



Big Analytics: Transforming Big Data into Meaningful Information

By Serhiy Haziyeu, VP of Software Architecture, SoftServe

Big Analytics: Transforming Big Data into Meaningful Information



Thanks to widely adopted smart devices and the proliferation of cloud computing and SaaS business models, there's never been a time when more is known about the behaviors and habits of customers, employees, sult is Big Data.

We're creating and storing more data than ever about every conceivable aspect of business and leisure. These vast volumes of data are the outcome of ongoing advancements in computing and data warehousing technologies, driven by an ever increasing wired and online population. The result is Big Data.

Big Data attracts great interest across organizations and industries because of its potential used to derive greater understanding of past and present events to predict the future and prescribe actions to create desired outcomes.

But Big Data alone can't achieve these things; having the data isn't enough. This is where analytics enters the discussion.

Analytics are mathematic and statistic models used to gain insights to past, present, and future events through the analysis of data. Similar to Big Data having marginal use as a standalone solution, analytics has marginal use until its advanced models are applied to volumes of data. Analytics feeds on large volumes of data to create highly insightful and reliable information. Big Data and Advanced Analytics need each other to achieve great things.

Thanks to widely adopted smart devices and the proliferation of cloud computing and SaaS business models, there's never been a time when more is known about the behaviors and habits of customers, employees, patients, and stakeholders. We're creating and storing more data than ever about every conceivable aspect of business and leisure. These vast volumes of data are the outcome of ongoing advancements in computing and data warehousing technologies, driven by an ever increasing wired and online population. The result is Big Data.

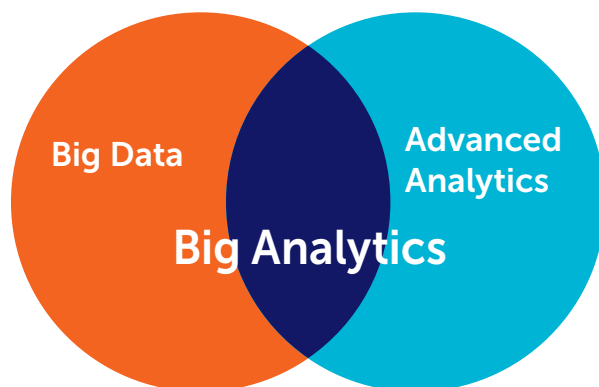
Big Data attracts great interest across organizations and industries because of its potential used to derive greater understanding of past and present events to predict the future and prescribe actions to create desired outcomes.

But Big Data alone can't achieve these things; having the data isn't enough. This is where analytics enters the discussion.

Analytics are mathematic and statistic models used to gain insights to past, present, and future events through the analysis of data. Similar to Big Data having marginal use as a standalone solution, analytics has marginal use until its advanced models are applied to volumes of data. Analytics feeds on large volumes of data to create highly insightful and reliable information. Big Data and Advanced Analytics need each other to achieve great things.

The Power of Big Analytics

Big Data and Advanced Analytics combine to form Big Analytics; capturing and transforming vast amounts of data into meaningful and highly useful information to create more impactful decisions, actions, and outcomes:



Retail. Retailers are using Big Analytics to create personalized shopping experiences and enhance loyalty programs by initiating tailored customer conversions to present unique offers to upsell and cross-sell products based on a shopper's online and in-store browsing, as well as similar consumer behaviors.

Healthcare. Healthcare administrators can use Big Analytics to leverage operational, economic, and demographic and health trends, to plan for investments in new facilities and equipment. Doctors and clinicians can leverage Big Analytics to identify at risk patients, tailor preventative health care activities, and prescribe treatment based on similar diagnosis and patient history.

Logistics. Big Analytics are being used to find the best route for deliveries and backhaul; updating real-time current conditions and calculating load availability against vehicle specifications.

Sales and Service. Big Analytics can be used by inside sales and customer support centers to make specific offers and modifications to subscriptions or buying plans based upon the nature and progression of a customer interaction.

