

# Mathematical patients privacy protection techniques in medical databases

*Centre of Health Information Systems*

## Background

With the development of information technology, more and more data is accumulated in various systems. The data, gathered by some companies and organizations, is not only stored but also published. This leads to an increasing amount of publicly available data, consequently increasing the risk of privacy violation.

Privacy is more and more important every day. However, it seems that it may be understood in different ways, so its meaning may not be the same for different people. This is particularly important in the case of medical databases, where there are many stakeholders whose privacy may be at risk.

Databases may be accessible both publicly or for limited audience (e.g. medical doctors), for which we can make additional assumptions. Therefore, various target groups should be taken into consideration.

## Problem description

The desired solution should help to reduce the risk of private-data leak from statistical databases. It should help the database owner to make appropriate decisions, like:

- What should be the form of available data (fixed statistical data sheet, dynamically generated statistics, etc.)?
- What should be the scope of available statistics?
- What additional parameters should be taken into account?
- ...

Furthermore, the proposed solution should meet the following criteria:

- be easy to understand (as a criterion of evaluation or as the information presented to the public);
- be easy to implement;
- be verifiable.

The project work may focus on the following aspects:

- possible measures of privacy in terms of medical databases, emphasizing statistical databases;
- analysis of methods of data protection adequate for statistical databases;
- propose some ideas for further exploration.