April 1st LECTURE

The Fourth Paradigm - Data-Intensive Scientific Discovery and Open Science

Tony Hey
Chief Data Scientist at Science and Technology Facilities Council, UK

Abstract

There is broad recognition within the scientific community that the ongoing deluge of scientific data is fundamentally transforming almost all areas of academic research. A wide variety of researchers —from scientists and engineers to social scientists and humanities researchers – now require tools, technologies, and platforms that seamlessly integrate into standard scientific methodologies and processes. "The Fourth Paradigm" refers to the data management techniques and the computational systems needed to manipulate, analyze, visualize, and manage large amounts of research data. This talk will illustrate the challenges researchers will face, the opportunities these changes will afford, and the resulting implications for data-intensive researchers. In addition, the talk will discuss the issues of open access, open data and open science and research reproducibility.

Short Bio

Tony Hey began his career as a theoretical physicist with a doctorate in particle physics from the University of Oxford in the UK. After a career in

physics that included research positions at Caltech and CERN, and a professorship at the University of Southampton in England, he became interested in parallel computing and moved into computer science. In the 1980's he was one of the pioneers of distributed memory message-passing computing and co-wrote the first draft of the successful MPI message-passing standard.

After being both Head of Department and Dean of Engineering at Southampton, Tony Hey escaped to lead the U.K.'s ground-breaking 'eScience' initiative in 2001. He recognized the importance of Big Data for science and wrote one of the first papers on the 'Data Deluge' in 2003. He joined Microsoft in 2005 as a Vice President and was responsible for Microsoft's global university research engagements. He worked with Jim Gray and his multidisciplinary eScience research group and edited a tribute to Jim called 'The Fourth Paradigm: Data-Intensive Scientific Discovery.' Hey left Microsoft in 2014 and spent a year as a Senior Data Science Fellow at the eScience Institute at the University of Washington. He returned to the UK in November 2015 and is now Chief Data Scientist at the Science and Technology Facilities Council.

In 1987 Tony Hey was asked by Caltech Nobel physicist Richard Feynman to write up his *'Lectures on Computation'*. This covered such unconventional topics as the thermodynamics of computing as well as an outline for a quantum computer. Feynman's introduction to the workings of a computer in terms of the actions of a 'dumb file clerk' was the inspiration for Tony Hey's attempt to write *'The Computing Universe*', a popular book about computer science. Tony Hey is a fellow of the AAAS and of the UK's Royal Academy of Engineering. In 2005, he was awarded a CBE by Prince Charles for his 'services to science.'