SMARTER, FASTER, BIGGER

HOW BIG DATA AND ANALYTICS ARE TRANSFORMING BUSINESS
INDUSTRY OVERVIEW
Data is becoming the basis of competition and growth

INDUSTRY PERFORMANCE
Big Data Market
Big Data and Business Analytics Market
Regional Markets for Big Data
Big Data Revenue by Product Segment

PRODUCTS AND MARKETS
Big Data Technology and Infrastructure
Big Data Analytics
Key Sectors

INDUSTRY TRENDS
Key Drivers
Technology Trends
Secular Trends
Political and Regulatory Trends

COMPETITIVE LANDSCAPE
Major Players
SWOT Analysis
Notable Exits

WHAT’S COMING?
As we enter the second decade of the 21st century, a new production factor is becoming increasingly important alongside capital and labor: data.

As recently as 2000, only 25% of data was stored in digital form. A decade later, that percentage had reached 94%. In the last two years, over 90% of all the world’s data was created. Internet users generate about 2.5 quintillion bytes of data each day. 294 billion emails are sent and over 1 billion Google searches are conducted every day.

These already mind-boggling statistics are growing at an exponential rate. Trillions of sensors monitor, track, and communicate with each other, populating the Internet of Things (IoT) with real-time data. IoT and smart device usage is exploding. In 2006 there were 2 billion smart devices. In 2020 that number is expected to reach 200 billion.
Big Data has the ability to create value in a variety of ways:

- Data can make information transparent and usable at a much higher frequency.
- Organizations can greatly improve their performance by creating and gathering transactional data in digital form.
- Firms can have much clearer information on everything from product inventories to sick days, and consequently expose variability and boost performance.

Leading companies are already using such data practices and even conducting controlled experiments to make better management decisions. Big data allows companies to segment their customers and produce more precisely tailored goods and services. McKinsey and Company estimates that a firm using big data to its full capacity has the potential to increase its operating margin by more than 60%. The numerous benefits of big data are widespread and cannot be overstated.
BIG DATA: WHAT IT IS AND HOW IT WORKS

DATA NEVER SLEEPS 6.0

How much data is generated every minute?

There’s no way around it: big data just keeps getting bigger. The numbers are staggering, but they’re not slowing down. By 2020, it’s estimated that for every person on earth, 1.7 MB of data will be created every second. In our 6th edition of Data Never Sleeps, we once again take a look at how much data is being created all around us every single minute of the day—and we have a feeling things are just getting started.

2018 every
MINUTE
of the
DAY

THE WEATHER CHANNEL
AMAZON
TUMBLR
REDIT
GOOGLE
TINDER
VENMO
SPOTIFY
AMERICANS

18,055,555 REQUESTS
1,111 PACKAGES
79,740 POSTS
1,944 NEW COMMENTS
3,877,140 SEARCHES
226 TIMES
91,218 TIMES
1.388,889 GIFS
97,222 LIVE STREAMS
2,083,333 SNAPCHATS
20+ NEW PROFESSIONALS
120+
4,333,560 VIDEOS
473,400 TWEETS
12,986,111 TEXTS
176,220 CALLS
49,380 PHOTOS

NETFLIX USERS STREAM 97,222 LIVES OF VIDEO
SNAPCHAT USERS SHARE 2,083,333 SNAPCHATS
LINKEDIN
YOUTUBE
TWITTER
SKYPE
INSTAGRAM

1.25 NEW BITCOIN ARE CLEARED
1.389 RIDES
3.4 BILLION OF INTERNET DATA
2.5 BILLION OF INTERNET DATA
1.3 BILLION OF INTERNET DATA

The world’s internet population is growing significantly year-on-year. In 2017, internet usage reached 47% of the world’s population and now represents 3.8 billion people.

The ability to make data-driven decisions is crucial to any business. With each click, swipe, share, and like, a world of valuable information is created. Domo puts the power to make those decisions right into the palm of your hand by connecting your data and your people at any moment, on any device, so they can make the kind of decisions that make an impact.

Learn more at domo.com

Source: Domo

GLOBAL INTERNET POPULATION GROWTH 2012-2017

(SOURCES: STATISTA, LINKEDIN, INTERNET LIVE STATS, DELOITE HANBINGS, GLASS MOU RIAA, BUSINESS OF APPS, INTERNATIONAL TELECOMMUNICATIONS UNION, INTERNATIONAL DATA CORPORATION)
In general, Big Data refers to datasets that are too large or complex for traditional data applications to process.

What makes data big? High volume, high velocity, and high variety are characteristics associated with Big Data. But Big Data can also refer to the set of different components that can be controlled, accessed and used at any time using various platforms deployed either on premise or in the cloud. These components can include machines, devices, assets, end products, sensors, thermostats, data generating, storing and retrieving software, Big Data analytical tools, and data output visuals.¹

**DATA IS BECOMING THE BASIS OF COMPETITION AND GROWTH**

In the coming years, firms will need to adopt Big Data practices in order to survive. The ubiquity of mobile interfaces and the power of artificial intelligence are just a couple of the technological trends that are reshaping customer expectations and creating the potential for virtually any sector with a distribution component to have its boundaries redefined faster than we have ever experienced.

