

Trusted Analytics Platform: Simplify Big Data Analytics with Open Source Software

Sept. 30, 2015 — At Strata + Hadoop World in New York City, Intel discussed the continued growth of an ecosystem around the Trusted Analytics Platform (TAP). TAP is an open source project that Intel developed to make it easier for developers and data scientists to deploy custom big data analytics solutions in the cloud as well as reduce development costs and time to market. The company also disclosed pilots with Penn Medicine*, Icahn School of Medicine at Mount Sinai*, and Oregon Health & Science University* (OHSU).

Every enterprise can gain competitive advantage, increase operational efficiency, or improve customer loyalty by analyzing data. However, many organizations are hindered by the lack of data science skills required to implement a big data analytics solution and manage the complexity of a big data infrastructure.

TAP includes the necessary tools, algorithms and engines to make it easier for developers to collaborate with data scientists in a shared environment to conduct advanced analytics. TAP includes open source software with hardware-enhanced performance and security features. The platform provides an end-to-end solution with three key layers:

- Data layer that includes Apache Hadoop*, Spark* and other data components optimized for performance and security
- Analytics layer that includes a data science tool kit to simplify model development and an extensible framework to generate predictive APIs
- Application layer that includes a managed runtime environment for cloud-native apps

Industry Deployments

TAP has been tested by data scientists in pilot projects at various organizations and is ideal for a range of industries including healthcare, retail, oil and gas, energy, and industrial.

- **Penn Medicine:** Penn Medicine is working with Intel to advance healthcare analytics with a solution that combines patient vitals, lab records and medications to develop predictive models that can forecast risk of disease or readmission. Using TAP, Penn Medicine will be able to build better models for predicting risk and is currently evaluating the integration of TAP with Penn Signals.
- **Icahn School of Medicine at Mount Sinai:** The Icahn School of Medicine at Mount Sinai is exploring the applicability of data science advances to therapeutic drug discovery through the program DrugGraph. Using TAP, the school can greatly reduce the complexity of the big data analytics process and can lead to faster, less expensive discovery of new drug therapies; improve patient outcomes by reducing toxic drug reactions; and reduce cost of treatment by predicting novel uses for existing compounds.
- **Oregon Health & Science University (OHSU):** Intel and OHSU are developing the “Collaborative Cancer Cloud,” a big data analytics solution for precision medicine that allows hospitals to securely share patient genomic data to enable potentially lifesaving discoveries. OHSU is using TAP to securely manage patient data gathered from wearable devices, labs and surveys in a

central location. The deployment takes in more than 3 million records every day with more than 360 million records already in the system. OHSU data scientists are using TAP to analyze the data to find new ways of determining a subject's overall health.

TAP features support from a range of partners including Accenture*, Arcadia*, Cask*, Cloudera*, Conduce*, DataRobot*, DeepSense*, H2O, Infosys*, Koverse*, Objectivity*, Silicon Valley Data Science*, SkyMind*, Talena*, and Typesafe*. TAP is also tested for deployment on a choice of public cloud infrastructures including Amazon Web Services*, Rackspace* and OVH.com*.

Supporting Quotes

"Penn Medicine is working with Intel to revolutionize health care by using advanced analytics to predict patient illness before clinicians know to look," said Mike Draugelis, chief data scientist, Penn Medicine. "Today we are always iterating, always coming back to say can't we improve on this model, and asking if it meets the needs of our clinicians. With open source and cloud-based solutions like the Trusted Analytics Platform, we have the potential to provide our clinicians and data scientists with a platform for development and testing where we can quickly build, explore, and deploy our analytic applications."

"At Mount Sinai, we're good at building big data algorithms for graph modeling, but to get these to run at scale is very difficult," said Dr. Joel Dudley, director of Biomedical Informatics, Icahn School of Medicine at Mount Sinai. "Using the Trusted Analytics Platform with Intel really solved that problem, enabling us to implement these algorithms in a very scalable way. That's a unique value proposition. We could stay focused on the scientific aspect of our DrugGraph project without worrying about scaling and the limitations of the data size."

"OHSU and Intel are committed to working together to drive scientific progress in understanding the origins of disease," said Adam Margolin, Ph.D., Director Computational Biology at Oregon Health & Science University. "To do that, we need secure, robust information tools that can handle the mind-boggling volumes of data generated in the process. The Trusted Analytics Platform is an example of the kind of next generation tool that will be essential to our work."

Intel and the Intel logo are trademarks of Intel Corporation in the United States and other countries.

*Other names and brands may be claimed as the property of others.