P1

Auditable privacy for clinical synthetic data

Closing the gap between formal guarantees and what attacks achieve in practice.

PhD direction | Privacy · Security | Suits: KTH · SU · co-supervised

The question

Can we build audit methods that let an independent party verify, with statistical confidence, that a synthetic clinical release behaves as its stated privacy guarantee says, without seeing the producer's internals?

Why it matters

Privacy auditing is an active research area, but clinical synthetic data has its own structure, such as linked tables and coded clinical events. Third-party verifiability is what turns a vendor claim into a trust standard.

What you would do

01 Develop and evaluate black-box audit methods for synthetic tabular and relational clinical data.

02 Study how tight empirical lower bounds on privacy loss can be made for realistic data.

03 Validate on public data and, where approved, inside partner secure environments.

04 Publish open audit tooling.

What you would produce

- A research programme of three to four papers
- Open audit tooling
- Collaboration across academia, hospitals and industry

Good fit if you have

Strong background in ML, statistics or security; master's degree in a relevant field.

Data and access

Public seed data (CMS DE-SynPUF) and synthetic releases from day one. Work on real patient records only inside a partner's approved secure environment, with ethics approval. We hold no Swedish patient data.

Timing and place

Start by agreement. We are an early-stage company and cannot fund a PhD position alone; this is a direction for a co-supervised project if a research group and funding (e.g. an industrial doctorate) align.

Supervision

Gaurav Raj Chattarki, Founder, Synthetix Health, together with your academic supervisor.

Questions or a proposal of your own? gaurav@synthetixhealth.com · synthetixhealth.com/careers#thesis

All projects run under a confidentiality and IP agreement agreed before work begins. We review drafts only for confidential information, never for results. Negative results are publishable.