CELSA
Central Europe Leuven Strategic Alliance

CELSA
Central Europe Leuven Strategic Alliance

CELSA - Collaborative research project - Application form - COVER PAGE

1. Identification of the principal investigator – co-ordinator
Full name: Yves Moreau
Faculty/Department: Faculty of Engineering Science/Department of Electrical Engineering (ESAT)
Research unit within Faculty/Department: ESAT - STADIUS, Stadius Centre for Dynamical Systems, Signal Processing and Data Analytics
Address: Kasteelpark Arenberg 10 - box 2446 3001 Leuven
University: K.U.Leuven
Tel: +32 16 32 86 45
Fax : +32 16 3 21970
email : Yves.Moreau@esat.kuleuven.be
Signature[1]: 

2. Identification of the second investigator
Full name: Peter Antal
Faculty/Department: Faculty of Electrical Engineering and Informatics/Department of Measurement and Information Systems
Research unit within Faculty/Department: Computational Biomedicine (ComBine) workgroup
Address: Magyar tudósok krt. 2. 1117 Budapest, Hungary
University: Budapest University of Technology and Economics (BME)
Tel: +36 1 463 43 94
Fax : +36 1 463 41 12
email : antal@mit.bme.hu
Signature[1]: 

3. Identification of third and fourth co-investigator(s) (if applicable) Expand table if more than four research units are involved.
Third co-investigator
Full name: Levente Buttyán
Faculty/Department: Faculty of Electrical Engineering and Informatics/Department of Networked Systems and Services
Research unit within Faculty/Department: Laboratory of Cryptography and System Security (CrySyS Lab)
Address: Magyar tudósok krt. 2. 1117 Budapest, Hungary
University: Budapest University of Technology and Economics (BME)
Tel: +36 1 463 1803
Fax: +36 1 463 3263
email: buttyan@hit.bme.hu
Signature[1]: 

[1] Scanned signatures will be accepted.

3. Non confidential and public friendly summary (max. 2000 characters)
New molecular measurement technologies, particularly in genetics, have fundamentally changed biomedical research in the past decades through the generation of massive amount of data. In parallel, it was also realized that further development depends at least as much on methods for the fusion of data and knowledge that create actual clinical value, as on further technological breakthroughs in data generation. These "big health" data sets could provide an unprecedented opportunity for the investigation of genetic, personal, environmental and societal aspects of health and diseases. However, the privacy of the data at a personal level and the confidential multi-centric distribution of the data at the institutional level pose a fundamental challenge for the joint analysis of these data sets. Such personal, intimate data cannot be distributed without appropriate safeguards. Classical IT solutions to these problems (primarily the management of access control) as important limitations deciding who can access which data under which conditions creates deep bureaucratic swamps. In our research, we will develop privacy-preserving machine learning methods that allow performing complex analyses without revealing sensitive personal data. Only nonpersonal derived data is shared

between parties, while personal data remains solely with its original controller at all times. Such approaches greatly lower the threshold for cooperation between parties and increase the confidentiality of the data. These methods will make it possible to explore the molecular background of diseases using distributed genetic-clinical data repositories and biomedical knowledge bases about the effect genetic variants and biochemical pathways. Thus, the combination of privacy-preserving machine learning and information fusion will be the key innovative element in our project. To cope with the uncertainty of the measurements and scientific knowledge, we will also adopt probabilistic approaches from the field of artificial intelligence. The information fusion methods, through various abstraction levels and probabilistic approaches, will be integrated with privacy-preserving solutions that allow distributed learning from federated sources, while guaranteeing personal privacy.

4. 5 keywords
Bayesian statistics, probabilistic graphical models, kernel methods, privacy preserving, data and knowledge fusion in biomedicine