Chapter 11
Data Management

Big Data

The term "big data" was coined in mid-1990s and is defined as collections of data so large, complex, and dynamic that they exceed the processing capacity of the conventional database architectures of organizations (Weiss and Indurkhya 1998). According to Gartner, the world's leading information technology research and advisory company, big data is comprised of high-volume, high-velocity, and high-variety data (the '3 Vs', as shown in Fig. 11.1 (Gartner IT Glossary 2019). These data sets are too large to be handled easily and flow in and out with excessive speed, making them difficult to analyze. The range and type of data sources are too great to assimilate (Diebold 2012).

The typical organization is therefore challenged in managing big data effectively, as it simply does not fit into the strictures of current database architectures. At the same time, big data draws from multiple sources and transactions and contains valuable patterns and information. The act of gathering and storing large amounts of data for eventual analysis is not new. Since the 1950s, businesses have been using basic analytics to uncover hidden patterns and trends, show changes over time, and confirm or challenge theories (Asllani 2015). As enterprises amass broader pools of data in big data platforms, they have increased opportunities to mine those data for predictive insights. As they cannot manage the data effectively with their current database architecture, they need to seek alternative ways to process the volume (Bayrak 2015). A well-defined data management strategy is essential for the successful use of big data in corporations (Bughin 2016). Data and analytics are playing increasingly important roles in improving competitive advantage (Taylor 2012), and corporations see big data and the ability to analyze it as an important driver of innovation and a significant source of value creation (Tan et al. 2015).

M. Attaran, A. Gunasekaran, *Applications of Blockchain Technology in Business*, SpringerBriefs in Operations Management,
https://doi.org/10.1007/978-3-030-27798-7_11

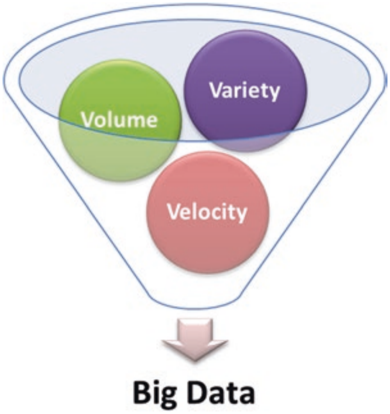

Fig. 11.1 Big data defined

Blockchain for Big Data

Blockchain is changing how the world approaches big data, and its technology could be well combined with big data. Blockchain could be used to store important data and ensure that the data is original as it is distributed and secured. Additionally, user trading patterns and transactions on blockchain could be extracted and used for big data analytics. Finally, as more and more blockchain applications for different fields are appearing, traditional industries could make use of blockchain to improve performance.

Blockchain can be used internally as a database for applications like managing physical and digital assets, recording internal transactions, and verifying identities. This may be an especially useful solution for companies struggling to reconcile multiple internal databases.

Blockchain is one of the fastest growing technologies to help secure and protect data through cryptography. The technology can be used in any market, sector, or application that needs to securely exchange data in a decentralized format.

Contracts, transactions, and the records of them are among the defining structures in our business, legal, and political systems. They govern interactions among nations, organizations, communities, and individuals. These critical tools—and the bureaucracies formed to manage them—have not kept up with the economy's digital transformation. Blockchain technology has the potential to solve this problem (Iansiti and Lakhani 2017).

International Data Corporation (IDC) estimated that worldwide revenues for big data and business analytics will grow from $130.1 billion in 2016 to more than $203 billion in 2020, at a compound annual growth rate (CAGR) of 11.7%. According to IDC, blockchain spending will rapidly expand during 2017–2022 forecast period, with a five-year CAGR of 73.2%. IDC further expected worldwide blockchain spending to reach $1.5 billion in 2018—double the amount spent on the emerging technology during 2017 (Rengegowda 2018).

The real value of data on the blockchain is the quality of data and how it is stored on the public ledger. *Blockchain technology* gives greater confidence in the integrity of the data we use. The technology provides immutable entries, consensus-driven timestamping, audit trails, and certainty about the origin of data. That means data will become far more valuable since it is being captured and validated on the blockchain. Path is an example of a blockchain platform that helps companies improve the data sets they analyze and leverage for their products and services. Using a Path platform, a company is able to get deeper insights into how its website performs around the world, or how long it takes for its application to load, how cluttered its network gets at certain hours of the day, etc. (Cole 2018).

