In: The Palgrave Handbook of Supply Chain Management, Palgrave Macmillan, Cham, pp. 1-25
DOI:10.1007/978-3-030-89822-9_74-1

Big Data Applications in Supply Chain Management

Emel Aktas*

Abstract

This chapter overviews emerging applications of big data analytics in supply chain management. The academic attention on big data applications and their practitioner uptake is growing. Many recent papers showcase descriptive, predictive, and prescriptive analytics applications where multiple benefits emerge from applying big data analytics to managerial problems. Such benefits include cost reduction, increases in revenues and profits, and minimization of the environmental impact of operations. Current concerns include the transition from traditional to digital supply chains and what can realistically be achieved over the next two decades. While we evidence excellent applications of big data analytics for supply chain planning and management problems, the issue of working in siloes persists. For an organization to fully exploit big data applications, data should be perceived as an asset. When deploying novel artificial intelligence algorithms, the explainability of these algorithms should be at the forefront of an implementation strategy. Future research directions should be aimed at devising a connected and coordinated analytics approach that will enable the benefits of big data applications to go beyond what is currently realized.

Keywords Data capture, Descriptive Analytics, Predictive Analytics, Prescriptive Analytics, Logistics, Distribution, Warehousing, Retail

Introduction

Together with the fourth industrial revolution, big data and analytics have entered our everyday conversations with great potential to improve how we run our operations in terms of costs, environmental impact, and social implications. One of the contributing factors to big data being the conversation topic from marketing to supply chain is the digitalization of our supply chains, relocation of our enterprise data to cloud servers, so they are accurate and accessible real-time, and an astonishing increase of connected devices, mobiles, and wearables as well as sensors deployed on physical networks. All these developments are exciting for supply chain planning and operation purposes. This chapter aims to outline existing and future applications of big data and analytics.

IBM defines big data as "data sets whose size or type is beyond the ability of traditional relational databases to capture, manage, and process the data with low-latency" (https://www.ibm.com/uk-en/analytics/hadoop/big-data-analytics). Big data are digitally and passively produced, automatically collected, geographically and temporally traceable, and available in real-time or almost real-time. Sensors, connected devices, social media posts, and mobile apps make supply chain management a relevant testbed for big data tools and applications.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

Published by Springer. This is the Author Accepted Manuscript.
The final published version (version of record) is available online at DOI:10.1007/978-3-030-89822-9_74-1. Please refer to any applicable publisher terms of use.

Big data are typically characterized by 5Vs: volume, velocity, variety, variability/veracity, and value. Volume refers to the size of the data sets. Velocity represents the speed of data generation. Variety indicates the diversity of data. Variability/Veracity is concerned with the value or the meaning of a variable depending on the context. Value, which has recently started to be explicitly mentioned, highlights the benefits that could be achieved from using big data to address a strategic, tactical, or operational decision problem. Since handling big data requires hardware, software, and users with data science skills, applications should generate a value higher than the costs of collecting and analyzing data.

With the advent of computers in the 1960s, data were stored without structure, and many efforts were needed to create value from them. Relational databases that followed structured the data, making it easier to organize and elicit value. As the volume of structured data grew faster than the computing power, data warehouses and data marts came into place to manage the data and generate insights. Everything has changed with the Internet, as the volume of unstructured data started to grow exponentially, and data format was no longer limited to data tables. In addition to the web content, data started to flow from connected devices, sensors, cameras, GPS transceivers, leading to only a tiny amount of data collected being purposefully analyzed. With the costs of collecting, storing, processing, and analyzing data decreasing, distributed and cloud computing allowed massive amounts of data to be analyzed almost in real-time.

Popularity of Big Data and Supply Chain

Figure 1 shows a Google Trend chart comparing the terms “big data” and “supply chain” from the beginning of 2004 onwards. While the “supply chain” term was at its highest popularity in 2004, the “big data” did not appear in searches on a regular basis until after the end of 2010. The significant growth in the interest in “big data” is visible until the end of 2016 and stable until the beginning of the pandemic in March 2020. While there is a slowing of searches on big data from March 2020 onwards, the growth in supply chain searches continues to date.

The growth in big data applications can be attributed to the growth of data collected by the companies through the Internet (e.g., web searches, social media, electronic documents), mobile phones, wearable devices, and the Internet of Things (sensors). With the disruptions caused by the COVID-19 Pandemic, the US-China Trade War, and Brexit, many companies are already investing or expected to invest soon in digital tracking of data along the supply chain using sensors and other connected devices, analyzing supply chain interdependencies, and developing contingency plans.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

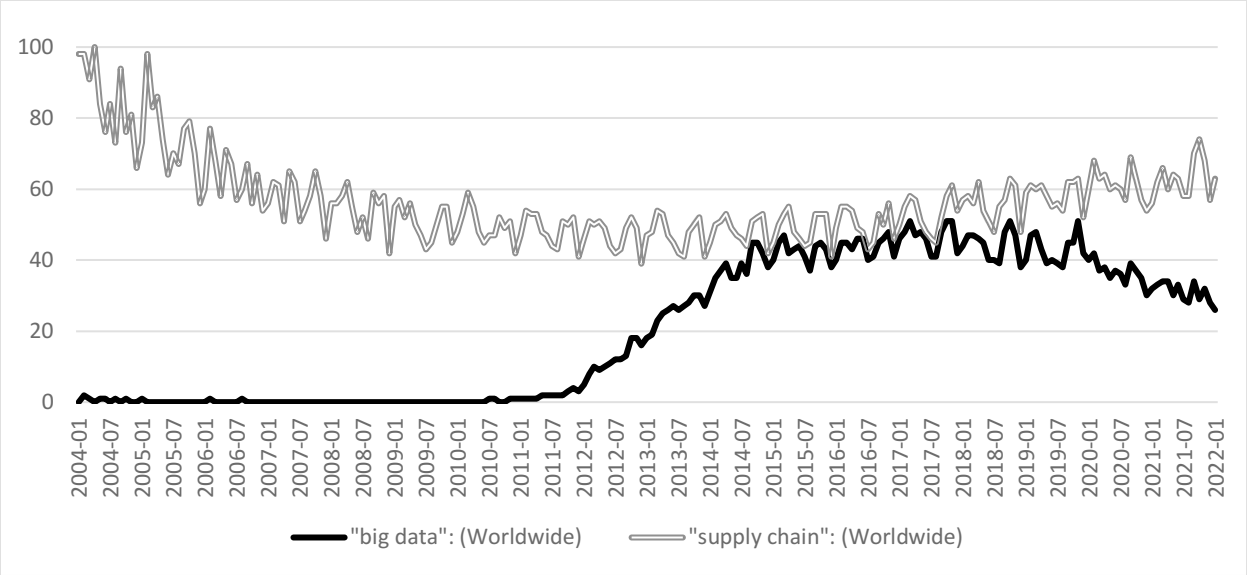

Figure 1 Comparison of "big data" and "supply chain" terms on Google Trends (https://trends.google.com/) date: Jan 2022

According to Research and Markets (2022), the big data market size is projected to grow from USD 162.6 billion in 2021 to USD 273.4 billion in 2026. This growth corresponds to a compound annual growth rate (CAGR) of 11.0%. Big data is perceived to have the potential to improve many supply chain operations such as order picking and stock-taking. Operational efficiency and revenue growth are expected to rely heavily on data-driven decisions at strategic, tactical, and operational levels. Many of the big data solutions providers are located in the Asia Pacific with expertise in India, China, and Japan.

Cloud platforms such as Microsoft Azure, Amazon Web Services, and Snowflake are the go-to platforms for storing data for many organizations. New terms such as "data fabric" and "data mesh" represent emerging architectures that access, manage, and integrate data across multiple platforms and technologies. While in traditional supply chains information flows in a linear fashion and is usually shared by the two sides (originator and receiver), the idea behind data fabric is to connect all entities in the supply chain so questions such as "who are the suppliers and the raw materials they provide for the finished good X?" can be queried within seconds (Megarbane, 2020).

