

Research data management pilot project proposal 2016 – cover sheet

Acronym	ActionableBiomarkerDK		
Title of proposal	Danish Infrastructure for Storing and Updating Actionable Biomarkers in Human Disease		
Proposal summary	<p>Biology and medicine have in the last decades been turned into data-driven disciplines – a change that at the same time affects healthcare and the pharmaceutical industry. It has become financially feasible to sequence entire human genomes and proteomes and investigate how human individuals differ from one another in terms of sequence variation. The proposed infrastructure address the important problem of turning a complete genome into a small set of actionable biomarkers that can be used in diagnostics, in dosage calibration of drugs and in the context of drug development. A major data management problem in that respect is that biomarker information grows significantly over time and that it resides in many different repositories world-wide in a fragmented way. The Danish Infrastructure for Storing and Updating Actionable Biomarkers in Human Disease aims to offer a comprehensive solution to this problem, to avoid duplication of effort and to build competence and capacity in this cross-disciplinary field at the interface between biology, genetics, biotechnology, bioinformatics and high performance computing. The stakeholder groups are considerable across academia, healthcare and industry. The pilot project builds on existing DeIC supported supercomputer resources at DTU/UCPH and SDU and aims to turn many existing workflows and resources into robust infrastructure using secure private cloud technologies. The pilot phase will not work with person-sensitive data that will require additional legal and ethical approval, however, a permanent infrastructure is meant to handle, in a highly secure manner, such data. The need in Denmark is likely to be large as the Danish Regions have announced projects in the 100,000+ patient range, projects that with time are likely to reach much larger parts of the Danish population.</p>		
Part(s) of the research data life cycle which are addressed	The pilot projects address all phases, but in particular part 5, “Data deposited”, and part 6, “Paper + data published”, of the life cycle model.		
Main area (tick more than one if applicable)	<input type="checkbox"/> Humanities <input type="checkbox"/> Social Sciences	<input checked="" type="checkbox"/> Health Sciences <input checked="" type="checkbox"/> Natural Sciences	<input checked="" type="checkbox"/> Technical sciences
Types of infrastructure	<input checked="" type="checkbox"/> New infrastructure <input type="checkbox"/> Already realised data management infrastructure, where the proposal concerns significant modifications or extensions <input checked="" type="checkbox"/> Already realised e-infrastructure, for example national HPC installations, where the proposal concerns significant data management related extensions <input type="checkbox"/> Digital tools that support and promote the use of already realised data management infrastructure		

Proposers

Main proposer and contact (title, name, institution, mobile number and email address)	Prof. Søren Brunak, Ph.D. Research Director at Novo Nordisk Foundation Center for Protein Research, UCPH +45 2067 2477, soren.brunak@cpr.ku.dk
Co-proposers (title, name and institution)	Peter Løngreen, M.Sc., Head of Supercomputing, Dept. of Systems Biology, DTU Ole Nørregaard Jensen, prof., head of department, Department of Biochemistry & Molecular Biology, SDU Stig Ejdrup Andersen, Consultant, Head of Unit, Clinical Pharmacology, Zealand University Hospital, Roskilde

Project duration and budget

Project start	Project end
01.05.2016	30.04.2018

Total budget	Self-financed share	Share applied for
4000000 DKK	2000000 DKK	2000000 DKK