

# Methods and tools for health data anonymization to allow open secondary usage and big-data analysis of health data

---

## Summary

---

*To develop and justify the possible methods and tools that allow automated and rule based data anonymization of health data so that these anonymizations (issue is also reported in IMI future research topics ) allow secure and dependable open secondary usage and big-data analysis of health data, without the risk of personal data leakage and breach of data protection rules, but still useful and reliable from the perspectives of data analysis and medical research.*

Research field:	Information and communication technology
Supervisors:	Gunnar Piho Peeter Ross
Availability:	This position is available.
Offered by:	School of Information Technologies Department of Health Technologies
Application deadline:	Applications are accepted between June 01, 2020 00:00 and July 03, 2020 23:59 (Europe/Zurich)

## Description

---

The routine clinical data are considered precious[1], and their secondary use[2] is considered beneficial for policy-makers, public health officers, scientists, clinicians, citizens and industry[3]. Different initiatives, including European Health Data Network[4] and Clinical Trial Data<sup>2</sup> initiative initiated by the European Commission and the EFPIA (European Federation of Pharmaceutical Industries and Associations), are searching for the possible solutions. Still, most of the medical data are locked behind the security systems of healthcare institutions and the regulations of the GDPR, and mostly divide-and-conquer methods[5] are in use. However, despite different interoperability standards in the medical domain, but without optimised data sharing and reuse, no big-data analytics will happen[6]. This PhD project is looking for different data anonymization methods (issue is also stated in the IMI future research topics<sup>2</sup>) so that secure and dependable open secondary usage and big-data analysis of health data is possible, without the risk of personal data leakage and breach of data protection rules. These anonymized data must still be useful and reliable from the perspectives of big-data analysis and medical research. Results are utilized in collaboration projects with TEHIK and international partners[7].

**Responsibilities and tasks:** To investigate the state of the art of the different data anonymization methods and evaluate them from the perspectives of real healthcare systems and medical researchers and practitioners. Based on the acquired knowledge, develop and evaluate rule based methods and tools so that automated anonymization is possible. Evaluate proposed tools and methods from the perspectives of GDPR, medical research, open data and open science.

**The applicants should fulfil the following requirements:**

- MSc in Software Engineering or related fields like Informatics, Computer Science or Medical Informatics.
- Excellent software engineering skills.
- Competence in medical informatics and healthcare systems interoperability is a plus but not mandatory.

[1] T.D.Wade, Refining gold from existing data, Curr Opin Allergy Clin Immunol, 2014; 14(3): 181-185

[2] PricewaterhouseCoopers, 2009, Transforming healthcare through secondary use of health data.

[3] W.O.Hackle and E.Ammenwerth, SPIRIT; systematic panning of intelligent reuse of integrated clinical routine data – a conceptual best-practice framework and procedure model, Methods of information in medicine, vol 55, no 02, pp.114-124.

[4] IMI (Innovative Medicines Initiative), 2017, 12<sup>th</sup> Call for Proposals, document reference IMI2/INT/2017-02169

[5] B.Shickel, P.J.Tighe, A.Bihorac, and P.Rashidi; Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis, 2018, IEEE Journal of Biomedical and Health Informatics, vol 22, no 5, pp 1589-1604

[6] X, Gansel, M. Mary, and A. van Belkum; Semantic data interoperability, digital medicine, and e-health in infectious disease management: a review; 2019, European Journal for Clinical Microbiology and Infectious Diseases, 38: 1023-1034

[7] Prof. Martin Leucker ( <https://www.isp.uni-luebeck.de/leucker>, Institute for Software Engineering and Programming Languages at University of Lübeck; google hi – 35), Prof. Yngve Lamo (<https://www.isp.uni-luebeck.de/leucker>, Department of Computing, Mathematics and Physics at Western Norway University of Applied Science; google hi – 16), University Medical Centre of Schleswig-Holstein (Germany), Houkeland University Hospital (Norway), and Zealand University Hospital (Denmark)



To get more information or to apply online, visit <https://taltech.glowbase.com/positions/76> or scan the the code on the left with your smartphone.