and comparison with the national benchmark

The complication registry in LROI records complications following joint prostheses. The national benchmark allows a hospital to compare its own complication ratio with the national average, adjusted for case-mix. Synthetic data can be used for methodological development and for situations where a hospital wants to understand the benchmark logic without direct access to the full national dataset.

Applying knowledge from the LROI registry within other registries

The approach LROI is developing for synthetic data is in principle transferable to other quality registries. Prerequisites are comparable: sufficient source data, clear governance, and a validation strategy aligned with the use case. LROI can play a pioneering role here and provide a reference model for other registries within the Dutch healthcare sector.

What does the use of AI within LROI mean for scientific research in hospitals?

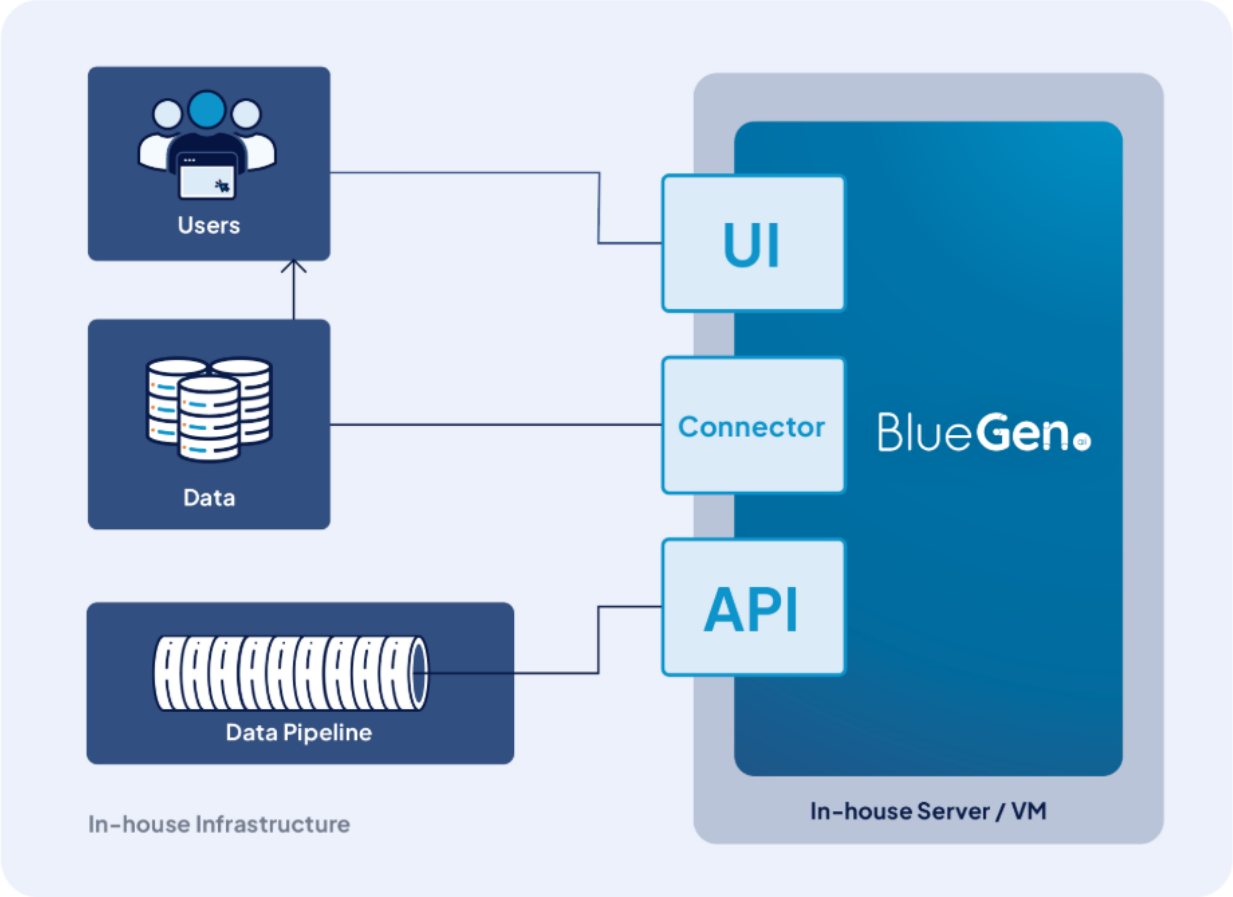
For hospitals, this means concretely that they can use LROI synthetic data to prepare and validate analyses locally without first submitting a formal data request. This lowers the threshold for smaller centers that have less capacity to complete the full data request process. The confirmatory step, and thus the actual LROI data, remains required for publications, but the path to it becomes considerably shorter.

Still have questions?

If you have further questions after reading this document, or would like to explore what synthetic data could mean for your specific situation, please feel free to contact us via iman@bluegen.ai or edwin@bluegen.ai.

Appendix — Visual Support

The visual below supports the content of the LROI case.

BlueGen.ai data flow architecture

This visual shows how source data, generation, and synthetic output remain strictly separated. BlueGen's software runs entirely locally at LROI. BlueGen receives only meta-data logs, never the source data or the synthetic output itself.