



**Dublin Institute of Technology (DIT)** is one of Ireland's largest and most innovative university-level institutions. DIT participates as a full member of the European University Association (EUA) and the International Association of Universities (IAU). Dublin Institute of Technology is ranked in the top 3% of universities worldwide. There are 20,000 students registered in DIT, and over 2000 members of staff. The project will be hosted by the Applied Intelligence Research Centre (AIRC) at the DIT School of Computing. The DIT School of Computing is the largest of its kind in Ireland and offers undergraduate, postgraduate and PhD programmes which are characterised by strong industry engagement through internships, industry projects and industry accreditation. The AIRC has strong track record in data and predictive analytics and machine learning. Currently, the membership of the AIRC includes 10 full time members of staff, 4 post-doctoral researchers, and 36 PhD students. The AIRC has received significant funding in the last 5 years for research in these areas.

**Main tasks in Project**

DIT cooperates in predictive modelling and contributes to the project in terms of machine learning and data science expertise. The major scientific challenges that DIT will address, include: data exploration research to identify risk factors of stroke, developing a framework for training and implementing extensible hybrid predictive models that integrate mechanistic and data driven models, and creating predictive models that are capable of predicting complex structured objects that model the multifaceted nature of a patient's quality of life. DIT has a long history in managing the development and delivering predictive models. DIT also has experience in modelling quality of experience and has expertise in natural language processing. As a result DIT is well positioned to engage with a range of partners with different expertise and backgrounds.



The **Estonian Genome Centre** integrating the Estonian Biobank is a research institute of the University of Tartu whose aim is to promote the development of human genetic research by collecting information on the health and genetics of the Estonian population. The activities of the EGC are focused on using the results of the latest genomic research to improve public health and developing personalised medicine in Estonia. The EGC is a centre identified for the Estonian and European Research (Infrastructure) Roadmap and a partner for the Centre of Excellence in Genomics. The EGC is participating in the European Strategy Forum on the Research Infrastructures project Biobanking and Biomolecular Resources Research Infrastructure (BBMRI – head of the National Node is Prof. A. Metspalu) and is also one of the founding Charter members of the international biobank consortium "Public Population Project in Genomics and Society" (P3G). From 2015, UOT/EGC is member of the EIT Knowledge and Innovation Community (KIC) consortium EIT-Health. The Estonian Biobank is the population-based biobank of the Estonian Genome Centre of University of Tartu. The biobank is conducted according to the Estonian Human Genes Research Act and all participants have signed a broad informed consent. The cohort size is currently 51,535 participants (age 18+). Metabolite concentrations (over 110; NMR and MS) are available for about 12,000 subjects; metabolic follow-up has been done for more than 2,000 donors. Over 50,000 subjects have been genotyped with different genome-wide arrays from Illumina, 2500 whole exomes and 2500 whole genomes have been sequenced. We have linked our database with the national registries and hospital databases for the collection of up-to-date phenotypic information. The law enables re-contacting the participants for further studies and the rules to access data and samples are clear and transparent (<http://www.biobank.ee/en/>). All procedures are run according to ISO 9000-2008 and using a custom laboratory information management system. There is a team of 51 people (40 FTE) working at the research unit, at the biobank, and at the core lab.

**Main tasks in Project**

The Estonian Genome Center contributes with both data and expertise for analysis of health records. Baseline and follow-up health related data and objective measurements of 1000 prevalent and 1100 incident stroke cases from the Estonian biobank will be used for the risk modelling in the project. The developed models for risk calculations can be validated within the Estonian personalized medicine programme.



Founded in 1988, the **DFKI (Deutsches Forschungszentrum für Künstliche Intelligenz)** today is one of the largest nonprofit contract research institutes in the field of innovative software technology based on Artificial Intelligence (AI) methods. DFKI is focusing on the complete cycle of innovation – from world-class basic research and technology development through leading-edge demonstrators and prototypes to product functions and commercialization. Based in Kaiserslautern, Saarbrücken, Bremen, and Berlin, DFKI ranks among the important "Centers of Excellence" worldwide. An important element of DFKI's mission is to move innovations as quickly as possible from the lab into the marketplace. Only by maintaining research projects at the forefront of science DFKI has the strength to meet its technology transfer goals. The research centre is set up as a public-private partnership, i.e., it is jointly owned by some of the largest German industrial corporations, two universities, three federal states and the Fraunhofer Society. DFKI benefits from interaction with the Informatics Faculties of the Universities of Kaiserslautern, Saarbrücken, and Bremen and in turn provides opportunities for research and PhD thesis supervision to students from these universities, which have an outstanding reputation in Computer Science and in Computational Linguistics. The Lab has vast experience in modeling areas in such a way that they are highly amenable to machine learning and classification tasks. Applications centered around recognizing different types of documents have been successfully fielded. As a sign of successful technology transfer, several spin-off companies have emerged from the LT Lab. Xtramind (now Attensity Europe), Acrolinx, and Yocoy each focus on different aspects of statistical language processing and show the practical relevance and strength of the solutions developed within the Lab.

**Main tasks in Project**

Relevant to the project is DFKI's team experience in computational linguistics, multi-lingual natural language parsing, artificial intelligence, text analysis, information retrieval. DFKI's main focus within the proposal will be on the text analysis of clinical documents and the mapping of parsed text and structured data to ontologies.



The **Department of Biomedical Engineering at Linköping University** was founded in 1973, and is Sweden's oldest department in Biomedical Engineering. The department lies at the epicenter of Linköping's medical, engineering, and biological research, and is the only department from the technical faculty that is physically situated in the University hospital area. This location ensures tight connections between healthcare and research, in areas including e.g. biomedical devices, medical informatics, and mathematical modelling for decision support systems. Recent EU projects of relevance to this project where LIU has participated include ASSESS CT, SemanticHealthNet, DebugIT, as well as SemanticMining, which also was coordinated from LIU. Of particular importance to this project LIU's participation in the DebugIT project included the delivery of semantically normalized health data from the Swedish ICU quality registry. The health informatics group, headed by Daniel Karlsson, also participates in national and international standardization projects, for example leading the development of models for health care observation SNOMED CT and other health care terminologies. The integrative systems biology group, headed by Gunnar Cedersund, lies at the epicenter of almost all biomedically relevant modelling done at Linköping university, and features 7 Ph.D. students, one associate and one assistant professor, and 5-10 M.Sc. and internship students.

**Main tasks in Project**

Linköping University participates in 2 main roles in the project: 1. as contact to and interfacing with clinical studies and data sources and 2. as researchers in the fields of predictive modeling. Linköping University will contribute with data from 4 extensive clinical trials: CARDIPP (670 patients), the Swedish Stroke Registry (>400.000 stroke events), SCAPIS (30.000 patients), and UKBIOBANK (500.000 patients). We will cooperate for integration of multi-level short-term models with long-term models, to create truly multi-scale mechanistic models. These models can simulate the time-evolution of key physiological parameters in response to treatments, such as diet, exercise, and certain medications, in all four clinical phases: prevention, acute, recovery, and re-integration. In collaboration we will be continuing the work on the combination of all available biomarkers into risk prediction models, using Bayesian graphical networks and machine learning.