

Ph.D. research topic

- Title: **Machine Learning Methods for the Anonymization of health Data**
 - Research axis of the 3iA: AI for Computational Medicine (Axis 2)
 - **Supervisor: Herve Delingette (Inria, herve.delingette@inria.fr)**
 - co-supervisor: Nicholas Ayache (Inria, nicholas.ayache@inria.fr)
 - The research group: Project-Team **Epione** at **Inria Sophia Antipolis**.
-

Apply by sending an email directly to the supervisor and co-supervisor.

The application will include:

- **Letter of recommendation of the supervisor indicated above**
 - Curriculum vitæ.
 - Motivation Letter.
 - Academic transcripts of a master's degree(s) or equivalent.
 - At least, one letter of recommendation.
 - Internship report, if possible.
-

Short description of the topic:

Anonymization of medical data is required for any storage or processing and for data sharing outside hospitals and dedicated medical data centers. Yet providing such level of anonymization goes well beyond the removal of identifying metadata and the challenge is to determine what level of risk of re-identification is acceptable in order to deliver the potential benefits of data sharing.

The goal of this thesis is to develop AI methods for producing anonymized databases of health data from existing pseudonymized ones in particular for medical images. To this end, methods for quantifying the risks of individualization, correlation and inference in medical databases will be first tested and techniques to overcome those risks will be proposed in a second stage. Various methodologies may be explored ranging from low level data processing, to generative probabilistic models, and differential privacy[1,2].

[1] Ziller, Alexander & Passerat-Palmbach, Jonathan & Ryffel, Théo & Usynin, Dmitrii & Trask, Andrew & Junior, Ionésio & Mancuso, Jason & Makowski, Marcus & Rueckert, Daniel & Braren, Rickmer & Kaissis, Georgios. (2020). Privacy-preserving medical image analysis.

[2] Kaissis, G.A., Makowski, M.R., Rückert, D. *et al.* Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence* **2**, 305–311 (2020).