The Computerworld Honors Program

Honoring those who use Information Technology to benefit society

Final Copy of Case Study

Status:
Laureate

Year:
2013

Organization Name:
Mitsui Knowledge Industry Co., LTD.

Organization URL:
http://www.mki.co.jp/english/

Project Name:
POC for accelerating cancer research through real-time, big data genome analysis

Please select the category in which you are submitting your entry:
Innovation

Please provide an overview of the nominated project. Describe the problem it was intended to solve, the technology or approach used, how it was innovative and any technical or other challenges that had to be overcome for successful implementation and adoption. (In 300 words or less.)

Mitsui Knowledge Industry (MKI) strives to deliver advanced medical treatment to as many patients as possible by introducing real-time, big data genome analysis into healthcare. This capability is becoming particularly useful to healthcare researchers. Now that DNA sequencers have become highly advanced, the amount of time and cost for DNA extraction has become significantly smaller. It used to cost $1 million for genome extraction whereas in the near future, the cost could be a low as $1,000. This will enable more DNA research, but genome analysis can still require a several days to complete and deliver results to patients. To accelerate the analysis process so that doctors and their patients

can know the results of DNA analysis faster, MKI partnered with SAP and its HANA in-memory database technology. SAP HANA enables MKI to analyze big data quickly to help drive rapid innovation. Tested against a conventional database running the same queries, HANA improved query responses by as much as 400,000 times. Applying SAP HANA to bio-science efforts gives MKI a reliable real-time platform for supporting mission critical core systems that analyze DNA. MKI has integrated SAP HANA with Hadoop and R from Revolution Analytics to create a single, real-time data platform prototype. Genome analysis is executed by SAP HANA while data mining is managed by R and HANA jointly. Hadoop manages the data pre-processing prior to data analysis and also handles high-speed storage. With this technology combination prototype, MKI has found a way to shorten the genome analysis time from several days to only 20 minutes. Having this capability is highly beneficial to the healthcare industry and patients. If doctors can deliver analysis by end of the day before patients leave the hospital, this can revolutionize cancer treatment.

When was this project implemented or last updated? (Please specify month and year.) Has it incorporated new technologies and/or other innovations since its initial deployment? (In 300 words or less.)

Completed in 2012.

Is implementation of the project complete? If no, please describe the project's phases and which phase the project is now in. (In 300 words or less.)

MKI has finished its pilot case and is now in the process of building a business case

Please provide at least one example of how the technology project has benefited a specific individual or organization. Feel free to include personal quotes from individuals who have directly benefited from the work. (In 300 words or less.)

"We're entering what I call the information age of genomics where the greatest challenges actually have to do with data interpretation, how we manage and think about genetic data, and how we use that data to improve human lives. We're fundamentally very good at writing statistical software and thinking mathematically about the problem we want to solve. But we're not always as good at implementing technology to support these processes. That is why we are so excited to work with SAP and its HANA product, which could revolutionize the speed at which we analyze the data. The ability to rapidly take a model and use HANA to apply it to a genome-wide data set seems like a dream come true. The democratizing of genome sequencing is critical. If it comes fast and cheap, then

there is no barrier -- people all over the world can do this." -- Dr. Carlos Bustamonte, Professor of Genetics, Stanford University "Genomic DNA analysis in real-time transforms how we enable comprehensive patient care to fight against cancer. SAP HANA will be the mission-critical and reliable data platform to make real-time cancer analytics into a reality." -- Yukihisa Kato, Director & Executive Officer, CTO, Research and Development Center, MKI

Would this project be considered an innovation, a best practice or other notable advancement that could be adopted by or tailored for other organizations and uses? If yes, please describe that here. (In 300 words or less.)

The genome analysis capabilities that MKI has developed by integrating SAP HANA technologies with Hadoop and R creates tremendous new possibilities for the entire healthcare research industry. Many researchers face the challenge of trying to reconcile huge volumes of genome data. Their ability to identify biomarkers is limited, and the pace to associate genes with diseases can be very slow. But with the data analysis process created by MKI, researchers can now more expediently identify gene biomarkers that they would never have been able to consider before. This allows them to consistently associate target genes with diseases and discover cancer-causing abnormalities in minutes rather than days. Doctors and their patients can now have peace-of-mind earlier than before by knowing sooner the status of a prognosis. Even more importantly, researchers and doctors may be able to treat and perhaps cure more patients and improve their quality of life than was previously possible.

If there are any other details that the judges should know about this project, please note them here. (In 300 words or less.)

Video: https://www.youtube.com/v/U6dA41_ulxo?autoplay=1&rel=0 Video: https://www.youtube.com/v/YfKq6oohUFM?autoplay=1&rel=0