Consequently, business leaders need to have cross-sector dynamics at the top of their minds. A company from one industry may begin to understand customers from another industry better than that industry’s leader, simply by using Big Data. It will be crucial for companies to defend their positions within an industry; but even more importantly, companies will need to attack and capture cross-sector opportunities before competitors get there first.²

¹ Mind Commerce
² McKinsey
## Big Data Market 2018 - 2023

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Big Data Markets</td>
<td>24</td>
<td>28</td>
<td>36</td>
<td>41</td>
<td>53</td>
<td>64</td>
<td>72</td>
<td>81</td>
<td>17.6%</td>
</tr>
</tbody>
</table>

*Global Big Data Markets (2018-2023)*
(Copyright: Mind Commence)
BIG DATA AND BUSINESS ANALYTICS MARKET

Statista finds that the global big data and business analytics market brought in revenue of $168.8 billion in 2018, and is forecasted to grow to $274.3 billion by 2022 with a five-year compound annual growth rate of 13.2%. The $189.1 billion forecast for 2019 is a 12.0% increase over 2018.

REGIONAL MARKETS FOR BIG DATA, 2018-2023

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NORTH AMERICA</td>
<td>7.2</td>
<td>8.4</td>
<td>10.8</td>
<td>12.3</td>
<td>14.8</td>
<td>17.9</td>
<td>19.4</td>
<td>21.9</td>
<td>15.2%</td>
</tr>
<tr>
<td>LATIN AND CENTRAL AMERICA</td>
<td>1.0</td>
<td>1.1</td>
<td>1.4</td>
<td>1.6</td>
<td>2.1</td>
<td>2.6</td>
<td>3.6</td>
<td>4.1</td>
<td>23.0%</td>
</tr>
<tr>
<td>EASTERN EUROPE</td>
<td>2.9</td>
<td>3.4</td>
<td>4.3</td>
<td>4.9</td>
<td>6.4</td>
<td>7.7</td>
<td>9.4</td>
<td>10.5</td>
<td>19.5%</td>
</tr>
<tr>
<td>WESTERN EUROPE</td>
<td>6.7</td>
<td>7.8</td>
<td>10.1</td>
<td>11.5</td>
<td>15.4</td>
<td>18.6</td>
<td>20.9</td>
<td>23.5</td>
<td>184%</td>
</tr>
<tr>
<td>ASIA PACIFIC</td>
<td>4.1</td>
<td>4.8</td>
<td>6.1</td>
<td>7.0</td>
<td>9.5</td>
<td>11.5</td>
<td>13.0</td>
<td>14.6</td>
<td>19.0%</td>
</tr>
<tr>
<td>MEA</td>
<td>2.2</td>
<td>2.5</td>
<td>3.2</td>
<td>3.7</td>
<td>4.8</td>
<td>5.8</td>
<td>5.8</td>
<td>6.5</td>
<td>14.9%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>24</td>
<td>28</td>
<td>38</td>
<td>41</td>
<td>53</td>
<td>65</td>
<td>72</td>
<td>81</td>
<td>17.6%</td>
</tr>
</tbody>
</table>
## BIG DATA REVENUE BY PRODUCT SEGMENT, 2018-2023

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DATABASE MANAGEMENT SYSTEMS</td>
<td>3,120</td>
<td>3,640</td>
<td>5,400</td>
<td>6,150</td>
<td>7,950</td>
<td>9,600</td>
<td>10,080</td>
<td>11,340</td>
<td>16.0%</td>
</tr>
<tr>
<td>DATA INTEGRATION TOOLS</td>
<td>6,240</td>
<td>7,280</td>
<td>7,920</td>
<td>8,200</td>
<td>10,070</td>
<td>11,520</td>
<td>12,240</td>
<td>13,770</td>
<td>11.7%</td>
</tr>
<tr>
<td>APPLICATION INFRASTRUCTURE AND MIDDLEWARE</td>
<td>2,880</td>
<td>3,360</td>
<td>4,680</td>
<td>5,330</td>
<td>7,420</td>
<td>8,960</td>
<td>10,800</td>
<td>12,150</td>
<td>21.0%</td>
</tr>
<tr>
<td>BUSINESS INTELLIGENCE TOOLS AND ANALYTICS PLATFORM</td>
<td>4,320</td>
<td>5,040</td>
<td>6,840</td>
<td>7,790</td>
<td>10,600</td>
<td>13,440</td>
<td>15,840</td>
<td>17,820</td>
<td>21.1%</td>
</tr>
<tr>
<td>OTHER MANAGEMENT TOOLS</td>
<td>1,920</td>
<td>2,240</td>
<td>2,880</td>
<td>3,280</td>
<td>3,710</td>
<td>3,840</td>
<td>4,320</td>
<td>4,860</td>
<td>11.0%</td>
</tr>
<tr>
<td>PROFESSIONAL SERVICES</td>
<td>5,520</td>
<td>6,440</td>
<td>8,280</td>
<td>10,250</td>
<td>13,250</td>
<td>16,640</td>
<td>18,720</td>
<td>21,060</td>
<td>20.5%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>24,000</td>
<td>28,000</td>
<td>36,000</td>
<td>41,000</td>
<td>53,000</td>
<td>64,000</td>
<td>72,000</td>
<td>81,000</td>
<td>17.8%</td>
</tr>
</tbody>
</table>
“Big Data” is a broad term for a complex stack including physical infrastructure, data collection and integration, data security infrastructure, data storage management, data mining and analytics, and data consumption.
The two technologies most commonly used to implement Big Data solutions are Hadoop (HDS, MapReduce) and NoSQL.