Fraud Detection. Big Analytics can quickly detect changing human behaviors and network anomalies signaling a threat, as well as take immediate action to thwart an attack.

The potential for information businesses, powered by Big Analytics, is undeniably great. And we're merely at the beginning of uncovering its potential.

Four Drivers Leading the Growth of Big Analytics

While early experiences with Big Analytics are exciting, their greatest value may be the way they are changing consumer and business behaviors. As new advances are made, perceptions of what is possible and expected in terms of business opportunities and customer interactions evolves. The result is a marketplace and technology environment that's aggressively pushing Big Analytics forward.

Driver #1: The Power of Real-time Personal Experience

In July 2013, SAP commissioned Forrester Consulting to understand the **current state of personalization**, identify how to expand its use in consumer interactions, and identify barriers to achieving the information needs of marketers. Results from the survey concluded marketers want to greatly expand their personalization capabilities, with more than 70% of respondents stating personalization has a significant impact on customer retention rates, customer lifetime value, customer advocacy rates, and promotion conversion rates.

The survey concluded there's a clear need for advanced business intelligence (BI) solutions to create real-time consumer profiles, capable of delivering the level of personalization consumers increasingly demand.

Time, distance and location are losing their impact on purchase decisions as consumers are likely to shop, purchase and demand service from anywhere, anytime. A purchase today may need to be delivered tomorrow or available for pick-up from a local store the same date of purchase.

Products purchased in a store may be returned through the mail or items purchased online may be exchanged locally. Consumers favor and reward retailers, enterprises, and service organizations of all type who can meet their needs, predict the next service need, and overall make it easy to do business with. Customers increasingly expect tailored services that meet their unique needs and the way to deliver them is through insights gained through Big Analytics.

Driver #2: Actionable Insights

Enterprises and organizations of all kind can benefit from Big Analytics through customer, marketing, operations, and risk management insights they can leverage to improve their performance. These insights are creating new opportunities to innovate and introduce first-in-market positions that not only serve corporate interest, but allows businesses and organizations to respond faster and more cost effectively to customer and stakeholder needs.

The promise of an information-driven business through Big Analytics is better informed, insightful and reasoned decision making, based on the fact the more data we collect and analyze, the greater our potential to achieve a positive impact from the decisions we make. Big Analytics is the answer to the question of how to improve future business decisions and reinvigorate tired industries.

Big Analytics has the potential to change the shape, if not the face of business as we know it – empowering the next evolution of the information age through information appliances, best practices, and industrial optimization.

Actionable insights drive competitive advantage and optimal returns.

Driver #3: Data Growth

According to Bernard Marr, a globally recognized expert in strategic metrics and data, **Big Data is exploding**. Today, every two minutes we are generating the same amount of data that was created from the beginning

of time until the year 2000 – the digital universe is expected to double every two years. The boom of the Internet of Things means the amount of devices connected to the Internet will rise from about 13 billion today to 50 billion by 2020.

The velocity, volume, and variety of data grows exponentially, which requires new technology solutions to capture and analyze data to create actionable information. This is giving rise to the Data Lake concept of warehousing data, combining volumes of data from multiple silos, including internal and external data sources, to deliver deep insights and even more valuable information to leverage in business, medicine, and public service.

As the volume of data grows, so does the need for new and innovative technologies to capture it, warehouse it, and analyze new information for decision makers to leverage in their business or organization. In turn, Big Analytics is fueled by the growth of data and the associated need to analyze it in a timely manner to use insights in the timeframe necessary to maximize the positive impact of decision making.

