



Case Study: Applying Particle Physics expertise in big data to Next Generation Proteomics

STFC Innovation Fund

This project was funded by STFC through the Impact Accelerator Account administered by the Innovation Fund. The Innovation Fund aims to facilitate the acceleration of research along the pathway to impact or knowledge exchange. The fund provides flexible support (e.g. develop prototypes, establish new collaborations and enable secondments) for a wide variety of projects from across QMUL.

The Partnership

Using his Innovation Fund award, Dr Jon Hays developed collaborations with academics who specialise in proteomics in both the School of Biological and Chemical Sciences and the School of Medicine and Dentistry at QMUL as part of this project. The collaborators helped inform the way particle physics analysis techniques could be applied to liquid-chromatography mass spectrometry (LC-MS) based proteomics.

The Idea

Dr Hays is a lecturer in Particle Physics and his research has covered a wide variety of aspects of modern particle physics experimentation including: detector design, testing and operation; software development; and data analysis and statistics.

Jon is interested in ways to apply the techniques derived from the challenges of data analysis at the Large Hadron Collider to areas beyond particle physics in this instance high throughput proteomics.

Proteomics technologies have proved effective in the discovery of potential biomarkers and drug targets. In particular, mass spectrometry-based proteomics enables the robust recognition and quantification of nearly all endogenous proteins within a single experiment.

However, current methods of mass spectrometry proteomics are hampered by bottlenecks in data analysis, an issue exacerbated by ever increasing volumes of data recorded with increasingly sophisticated mass spectrometers.

The Outputs

- Analysis software has been developed as a demonstration of what can be achieved through the optimization of data storage.
- This project suggests huge gains can be made in analysis run times.
- These prototypes could potentially revolutionise the proteomics field leading to 'Next Generation Proteomics'.
- Collaborations with companies are now being explored who wish to develop these prototypes towards software products.
- These techniques could also be applied to other industry sectors with big data challenges.

Partner quote:

"Our fast feature finding project was born from a coffee-break conversation with colleagues in the School of Biological and Chemical Sciences at Queen Mary.

The funding that we have received through the STFC Impact Accelerator, by funding an interdisciplinary postdoctoral research position, has enabled us to develop our ideas into a prototype computer application. This brings ideas and techniques from particle physics to dramatically decrease processing times for large proteomics data sets."

Dr Jon Hays, QMUL