**HADOOP**

Hadoop -- a "high-availability distributed object-oriented platform" -- is a software framework that analyzes structured and unstructured data and distributes applications across different servers. Hadoop is an open-source software framework for storage and large-scale processing of data on clusters of commodity hardware.

The software is designed to scale up from a single server to thousands of machines, with a very high degree of fault tolerance (the ability of a system to continue operating properly in the event of the failure of some of its components). Hadoop is licensed through Apache 2.0.

Hadoop is typically used for large-scale analytic applications, including predictive analytics, fraud detection, and user recommendations.

**NoSQL**

NoSQL was created to address the limitations of typical relational databases designed before the advent of today’s Big Data solutions. Relational databases were usually not designed to handle the scale and agility challenges that modern applications now face, and were not built to make full use of the commodity storage and processing power available today.

NoSQL deals with unstructured and unpredictable datasets. The NoSQL database provides a data storage and retrieval mechanism that is modeled on methodologies other than the tabular relations used in relational databases.

Data in NoSQL is typically information that is updated rapidly: for instance, include user transactions, sensor data, and customer profiles.
NoSQL VS HADOOP

Both technologies are incredibly useful for handling large and growing volumes of data, as well as varying data formats.

They both support clusters of commodity hardware that a business can incrementally and horizontally scale out by simply adding more hardware to their system.

Both technologies have a lot of overlap, yet are not direct competitors. They work together quite well, as they are built to handle different workloads.

NoSQL is designed for real-time and interactive data accesses. It is typically used in an interactive end-user environment. More broadly, it supports rapid reading and writing of data.

Hadoop is more batch-oriented and deals with large-scale processing. To run big processing tasks on large volumes of data, the work needs to be spread across many servers in parallel. Hadoop manages that process in a divide-and-conquer methodology known as MapReduce, which can process close to the data so that it’s not accessing datasets across the network and slowing down the overall task.

Companies typically create clustered architectures that enable NoSQL and Hadoop to work together. The NoSQL cluster handles the read and write operations, while the Hadoop cluster handles large-scale analytics.

BIG DATA ANALYTICS

So why collect terabytes of data? Raw data by itself is useless. A major driver of the industry will come from the advancements in new analytic techniques.

The greatest value of Big Data is the insights that we can create by analyzing it. New technological advancements like machine learning are making it possible to tackle the most difficult and challenging business problems. But while many companies will soon have the ability to gather large and diverse volumes of data, what will distinguish one business from another is the ability to ask new questions and formulate new hypotheses about data.
BIG DATA ANALYTICS TECHNOLOGY AND IMPLEMENTATION APPROACHES

**Grid Computing**

Grid Computing is a special kind of distributed computing. It relies on a computer network in which each computer’s resources are shared with every other computer in the system, and it can be used to more effectively manage and analyze large volumes of data.

Hadoop is an example of Grid Computing.

**In-Database Processing**

This technology allows data to be processed within the database by building analytic logic into the database itself. This eliminates the need to transform data and move it between a database and a separate analytics application.

**In-Memory Analytics**

This enterprise architecture framework solution is used to solve complex and time-sensitive business scenarios by querying data from system memory (RAM), versus the traditional hard disk drive medium.

**Data Mining**

Data mining is an analytic process used to explore data and find consistent patterns and systematic relationships between variables. These patterns and relationships are then applied to new data subsets for verification.

**Predictive Analytics**

Predictive analytics refers to any kind of statistical technique that uses current and historical facts to make predictions about hypothetical future scenarios.

**Natural Language Processing**

NLP allows computers to read, understand, and extract meaning from human languages.

Natural language processing can be used to understand market sentiment. Voice recognition software is already used by various businesses in call centers. Companies are increasingly using it to gain a general understanding of how customers feel and what competitors are doing, typically by detecting how often competitors’ names are mentioned.
**Text Analytics**

This process can be used to analyze the content of emails, blogs, tweets, forums, and other forms of unstructured data.

**Visual Analytics**

Visual analytics use interactive graphical data displays to generate analytical results and insights. This method uses data visualization techniques to help data scientists identify trends, patterns, and relationships in the data they are working with.

**Association Rule Learning**

Association rule learning is a method for identifying useful and insightful correlations between variables in large databases.

**Classification Tree Analysis**

Statistical classification is a way of identifying categories that new pieces of data belong to.

**Machine Learning**

Machine learning refers to software that can learn from data without being explicitly programmed. Machine learning algorithms make predictions based on known properties learned from sets of “training data.”

**Neural Networks**

Neural networks are computing systems that are modeled after the biological neural networks of the brain. Neural networks are used when the exact nature of the relationship between inputs and outputs is not known, and are typically trained to analyze these ambiguous relationships.