Currently, the most transformative blockchain application is smart contracts. Smart contracts are self-automated computer programs that can carry out the terms of any contract. They are automated contracts and self-executing with specific instructions written in its code that get executed when certain conditions are made. Although smart contracts come with many advantages, we are decades away from widespread adoption. They will not be effective without institutional buy-in and a high degree of coordination and clarity on how smart contracts are designed, verified, implemented, and enforced (Iansiti and Lakhani 2017).

Cloud Computing Technology (CCT)

The term "cloud" is used to refer to different types of platforms for distributed computing—a cluster of servers, network, software, interface, etc. that users require to execute a particular task. "Computing" refers to the delivery of this package as a service that users can utilize as they wish (Attaran 2017). The user does not need to own a massive computing infrastructure. Rather, the user can utilize a similar infrastructure owned by a third party, and pay only for the amount of computing needed. This pay-per-use model allows convenient, on-demand network access and timesaving in building huge computing infrastructure. This enables the user to concentrate efforts on critical business activities (Mell and Grance 2011). The user accesses information online in a 24/7 format from a variety of devices—desktop, laptop, tablet, and smartphone (Bask 2015)

The National Institute of Standards and Technology (NIST) describes five characteristics of a cloud-computing model. (Mell and Grance 2011):

1. **On-demand self-service.** Server time, network storage, and other computing resources are obtained and configured, as needed. No human interaction with the service provider is required
2. **Broad network access,** The service is accessed over a network using multiple platforms (e.g., mobile phones, tablets, laptops, and workstations)
3. **Resource pooling.** Resources are pooled to serve multiple users. No user has exclusive use of the underlying hardware or software resources. According to demand, physical and virtual resources are assigned and reassigned dynamically

4. **Rapid elasticity.** Resources scale rapidly up and down with demand, since they are elastically provisioned and released
5. **Measured service.** Metering is used to automatically optimize resource use (e.g., storage, processing, bandwidth, and active user accounts). There is a direct relationship between use and cost.

According to a 2016 Gartner report, CCT is perhaps the most promising and anticipated technology to come around in a number of years (Smith 2016). For some businesses, making a heavy move toward a cloud structure is a way to significantly cut hardware costs. For others, CCT streamlines operations and speeds up development cycles. Properly planned and implemented, CCT has the potential to drastically improve operational efficiency of businesses.

Cloud-Based Blockchain

In the past few years, centralized proprietary services offered on the Internet are being replaced with decentralized open ones—inefficient monolithic services replaced with peer-to-peer algorithmic markets. Blockchain technology first decentralized money with bitcoin and is now moving ahead to decentralize other aspects of business process. The technology is being used to scrutinize every process and application in various industries in order to improve them. As a result, decentralized storage networks (DSN) are emerging to challenge the giant traditional cloud storage companies like DropBox, OneDrive, and Amazon. These companies have relied on centralized services where clients give up sovereignty of their data and are at the companies' mercy. DSN aggregates storage offered by multiple independent storage providers and self-coordinates to provide data storage and data retrieval to clients. With the rise of decentralized cloud storage services powered by blockchain, the days of stress over data theft, hacking, and copying clients' data and selling it elsewhere will be gone (Khatwani 2018).

Blockchain-based technologies have proven the utility of decentralized transaction ledgers. They liberate data from silos, work offline, route around censorship, and gives permanence to digital information. Blockchain technology can help business simplify complex business processes and stay ahead in the industry. Blockchain enables financial services, manufacturing, healthcare, government, and retail to realize the potential of blockchain capabilities in the cloud. Those include faster and more secure transactions, streamlined and automated back office operations, and reduced costs.

Blockchain technology offers great data integrity advantages for big data, including:

- Greater confidence in the integrity of the data
- Immutable entries
- Consensus-driven timestamping

- Audit trails
- Certainty about the origin of data

Beyond data integrity, the inherent immutability that blockchains offers leads to more confidence in training and testing data and the models they produce. Additionally, blockchain-based technology could make an impact in the cost of storing data and in the amount (and quality) of data available. Cost savings in data storage will come from the disintermediation of centralized cloud storage providers. This should also create downward pricing pressure on software-as-a-service (SaaS) providers as they move to decentralized storage providers. This decentralized approach could reduce the costs of storing data by 90% compared to Amazon Web Services (AWS) (Epstein 2017). Beyond data privacy and security, if the data center's power goes down or data become corrupted, the blockchain's algorithms ensure data is distributed widely enough to maintain high availability (Mallon 2018).