Innovation Potential

Anyone studying supply chain management or conducting research in supply chain management cannot ignore the increasing relevance and significance of big data applications to innovation. Figure 2 shows the linear growth in patents issued with big data and supply chain focus, with an additional 200 patents every year over the past years. The topics of these patents include logistics storage bin allocation based on big data, intelligent information management systems for logistics transportation data, supply chain big data service systems, and computer architectures for big data calculations, to name a few. Many of the patents are focusing on big data acquisition and increasing visibility of product flows across the supply chain, which then lends itself to higher-level analytical approaches such as prediction and optimization.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

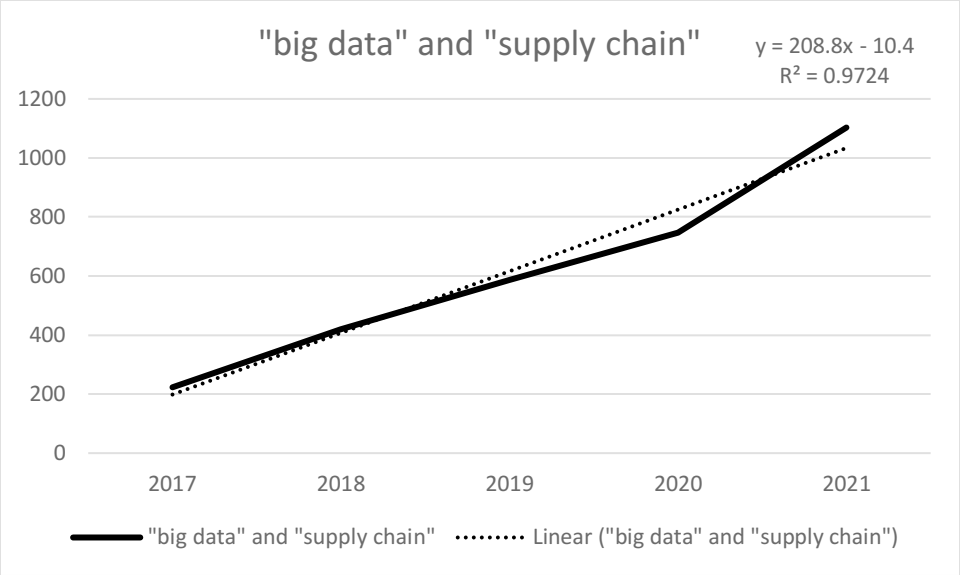

Figure 2 Patents published with big data and supply chain keywords over the past five years (data source: European Patent Office)

The United States has one of the most innovative economies in the world, with innovations emerging from US-based companies affecting people's lives globally. Since 2018, Innovation Roundtables have been held by the Bureau of Economic and Business Affairs with the US private sector to explore challenges and opportunities emerging in the information and communication technology sector. Among these roundtables, the Internet of Things is perceived as an increasingly high-profile area of technological development with software, devices, and operating systems. Similarly, Cloud Computing is recognized as the enabler of many services and technologies. The US Department of State is working on creating policy environments and regulatory frameworks that support the adoption of blockchain and distributed ledger technologies in manufacturing supply chains (US Department of State, 2022).

Digital transformation is at the heart of the European Union's research and innovation strategy, which strives to move the society towards a sustainable and prosperous future faster and in a way that is respecting European values (European Union, 2022). In the Strategic Plan for 2020-2024, making Europe fit for the digital age is recognized as one of the seven specific research and innovation objectives. Ethical use of artificial intelligence is highlighted along with data governance. Data-driven manufacturing of tailor-made products and the delivery of personalized services emphasize the need for big data applications in supply chains.

The UK government has the vision to make the UK a global hub for innovation by 2035 (UK Department for Business, Energy, and Industrial Strategy, 2022). For this purpose, annual public investment in research and development is set at around £22 billion. There is a strong emphasis on protecting data, research, and intellectual property to maintain competitiveness. Specific technologies related to big data and underlined in the national data strategy are artificial intelligence, computing, and augmented and virtual reality. The National Cyber Security Centre provides guidance on secure innovation to startups and

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

companies growing with emerging technologies. Cyber security company Darktrace is mentioned as the most innovative cyber defense company globally. Darktrace investigates defensive and adversarial artificial intelligence, social engineering, natural language processing, and graph theory to fight back against ransomware that disrupts businesses.

Job Prospects

Similar observations can be made for the job market. With a naïve search of "big data" terms in the query database, thousands of jobs were identified in top job portals such as Indeed, LinkedIn, and SimplyHired (Figure 3). The number of hits is lower but still in the order of hundreds in more UK-specific job portals such as the Gov.UK's job search engine, Reed, or CV Library. The search identified jobs in the UK that have "big data" in the title or the description. Job titles included "big data engineer", "big data architect", "big data developer", and "big data solution designer".

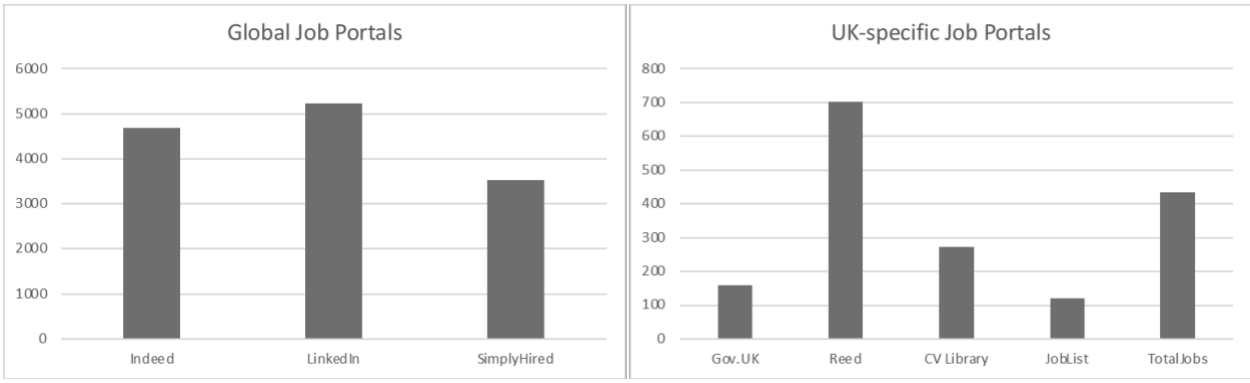

Figure 3 Jobs with "big data" skills in the UK. Author's search on Jan 20, 2022.

According to the Masters Portal (2022), there are 385 live data science and big data masters programs in the UK and 429 programs in the US as of February 2022. On the other hand, Germany has 56 degrees, France has 51, and Canada has 31. Many of these programs offer online learning as an alternative. Typical modules taught involve machine learning, neural networks and deep learning, big data processing, data mining, and natural language processing. Electives offered highlight sustainability and how data and models could help improve the environmental and social impacts of businesses.

Among the largest private companies in the UK, British Petrol is looking for big data engineers, data scientists, and cyber security experts. There are tens of open positions in HSBC UK that are related to big data expertise. Similarly, Unilever is recruiting hundreds of big data professionals globally. Typical roles involve business insights analyst, product data scientist, consumer data scientist, tax data operations manager, and data lake engineer. In the pharmaceuticals sector, Glaxo Smith Kline has tens of vacancies that are broadly categorized under digital data and analytics, with a specific focus on artificial intelligence and machine learning.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

Chapter Outline

The rest of the chapter is organized as follows. The background section provides an overview of the key successes in supporting supply chain management with big data applications, commenting on their links with other Industry 4.0 technologies such as the Internet of things, cloud technologies, and artificial intelligence. The section on current concerns highlights the state-of-the-art big data applications to typical supply chain problems strategic, tactical, and operational decision levels. The emergent concerns section provides directions for taking big data analytics forward, exemplifying open research areas in several domains. The managerial implications section that follows provides key deliverables expected from a big data application, such as benchmarks against current status, alternative solutions to the proposed solutions, and sensitivity of results to the changes in input parameters or external factors that may not have been reflected in the application. The final section summarizes and concludes the chapter.

Background

One thing to note is the difference between "large" and "big" data: where big data implies higher complexity of data structure and technologies to process it, in addition to the size of the data (van de Klundert, 2016). Another case in point is the methodologies needed to process the data. It is unlikely that the big data can be analyzed with standard statistical methods.

The growing availability of big data has facilitated the development of new software packages and inspired new algorithms and approaches to big data analytics. R and Python are the two open-source software that are most widely used for analytics, as well as big data analytics. While for most analytics tasks each software performs reasonably well, R is preferred for statistical analysis and data visualization, whereas Python is preferred for machine learning. There is growing compatibility and sharing of tools between the two software such as Keras and TensorFlow. Keras is a library that provides a Python interface for building artificial neural networks, whereas TensorFlow is a library specifically used for the training and inference of deep neural networks.