**Radial Basis Functions (RBF)**

RBF is a function built with a distance criterion with respect to a center. RBF functions can efficiently interpolate and smooth data.
Geospatial Predictive Modeling

This method is used to analyze events of any kind through a geographic lens in order to make observations about the probability for event occurrence or emergence. Spatial geographic factors include such things as the infrastructure, topography, or even socio-cultural make-up of a geographic region.

Regression Analysis

Regression analysis is a long-standing technique for data scientists. It simply involves manipulating an independent variable to see how it influences a dependent variable.

Social Network Analysis

Social network analysis broadly refers to the analysis of relationships among people in many fields and commercial activities. By using nodes to represent individuals, the ties between these nodes can be studied to understand shifting social dynamics.
KEY SECTORS

MACHINE TO MACHINE (M2M)

M2M refers to the direct communication between devices using any communications channel, including wired and wireless.

IT and business leaders are beginning to implement M2M in areas ranging from smart grid energy networks, to manufacturing and industrial plant monitoring, to patient monitoring in healthcare. For instance, healthcare practitioners can use Big Data techniques on M2M subsystems to track patients’ drug interactions and other aspects of patient monitoring.

RETAIL AND HOSPITALITY

The power of Big Data can be leveraged to determine buying patterns, optimize pricing, assess customer sentiment, and optimize the supply chain to provide better business intelligence to retail and hospitality businesses. For instance, according to Mind Commerce, Big Data can help businesses answer questions like these:

- Who are your customers?
- How do they feel about your products and services?
- Why did they choose you?
- Are they satisfied with you and your products?
- How do they behave? What drives their purchasing decisions?

Hospitality

The hospitality industry already identifies consumer preferences and acts on those preferences on a multinational basis. This presents a massive opportunity for Big Data.

Consumer preferences are primarily based on market data. Market data can be used to help define web search keywords that will help funnel customers toward products and services that meet their specific needs. Internet keyword search will be discussed in further detail in the “Technology Trends” section of this report.
Social Media
Social media sites are massive sources of data. Much of it is unstructured information on personal views, tastes, interests, etc. For instance, Facebook uses facial recognition tools to compare photos you have uploaded with those of other users to find potential friends of yours.

Social Gaming Analytics
Social gaming companies have begun using Big Data techniques to identify common user characteristics than can then be used to target these users with the right advertising placement and content. These techniques help retain and monetize more users.

Use of Social Media Analytics by Other Verticals
Consumer product companies and retail organizations analyze social media like Facebook and Twitter to assess consumer preferences, behavior, and product perception.

Utilities
Big Data can be greatly beneficial for utility companies. It can turn the information from smart meter and smart grid projects into meaningful operational insights about customer behavior.

Within the Utilities and Energy sector, Big Data analytics is increasingly being used for customer retention, forecasting energy usage, billing analytics predictive generation equipment, and performance optimization of power generation equipment.
FINANCIAL SERVICES

Fraud Analysis, Mitigation & Risk Profiling
Bulge bracket banks such as JP Morgan Chase and Morgan Stanley are using Big Data to conduct fraud detection, fraud analysis, and risk profiling.

Merchant-funded Reward Programs
Banks can analyze behavioral data to deliver highly targeted offers to customers.

Customer Segmentation
Big Data analytics can help marketers better understand their customers and target promotions by drawing correlations between purchase history and social media data.

Customer Retention & Personalized Product Offering
According to a 2018 Mind Commerce research report, GE Capital Retail Bank has begun using Big Data to offer credit card products and services customized to their clients’ needs. The goal of the initiative is to capture a larger percent of their credit card users’ spending.

Through this initiative, GE Capital has accumulated the following types of data from their clients:

- Customer data (demographics, social, and on-line behaviors)
- Industry benchmarks (trending and metrics)
- Channel interactions (GE Capital customer care, marketing, and billing data)
- Client data (shopping patterns)
- Transaction data (credit card transactions, account attributes)

With this data, GE Capital has the ability to develop programs and services to:

- Allow their marketing department to perform analytics, determine how to get customers to use their GE credit card more, and define marketing programs to grow business.
- Provide their clients with information and intelligence to shift spending from their competition to GE Capital
Insurance Companies

Insurance companies are using Big Data analysis to see which home insurance applications can get processed on the spot and which ones need a validating in-person visit from an agent.

HEALTHCARE AND PHARMACEUTICAL

This sector has primarily focused their Big Data investment on managing population health efficiently, improving patient care with medical data analytics, improving clinical development and trials, and improving time to market for pharmaceutical products.

Applications include:

Drug development

Medical Data Analytics: Hospitals use Big Data analytics to predict which patients will likely seek readmission after a few months of being discharged. This enables hospitals to save tremendous costs by intervening before patients seek readmission.

TELECOMMUNICATIONS

Key Big Data applications for telecom companies include network performance and coverage optimization, customer churn prevention, personalized marketing, location-based services and fraud detection.

Telco Analytics: Customer/Usage Profiling and Service Optimization
With Big Data, telecommunication companies can analyze call logs and complex data from multiple sources. The goal is to use log data to build customer profiles (popular mobile devices, popular websites etc.), segment customers, and optimize products and services accordingly.