Cloud-based blockchain services are emerging from both start-ups and large platforms like Amazon and Microsoft (Iansiti and Lakhani 2017). Microsoft's blockchain as a service (BaaS) platform is backed by Microsoft Azure, which provides hybrid cloud capabilities, extensive compliance certification portfolio, and enterprise-grade security. Other decentralized solutions like Storj, Sia, MaidSafe, and FileCoin are starting to provide services in the enterprise storage space (Epstein 2017). Some of these decentralized cloud storage solutions offer a single trading platform and a marketplace for hosts to rent out extra space on their personal computers to clients who wants to use it.

In the future, public blockchains will ultimately defeat private blockchains. The technology has the potential to move us from proprietary data silos to blockchain-enabled shared data layers. In this scenario, the power of those who owned the data will shift to those who can access the most data and who can gain the most insights most rapidly (Epstein 2017). When data moves onto open blockchains, having the data is no longer a competitive advantage. Interpreting the data becomes the advantage. Additionally, customer data will not belong to organizations, it will belong to each individual and the customer of the future will grant access to others as necessary (Epstein 2017). A massive amount of storage space is unused in data centers and hard drives around the world. Using blockchain technology, companies can tune their storage strategy to suit their needs, creating a custom balance between redundancy, speed of retrieval, and cost.

Some companies specialize in combining big data and blockchain to bring all the benefits of decentralized systems, while also offering new insights and increased profits to businesses. Table 11.1 summarizes examples of blockchain-based decentralized data storage providers that are delivering savings for big data and competing with traditional cloud storage solutions (Mallon 2018; Kh 2019).

Table 11.1 Big data blockchain-based projects

Company	Services Provided
Omnilytics	• Combines blockchain with big data analytics • It uses artificial intelligence (AI) and Machine Learning (ML) with marketing, trend forecasting, auditing, and applications
Datum	• It is a decentralized storage network • Enables individuals to monetize their own data in an open marketplace
Storj	• It is a decentralized data storage provider • Provides saving for big data in comparison to traditional cloud storage companies • It is the largest, cheapest, and most secure cloud service • Does not create a free market space for lending client extra space
Sia	• It is a decentralized peer-to-peer cloud storage solution that aims at reducing cost and increasing the security of stored data • It creates a marketplace for hosts and users via the token economics of Siacoin • Hosts can rent out extra space on their PC to the Sia decentralized network and get paid in Siacoin. Similarly, client who want to use space pay Siacoins to a host
Filecoin	• It is a decentralized storage network that turns cloud storage into an algorithmic market • Miners earn Filecoin by providing storage to clients. Clients spend Filecoin hiring miners to store or distribute data
Rublix	• Aims to unite cryptocurrency investors across the world • Offers a simple trading platform that verifies the authenticity and credibility of traders • Provides access to market information to reduce the current market confusion
InboundMuse	• It provides a wide range of services related to artificial intelligence and blockchain • The company has domain expertise in finance, digital marketing, and recruitment
Hidden Brains	• Provides expertise in healthcare, real estate, logistics, education, energy, and other industries

Source: Mallon 2018; Kh 2019

Monetizing Big Data

Many of the world's largest tech companies are looking to collect behavioral data during regular customer interaction with their products. This information is collected from monitoring social media and the way customers use their web pages. Companies perform analytics to draw insight from this data and to develop a plan of action to streamline the client's experience. Blockchain can also be a link for big data processors. The technology created a system where content creators are able to sell content without needing a central authority to monitor it. Companies that create large amounts of data are able to profit from it directly by setting their own prices and selling it to firms who need it (Piletic 2018).

Big Data Analytics

Analytics, in the form of business intelligence, is defined as a set of technologies, processes, and tools that use data to predict likely behavior by individuals, machinery, or other entities. By using the right analytics, big data can deliver richer insights and uncover hidden patterns and relationships. More data could translate into more possibilities for a business only if it can discover the meaning inside of it (Minelli et al. 2013).