A widely accepted terminology in big data applications refers to descriptive, predictive, and prescriptive analytics Figure 4. Descriptive analytics, also diagnostic analytics, is a summary of the current situation. It would inform what has happened in the past or the current level of performance across indicators of interest. For example, we may have a descriptive statistic of an "on time in full" performance as of 11:00 am today, for the last week, or the previous month. It may be used for diagnostic purposes, for example, for identifying deviations from expected performance or outliers in a set of similar units of analysis. A retailer interested in store profitability may monitor the sales performance of each store and identify those that are significantly overperforming or underperforming within hundreds of stores across the country. Descriptive analytics information is usually provided in dashboards, most often with some kind of a traffic light system that turns from green to orange for warning and from orange to red to indicate severe problems in the running of the supply chain. It is not concerned with the future or what the company needs to do to achieve a given objective, such as minimizing costs. However, it is the foundation of any further data modeling that may follow in the predictive or prescriptive stages.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

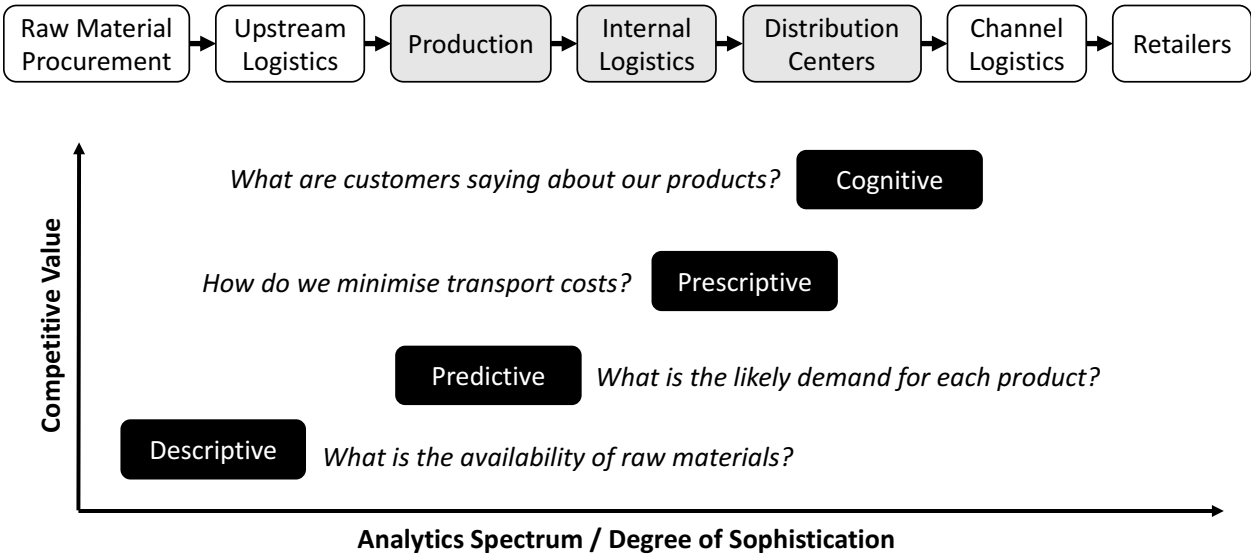

Figure 4 Analytics approaches applicable along the supply chain

Predictive analytics is only concerned with producing likely future values of a variable of interest. As per the example in Figure 4, a typical variable of interest that affects distribution and production planning, as well as procurement activities, is the demand or sales of the products and services offered by the supply chain. Other typical variables on which predictive models are used to produce forecasts are the number of customers arriving at a store or the number of patients admitted to a hospital, the number of wind turbines that are likely to fail in the next few weeks to inform preventive maintenance, or the expected time of arrival for logistics operations. A vast range of methods and models are available from simple time series to neural networks, some of which are better at handling big data.

Prescriptive analytics is the gold standard that provides decision support to reach the decision-maker's objectives. Prescriptive analytics approaches comprise mathematical modeling and sometimes simulation when the aim is to identify the best course of action for the company. The best course of action varies from strategic to tactical and operational decisions in the supply chain. Strategic decisions such as facility location, make or buy, or plant capacity would be difficult and costly to change in the short term. Operational decisions are those that are repeated on a daily or weekly basis, such as production and inventory routing decisions where the quantities to be produced and distributed across the network every day would be decided to minimize costs given a customer service level.

Recently cognitive analytics has also started to appear in the discussions, which indicates the capacity of computer hardware and software to simultaneously analyze and synthesize big data flowing from a range of sources. Most user-generated data such as expert opinions, ratings, recommendations, and testimonials are readily available in an unstructured format for mining. The abundance of user-generated text data inspired new automation tools comprised of machine learning methods and deep learning algorithms that can parse and understand natural language, paving the way for sentiment analysis

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

applications with increasing accuracy. For example, over 850,000 receipts collected over a year by a retailer were analyzed to estimate the customer's lifetime value (de Marco et al., 2021).

Cognitive analytics is capable of delivering descriptive, predictive, and prescriptive analytics based on big data collected and analyzed in real-time or near-real-time. Data used for cognitive analytics purposes would originate from multiple sources in a structured and unstructured way. Typical elements of a cognitive analytics platform would be a layer responsible for data acquisition, cleaning, and preprocessing, where the data are prepared for modeling. Another layer in a cognitive analytics platform would be responsible for model building and knowledge management with a range of predictive algorithms, which would serve as a repository of tools and models from which the required model could be selected and deployed to provide predictions of a variable of interest (Rousopoulou et al., 2022).

Description

Dashboards are where key descriptive analytics information is captured and presented to the users in a simple way. The main purpose is monitoring, and many dashboards follow an indicator to draw attention to deviations from long-term averages in the metrics. Typical information that can be found on descriptive analytics dashboards is the proportion of on-time and in-full deliveries for transportation activities. For a warehouse, the dashboard is likely to have the unit, line, and order fill proportions where unit fill monitors the proportion of items delivered, line fill monitors the proportion of order lines delivered, and the order fill monitors the proportion of orders delivered. Usually, monitoring all three is redundant, and the indicator to monitor could be selected in line with the service level agreement. From an inventory management point of view, the relevant indicators are inventory days of supply, days sales outstanding, and inventory turnover. Procurement-related descriptive indicators could involve contract spend to date compared against the spending commitment, supplier performance indicators around delivery reliability and quality, and distribution of spend over categories. With the increasing e-commerce, the returns have also increased significantly. Therefore, a dashboard reporting the proportion of orders returned and return reasons could inform actions to reduce the returns.

Prediction

A typical problem in supply chain management, forecasting of demand, is increasingly using higher frequency data such as daily or weekly series. Daily or hourly data are relevant to forecasting the sales of perishable items for replenishment decisions (Makridakis et al., 2021). However, the frequency should be linked with the decision level rather than the availability of granular data. More granular data should be used for operational, day-to-day decisions such as transport schedules, whereas monthly or yearly data should inform strategic decisions such as supply network design.

Figure 5 shows the broad categorization of machine learning methods. Supervised learning algorithms require matched inputs and outputs (targets), where the data are divided into training, validation, and test sets. The algorithm is trained on the training set to minimize the errors based on the matched inputs and targets. The parameters of the algorithm are tuned with the validation set, which is separate from the training set. Finally, the model performance is assessed on the test set where no further adjustments

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

are made to the model. Typical applications are classification and regression, where the algorithm tries to detect fraudulent transactions or spam emails in the classification problem and forecasts the market size or the sales of a product in the regression problem.

In unsupervised learning, the algorithms are presented with a data structure and tasked with finding patterns in data. Typical applications are dimension reduction, for example, to identify the most useful features in a multidimensional problem to focus later on fewer features or clustering to segment customers and present them with customized marketing strategies.

Reinforcement learning is different from supervised and unsupervised learning in the sense that the algorithm learns what it needs to do based on rewards and penalties. It does not necessarily know the right answer but infers the right behavior to reach the right answer by maximizing the reward and minimizing the penalty. A typical application in the supply chain is automated inventory management decisions. The algorithm decides on the right quantities to order from penalties associated with having too much (excess stocks and possibly waste if items are perishable) and not having enough (shortages). In the warehousing context, the robot navigation application is exemplified in obstacle avoidance, i.e., the robot is rewarded for traveling on free paths.