Big Data Analytic Tools
Telecom operators are developing a Big Data products and environments that support a range of tools and deployment mechanisms, giving them the ability to modularly add internally developed or third-party analytic tools to their Big Data toolkit.
Speech Analytics
Telecom companies can use Big Data to analyze phone calls to help provide more tailored offers to customers.

New Products
The Big Data techniques discussed here are still in the early stages. There is still much to come within the telecom sector.

GOVERNMENT AND HOMELAND SECURITY

Multiple government sectors are working on Big Data projects. The Defense Department is currently working on techniques for intelligence gathering and preventing injuries on the battlefield.

Numerous government agencies are developing urban planning projects in an effort to create smart cities and intelligent buildings that perform energy optimization and intelligent building analytics. These same agencies are learning how to use Big Data in urban transportation management, optimizing energy production, water management, and urban waste management.

Big Data is being used in the Public Safety and Homeland Security sector to provide cyber-crime mitigation, crime prediction analytics, video analytics, and situational awareness.

Within the Public Services sector, Big Data analytics are being used to conduct public sentiment analysis, fraud detection and prevention, and economic analysis.
OTHER SECTORS

Aviation

Big Data can be used by air traffic controllers to receive early warnings or storm updates.

Big Data gives airport managers access to real-time data on performance indicators such as average luggage delivery times, delays, and airport security levels.

Transportation and Logistics: Optimizing Fleet Usage

Logistics companies can collect various data elements, such as tire and fuel usage, engine operation, and geospatial data for fleet tracking. This data can be used to optimize fleet usage and save logistics companies millions of dollars a year.

Real-time Processing of Sports Statistics

Statistical data can be recorded and analyzed in real time during sporting events. Teams and coaching staffs can use historical data and predictive models to make timely corrections to their game strategy.

Education

Big Data will be used to enable student directed learning. The use of computerized testing or tablets for learning exercises allows school to collect more data on how children are doing. Teachers can analyze every answer on an assessment to identify the strengths and weaknesses of individual students and classes as a whole.

With Big Data, wrong answers no longer constitute a bad grade, but rather can allow teachers to gain a deeper understanding of why students picked a wrong answer. Consequently, this data can be relayed onto publishers to understand which chapters of their textbooks are effective and which are not.
Manufacturing

Manufacturers are experiencing pressure to eliminate defects and squeeze out costs, which is forcing them to put more focus on data.

Modern factory equipment has also become smarter and more sophisticated. Manufacturing equipment now comes with computerized controls that make the capture, analysis, and sharing of data more seamless.

Manufacturers are using Big Data in manufacturing execution systems (MES) to identify procedural mistakes in manufacturing plants so that production can halt itself in the case of an error.

Software keeps constant track of the smallest details of production, so that if any variable drifts away from the prescribed setting, the machinery is automatically adjusted.

The use of Big Data will also be crucial in the post-production phase. Millions of dollars have been spent to rework and repair manufactured parts. Manufacturers can now keep data for all of their manufactured products automatically, including the names of all the machine operators who worked on any product and the humidity and temperature at each stage of production. This data can help companies go back and figure out what went wrong if flaws emerge.

Mind Commerce identifies the following benefits for Big Data in manufacturing:

- Operating margins increase by 16%
- More high-value customers with more responsive service and greater consistency in quality.
- Unscheduled downtime reduced 36%
- Manufacturers are vastly more connected with each other. Manufacturing leaders can make use of an extended network to anticipate the availability of materials and the impact of factors that may influence supply.
- Cash flow increase of 10%
- Inventory reductions between 15% and 50%

Financial and Support Services

Leaders are able to communicate financial, safety, compliance, and product data with regulators more efficiently and accurately. Manufacturers can also immediately verify the credit worthiness of new customers.
DATA VOLUME AND VARIETY
The rapid growth in global network traffic volume and data variety is the primary driver of Big Data.

INCREASING ADOPTION OF BIG DATA BY ENTERPRISES AND TELECOM
Enterprise and telecom organizations are focusing considerable attention on Big Data infrastructure in an effort to bring together data from diverse sources, including operations, services, marketing, and vertical sectors such as Entertainment and Healthcare.

MATURATION OF BIG DATA SOFTWARE
As Hadoop, NoSQL, and other Big Data software are increasingly adopted by firms, they will mature and further propel the growth of the Big Data market.

CONTINUED INVESTMENT IN BIG DATA BY WEB GIANTS
Major fortune 500 companies are fueling the growth of the industry by devoting substantial investment to Big Data infrastructure. Google recently spent $3 billion on Big Data technology and Facebook invested $1 billion in the infrastructure that powers its social network.
TECHNOLOGY TRENDS

IOT (INTERNET OF THINGS)
IoT-connected devices and edge computing will generate substantial amounts of unstructured data, which will create substantial technological challenges for operating companies and their vendors.

STREAMING ANALYTICS
Industry leaders in telecommunications are beginning to explore new methods of performing analytics on Big Data. Streaming analytics involves capturing data in real time, performing analytics on the necessary and useful data, and then throwing away the remainder of the data and saving the analytic results.

The purpose of this method is to perform analytics on source data in real time, while saving disk storage and processing time.