Since the 1950s, decades before anyone used the term big data, businesses were using basic analytics to uncover hidden patterns and trends, show changes over time, and confirm or challenge theories. The new benefits that modern data analytics brings to the table are speed and efficiency. The ability to work faster—and stay agile—gives organizations a competitive edge. Over the past decade, data exploded and became big, and business intelligence has been revolutionized. Widespread access to the cloud, insightful data visualizations, interactive business dashboards, and rise of self-service analytics made the technology available and affordable for businesses of all sizes. Suddenly advanced analytics is not just for the analysts (Gaitho 2017). Analytics is commonly used in finance, marketing, human resources, healthcare, government policies, and every industry where data is generated (DeAngelis 2015; Decker 2017).

The past few years have seen an explosion in business use of analytics. Corporations around the world are using analytical tools, including business intelligence (BI), dashboards, and data mining to gain a better understanding of their present customers and to predict the needs of those who will potentially become customers. With the help of new analytics tools, enterprises can leverage big data analytics to drive a host of business objectives, from streamlined operations to improved customer relations (Henke et al. 2016). In fact, big data analytics will transform virtually every business activity, and bring businesses benefits including enhanced customer service, optimized production levels, superior capacity planning, reduced repair and maintenance costs, and improved working capital utilization (Bughin 2016). According to a 2016 Forester study, the top three tangible analytics benefits were identified as: increased margin, profitability, and increased gross sales (Evelson and Bennett 2015). Several research studies have documented the advantages and widespread applications of analytics tools in corporations around the world (Evelson and Bennett 2015; Gaitho 2017; Lebied 2016; Eckerson 2016; Henke et al. 2016; Minelli et al. 2013).

Analytics is constantly evolving. It has changed dramatically over the years and is advancing rapidly today. There are four categories of analytics: descriptive, diagnostics, predictive, and prescriptive (Attaran and Attaran 2018). These categories build on each other and enable enterprises to make faster and smarter decisions. As organizations evolve, they move from focusing on historical "what" and "why" questions to a more forward-looking predictive and prescriptive predictions. Descriptive analytics is the simplest class of analytics. It allows you to condense big data into smaller, more useful nuggets of information. The purpose of descriptive analytics is to summarize what happened in the past and uncover patterns that offer insights. It uses data modeling, reporting, visualization, and regression to collect and store data in an efficient way, to create reports and presentation information, and to identify trends. Since data is scattered in large numbers of disparate data sources, analyzing all relevant information can be a challenge for most organizations. Most descriptive analytics are exact (number of likes, number of clicks, etc.) because they are defined by a single, deterministic model that does not allow contradicting results. It is estimated that more than 80% of business analytics, most notably social analytics, are descriptive (Attaran and Attaran 2018).

Diagnostic analytics is the next step up in data reduction and it is used for discovery. It examines data or content to answer the question: Why did it happen? Diagnostic analytics takes a deeper look at data to attempt to understand the root causes of events and behaviors in an organization. To optimize diagnostic analytics, it needs to be extended to operational employees of the organization. The result of the diagnostic analytics is often an analytic dashboard that is used for discovery or to determine why something happened (Attaran and Attaran 2018).

Predictive analytics analyze current and historical data to provide insights into what will happen and why it will happen in the future with an acceptable level of reliability. It attempts to accurately project the future conditions and states. It does not predict one possible future, but rather multiple futures based on the decision-maker's actions. It utilizes a variety of statistical, modeling, data mining, text, media mining, forecasting, and predictive modeling to identify probabilities of potential outcomes and/or likely results of specific operations. Predictive analytics can only forecast what might happen in the future, because all predictive analytics are probabilistic in nature. Predictive analytics can help businesses with a wide range of problems. Businesses are using predictive analytics to analyze historical data and facts in order to better understand clients' needs, market potentials, products, suppliers, and partners, and to identify potential risks and opportunities for a company (Attaran and Attaran 2018; Lebied 2016). Other businesses use this type of analytics for predicting which customers are most valuable, scheduling preventive maintenance, and detecting fraud. Airlines use these analytics to decide how many tickets to sell at what discounted price for a flight. Similarly, hotels use it to predict the number of guests they can expect to book on any given night to maximize revenue.