Driver #4: Cost of Data Systems and the Impact of Hadoop

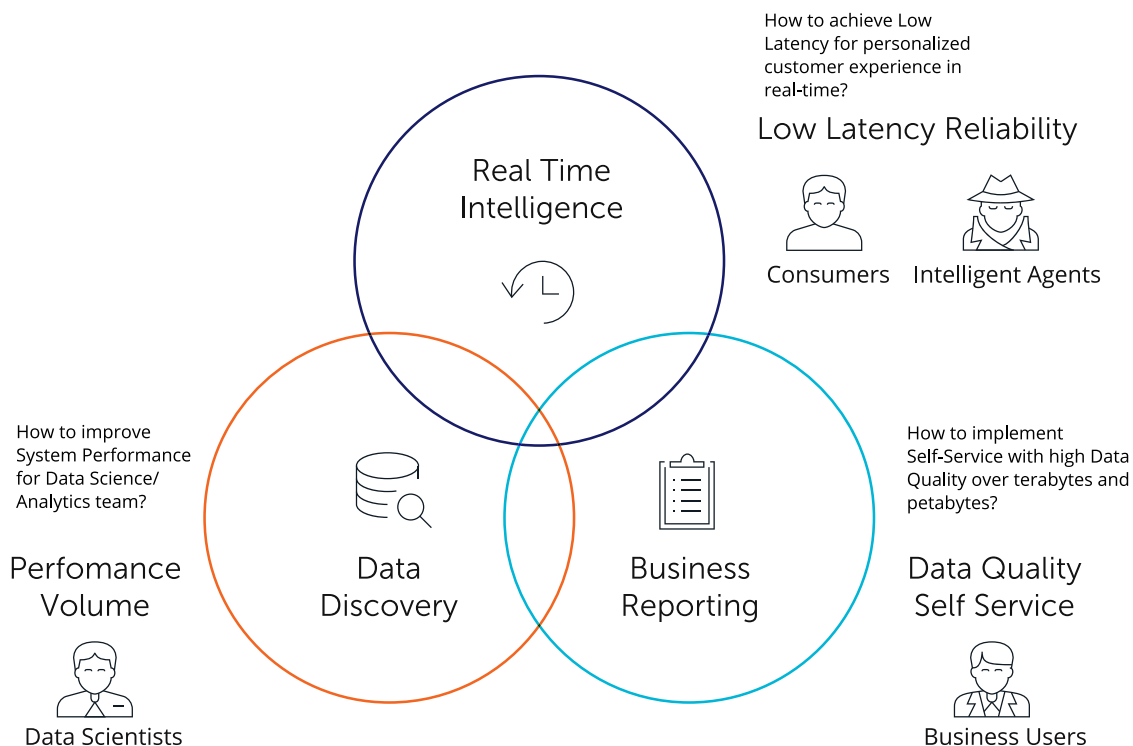
Hadoop and next generation databases enable high performance unstructured data analysis at a lower cost. The difference of Cost/Terabyte between traditional and modern architectures reaches 30x times.

Hadoop allows enterprises to process and analyze large volumes of unstructured and semi-structured data, otherwise inaccessible in a cost and time-effective manner. Because Hadoop clusters can scale to petabytes and even exabytes of data, enterprises no longer must rely on sample data sets, but can for the first time cost effectively process and analyze all relevant data.

Big Analytics Challenges

The potential for information and analytic business models are undeniably great. But while some companies are enjoying this potential, many remain reluctant or merely dabble with analytics and Big Data on a small scale to achieve a specific, narrowly defined outcome as opposed to a strategic implementation capable of transforming their business.

Given the success and potential of Big Analytics, the question is why aren't more companies and organizations embracing information-driven business models?



Real-time Intelligence

Real-time intelligence is the delivery of information about business or organizational operations as they occur. Real-time means near zero latency to access information wherever and whenever it is required. This is especially important serving fraud protection, customer interactions, and many healthcare applications of Big Analytics.

While all intelligence systems have some level of latency, the goal is to minimize the time from an event as it happens to a corrective action. The business and technical challenge is how can we achieve near zero latency to deliver personalized customer experiences in real-time?

