

BIG DATA

Cern: 'We generate 1PB each second. Here's what that's taught us about big data'

The big data pioneer on how to use machine learning to mine gold from your company's data stores.

By Nick Heath | February 21, 2017, 8:16 AM PST

Recommended Content:

Training: Data Science and Machine Learning with Python!

If you've got some programming or scripting experience, this course will teach you the techniques used by real tech industry - and prepare you for a move into this hot career path. This comprehensive course includes 68 alr



Cern's experiments probing the fundamental nature of the universe create a lot of data—roughly four times that held in the US Library of Congress—every second.

The EU research institution has a long history of generating more data each day than most companies do in a lifetime, during its experiments colliding particles travelling at close to the speed of light.

While the bulk of this data is thrown away, Cern shares the remaining information with researchers around the world and has data on the workings of its particle accelerators—the most

More about Big Data

[An inside look at why Apache Kafka adoption is exploding](#)

famous of which is the Large Hadron Collider (LHC)—stretching back decades.

Manuel Martin Marquez, lead data scientist at Cern, talked to TechRepublic about how, over the past five years, Europe's largest particle physics lab has begun employing machine learning to discover new ways to filter noise in its experiments and to keep its accelerators running smoothly.

Marquez says Cern's machine learning models are already helping engineers carry out predictive maintenance on the cryogenics system that cools the superconducting magnets used in the LHC to close to absolute zero

(<https://home.cern/about/engineering/cryogenics-low-temperatures-high-performance>), reducing the risk of a costly failure.

(<http://www.techrepublic.com/article/inside-look-at-why-apache-kafka-adoption-is-exploding/>)

Video: How Softbank's Pepper robot captures analytics on human interactions

(<http://www.techrepublic.com/article/how-softbanks-pepper-robot-captures-analytics-on-human-interactions/>)

Hiring kit: Data architect

(<http://www.techproresearch.com/do-kit-data-architect/>)

Subscribe to our Big Data Essentials newsletter

(<http://www.techrepublic.com/newsle>

With Cern's research budget being largely flat, finding ways to make operations more efficient is increasingly important to the institution, he said, stressing how the facility had benefited from identifying machine learning projects with a clear scope that provide 'quick wins'.

Watch the video to hear how Cern is using big data and machine learning, and what your company can learn from an institution used to making sense of vast amounts of data.

For more big data advice and tips, subscribe to our Big Data Essentials newsletter.

SUBSCRIBE

Also see

- **Before Big Data, clean data** (<http://www.techrepublic.com/blog/big-data-analytics/before-big-data-clean-data/>) (TechRepublic)
- **Farm out big data chores so employees can focus on analytics** (<http://www.techrepublic.com/article/farm-out-big-data-chores-so-employees-can-focus-on-analytics/>) (TechRepublic)
- **Leadership challenges of a data cleansing effort** (<http://www.techrepublic.com/blog/tech-decision-maker/leadership-challenges-of-a-data-cleansing-effort/>) (TechRepublic)
- **6 myths about big data** (<http://www.techrepublic.com/article/6-myths-about-big-data/>) (TechRepublic)
- **Data to analytics to AI: From descriptive to predictive analytics** (<http://www.zdnet.com/article/data-to-analytics-to-ai-from-descriptive-to-predictive-analytics/>) (ZDNet)



About Nick Heath

Nick Heath is chief reporter for TechRepublic. He writes about the technology that IT decision makers need to know about, and the latest happenings in the European tech scene.

Recommended

Promoted Links by Taboola

An Easy Way to Restore Your Vision

Vision Protocol Guide

The Flashlight That Won The Best of 2016

Gladiator Flashlights

360 Video Experiences Of Abu Dhabi-You'll Wonder Why You Haven't Visited Yet

CNN for Abu Dhabi Tourism

Video: Cern's data expert on how to get the most from your company's big data

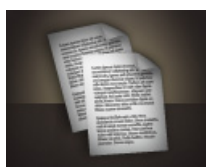
Google's DeepMind 'Lab' opens up source code, joins race to develop artificial general intelligence

Raspberry Pi in 2017: New boards, new OSes and more

WHITE PAPERS, WEBCASTS, AND DOWNLOADS

White Papers // From Riverbed

The Business Value of Riverbed SteelHead



IDC conducted detailed interviews to understand the impact of optimizing the delivery of business applications with Riverbed SteelHead. They reported SteelHead provided significant value by cost effectively improving network and application performance.

DOWNLOAD NOW

Downloads // From SolarWinds

Network troubleshooting gets real

DOWNLOAD NOW

Training // From Udemy

UX & Web Design Master Course: Strategy, Design, Development

DOWNLOAD NOW

Webcasts // From Softerra

Adaxes. Active Directory at Full Throttle

[LEARN MORE](#)[Downloads](#) // [From SolarWinds](#)

Find Your Top 5 Storage Performance Issues in Minutes

[DOWNLOAD NOW](#)

RECOMMENDED FOR YOU

Data Science and Machine Learning with Python!

Training provided by Udemy

[DOWNLOAD NOW](#)

EDITOR'S PICKS



Google Fiber 2.0 targets where it will stage its comeback, as AT&T Fiber prepares to go nuclear



How Mark Shuttleworth became the first African in space and launched a software revolution



Inside Amazon's clickworker platform: How half a million people are paid pennies to train AI



How K9s became a secret weapon for solving high-tech crimes

RECOMMENDED FOR YOU

4 Reasons to Use All-Flash Storage to Consolidate Cloud Applications

White Papers provided by [Pure Storage](#)

DOWNLOAD NOW

RECOMMENDED

Sponsored Links by Taboola

Crazy Cheap Tracking Device - Selling Like Hotcakes!

TrackR Bravo

Grepolis: Are you ready to conquer Ancient Greece? Already 35 million players!

Grepolis - Online Free Game

RECOMMENDED FOR YOU

Democratisation Of All Flash Storage

White Papers provided by [Pure Storage](#)

DOWNLOAD NOW

Recommended Content:

White Papers: The True Costs of Public versus Hosted Private Cloud Infrastruc

If you're considering moving highly virtualized workloads off-premises, use the lessons learned here to compare a private cloud at a colocation facility with the costs of renting a public cloud.