The emerging technology of prescriptive analytics goes beyond descriptive and predictive models and shows the likely outcome of each decision. It goes a step further in the future and attempts to answers what should be done and why. It employs data techniques such as decision modeling, graph analysis, simulation, neural networks, heuristics, and ML to suggest actions that the organization could take to achieve the desired outcome (Attaran and Deb 2018). Prescriptive analysis tries to evaluate the effect of future decisions in order to adjust the decisions before they are actually made. Future outcomes are taken into consideration in the prediction. Prescriptive analytics are commonly used in organizations to optimize scheduling, production, inventory and supply chain design, and other organizational activities to deliver what the customers want, and meet and exceed customers' expectations. Prescriptive analytics is the most valuable kind of analytics and usually results in rules and recommendations for next steps. However, it is largely not used (Attaran and Attaran 2018). A 2012 survey by Gartner shows most large retailers still focus on measurement of the past, with only 13% of them making extensive use of predictive analytics. Less than 3% of large retailers are using prescriptive analytics tools, such as decision/mathematical modeling, simulation, and optimization (Hetu 2015). Another study shows that by 2020, 40% of new investment in analytics tools will be in predictive and prescriptive analytics.

Blockchain and Big Data Analytics

Both blockchain and big data analytics deal with data. The latter analyses data for predictions and actionable insights while the former records and validates data for integrity. They both use algorithms to govern the interactions with different data segments. Big data analytics transform raw data into easy-to-understand and actionable insights and, like any other technology, have its own challenges and limitations. Some of these challenges include inaccessible data, dirty data (duplicate and incorrect data), and privacy issues. Dirty data is considered as the biggest challenge to big data analytics. According to Kaggle in a 2017 industry-wide survey of 16,000 data professionals, dirty data was the most common problem for workers in the data science realm (Kaggle 2017). Big data analytics is focused on making predictions from large amounts of data while blockchain is concerned with validating data. Blockchain manages and operates data not in a centralized manner where all data is brought together. It does it in a decentralized manner where data may be analyzed right off the edges of individual devices. Blockchain integrates with other technologies, like CCT, AI and the IoT. Additionally, blockchain generates immutable data that are validated, structured, and complete.

There are several ways blockchain can bring a whole new way of managing and operating big data analytics (Sarikaya 2019):

1. **Data integrity and trust**. Data recorded on the blockchain have gone through a verification process that ensures its quality and reliability. Blockchain technology generates validated data that are structured, complete, and transparent to ensure data integrity and trust, and prevent malicious activities.
2. **Preventing hacking.** Blockchain peer-to-peer distributed networks use consensus algorithm to verify transactions. It is impossible for a single node or unit to

pose a threat to the data network. Furthermore, it is almost impossible for a single party to alter the validation criteria and allow unwanted data in the system.

3. **Performing predictive analytics.** As discussed earlier in this chapter, predictive analysis uses large sets of data to determine with good accuracy the outcome of social events, such as customer preferences, customer lifetime value, dynamic prices, etc. Blockchain provides structured data gathered from individuals or individual devices. Data recorded on the blockchain can be analyzed to reveal valuable insights into the behaviors and trends and can be used to predict future outcomes. Because blockchain uses distributed network and huge computational power even smaller organizations can undertake extensive predictive analysis projects. Organizations can use the computational power of numerous computers connected on a blockchain network as a cloud-based service to analyze big data in a scale that would not have been otherwise possible.
4. **Real-time data analysis.** Blockchain transactions are made in real-time and are settled and verifiable within seconds. In the same manner, a blockchain-enabled system can help organizations achieve real-time analysis of large-scale data. The speed and accuracy of blockchain allows organizations to observe changes in data and enables them to make quick decisions.
5. **Data-sharing management.** Blockchain securely stores data obtained from data studies in its network. This prevents project teams from repeating data analysis already carried out by other teams or reusing data. Also, a blockchain platform can help businesses trade outcome of their analysis and monetize their work.
6. **Audit trails.** Activities and transactions that take place on the blockchain network can be traced for transparency.