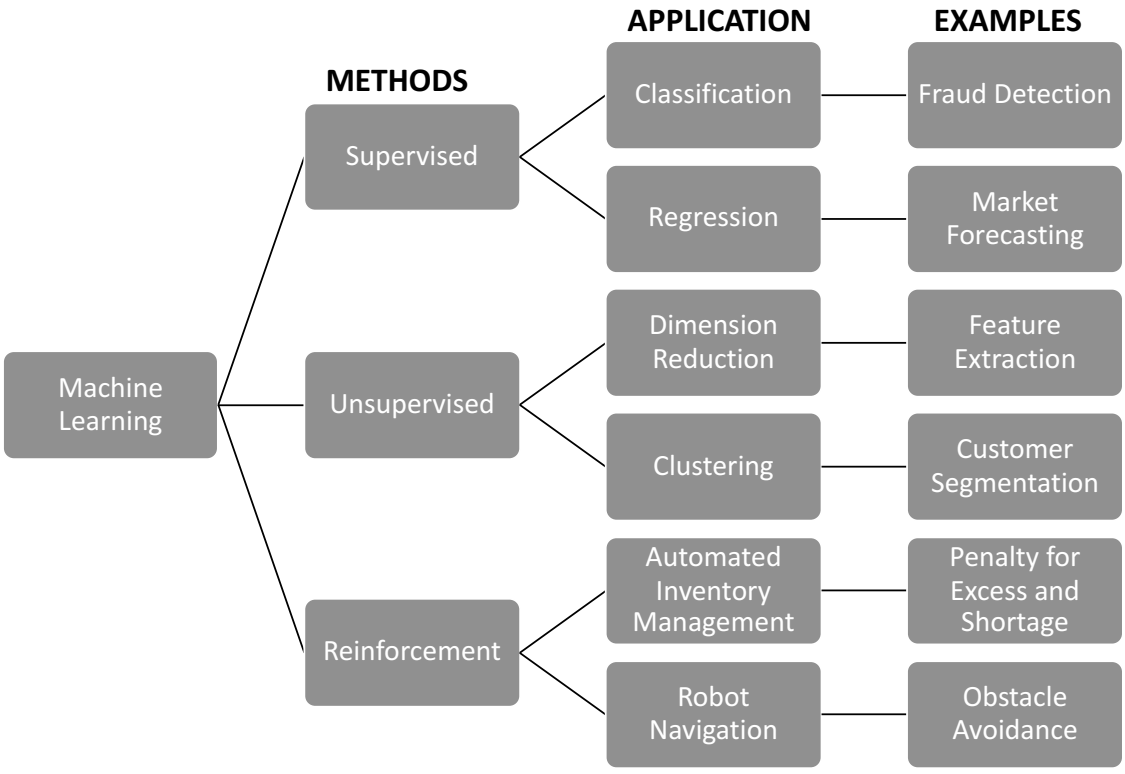

Figure 5 A representative categorization of machine learning methods with examples of specific applications

Machine learning systems can predict service-level failures a few weeks in advance to alert the planners (Melançon et al., 2021). Service levels fall below targets due to a range of reasons, one of which is machine

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

failures. Melançon et al. (2021) model service level failures as binary variables and build Gradient Boosted Decision Trees that produce a probability of failure over a prediction horizon of 14 days. Comparable other methods are artificial neural networks, logistic regression, and classification and regression trees (CART), and random forests. Using the authors' model, the planners in Michelin are able to improve the service level by 10 percentage points and to recognize safety-stock calculation problems.

In a forecasting exercise with temporal big data, the least absolute shrinkage selection operator (LASSO) regression was used to predict 12 months into the future using 814,212 variables and 66 data points, which cannot be solved with conventional causal-regression predictive modeling (Sagaert et al., 2018). The authors automated the identification of key leading macro-economic indicators from a set of 67,851 to predict tire sales. Macro-economic indicators were considered in the tactical tire sales prediction as they drive sales across different regions and markets. The LASSO method improved the accuracy of prediction by 16.1% compared with the company's existing practice.

Prescription

We can find many applications of big data analytics for supply chain and operations management problems. For example, scheduling tugboats in a seaport is a common problem observed across the world in the context of seaborne cargo transportation. Many supply chains got disrupted when Ever Given ran aground in the Suez Canal on 23 March 2021. The ship did not have a tugboat and was behind two containerships that were paired with a tugboat. While the reasons for Ever Given's not being assisted by a tugboat remain unclear, the research is underway for scheduling tugboats with instances from a busy container port in Shanghai (Jia et al., 2021). In scheduling problems such as this one concerning tugboats for ports, a range of models are used (e.g., mixed integer linear programming, heuristics, and metaheuristics) where the computational complexity of the problem requires relaxation and decomposition approaches.

Production and distribution planning decisions are interdependent and difficult to solve simultaneously. A decision support system executing operations research tools on big data collected by Danone Waters to optimize production and distribution decisions is reported to have reduced the total cost of operation by 4.7%, by approximately $5.2 million (Zhang & Song, 2018). The big data application involved a mixed integer linear programming model that minimizes production, transportation, and storage costs under operating constraints. In addition to the direct economic benefits of using this decision support system, managers at Danone started making data-driven decisions with lower complexity.

Prediction and Prescription

The typical workflow of a big data analytics project would be from descriptive to predictive analytics, followed by prescriptive analytics depending on the needs of the business problem addressed. An exemplary execution of this flow is applied to distribution network design (Liu et al., 2021). The prediction stage leverages artificial neural networks to predict customer demand, as affected by the logistics service quality of a network and the city-level purchasing power of customers based on demographic characteristics of residential areas. The optimization in the prescriptive stage solves a mixed integer linear

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

programming model to choose facility locations so that the total costs of transportation, facility setup, and package processing are minimized.

Four themes for big data applications in retail operations are availability, assortment, pricing, and layout planning. Interviews with experts suggest that historical sales data and loyalty schemes help obtain customer insights for operational planning, but granular sales data can be used to increase availability and facilitate assortment decisions (Aktas & Meng, 2017). Demand forecasting and pricing decisions can benefit from models that incorporate external data such as competitors' prices and weather conditions. Retailers could leverage big data to improve inventory decisions by stocking the correct kinds of products in right quantities. They could make pricing decisions for the season-end products or promotions by discounting prices at prescribed percentages.

Assortment planning is a key part of retail operations management. It involves making decisions on which products to carry (assortment breadth) and how many of each product to carry (assortment depth) in store. Big data applications have great potential to improve assortment planning decisions, as is evidenced in one of the leading outdoor fashion brands in South Korea, Kolon Sport. The company teamed up with a business analytics team to improve their distribution decisions: what set of items to pack in a box and how to allocate boxes to retail stores (Sung et al., 2017). The assort packing and distribution decisions need to be taken for more than 4,000 articles sold at more than 250 stores, making the problem extremely difficult to solve for human decision makers. The authors develop a two-stage prediction – optimization methodology, where they use multiple linear regression in the first stage to predict the demand for articles sold at stores, followed by an optimization model that assigns articles at different sizes and quantities into shipment boxes to minimize the time needed for distribution in the second stage. Their method is anticipated to increase sales by eight percent due to the increased availability of items in stores.

Equally relevant is managing and monitoring thousands of suppliers and controlling spend. Strategic sourcing teams hold lengthy, labor-intensive negotiations. In an impressive application of big data analytics, Verizon built a supplier rationalization tool based on operations research, machine learning, text mining, natural language processing, and artificial intelligence and reduced spend by millions of dollars while obtaining the lowest price per unit for the sourced products and services (Abdollahnejadbarough et al., 2020). Other benefits reported are a centralized and transparent contract and supplier relationship management, overhead cost reduction, decreased contract execution lead time, and service quality improvement for strategic sourcing teams.

Big data applications are expected to support assessing risk exposures and developing supply chain risk management practices while managing the transfer of risk along the supply chain. Zurich Insurance combined legacy claims databases with the large claims database and produced a data lake of historical claims from structured and semi-structured regional databases to understand risk exposures based on its customers' industries (Mizgier et al., 2018). Hence Zurich Insurance is not only able to better assess risk exposure but also can offer its customers customized insurance products as well as risk mitigation strategies such as operational hedging (design of agile and flexible supply chains) and the reduction of supply chain complexity.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

An award-winning implementation of predictive and prescriptive analytics in Walmart helps clear excess inventory in stores by a certain time period, improves revenue by minimizing the discounts needed to move inventory and clear shelves, and reduces the operational cost of relabeling discounted products (Chen et al., 2021). The integrated system developed by the authors first builds a forecasting model that produces the price-to-demand forecast based on items that need discounting, store demographics variables, store inventory, and discount depth. Second, the output of the forecast model is inputted into an optimization model along with operations costs, markdown period to recommend when the products should be marked down and their corresponding prices. The benefits of the model include a 7% reduction in costs, a 21% increase in sell-through rate (the proportion of items sold by the end of the markdown period), and a store-specific price-adjustment policy.

