100 Most Promising Big Data Solution Providers 2016

'Big Data' is no more a buzzword. Now that organizations have already put their wide arms around Big Data, the next impediment lies in refining the data to bring out insightful and meaningful results. With every bit of '0' and '1' being harnessed as meaningful "collections," enterprises are sure to achieve complete utilization of their concrete information, and attain significant outcomes. Another trend that has gained notable traction is capitalization of cloud for storing invaluable sets of data. While majority of enterprises consider it as a risky venture owing to the concerns of security, others embark upon it for the sake of business agility. With myriad of such transformations, enterprises are confident to embrace innovative ways to hold together the regulations of corporate world, and ensure their participation

in the realm of Big Data. In order to uphold a fine balance, it has become critical for the CIOs to choose proper technology and select best vendors that are at the forefront of effectively tacking the impediments across the Big Data realm. To help them accomplish their objective, CIO Review presents "100 Most Promising Big Data Solution Providers 2016." A distinguished panel comprising of CEOs, CIOs, VCs, and analysts including CIO Review's editorial board has decided the top Big Data Solution Providers from over thousand companies. The companies featured in this list provide a look into how their products work in the real world, so that you can gain a comprehensive understanding of the solutions available and how they stand against competition. We present to you CIO Review's 100 Most Promising Big Data Solution Providers 2016.

Rasdaman

recognized by **CIOReview** *magazine as*

100 MOST PROMISING BIG DATA SOLUTION PROVIDERS 2016

An annual listing of 100 companies that are at the forefront of providing bigdata solutions and impacting the marketplace

Company:
Rasdaman

Key Person:
Peter Baumann,
CEO

Description:
Agile analytics on massive multidimensional arrays such as sensor, image, simulation, and statistics data

Website:
rasdaman.com

---

rasdaman
Pioneering Agile Array Analytics

In our quest for disruptive big data technology we crossed the Atlantic to find German-based rasdaman, which truly has the potential for becoming a game changer right tomorrow. Turns out the Caribbean sounding name actually stands for "raster data manager".

Peter Baumann, inventor CEO of the company, early on has recognized a gap in big data analytics: missing support for massive matrices and datacubes, also known as multi-dimensional arrays. We find these arrays everywhere: in business, such as stock risk analysis and OLAP; in Life Science; in exploration data; and in industrial simulation. Even analyzing large graphs, like the Facebook One, can be done through array operations. In fact, datacubes are recognized as a key paradigm that is simple to grasp and easy to query. As we see the potential for array engines, we wonder: How could that go undetected?

"SQL has missed a train," reveals Baumann, "its designers didn't like what didn't fit table world. NoSQL ultimately is a response to this, and array engines are part of it." Left alone, data centers and companies created homegrown solutions which typically ended up in silo solutions with a rapidly eroding architecture. Rasdaman, conversely, started with a generic, clean slate architecture. The core idea is as simple as compelling: combine the flexibility of SQL with the power of array manipulation.

"Good old SQL is a great tool with its flexibility of 'any query, anytime' – just not on datacubes," explains Baumann. Constructed as a NoSQL database, rasdaman connects itself with standard SQL and extends it with array operations. This opens up additional vistas: suddenly the big, clumsy data and the small, smart metadata can be queried together. Closing

this age-old gap liberates users from keeping in mind different access and retrieval techniques – a single common information space emerges.

rasdaman can handle really large volumes quickly, and can combine data sources distributed on planetary scale. We found installations exceeding 130 Terabyte at publicly accessible services, and researchers at superscale data centers, like the European Space Agency (ESA) and the European Centre for Medium-Range Weather Forecast (ECMWF) with its 87 Petabyte climate archive, are feeding rasdaman to go beyond the Petabyte frontier.

> " With rasdaman's unique adaptive data partitioning and parallelization, data cubes are analyzed and combined in a straightforward and ultrafast manner "

How rasdaman does this job? "The whole architecture is crafted from scratch, optimized for array handling," outlines Baumann. There is a series of strong optimizations which altogether make rasdaman fast. Adaptive data partitioning and distribution is one element, augmented with effective compression on dense and sparse datacubes; intelligent processing utilizes all silicon it can get hold of, within and across nodes and even data centers while respecting security. "Query optimization and parallelization is done individually for each incoming query, as opposed to static parallelization like in Spark," describes Baumann. And confronts us with a demo where we see a Terabyte analyzed in less than 100 milliseconds.

Peter Baumann

And how do data get in? Ingestion pipelines are configured rapidly – or you avoid copying at all, as rasdaman can be adjusted to tap directly into any preexisting archive structure. "This allows us to be uniquely customer oriented and fast while accommodating any change request swiftly," says Baumann. "Customers deploy rasdaman as a cost-effective solution for value-adding services, essentially enabling users to build their own data product on the go," he adds.

This lead got rasdaman on the radar of standardization bodies. ISO is finalizing an SQL extension called MDA (for Multi-Dimensional Arrays) – using rasdaman as its blueprint. "In 2017, ISO SQL/MDA will become the standard for large-scale array analytics, and rasdaman will be the first fully scalable implementation," notes Baumann.

"The net effect is a tremendous boost in productivity of analysts, scientists, and engineers, achieved on commodity hardware and clouds," concludes Baumann. We follow this, and wholeheartedly include this pick in our list of 100 Most Promising Big Data technologies. **CR**