Challenges for Blockchain in Data Analytics

As blockchain technology matures, big data analytics will be one of the areas that will benefit the most. Currently, existing peer-to-peer blockchain networks can only process small amount of data and a handful of transactions per second. Storage methods and processing capability of blockchain networks will need to be improved to deal with large volumes of data collected per second for big data and other data analysis tasks. Additionally, data storage on a blockchain is expensive compared to traditional means (Sarikaya 2019).

Blockchain and Ricardian Contracts

Ian Grigg, a pioneer in financial cryptography, introduced Ricardian contracts in 1995. A Ricardian contract uses cryptographic signatures and is a “form of digital documents that act as an agreement between two parties on the terms and condition for an interaction between the agreed parties” (Alam 2018). Ricardian contracts not only define intentions, but also execute instructions automatically when certain pre-conditions have been met. Ricardian contracts have several advantages (Alam 2018; Rijmenam 2018):

- They are easy-to-read legal contract for lawyers and the parties involved
- They are both machine-readable, as well as human readable
- Ricardian contracts can be easily hashed, signed, and can be saved on the blockchain
- In case of a dispute among parties involved, the case can be decided in court. This is not possible with smart contracts since they are not legally binding agreements.
- Ricardian contracts are extremely secure.
- Ricardian contracts come with a legal framework, since lawyers are required to create and deploy them. This adds clarity for all stakeholders.

The above benefits of Ricardian contracts make them a prerequisite for fast, efficient, and secure transactions and will be blockchain's killer application. We will see more implementations of Ricardian contracts in decentralized applications in the future.

References

Alam I (2018) What are Ricardian contracts? a complete guide. *101 Blockchain*. October, 28. Retrieved May 02, 2019, from https://101blockchains.com/ricardian-contracts/?utm_source=datafloq&utm_medium=ref&utm_campaign=datafloq

Asllani A (2015) Business analytics with management science, models and methods. Pearson Education Inc, Upper Saddle River

Attaran M (2017) Cloud computing technology: leveraging the power of the internet to improve business performance. J Int Technol Inf Manag,. Summer, 26(1):112–137

Attaran M, Attaran S (2018) Opportunities and challenges of implementing predictive analytics for competitive advantage. Int J Bus Intell Res 9(2):1–26

Attaran M, Deb P (2018) Machine learning: the new 'big thing' for competitive advantage. Int J Knowl Eng Data Min 5(4):277–305

Bask I (2015) "Why cloud technology is the smart move right from start up." Entrepreneur, April 2. https://www.entrepreneur.com/article/241914

Bayrak T (2015) A review of business analytics: a business enabler or another passing fad. Procedia Soc Behav Sci. 195:230–239.

Bughin J (2016) Big data: getting a better read on performance. McKinsey Quarterly, February. Retrieved May 25, 2019, from https://www.mckinsey.com/industries/high-tech/our-insights/big-data-getting-a-better-read-on-performance

Cole N (2018) Here's why big data is going to love blockchain technology. Medium Business September 6. Retrieved April 12, 2019, from https://medium.com/@nicolascole77/heres-why-big-data-is-going-to-love-blockchain-technology-d82904cca8f8

DeAngelis SF (2015) Predictive analytics becoming a mainstream business tool. Enterra Solutions. Retrieved June 12, 2019 from https://www.enterrasolutions.com/blog/predictive-analytics-becoming-a-mainstream-business-tool/

Decker J (2017) Embedded analytics for dummies. Wiley

Diebold FX (2012) On the origin(s) and development of the term big data. April 30. (PIER Working Paper 12–037). Penn Institute for Economic Research, Department of Economics, University of Pennsylvania. Retrieved from https://economics.sas.upenn.edu/sites/economics.sas.upenn.edu/files/12-037.pdf

Eckerson W (2016) Embedded analytics: the future of business intelligence. Eckerson Group. April. Retrieved May 25, 2019, from https://www3.technologyevaluation.com/research/white-paper/embedded-analytics-the-futureof-business-intelligence.html

Epstein J (2017) When blockchain meets big data, the payoff will be huge. NEVER STOP MARKETING. July 30. Retrieved April 12, 2019, from https://venturebeat.com/2017/07/30/when-blockchain-meets-big-data-the-payoff-will-be-huge/