In a comparable study, Intel increased their profitability by over $25 billion using advanced analytics in their supply chain planning and product design (Heiney et al., 2021). As can be appreciated, Intel has a complex supply chain owing to its scale, products, and capital-intensive manufacturing processes. Heiney et al. (2021) developed an integrated approach that generates and optimizes product design alternatives using genetic algorithms and mixed integer programming. The approach not only brought financial benefits by increasing revenues and reducing costs but also provided organizational benefits where the users took ownership of the solution. The users advocated for its use owing to its environmental benefits that reduced the amount of water used and the wastewater generated in the system.

Cognition

Cognitive analytics for supply chain management using big data is in its infancy compared with other analytics approaches. Many studies focus on text mining of customer reviews (Varudharajulu & Ma, 2018) or incorporating weather and Twitter data into predicting the estimated time of arrival for trucks delivering certain goods (ben Miled et al., 2021). For example, A market leader in fruit juice manufacturing used big data in their new product development process (Jagtap & Duong, 2019), specifically by collecting data on sugar levels (nutrition data), prices (retailer data), quality (customer reviews), cost (own data), and environmental impacts (estimates informed by research reports) of own alternatives and competitors' existing products. Incorporating big data into the new product development reduced the associated costs by 33% and the development time by 11% for this fruit juice manufacturer.

Mobile crowdsensing is an approach that collects social data from a large number of users carrying mobile devices. Such data can then be used for cognition purposes, for making near-real-time predictions that can inform industrial processes. Typical research questions involve extrapolating from sparse data collected from a few users, learning nonlinear correlations, and completing the picture where sensor or mobile data are unavailable. Once the nonlinear correlations are learned, these correlations should be preserved and used for predicting the near future (E. Wang et al., 2021).

Real-time stock information can be captured using machine-to-machine communication and inform replenishment decisions so that stock-outs can be minimized. With the advent of the Internet of Things, sensors in vending machines can monitor the sales rate of various snacks and beverages in real-time. Machine-to-machine communication allows data flow between entities in the system without human

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

intervention. For example, the stock position in the vending machine is communicated to a Web API via an Internet of Things module. Web API then passes this information to an AppServer, which then communicates with a mobile app used by the owner. None of these data transactions require human intervention (Dijaya et al., 2019).

Current concerns

Using the machine-to-machine communication framework in Dijaya et al. (2019), real-time stocking information obtained by an IoT module is sent to a Web Application Programming Interface (Web API), which is then communicated to the vendor of the products via an Application Server (AppServer) on the computer or on a mobile device. Hence, the user can access real-time stock positions of products in the vending machine and make timely restocking decisions for individual products. The next step in this type of exercise would be an automatic reordering of low stock by the system, eliminating human supervision on stocking decisions.

As the technology is developing, empirical and conceptual works on big data analytics and applications in logistics and supply chain management are also evolving. Big data applications for logistics and supply chain strategy decisions comprise strategic sourcing, supply chain network design, and product design and development, while operational decisions involve demand planning, procurement, production, inventory, and logistics (G. Wang et al., 2016). We find statistical analysis, simulation, and optimization to be the driving analytical approaches, where new techniques need to be invented to cope with increasing multi-dimensionality and observations in data flowing from systems. Figure 6 shows the move towards simply sensing data and building individual models in the 2020s towards fully operational digital twins and the vision of end-to-end coordination of supply chains in the next two decades to come.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

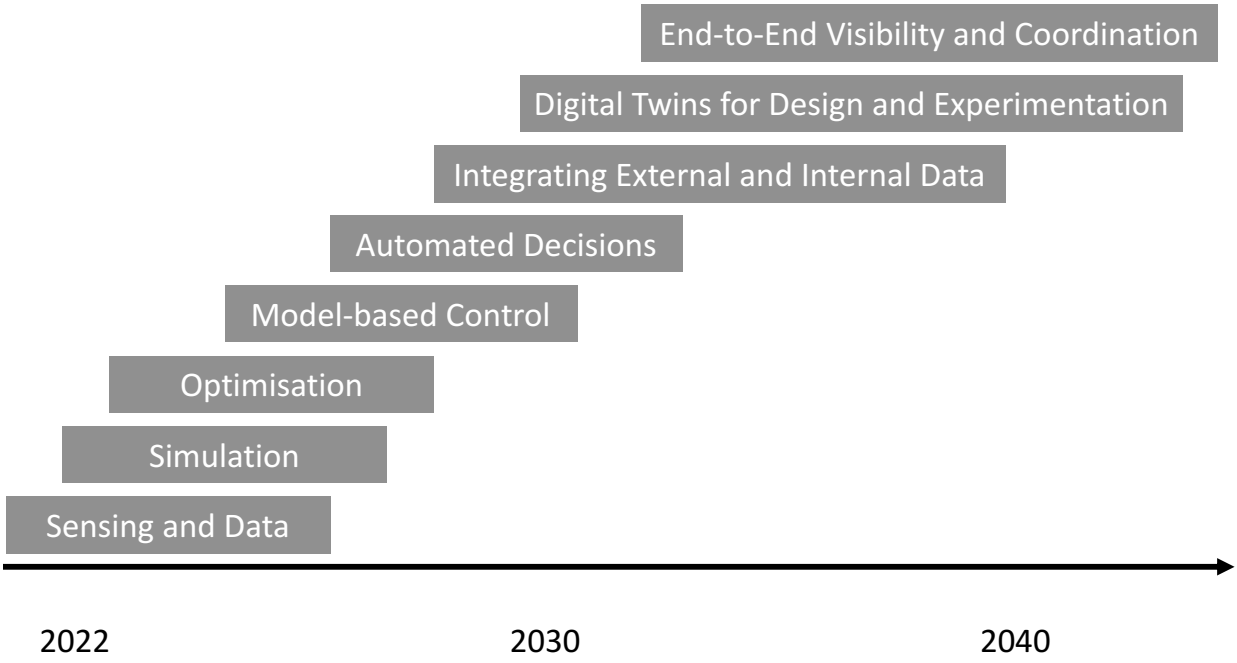

Figure 6 Anticipated big data analytics development over the next two decades

Today, we can build and run simulation and optimization models, but these need to be integrated with business functions and other supply chain partners using real-time data (sensing). Simulation and optimization models could provide higher value when data from external stakeholders are incorporated into the decision making. Such external data integration is expected to facilitate dynamic replanning in case of deviations from expectations. The future of decision making in logistics is expected to rely on model-based control, automated decisions, and the ultimate replica of the real-world operations with digital twins, which will enable experimenting with different configurations and observing the consequences of strategic, tactical, and operational decisions without changing the real-world operation.

Emergent Concerns, Outstanding Research, and Future Directions

In many of the applications reported in this chapter, the authors address a real-world case with some simplifying assumptions. The authors ignore some costs to keep the models at a practical complexity, or they consider some factors that may affect the costs out of scope. Many of the models solve a deterministic problem, and future research is expected to address the demand uncertainty as well as supply chain disruptions with stochastic modeling approaches. Figure 7 highlights the key concepts in transitioning from traditional to digital supply chains. Our factories are becoming smarter with the Internet of Things devices, mobile and fixed robots, and automation. Visibility of machine status has increased thanks to frequent readings from sensors which enable predictive and preventive maintenance.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