This type of data analytics requires complex Machine Learning algorithms. Machine Learning trains the data system to predict outcomes with reasonable accuracy. As the primary model adjusts and evolves, models at the edge or in the Cloud coordinate to match the changes as needed.

AI PLATFORMS
AI platforms will have significant impact in processing Big Data and improving business intelligence.

CLOUD TECHNOLOGY
Cloud technologies provide multi-processing environments that can be very useful on Big Data files.
GOOGLE SEARCH
The Search function in Google exemplifies all that NoSQL and Big Data technology are and can provide.

When a keyword search is made, Google uses Petabytes (one million gigabytes) of information via their multi-tier processor systems to provide sub-second response to the query.

Google leverages the fact that their data files are continually growing by constantly reviewing their products and services to improve their performance and efficiency.

CUSTOMIZE ANALYTICAL TOOLS
One of the most substantial benefits of Big Data for businesses will come from understanding their consumers and producing more tailored products. This will require developing analytical tools, business metrics, and scoring models to quickly identify trends, patterns, or gaps within a market or business.

Examples: Using analytical tools to identify, understand, and predict customers’ needs by their demographics, transactions, influence, and sentiment.

INTERNET KEYWORDS
An important application of Big Data is reacting to keyword searches to serve a consumer’s buying intent.

A good example of this is found in the travel industry: internet keywords could be used to differentiate between two people who are travelling to the same place, but for different reasons. One might be travelling for business and the other for pleasure. The one travelling for business will want accommodations close to where they have meetings and would most likely prefer to go back to their hotel at night and relax, whereas someone on vacation may prefer to be near the city’s entertainment/tourist/shopping spots. Internet keywords can help businesses funnel the right customers onto their platforms.
GAMIFICATION

This technique develops business intelligence by providing an environment in which users feel as if they are playing a game, while in reality they are supplying the game creator with the desired business intelligence. This technique turns companies’ insight-gathering efforts into an enjoyable experience for users.

AR/VR

AR and VR are among the most data-intensive technologies, and require significant amounts of bandwidth to operate.

SECULAR TRENDS

Big Data inevitably conveys significant competitive advantages, which is forcing enterprises to invest heavily in the space.

Big Data solution provider dynamics are evolving almost equally as fast as the data management technologies themselves. Some companies rely solely on proprietary solutions while others rely on open source Hadoop technology. Thus, a unique feature of the evolving Big Data market is the extent to which market leaders are choosing to either compete or collaborate with one another.
POLITICAL AND REGULATORY TRENDS

On May 25th, 2018, the European Union put the General Data Protection Regulation (GDPR) into effect. This commitment to privacy has not been followed as strictly by other corporations and organizations. Because the GDPR does not allow businesses working with Europeans to rely on “assumed consent” for all processing operations. Businesses have had to institute new procedures for notices and receiving consent, many businesses in the US are behind in this regard.

The implementation of the GDPR has given technology leaders in the US the opportunity to push for similar privacy legislation in the United States.

Big Data collection and use is not just under the purview of federal statutes or federal agency regulations; Big Data is affected by the state privacy laws that directly address online disclosures and record keeping. States have the power to go further in privacy regulation than the federal government, and large states such as California and New York are legal influencers because of their size and level of commercial activity.

Generally, US laws addressing data collection are outdated and inadequate. With the coming wave of regulation in the US, it will be critical for businesses to utilize Big Data in safe and secure ways. If companies whose businesses are dependent on data are found to be misusing that data, the consequences could be fatal to a business.
A 2018 Mind Commerce report identifies IBM as the industry leader of the Big Data market in terms of current investment.

<table>
<thead>
<tr>
<th>RANK</th>
<th>VENDOR</th>
<th>BIG DATA OFFERINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IBM</td>
<td>SERVICES, SOFTWARE, HARDWARE</td>
</tr>
<tr>
<td>2</td>
<td>HP</td>
<td>SERVICES, SOFTWARE, HARDWARE</td>
</tr>
<tr>
<td>3</td>
<td>TERADATA</td>
<td>SERVICES, SOFTWARE, HARDWARE</td>
</tr>
<tr>
<td>4</td>
<td>DELL</td>
<td>SERVICES, SOFTWARE, HARDWARE</td>
</tr>
<tr>
<td>5</td>
<td>ORACLE</td>
<td>SERVICES, SOFTWARE, HARDWARE</td>
</tr>
<tr>
<td>6</td>
<td>SAP</td>
<td>SERVICES, SOFTWARE</td>
</tr>
<tr>
<td>7</td>
<td>EMC</td>
<td>SERVICES, SOFTWARE, HARDWARE</td>
</tr>
<tr>
<td>8</td>
<td>CISCO</td>
<td>SERVICES, HARDWARE</td>
</tr>
<tr>
<td>9</td>
<td>MICROSOFT</td>
<td>SERVICES, SOFTWARE</td>
</tr>
<tr>
<td>10</td>
<td>ACCENTURE</td>
<td>SERVICES</td>
</tr>
</tbody>
</table>
**SWOT ANALYSIS: BIG DATA SOLUTION ADOPTION**

**STRENGTHS**
- Better decision making
- Increased productivity
- Reduced costs
- Improved customer service
- Fraud detection
- Increased revenue
- Increased agility
- Faster speed to market

**WEAKNESSES**
- Workforce Re-skilling and Organizational Resistance
  - Re-skilling refers to the training of IT personnel in the new technologies involved in supporting Big Data analytics and enabling a large part of the rest of a company to create or use Big Data analytics in key business functions.
  - Many enterprises are witnessing resistance towards adopting Big Data analytics instead of using human driven decisions.
Unclear Big Data Strategies
Many organizations around the globe lack a clear and formulated Big Data strategy. With overall slower revenue in the last decade, companies have focused more on driving profits by cutting costs and less on undertaking major new projects. This has impeded many organizations from creating a clear and optimal Big Data initiative.

