Online-Workshop on Privacy Preserving Analysis

14th July, 1-3pm

Link to videoconference:

https://hs-osnabrueck.zoom.us/j/92493563043?pwd=WkxodkFEVGVDbkhGZXZUNkxCWTUvdz09

The project Zukunftslabor Gesundheit (https://www.zdin.de/zukunftslabore/gesundheit) focuses on the potential of digitization in the healthcare sector to create a networked health care system. The subproject “TP1 - data analysis and data integration” develops a research platform to support comprehensive provision of health data and is developing and testing privacy preserving data analysis of health data. This workshop on privacy preserving analysis will be organized by the subproject TP1. Participants will learn about two current data analysis problems and solutions using privacy preserving analysis.

1. Privacy risks: From anonymization to machine learning

Mining sensitive data such as health data can lead to faster medical decisions, improvement in the quality of treatment, disease prevention and innovative solutions. However, health data is highly sensitive and subject to regulations such as the General Data Protection Regulation (GDPR), which aims to ensure patient's privacy. Anonymization or removal of patient identifiable information, though the most conventional way, is the first important step to adhere to the regulations and incorporate privacy concerns. Nonetheless, anonymization alone is not sufficient.
In this first part of the talk, we will see a reconstruction attack on anonymized data that can retrieve the original private data with a high accuracy.

Similarly, when machine learning models are trained on sensitive data, the released model can still leak information on the data it was trained on.
In the second part of the talk, we will discuss membership inference attack on graph neural networks (GNNs) where the goal of the adversary is to determine whether a particular data was used in training the target model.

2. Differentially private outlier detection

Outlier detection is often an important objective in data analysis. Especially in the field of medical informatics, however, we need to consider data privacy in order to protect sensitive data. The challenge is to find mechanisms that guarantee privacy up to a certain degree, without negatively impacting performance of the data analysis too much. In this talk, we will look at two approaches from both domains, outlier detection and privacy preserving data analysis, and how to combine them.

Agenda	
1:00 - 1:15pm	**Welcome** Prof. Dagmar Krefting (UMG, Head of Zukunftslabor Gesundheit)
1:15 - 2:00pm	**Presentation: Privacy risks: From anonymization to machine learning** Emmanuel Iyiola Olatunji (L3S Hannover)
2:00 - 2:45pm	**Presentation: Differentially private outlier detection** Dr. Jens Rauch (Health Informatics Reseach Group, Hochschule Osnabrück)
2:45 - 3:00pm	**Discussion** Prof. Dagmar Krefting (UMG, Head of Zukunftslabor Gesundheit) / Dr. Klaus-Hendrik Wolf (PLRI TU Braunschweig und MHH)