CERN Accelerating science

CERN Bulletin

News Articles   **Training**   **Announcements**   **Events**   **Staff Association**

english | français

Issue No. 24-25/2014 - Monday 9 June 2014
No printable version available - Subscribe:

Subscribe by RSS

Subscribe by RSS for this category only

A SEARCH ENGINE TO FIND THE BEST DATA?

What if you could see your experiment's results in a "page rank" style? How would your workflow change if you could collaborate with your colleagues on a single platform? What if you could search all your event data for certain specifications? All of these ideas (and more) are being explored at the LHCb experiment in collaboration with Internet giant Yandex.
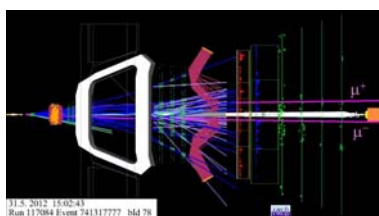
An extremely rare $B^o{}_s \rightarrow \mu\mu$ decay candidate event observed in the LHCb detector.

As the leading search provider in Russia, with over 60% of the market share, Yandex is to East what Google is to West. Their collaboration with CERN began back in 2011, when Yandex co-founder Ilya Segalovich was approached by then-LHCb spokesperson Andrei Golutvin. "Just as Yandex's search engines sift through thousands of websites to find the right page, our experimentalists apply algorithms to find the best result in our data," says Andrei Golutvin. "Perhaps the technique used to rank webpages could also be applied to ranking data?"

It was an idea that Yandex decided to put to the test, and they are now collaborating with the Organization under the auspices of CERN openlab. Yandex has developed an event search and selection algorithm in collaboration with the LHCb experiment. This algorithm uses the patented MatrixNet machine learning technology, building upon previous experiences to create more relevant results. The algorithm appears particularly suited to searching for extremely rare events (like the one shown in the picture) and is now being used in several analyses to help improve selection performance, challenging standard statistical techniques.

However, Yandex's most useful development came as a surprise: "We found that it was not the algorithm itself that gave an advantage, but rather the user-friendly interface we developed to go with it," says Andrey Ustyuzhanin, Yandex employee and member of the LHCb collaboration. "It allows scientists to easily interact as they work together on the same data set. The platform is a functional Wikipedia, if you will, where you can perform complicated computational tasks and share the results with others." Furthermore, the interactive platform is not limited to Yandex algorithms, as any event selection process can be used.

Although still at an early stage, Yandex encourages CERN's experiments to explore the potential of their platform. "Such a platform can be a much more efficient way to collaborate," explains Ustyuzhanin. "By uniting the analysis process, in the spirit of open science, scientists can share ideas to improve codes or even re-use the same analysis software on a different data set. Even if our particular platform isn't used, our hope is that more experiments consider this virtual model of collaboration."

But that's not all. As Yandex's collaboration with the Organization continues to expand, more and more avant-garde ideas are being explored. Could we create a search tool that scans data for a particular type of event? How about a platform that demonstrates how

results can improve based on different analyses? Could we automate the improvement of analysis algorithms to reflect ever-changing conditions? All that and more is on the drawing board.

by Katarina Anthony
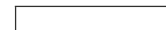
CERN BULLETIN INFO

Past issues

Contact us

RSS feed

Sign up

MORE

CERN Courier

Staff Association

Press Office

Training

CERN & HEP events

Clubs

© Copyright CERN 2014
CERN Publications, DG-CO