Technical challenges: Scalability & maintenance
Maximizing data use may help companies in many ways but could also become increasingly burdensome, as they must be responsible for managing much greater volumes of information.

Big Data Development Expertise
Essential to the effective execution of Big Data strategies will be the Data Scientists acquired to do the work. However, this field is very new, and educational institutions are still in the process of developing and maturing their data science curriculums. Hiring well trained Data Scientists will be key for the success of Big Data projects.

Businesses also face the issue of scarce data scientists: they are extremely expensive which puts smaller businesses at a significant disadvantage to leverage analytics. Additionally, data scientists may have significant data expertise, but lack business knowledge. Ultimately, data scientists are human and their attention span and ability to do repetitive work are limited. Human data scientists typically analyze 10% of a business’s data, based on which they think is most important.

Difficulty Integrating Legacy Systems
Many companies that have been at the top of their industries for decades have siloed their data in a variety of applications. These companies face major challenges integrating disparate data sources and moving data to new solutions.
OPPORTUNITIES

The following security needs will create persistent and growing demand for cybersecurity solutions:

**Securing State Data**
State data that is currently stored in different databases like NAS (Network Attached Storage) and SAN (Storage Area Network) need to be secured. This includes data such as credit card information, social security numbers, financial information, intellectual property, and others. This data can be secured through data encryption or data federation. These two cyber security techniques protect data from theft, unauthorized access, and visibility.

**Securing APIs**
It will be crucial to secure APIs (the set of functions which allow the creation of applications that access the features or data of an operating system or other service). Industry-standard authentication protocols will have to be used such as basic authentication, Oauth protocols, custom authentication protocols, and JSON Web Token.

**Securing Applications**
Application security starts with protecting the coding of confidentiality, integrity, and accessibility of data and information. Securing code will protect applications from targeted attacks.

**Securing Data for Analysis**
Securing data will prevent unauthorized access to databases and potential corruption or infection.

**Securing User Privileges**
The following steps can be implemented to secure user privileges:

- Define user roles and map them
- Grant and revoke from public role
- Grant request as per defined and assigned user roles
- Maintain an authorized list of users
- Map resources with user profiles
- Revoke user privileges and roles whenever necessary
- Store lists in database along with full user profiles
- Use operating system or network to grant or revoke user roles
Securing Enterprise Data

To protect their information, enterprises must develop enterprise information security architecture (EISA) with defined roles.

THREATS

Privacy and Security

As businesses become more reliant on Big Data, the stakes for protecting it will also increase and cyberattacks will occur more frequently. Cyberattacks will have the potential to be devastating to a company, even fatal. Companies will possess vast amounts of sensitive information; if sensitive personal consumer data were to be leaked, companies could be legally liable and lose the majority of their user bases.

Distributed frameworks such as Hadoop allow data processing to be distributed throughout different systems, thus putting less strain on individual systems. The drawback is that there are more systems where security issues can crop up.

Given the lack of US data regulation, the space resembles the Wild West. The companies who abuse this lack of regulation could face severe repercussions in the future: data gathering and analytical techniques which may be the core of a business today could be subject to government regulation in the tomorrow.
NOTABLE EXITS

IPOS

Pivotal Software
Pivotal is a software and services company that provides a cloud-native platform and Big Data products. They raised $638 million in their IPO on April 20th of 2018. A total of 42,533,333 shares were sold at $15 per share.

Domo
Domo offers a cloud-based operating system that integrates every component of a business with real-time business dashboard visualization and key performance indicator reporting. The platform unifies data, systems, and people for a digitally connected business. The company raised $193.2 million in its IPO on June 29, 2018. They sold 9,200,000 shares at $21 per share. After the offering, there was a total of 24,953,806 outstanding shares.

M&A

Dell VMware
VMware is a publicly listed Data Center and Cloud Infrastructure company that was acquired for $24 billion in a reverse merger by Dell technologies in December of 2018. Dell Technologies paid $14 billion in cash and issued 149,387,617 shares of its Class C common stock.

Mobileye and Intel
Mobileye develops collision avoidance systems designed to reduce vehicle injuries and fatalities. The company was acquired by Intel for $15.3 billion on September 1, 2017. The acquisition will allow Mobileye’s leading computer vision technology and expertise to combine with Intel’s high-performance computing and connectivity expertise to create automated driving solutions from cloud to car. The integrated company plans on building a fleet of 100 autonomous test cars to then present to prospective customers in a real-world landscape.
Hewlett Packard and Micro Focus Group
The software division of Hewlett Packard Enterprise was acquired by Micro Focus Group for $8.8 billion on September 1, 2017.

