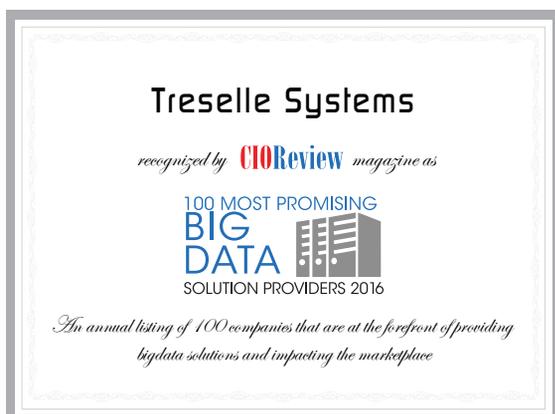# 100 Most Promising Big Data Solution Providers 2016

'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.

Treselle Systems

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:
Treselle Systems

Key Person:
Sudharsan Madabusi,
President and CEO

Description:
Provides Big Data strategy, architecture, product development, and project execution services to assist customers gain strategic business insights by capturing, supporting, managing, and seamlessly accessing variety of data across multiple platforms

Website:
treselle.com

Treselle Systems
Transforming Data into Business Insights

With the rapid data explosion, today, the complexity of data is changing from structured to unstructured—from one-dimensional transactional data flow to multi-dimensional interaction data flow. Additionally, the variety of data types that are managed and analyzed is changing at a rapid rate. In such a scenario, understanding where big data can drive competitive advantage and realizing its value can be challenging for data managers. Defining how businesses can and should utilize big data, Treselle Systems provides big data strategy & architecture service that assists businesses to capture, support, manage, and seamlessly access all the data from multiple data platforms.

> " We enable our clients to efficiently ingest, transform, aggregate, and gain better insights from variety of data sources "

The company's big data strategy, entity information management, and governance services generate business insights from the variety of data with proper master data management and data governance strategies. "We enable our clients to efficiently ingest and aggregate data from various sources using Hadoop & Spark ecosystem and other modern techniques—Talend and Camel for the ETL, mediation, and routing," says Sudharsan Madabusi, President and CEO,

Treselle Systems. Treselle has expertise in data visualization technologies and can perform time-series analysis, geo-spatial analysis, forecasting, classification, clustering, graph-based visualization and back-testing analysis. The company also has extensive experience with multiple NoSQL databases, Big Data SQL Engine, Big Data Cloud computing, Statistical & Quantitative analysis, Text Search & NLP, and Big Data Quality Assurance.

Treselle's Big Data R&D team constantly explores technologies that solve interesting use cases for its clients. For example, Treselle implemented OrientDB as a Polyglot persistence mechanism, for one of their clients, due to its multi-model capabilities that stores data in document database but still has relationship between documents to provide graphing capabilities. Furthermore, Treselle recommended Apache Drill, a flexible query execution framework that has self-describing data exploration behavior across different data stores to avoid too much data movement & syncing issues causing staleness across MongoDB, MySQL, S3, and flat files by keeping data at source.

In one instance, a client from healthcare sector, processing more than 4000 data sources was struggling with ineffective entity identification, disambiguation, and linkage, since same information about healthcare practitioners came in various formats. The client's existing data processing technique was too tedious, slow and error-prone to processing large data sets. The client required a system, which can effectively process their data and integrate with their data management platform, identifying and linking the entities within their massive 40 million nodes and relationships managed in Neo4j. Treselle's Big Data engineering team utilized various advanced technologies

Sudharsan Madabusi

and integration points to perform data manipulation, munging, cleansing, and transformation. The team used Hadoop ecosystem with Pig and other user defined functions to do the data transformation in batch mode, R's text engineering capabilities to perform cleansing, entity identification and linking, and integrated these backend systems with OpenRefine GUI with custom GREL macro expressions to provide excel-like features on the web for the client's data science team. This reduced their data scientist's time from days to hours to perform various data munging capabilities and enriched the data with client's internal APIs.

The average organization today collects more data than ever before, and the variety of data types that are stored, managed, and analyzed has increased exponentially. Engineering talent with different data skills is needed to ingest, transform, aggregate, model, analyze and create insights. "We help businesses by providing the talent to build strong data teams that include data engineers, modelers, scientists, and BI analysts," concludes Madabusi. CR