Dashboards and Business Reporting

Businesses and organizations that maintain a robust data infrastructure have many competitive advantages ranging from the ability to make better decisions, track the health of their supply chain, and manage sales relationships. Intelligent business reporting empowers businesses and organizations of all kind to timely identify new opportunities or head-off problems before they mature. Dashboards and visualization are extremely helpful in this area, supporting easy access to information and intelligence resulting from the analysis of large volumes of data.

But to provide such an information network requires consolidating, warehousing, and managing data quality as the pool of data dramatically increases over time. The challenge to overcome is how to serve the growing demands of decision makers and consumers for self-service access to information and intelligence, extended over terabytes and petabytes of data, while maintaining a high level of data quality?

Data Science and Discovery

Human nature is playing a role in technology adoption, with subject matter experts and administrators storing data in local silos as opposed to shared warehouses for reasons ranging from job security to instant gratification, resulting in anywhere from 50% to 80% of an organization's data sitting in spreadsheets, unavailable for advanced analysis. Centralized data warehouses and centralized BI platforms have many advantages, but often create layers of bureaucracy and approval in their design and implementation, as well as being costly and time consuming, hence data silos continue to be the norm. This sits at odds with the need of IT organizations to build flexible data warehouses and analytic systems quickly and inexpensively to meet existing and future requirements. Data needs to be democratized to support intelligence intensive businesses.

Looking at a broader perspective of data, only an approximate 10% of all data is structured. Nearly 90% of all digital data is unstructured or semi-structured -- social media, sensors, servers, wearable devices, cameras, mobile phones, etc. This unstructured data creates countless benefits to a company that can leverage it with analytics to increase profits and productivity. It's possible for a team of data scientists and analysts to analyze unstructured and semi-structured data with the help of multiple techniques including data mining, predictive analysis, and deep learning. The desired result is the discovery of hidden patterns, correlations, and other useful information otherwise unknown that can be used to make better informed decisions.

The business and engineering challenge is how can we enable the data discovery and improve data science and analytics teams performance to harness the potential of exercising advanced analytics across all possible data sources?

The SoftServe Solution

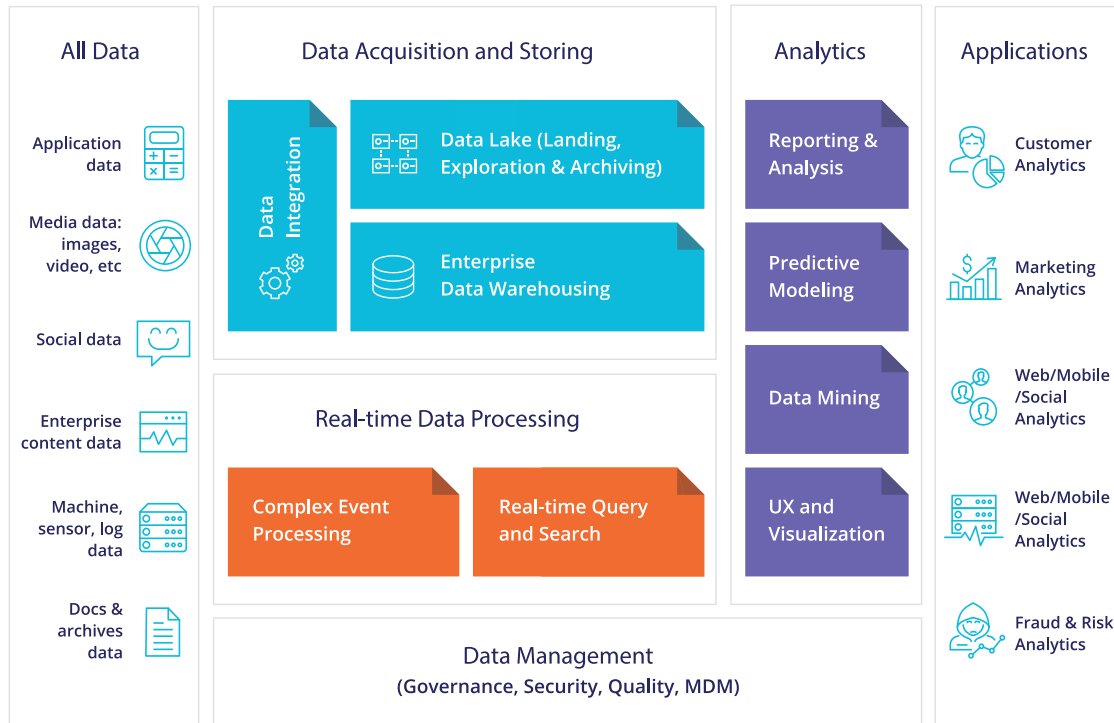
To unharness the potential of Big Analytics and tackle engineering challenges in its implementation, SoftServe has developed a number of services and solutions purpose-built to meet the needs of modern information and intelligence driven business and organizations.

