Machine Learning Methods for the Anonymisation of health data

**Supervisor:** Nicholas Ayache, nicholas.ayache@inria.fr  
**Co-supervisor:**  
Hervé Delingette, herve.delingette@inria.fr  
Xavier Pennec, xavier.pennec@inria.fr  
Marco Lorenzi, marco.lorenzi@inria.fr  
Maxime Sermesant, maxime.sermesant@inria.fr  
Olivier Humbert

**Laboratory or research group:** Epione, in partnership with CHU Nice,

**Context:** 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. More precisely, European regulating bodies have established guidelines for producing medical anonymised data including the need to handle three types of risks: individualization, correlation and inference. In the scope of the 3IA Côte d’Azur, an important challenge is to define guidelines and processes for making various medical databases accessible to the AI academic community.

**Post-doc Topics:** In this context, the goal of this post-doc is to develop AI methods for producing anonymized databases of health data from existing pseudonymised ones in particular for medical images, omics, and clinical reports data. 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

**Required Competences**
- Phd degree and strong experience in statistical learning, or biomedical imaging are required
- Solid programming and IT skills are necessary (Python and C++, bash scripting, version control systems).
- Strong communication abilities
- Fluent English (written and spoken)