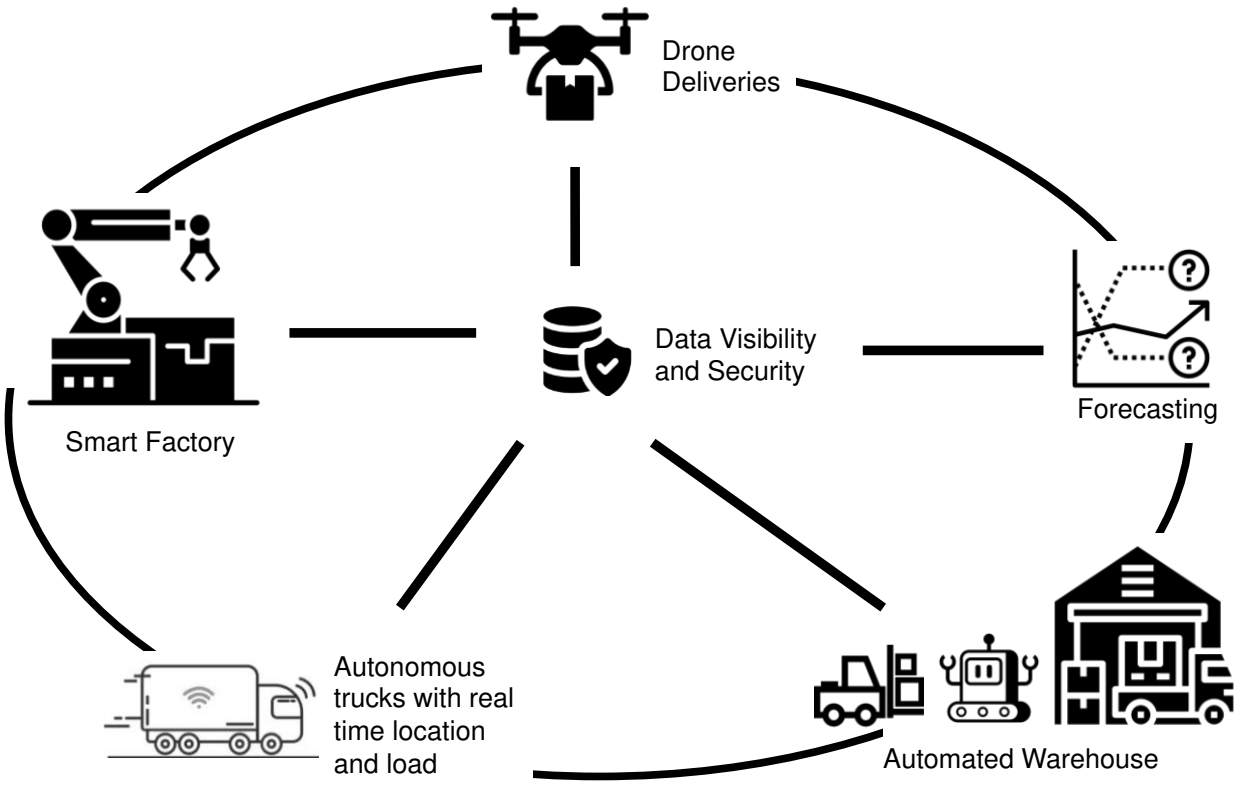

Figure 7 Transitioning into digital supply chains

We are fast approaching an age where autonomous trucks will be on the roads with real-time information on the location and the load (Simpson & Mishra, 2021; Talebian & Mishra, 2022), allowing the dynamic update of schedules in case of deviations from the expected journey (Cota et al., 2022; Nagarajan et al., 2022). However, many factors affect technology adoption, and larger organizations are likely to be slower in adopting new technologies. Having visibility of truck travel times and deviations from the plan would help minimize penalties caused by delays in fulfilling customers' orders.

Automation in warehouses is also an exciting area where we can incorporate multiple technologies, including autonomous robots. Warehouses are under increasing pressure from rapidly growing e-commerce activities, online purchasing, and same-day deliveries. Autonomous robots support human order pickers by reducing the non-productive walking time (Žulj et al., 2022). Current concerns around incorporating autonomous mobile robots into warehouses include robot path planning and navigation, obstacle avoidance, collision avoidance, and battery / recharging management.

With the advance of machine-to-machine communication, we have higher data visibility, but we need to consider data security issues in all applications. Services for conducting data breaches are increasingly shared by criminals in online forums. These services are tools and software sold at various prices to collect financial and personal information of victims, scam websites to acquire sensitive information, and compromise computers for malicious purposes (Li et al., 2016). Online forums such as cardmafia, cardpro,

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

or cardclub have tens of thousands of members with hundreds of thousands of threads and several hundreds of thousands of posts. Li et al. (2016) extracted service listings and their ratings by cybercriminals from these forums and proposed a text-mining approach to predict the service quality of these listings. They were able to identify key services used for data breaches. Comparably, Bayesian Belief Networks are used to detect cyber-attacks and improve cyber security (Yeboah-Ofori et al., 2019). Categories of malware recognized are viruses, trojans, rootkit, and botnet. The Bayesian Belief Network is capable of determining the likelihood of an attack and its corresponding impact.

Blockchain and smart contracts are exciting developments in this area (Omar et al., 2022; Tolmach et al., 2022). Smart contracts reside in a blockchain and self-execute upon instructions received to exchange money, property, or any other value over transactions recorded in a blockchain (Grida & Mostafa, 2022). However, the transparency offered by smart contracts needs further digitalization of supply chains to match real-world assets with asset records in the virtual world.

Thanks to higher volumes of structured and unstructured data available, we now have many options to improve forecasting models, increase service and reduce inventory costs. While the widely used inventory models rely on the normality assumption, new techniques such as bootstrapping, robust optimization, density forecasts, quantile estimation, and machine learning approaches exploit the increased computing power at the hands of the user and the big data flowing from enterprise information systems (Goltsos et al., 2022).

We shall not forget exciting developments in the last mile delivery process with drone deliveries and autonomous robots. Although drone deliveries are mostly in the experimental phase, they have been used to transport urgent medication during the COVID-19 pandemic (Constant, 2021). Comparably, autonomous delivery robots have been extensively used to deliver groceries when many countries in the world decided to introduce lockdowns to curb the spread of coronavirus. Using delivery robots in the last mile not only increases the delivery efficiency but also reduces the rate of infection transmission (Du, 2021) and serves those customers who cannot leave their homes to join long queues in supermarkets. COVID-19 stimulated the demand for unmanned and contactless deliveries, urging different types of autonomous e-commerce delivery vehicles and scenarios to be tested across the world (Buldeo Rai et al., 2022). The new business models developed and tested during the pandemic are promising further adoption of automation and efficiency increase in the last mile delivery problem.

While areas such as text mining and relationship network analysis have witnessed significant developments over the last decade, future research on them is still exciting and attracting many researchers and practitioners. Some of the prominent conferences where new research on text mining with supply chain applications regularly appear are as follows: IEEE International Conference on Digital Twins and Parallel Intelligence, IEEE Conference on Technologies and Applications of Artificial Intelligence, and IEEE International Conference on Big Data. These conferences provide a fresh overview of the current big data, text mining, and supply chain research, and they deserve to be regularly monitored for the most current applications of methodologies on supply chain big data.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

One of the emergent concerns around big data applications for supply chain management is the need for updates and innovations in curriculum design. New academic programs are opening on big data, analytics, data science, and a few on supply chain analytics. As the availability of data increases and the companies' hunger for converting big data into insights strengthens, the gap in academia is felt both by academic institutions and their collaborating industry partners. Some organizations, such as the United States Air Force Academy, have spent significant time and effort on developing an operations research program for practitioners (Armacost et al., 2018). In the program, big data analytics is highlighted with R and Python classes and Apache Spark final projects, which require the students use parallel processing and machine learning on big data. There is a growing need for considering changes and extensions to tools used on big data; an example could be natural language processing.

Big data applications increase supply chain visibility, reduce time and resources needed for coordination, and grow revenue streams by better targeting customers or offering new products or services built on the big data capabilities of an organization (Nguyen et al., 2018). Yet, there is scope for a systematic investigation of business models leveraging big data and for empirical research on maximizing the value captured from big data and how traditional businesses can transition into an operating environment enhanced by big data flowing from internal and external systems.

However, visualizing big data is not straightforward due to the complexity of the data. Some kind of data aggregation and dimension reduction need to be executed on the data before it can be visualized clearly. Recently, visualizing big data with augmented reality has been offered as a novel and interactive approach for users to engage with data. A representative example is a 3D visualization of stock levels overlaid on a geographical map (Ramaseri Chandra et al., 2019). Augmented reality could help improve the understanding of thermal characteristics of warehouses, the computational power demand of complex calculations, or the communication of supply chain risk.

Managerial implications

Preliminaries

The path to exploiting big data is not straightforward. Managers need to overcome shortages of people with the right set of data analytics skills, lack of support from other players in the supply chain, IT integration and data security issues, and physical capability to respond to real-time changes captured by big data. A data maturity profile can help businesses gauge their status and plan the way forward to embed big data analytics into their operational decision making.