Evelson B, Bennett M (2015) Quantify tangible business value of BI. *Forrester.* January 8. Retrieved Retrieved May 24, 2019, from http://www.lavastorm.com/assets/2015-Forrester-Quantify-Tangible-Business-Value-of-BI.pdf

Gaitho M (2017) How applications of big data drive industries. SimpliLearn August 8. Retrieved March 24, 2019, from https://www.simplilearn.com/big-data-applications-in-industries-article

Gartner IT Glossary (2019) Big data. Retrieved April 12, 2019, from https://www.gart ner.com/it-glossary/big-data

Henke N, Bughin J, Chui M, Manyika J, Saleh T, Wiseman B, Sethupathy G (2016) The age of analytics: competing in a data-driven world. Mckinsey Global Institute

Hetu R (2015) Retailers increasing predictive analytics capabilities. Gartner. Retrieved March 24, 2019 from https://blogs.gartner.com/robert-hetu/retailers-increasing-predictive-analytics-capabilities/

Iansiti M Lakhani K (2017) The truth about blockchain. Harvard Business Review. January–February Issue. Retrieved April 12, 2019, from https://hbr.org/2017/01/the-truth-about-blockchain?utm_source=datafloq&utm_medium=ref&utm_campaign=datafloq

Kaggle (2017) The state of data science & machine learning. Retrieved April 12, 2019, from https://www.kaggle.com/surveys/2017

Kh R (2019) 30 Top artificial intelligence and machine learning companies. *SmartDataCollective*. February, 21. Retrieved April 15, 2019, from https://www.smartdatacollective.com/30-top-artificial-intelligence-and-machine-learning-companies/

Khatwani S (2018) A look at the top decentralized storage networks (DSN) & their tokens *CoinSUTRA*. September 18. Retrieved April 15, 2019, from https://coinsutra.com/decentralized-storage-network-dsn/

Lebied M (2016) Top 11 business intelligence and analytics trends for 2017. Business Intelligence, December 15. Retrieved April 15, from http://www.datapine.com/blog/business-intelligence-trends-2017/

Mallon J (2018) 6 big data blockchain projects you should know about. SmartDataCollective August 29. Retrieved April 12, 2019, from https://www.smartdatacollective.com/6-big-data-blockchain-projects-you-should-know-about/

Mell P, & Grance T (2011) The NIST definition of cloud computing. Special Publication 800–145. http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

Minelli M, Chambers M, Dhiraj A (2013) Big data, big analytics: emerging business intelligence and analytic trends for today's businesses. Wiley, Hoboken

Piletic P (2018) How data monetization can add value to your analytics. SmartDataCollective. July 18. Retrieved April 15, 2019, from https://www.smartdatacollective.com/how-data-monetization-can-add-value-analytics/

Rengegowda D (2018) Blockchain: what ids it, how it works, and what it means for big data. September, 28. *Datafloq.com*. Retrieved April 12, 2019, from https://datafloq.com/read/blockchain-what-is-how-works-what-means-big-data/5532

Rijmenam (2018) Why blockchain is quickly becoming the gold standard for supply chain. November 21. Retrieved April 18, 2019, from https://vanrijmenam.nl/blockchain-becoming-gold-standard-supply-chains/?utm_source=datafloq&utm_medium=ref&utm_campaign=datafloq

Sarikaya S (2019) How blockchain will disrupt data science: 5 blockchain use case in big data. Towards Data Science. Jan 5. Retrieved April 18, 2019, from https://towardsdatascience.com/how-blockchain-will-disrupt-data-science-5-blockchain-use-cases-in-big-data-e2e254e3e0ab
Smith D (2016) Cloud computing deployments should begin with service definition. Gartner Report. Retrieved March 6, 2019, from https://www.gartner.com/doc/reprints?id=1-G2H8FE&ct=160826&st=sb
Tan KH, Zhan Y, Ji G, Ye F, Chang C (2015) Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. Int J Prod Econ 165:223–233.
Taylor P (2012) Crunch time for big data. The Financial Times, June 19
Weiss SM, Indurkhya N (1998) Predictive data mining: a practical guide. Morgan Kaufmann Publishers, Inc, San Francisco