Big Analytics Reference Architecture

We developed a Big Analytics Reference Architecture based on industry and in-house best practices to meet the needs of all enterprises and organizations. This technology agnostic architecture supports the processing of structured, unstructured, and semi-structured data, as well as historical and real-time analysis.

This reference architecture is designed to support the varied needs of Big Analytics to include analysis of customer behaviors, marketing analytics, operational analytics, fraud detection, and other Advanced Analytics

applications. Technology independence allows the architecture to be employed in a variety of environments, taking into consideration a variety of customer technology preferences and limitations.



Technology independence covers the entire spectrum of data integration, storage, processing and analytics platforms to flexibly support any desired solution.

Together with Rick Kazman, an SEI Carnegie-Mellon principal researcher, we created a gamification process of designing an architecture for a Big Analytics System. This “Smart Decisions” game (aka Architecture Poker) reduced the design time of a Big Data concept architecture from a typical couple week process to merely several hours.

Big Data Architecture Poker

Architecture Poker uses playing cards that represent technologies and architecture patterns. The game helps design complex data systems in a very short time, as well as teach best design practices. The game is based on a set of rules that reflect an Attribute Driven Design version 3 (ADD 3.0) which is an iterative systematic method for designing the software architecture of

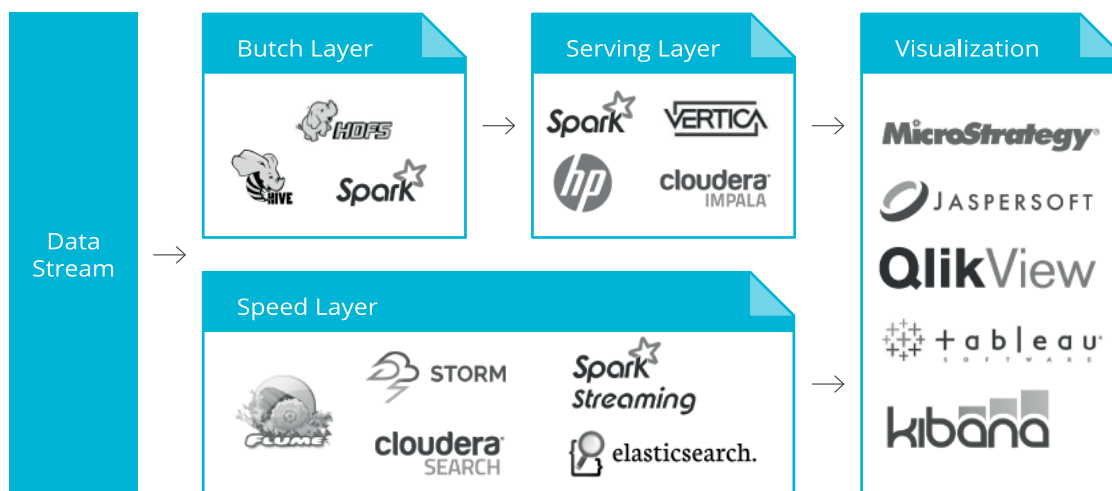
a software-intensive system. The cards are used to select the most optimal patterns and technologies to meet business and system requirements.

