



Università di Pavia
Dipartimento di Ingegneria Industriale e dell'Informazione
Laboratorio di Informatica Biomedica Mario Stefanelli

SEMINARIO

Practical Lessons from Using Synthetic Healthcare Data to Validate AI Models

Allan Tucker, Brunel University London, UK

Venerdì 2 Luglio 2021, 9.30-11.00

ONLINE: <https://us02web.zoom.us/j/6807412420>

Healthcare data holds huge societal and monetary value. It contains information about how disease manifests within populations over time, and therefore could be used to improve public health dramatically. To the growing AI in health industry, this data offers huge potential in generating markets for new technologies in healthcare. However, primary care data is extremely sensitive. It contains data on individuals that is of a highly personal nature. As a result, many countries are reluctant to release this resource. This seminar explores some key issues in the use of synthetic data as a substitute for real primary care data: Handling the complexities of real world data to transparently capture realistic distributions and relationships, modelling time, and minimizing the matching of real patients to synthetic data points. I will show that if the correct modelling approaches are used, then transparency and trust can be ensured in the underlying distributions and relationships of the resulting synthetic datasets. What is more, these datasets offer a strong level of privacy through lower risks of identifying real patients. However, there is still a number of open questions concerning bias, transparency, and risks when dealing with rare disease and personalized medicine.

Organizer

Prof. Riccardo Bellazzi and Dr. Arianna Dagliati

For more information: arianna.dagliati@unipv.it