



Improving data analytics for better science

Last week, representatives of major research institutions and leading IT companies came to CERN for a data analytics workshop organized by CERN openlab.

The

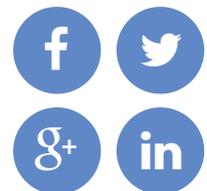


CERN Data Centre. Image courtesy CERN.

CERN Data Centre processes a staggering **1000 terabytes of data every day**, which is the equivalent of around 210,000 DVDs. And, at peak times, **the**

Posted on FEB 26
2014 10:00AM

Share this
story



Republish

Tags

Worldwide Large Hadron Collider (LHC)

[Computing Grid](#) may transfer as much as 10 gigabytes of data from its servers every second.

However, CERN is not the only research organization dealing with such 'big data'. Ever-cheaper sensing and data-storage technologies are driving an explosion in data across many fields of science, as well as for businesses of all sizes.

Yesterday, [CERN openlab](#) invited research institutions and leading IT companies to participate in a workshop focused on the latest advances in data analytics. "Data analytics brings the promise of further increasing the efficiency of LHC operations, accelerating physics analysis, and preparing us for even more ambitious initiatives," says Alberto Di Meglio of CERN openlab. "[The next CERN openlab phase](#) will lead us to the start of the next planned shutdown of the LHC (LS2) in 2018. This is an excellent opportunity for us to investigate new technical challenges with commercial partners and other European research laboratories."

Presentations were given at the event by representatives of the invited IT companies. Africa Perianez of [the German Meteorological Service](#) spoke about how big data is driving advancements in weather prediction, while at the same time posing analytical challenges. Meanwhile, Nenad Buncic of [the Swiss Federal Institute of Technology in Lausanne \(EPFL\)](#) gave a presentation on [the Human Brain Project](#), an ambitious attempt to improve understanding of the brain by building a completely new type of computing infrastructure.

analytics

big data

CERN

CERN openlab

data

Whereas at CERN the vast majority of scientific data is generated centrally by [the LHC experiments](#), the data for the Human Brain Project is produced by a host of smaller, distributed laboratories. "A lack of data sharing between labs is probably the biggest challenge in neuroscience," says Buncic. "We are incentivizing labs to share data... but we need to make sure that data conforms to standards and that researchers speak a common vocabulary."

Salim Ansari from [the European Space Research and Technology Centre](#) also gave a presentation on the vast amounts of data being generated by the European Space Agency's fleet of [Earth-observing](#) and [space-science](#) satellites. He discussed how the operational efficiency of satellites can be improved by analyzing orbital data and explained how the concept of 'predictive analytics' is changing how astrophysical data is handled. "Analytics can play a major role in simplifying scientific research," says Ansari. "Discovery tools can be automated, thus freeing the scientist from having to carry out much of the search for knowledge."

Finally, CERN's Johannes Gutleber gave a brief presentation [on long-term studies into a potential successor to the LHC](#). While any such machine wouldn't begin operations for at least another 20 years, it is important, says Gutleber, to think now about the computing infrastructures that will be required to analyze the huge data output. "We need to think about how the computing infrastructures are going to evolve. We need to build infrastructures that are going to survive the trends

and fashions of computing today, while being flexible enough to incorporate future advancements in technology."

- Andrew Purcell

Join the conversation

[Contribute](#)



Do you have story ideas or something to contribute? **Let us know!**

OUR UNDERWRITERS

Thank to you our underwriters, who have supported us since the transition from International Science Grid This Week (iSGTW) into Science Node in 2015. We are incredibly grateful.

[View all underwriters](#)

CATEGORIES

[Advanced computing](#)

[Research networks](#)

[Big data](#)

[Tech trends](#)

[Community building](#)

CONTACT

Science Node

Email:

editors@sciencenode.org

Website:

sciencenode.org



Copyright © 2022 Science Node™ | [Privacy Notice](#) | [Sitemap](#)

Disclaimer: While Science Node™ does its best to provide complete and up-to-date information, it does not warrant that the information is error-free and disclaims all liability with respect to results from the use of the information.