This approach dramatically decreases the time taken to design Big Data systems, in some cases reducing design time from weeks to hours.

SoftServe Lambda Architecture Accelerator

Lambda Architecture is an emerging industry standard – a highly scalable, and reliable data processing architecture based on Twitter’s successful experience in Big Data and Analytics.

Our Lambda Architecture Accelerator is a business accelerator that reduces time to market and implements a highly scalable and reliable data-processing solution designed to handle massive amounts of data by taking advantage of both batch and stream-processing methods. Event and log data enters our Lambda Architecture stack from various data sources and is processed by multiple layers using open source technologies to achieve the lowest cost per terabyte.



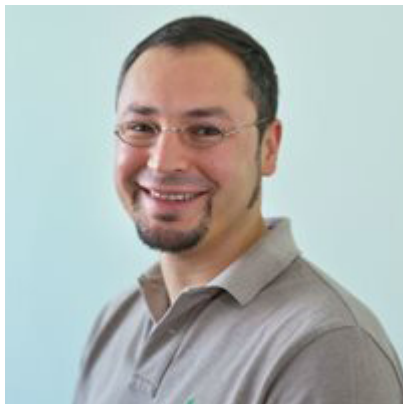
Consistent with our Big Analytics Reference Architecture, our Lambda Architecture Accelerator is technology agnostic, allowing us to recommend technologies based on a particular need or adapt to a unique customer environment.

Ready to create your Big Analytics solution with us?

We're at the beginning of the next technology evolution. Big Data and advanced analytics are rapidly changing, creating a great opportunity with Big Analytics to change the face of business and leisure as we know it. The advantages of data democratization, open systems, and shared warehousing have never been greater or more compelling – even the most proprietary and slowest markets to accept change, healthcare and the public sector, acknowledge the potential of Big Analytics.

The future of Big Analytics and intelligence-driven business is great. The world as we know it is changing and may never be the same. Is it time for you to change too? Contact us to learn how we may be able to unleash your hidden data and information to help create a bright, exciting, and prosperous future for your business or organization.

About the Author



Serhiy Haziyeu is a VP of Software Architecture at **SoftServe**. He has more than 17 years' experience in enterprise-level solutions including SaaS/Cloud, Big Data, Business Intelligence, SOA and Carrier-grade telecommunication services. He specializes in software architecture methodologies, architectural patterns and software development practices for large and complex projects in multiple industry verticals,

including healthcare. Serhiy holds various qualifications and is a frequent speaker at external and corporate conferences.

About SoftServe

SoftServe is a leading technology solutions company specializing in software development and consultancy services. Since 1993 we've been partnering with organizations from start-ups to large enterprises to help them accelerate growth and innovation, transform operational efficiency, and deliver new products to market.

To achieve this we've built a strong team of the brightest, most inquiring minds in the industry, and we form close, collaborative relationships with our clients so we can really understand their needs and deliver intuitive software that exceeds their expectations.

Our experience stretches from Cloud, Security and UX Design to Big Analytics and the Internet of Things, we have offices across the globe and development centers across Eastern Europe. For more information please visit www.softserveinc.com.

USA HQ

Toll Free: 866-687-3588

Tel: +1-512-516-8880

Ukraine HQ

Tel: +380-32-240-9090

Bulgaria

Tel: +359-2-902-3760

Germany

Tel: +49-69-2602-5857

Netherlands

Tel: +31-20-262-33-23

Poland

Tel: +48-71-382-2800

UK

Tel: +44-207-544-8414

EMAIL

info@softserveinc.com

WEBSITE:

www.softserveinc.com