

Data, storage & information management: collaboration at the frontiers of science and technology

*Jakub Moscicki*¹

¹The European Organization of Nuclear Research (CERN)

LHC — the world’s largest scientific instrument ever built — is a collaboration at the frontiers of science and technology. In addition to hosting the LHC accelerator and experiments, CERN runs an entire accelerator complex and hosts numerous other experimental facilities that have been developed for many decades. CERN is a place of collaboration between generations of scientists and between the nations working together towards a shared and peaceful goal: understanding of the very fundamental laws of nature which underpin the very existence of the universe and the humans within it. CERN is also a place of ground-breaking scientific discoveries, such as the Higgs Boson, as well as ground-breaking technical inventions, such as the World Wide Web.

Specifically, information technology for experiments — in all aspects of data&information management, storage and computing — is a challenging task due to the size of generated data, amount of processing power and distributed landscape of research collaborations. It is a coordinated effort of researchers, engineers and software developers across hundreds of universities and research institutes. As a result, massive processing infrastructure called the Worldwide LHC Computing Grid has been developed — a global network of data processing centers to support the LHC scientific programme. In addition, innovative services and open-source technologies for data storage, data analysis, information management and online collaboration have been created at CERN. In this presentation we will look into the larger picture and selected challenges of data storage and data&information management for CERN’s experiments.