To capture benefits from big data, managers should have flexible IT resources to integrate with their company's upstream and downstream supply chain partners. The operations in many supply chains are interconnected, and investment in IT resources alone does not bring sufficient gains. Excellence depends on data-driven supply chains. Actors in the sector can adopt cloud-based applications for accurate and timely sharing of data. Technology alone won't be sufficient to reap the benefits. Appropriate change management and technology adoption processes should be put in place, and transition should be governed carefully. When a new technology is being adopted, the affected parties' views on the

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

expected benefits and potential issues should be collected and reviewed prior to taking the decision to implement.

Data governance is a key data capability required to unlock and maximize the value captured from data. Many companies use business intelligence, digital products, analytics, and artificial intelligence to unlock the value of data. These are the most talked-about developments that feature news and presentations. However, for any of these tools to deliver value, the data that inform them must be trustable. Trusted data can only be obtained by formal data quality management, monitoring data collection, and building a data culture. Data culture entails roles and responsibilities for data owners and awareness of how data are leveraged within the business. There should be a single point of reference for everyone where data definitions, data sources, and data flows can be found and queried.

A manager needs to reconcile the growth in big data processing capabilities of a business with its decision-making processes, which are likely to be updated to accommodate new data streams while keeping a certain level of standardization for transparency and accountability purposes. Incorporating big data analytics applications into the supply chain management decisions allow managers to observe trade-offs between different decisions such as the production and transportation costs with the appropriate tools and methodologies that can estimate them accurately (Zhang & Song, 2018). Decision support systems leveraging big data analytics applications allow managers to decide objectively, based on evidence rather than guesswork.

Success Factors

For any big data analysis of supply chain management problems to inform managerial decisions effectively, the data should be presented in a way that can be effortlessly followed by its audience. A sensible recommendation is to present the managerial problem first, followed by the tensions caused by this problem in the business, and conclude with a resolution of the problem with a proposed solution approach (Camm, 2018). This type of structured communication of big data problems can facilitate better engagement with the stakeholders and alleviate the anxiety generated by the jargon in the big data analytics domain.

A good big data application would not only present solutions to the given managerial problem but also provides alternative optimal solutions along with a discussion of limitations. Such a discussion should note factors external to the model that may make one of the alternative solutions preferred over the others. Since applications are built on assumptions, the limitations should acknowledge how the solution may shift from one alternative to another when factor values change within an acceptable range.

Another element of successful big data applications is the explainability and reproducibility of results. The receivers of the solutions provided by the application should be confident why the application is suggesting these specific solutions. It is unlikely for a black-box application to be widely adopted in business.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

Any analytics project, irrespective of the size of the data that informs it, should start with a solid managerial problem that is identified and agreed upon by the stakeholders. The stakeholder buy-in and involvement in the project is a must for any decision support tool developed to address the problem to be accepted and implemented in the business. There is a strong aspect of change management to carry big data analytics projects in the supply chain to completion and exploitation.

Breaking Siloes

Many of the applied research reviewed in this chapter refer to working in siloes and suffering from an inability to have a holistic view of the problem. Our systems are generating big data, and we are capable of storing and processing them; there is scope for integrating various sources of data and how decisions informed by a set of data affect other decisions. While we are improving an element of the supply chain, we may be deteriorating another. The typical tension is between procurement and inventory management. Large quantities are incentivized with lower unit cost, improving the procurement performance metric while the items bought need to be stocked elsewhere in the supply chain, deteriorating days of inventory metric.

Traditionally, many of the big data applications in supply chain and operations management focus on a single or a few objectives, which makes it challenging to have a holistic approach to managerial issues. A way forward is to incorporate multi-objective decision analysis and explore the trade-offs between interdependent decisions. Any big data application on supply chain management thus shall try to incorporate stakeholders from different departments to the problem at hand and possibly produce a trade-off analysis for a range of decisions. Once the business priorities are set and understood by everybody involved in the process, the decision making could be left to automation in cases of repeat operational decisions.

Managerial challenges comprise shortages of people with the right set of skills, the lack of support from upstream and downstream supply chain partners, issues in information technology and systems integration, concerns around information sharing and security, and physical capability of the supply chain to respond to real-time changes captured by big data. A data maturity profile for businesses is proposed in Table 1.

Table 1 A big data analytics maturity proposal around people, supply chain, and information technology

Indicator	Level 1	Level 2	Level 3	Level 4
People Capabilities	People who can use the system	People who can interpret data and apply it to business	People with excellent analytical skills, using prescriptive models	People developing hypotheses for improving operations

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

Indicator	Level 1	Level 2	Level 3	Level 4
Information technology and systems integration	Data handled in isolation	The system is connected but not fully integrated	IT integrated, but the user requires frequent changes	The system is fully integrated, and decisions are automated
Supply Chain Operation	Surviving	Improving performance with downstream partners	Improving performance with upstream and downstream partners	The entire supply chain is synchronized

People's capabilities in Table 1 refer to a range of challenges that need to be overcome as organizations introduce and improve their big data applications. Organizations sometimes suffer from the lack of people who can use a big data application system effectively. At the very least, people could be trained to use the system, collect data, and produce complex analyses. However, further investment needs to be made to either develop skills to interpret analysis results and produce insights or further introduce hypotheses to improve current operations. It should be noted that people at the fourth level of maturity are difficult to find and expensive to keep. The organization needs to have in place career progression plans and both financial and nonfinancial incentives to attract and retain people with big data analytics capabilities.

Integrating the supply chain with downstream partners is relatively easy, for it is in the best interest of these partners (for example, retailers) to improve their predictive capabilities or reduce lead times to serve the customer better and increase the supply chain surplus. Then the experience in downstream integration could be extrapolated to upstream integration in a comparable way. The vision for the supply chain is to be synchronized entirely with end-to-end visibility and coordination of decisions. Reaching end-to-end synchronization requires investments and improvements in both hardware and software, along with skills development.

Summary and Conclusion

Big data applications in supply chain management draw heavily on statistical analysis and operations research methods where deterministic and probabilistic mathematical modeling have been widely utilized. Now many companies are at a transition phase where they are investigating the availability of data in their systems and matching the data with business problems that need immediate attention. More and more, we will see an increasing demand for people who have not only supply chain management knowledge but also has a good command of data science skills so they can bridge supply chain management with analytical approaches.

Encouraging attempts to leverage big data for forecasting have been highlighted in the chapter. It is possible to include new variables into predictor sets, considering not only macroeconomic indicators but also weather temperature or other incorporating unstructured data from customer reviews of the products into the prediction model.

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

It is now the time to treat data as an asset and investigate how it can be effectively used to improve the efficiency of operations and resources in the supply chain. It is not a trivial exercise to determine which data assets are valuable, but this process is supported by automated algorithms that sift terabytes of big data. The key idea behind assessing the value of data focuses on the use of data in applications and business integration, monitoring, and management.

Introducing big data applications could face resistance from people within the organization. Curiosity and openness to novelty, searching for innovative ways of running the business should be encouraged. A culture of 'fast fail' could help navigate the uncertain waters of big data technology adoption. Fast fail culture entails being open to new opportunities when success is not guaranteed and being able to discontinue failing endeavors swiftly without deploying further resources into them.

Limitations

While the content reported is novel at the time of publication, big data applications are expected to evolve and become more intelligent as our understanding of algorithms and computing capabilities increases. What may be new and exciting for 2021 is likely to become standard practice over the next decade.

Acknowledgments

I would like to thank the editorial team for their kind invitation to contribute to this reference book and their feedback on earlier versions of this chapter.

References

Abdollahnejadbarough, H., Mupparaju, K. S., Shah, S., Golding, C. P., Leites, A. C., Popp, T. D., Shroyer, E., Golany, Y. S., Robinson, A. G., & Akgun, V. (2020). Verizon uses advanced analytics to rationalize its tail spend suppliers. *Interfaces*, *50*(3). https://doi.org/10.1287/inte.2020.1038

Aktas, E., & Meng, Y. (2017). An Exploration of Big Data Practices in Retail Sector. *Logistics*, *1*(2). https://doi.org/10.3390/logistics1020012

Armacost, A., Lowe, J., Pietz, J., Martin, K., Wilck, J., & Ives, D. (2018). Developing operations research practitioners: United States Air Force Academy operations research program. In *Interfaces* (Vol. 48, Issue 6). https://doi.org/10.1287/inte.2018.0968

ben Miled, Z., Archbold, J., & Cochenour, B. R. (2021). Predicting distribution transit times: a case study of outbound logistics. In E. Aktas, M. Bourlakis, I. Minis, & V. Zeimpekis (Eds.), *Supply Chain 4.0: Improving Supply Chains with Analytics and Industry 4.0 Technologies* (pp. 189–208). Kogan Page.