Hortonworks and Cloudera
Hortonworks creates, distributes, and supports enterprise-ready open data platforms and modern data applications. Cloudera develops a software platform for data engineering, data warehousing, machine learning and analytics that runs in the cloud or on premises. The company acquired Hortonworks for $5.2 billion on January 3, 2019. Cloudera now owns 60% of the equity of the combined company and Hortonworks stockholders will own the other 40%. The acquisition will help Cloudera optimize sales and research and development.
WHAT IS COMING?

AUGMENTED ANALYTICS

Generating valuable insights from data still remains a huge challenge for almost all businesses. Data, on its own is completely useless. A business may look at an online data point which reveals that revenue has decreased 12% from the last month. What does this mean to their business? Is there a general declining industry trend? Are some or one of the businesses advertising channels not working?

These questions force analytics to assess ecommerce, social media, and other market indicators to understand the decline in revenue. These changes, once identified, need to be placed into a business context to identify which ones can be acted upon immediately. This is the point where data becomes actionable.

From raw data to insights, the following steps need to be executed:

1. Capture data from multiple sources
2. Clean data so it is ready for analysis
3. Conduct the analysis
4. Generate insights
5. Communicate those insights with the business and convert them into actionable strategies

Augmented Analytics will solve all of these challenges by relieving a company’s dependence on scarce data scientists and by automating insight generation in a company through the use of advanced machine learning and artificial intelligence algorithms. An augmented analytics engine will be able to go through a company’s data, clean it, analyze it, and convert insights into actionable steps for the executives or marketers with little to no supervision from a technical person. Currently, the most sophisticated analytics tools support analysis; they do not conduct it. But eventually, the need for data scientists and business analysts could essentially become obsolete.
Augmented Analytics are still nascent but will grow remarkably fast in the next couple of years. There are three main stages for the development of augmented analytics technologies:

**NOW: Data Preparation and Discovery**

Most augmented analytics technologies are currently in this stage. In this stage, augmented analytics algorithms serve as a great complement to existing data scientists, but do not yet have the ability to replace them.

Here, algorithm's primary purpose is to automate boring data preparation tasks such as data cleaning, data labeling, and data collection.

**2-3 YEARS: Signal Detection**

At this stage, algorithms can detect true signals in a company's data with extreme reliability but are unable to connect these discoveries with business situations or business actions. This phase will synthesize within the next 2-3 years.

**5-10 YEARS: Actionable Insight Generation**

At this stage, augmented analytics can directly interface with executives in the company with little or no need for input from a business analyst or data scientist.

There will be a large knowledge base of past business cases which will be developed to help augmented analytics systems to connect trends in the company's data to the larger context of business. It can then offer concrete actionable steps based on these insights.

The system will also be able to track the implementation of these actions and provide additional insights on what the company can do better next time to maximize its operational effectiveness. This stage will be reached within 5-10 years.
THANK YOU FOR READING

WE HOPE YOU FOUND THIS REPORT INSIGHTFUL

AT ADIT, we’re committed to keeping investors informed about the cultural and technological trends that are shifting the markets and shaping our world. Explore our website for more insightful reports on key emerging industries, and join our newsletter for regular updates on these important evolutions in the market.

We also offer access to equity in private companies that are driving these trends, including tailored opportunities for growth equity investments in what we believe are some of the world’s premiere emerging enterprises. To learn more -- and to start creating a portfolio that’s built for tomorrow’s marketplace -- contact our team of private equity investors.

WWW.ADITVENTURES.VC

1345 AVENUE OF THE AMERICAS, FL 33
NEW YORK, NY 10105

646.840.4944

DISCLAIMER

This is not an offer to purchase securities. Any offer to purchase securities must be made in accordance with all local and federal securities regulations, and by a private placement memorandum. All the information provided is deemed to be accurate and obtained from reliable sources. However, there can be no assurances that the information herein is correct and has not been independently verified by Adit Ventures, LLC or any of its affiliates. All investors should make their own investment decisions in conjunction with accountants, advisors and or counsel. There is a risk supplement specific to each individual company which includes information on each of the emerging growth equities we invest in, outlining the issues facing the firms. The information herein is believed to be reliable and has been obtained from public sources believed to be reliable. All information herein reflects the opinions of Adit Ventures, LLC and is subject to change without notice.

The graphs, charts and other visual aids are provided for informational purposes only. None of these graphs, charts or visual aids can of themselves be used to make investment decisions. No representation is made that these will assist any person in making investment decisions and no graph, chart or other visual aid can capture all factors and variables required in making such decisions.

Any projections, forecasts and estimates contained in this material are necessarily speculative in nature and are based upon certain assumptions. It can be expected that some or all of such assumptions will not materialize or will vary significantly from actual results. The companies identified were chosen for general commentary and should not be interpreted as any endorsement or recommendation to invest in such companies nor is it a representation of their services provided or profitability. Investments in any companies or in any fund that Adit Ventures, LLC advises is subject to substantial risk including the complete loss of the principal amount invested.