Buldeo Rai, H., Touami, S., & Dablanc, L. (2022). Autonomous e-commerce delivery in ordinary and exceptional circumstances. The French case. *Research in Transportation Business & Management*, 100774. https://doi.org/https://doi.org/10.1016/j.rtbm.2021.100774

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

Camm, J. D. (2018). How to Influence and Improve Decisions Through Optimization Models. In *Recent Advances in Optimization and Modeling of Contemporary Problems*. https://doi.org/10.1287/educ.2018.0180

Chen, Y., Mehrotra, P., Samala, N. K. S., Ahmadi, K., Jivane, V., Pang, L., Shrivastav, M., Lyman, N., & Pleiman, S. (2021). A multiobjective optimization for clearance in walmart brick-and-mortar stores. *Interfaces*, *51*(1). https://doi.org/10.1287/INTE.2020.1065

Constant, S. (2021, February 23). *NHS launches UK's first COVID test drone delivery service in Scotland*. https://skyports.net/2021/02/nhs-launches-uks-first-covid-test-drone-delivery-service-in-scotland/

Cota, P. M., Nogueira, T. H., Juan, A. A., & Ravetti, M. G. (2022). Integrating vehicle scheduling and open routing decisions in a cross-docking center with multiple docks. *Computers & Industrial Engineering*, *164*, 107869. https://doi.org/https://doi.org/10.1016/j.cie.2021.107869

de Marco, M., Fantozzi, P., Fornaro, C., Laura, L., & Miloso, A. (2021). Cognitive analytics management of the customer lifetime value: an artificial neural network approach. *Journal of Enterprise Information Management*, *34*(2), 679–696. https://doi.org/10.1108/JEIM-01-2020-0029

Dijaya, R., Suprayitno, E. A., & Wicaksono, A. (2019). Integrated Point of Sales and Snack Vending Machine based on Internet of Things for Self Service Scale Micro Enterprises. *Journal of Physics: Conference Series*, *1179*(1). https://doi.org/10.1088/1742-6596/1179/1/012098

Du, D. (2021). Research on the application of "last-mile" autonomous delivery vehicles in the context of epidemic prevention and control. *Proceedings - 2021 International Symposium on Artificial Intelligence and Its Application on Media, ISAIAM 2021*. https://doi.org/10.1109/ISAIAM53259.2021.00022

Goltsos, T. E., Syntetos, A. A., Glock, C. H., & Ioannou, G. (2022). Inventory – forecasting: Mind the gap. In *European Journal of Operational Research* (Vol. 299, Issue 2). https://doi.org/10.1016/j.ejor.2021.07.040

Grida, M., & Mostafa, N. A. (2022). Are smart contracts too smart for Supply Chain 4.0? A blockchain framework to mitigate challenges. *Journal of Manufacturing Technology Management*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/JMTM-09-2021-0359

Heiney, J., Lovrien, R., Mason, N., Ovacik, I., Rash, E., Sarkar, N., Travis, H., Zhao, Z., Ching, K., Shirodkar, S., & Kempf, K. (2021). Intel realizes $25 billion by applying advanced analytics from product architecture design through supply chain planning. *Interfaces*, *51*(1). https://doi.org/10.1287/INTE.2020.1067

Jia, S., Li, S., Lin, X., & Chen, X. (2021). Scheduling Tugboats in a Seaport. *Transportation Science*, *55*(6). https://doi.org/10.1287/trsc.2021.1079

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

Li, W., Yin, J., & Chen, H. (2016). Targeting key data breach services in underground supply chain. *IEEE International Conference on Intelligence and Security Informatics: Cybersecurity and Big Data, ISI 2016*. https://doi.org/10.1109/ISI.2016.7745501

Liu, J., Chen, W., Yang, J., Xiong, H., & Chen, C. (2021). Iterative Prediction-and-Optimization for E-Logistics Distribution Network Design. *INFORMS Journal on Computing*. https://doi.org/10.1287/ijoc.2021.1107

Makridakis, S., Fry, C., Petropoulos, F., & Spiliotis, E. (2021). The Future of Forecasting Competitions: Design Attributes and Principles. *INFORMS Journal on Data Science*. https://doi.org/10.1287/ijds.2021.0003

Megarbane, K. (2020, October 12). *What is a Data Fabric?* https://www.stardog.com/enterprise-data-fabric/

Melançon, G. G., Grangier, P., Prescott-Gagnon, E., Sabourin, E., & Rousseau, L. M. (2021). A machine learning-based system for predicting service-level failures in supply chains. *Interfaces*, *51*(3). https://doi.org/10.1287/INTE.2020.1055

Mizgier, K. J., Kocsis, O., & Wagner, S. M. (2018). Zurich insurance uses data analytics to leverage the BI insurance proposition. In *Interfaces* (Vol. 48, Issue 2). https://doi.org/10.1287/inte.2017.0928

Nagarajan, S. M., Deverajan, G. G., Chatterjee, P., Alnumay, W., & Muthukumaran, V. (2022). Integration of IoT based routing process for food supply chain management in sustainable smart cities. *Sustainable Cities and Society*, *76*. https://doi.org/10.1016/j.scs.2021.103448

Nguyen, T., ZHOU, L., Spiegler, V., Ieromonachou, P., & Lin, Y. (2018). Big data analytics in supply chain management: A state-of-the-art literature review. *Computers and Operations Research*, *98*. https://doi.org/10.1016/j.cor.2017.07.004

Omar, I. A., Jayaraman, R., Debe, M. S., Hasan, H. R., Salah, K., & Omar, M. (2022). Supply Chain Inventory Sharing Using Ethereum Blockchain and Smart Contracts. *IEEE Access*, *10*, 2345–2356. https://doi.org/10.1109/ACCESS.2021.3139829

Ramaseri Chandra, A. N., el Jamiy, F., & Reza, H. (2019). Augmented reality for big data visualization: A review. *Proceedings - 6th Annual Conference on Computational Science and Computational Intelligence, CSCI 2019*. https://doi.org/10.1109/CSCI49370.2019.00238

Rousopoulou, V., Vafeiadis, T., Nizamis, A., Iakovidis, I., Samaras, L., Kirtsoglou, A., Georgiadis, K., Ioannidis, D., & Tzovaras, D. (2022). Cognitive analytics platform with AI solutions for anomaly detection. *Computers in Industry*, *134*. https://doi.org/10.1016/j.compind.2021.103555

Sagaert, Y. R., Aghezzaf, E. H., Kourentzes, N., & Desmet, B. (2018). Temporal big data for tactical sales forecasting in the tire industry. In *Interfaces* (Vol. 48, Issue 2). https://doi.org/10.1287/inte.2017.0901

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

Simpson, J. R., & Mishra, S. (2021). Developing a methodology to predict the adoption rate of Connected Autonomous Trucks in transportation organizations using peer effects. *Research in Transportation Economics*, *90*. https://doi.org/10.1016/j.retrec.2020.100866

Sung, S. W., Jang, Y. J., Kim, J. H., & Lee, J. (2017). Business analytics for streamlined assort packing and distribution of fashion goods at kolon sport. *Interfaces*, *47*(6). https://doi.org/10.1287/inte.2017.0904

Talebian, A., & Mishra, S. (2022). Unfolding the state of the adoption of connected autonomous trucks by the commercial fleet owner industry. *Transportation Research Part E: Logistics and Transportation Review*, *158*, 102616. https://doi.org/https://doi.org/10.1016/j.tre.2022.102616

Tolmach, P., Li, Y., Lin, S. W., Liu, Y., & Li, Z. (2022). A Survey of Smart Contract Formal Specification and Verification. *ACM Computing Surveys*, *54*(7). https://doi.org/10.1145/3464421

van de Klundert, J. (2016). Healthcare Analytics: Big Data, Little Evidence. In *Optimization Challenges in Complex, Networked and Risky Systems*. https://doi.org/10.1287/educ.2016.0158

Varudharajulu, A. K., & Ma, Y. (2018). Feature-based restaurant customer reviews process model using data mining. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3277104.3277113

Wang, E., Zhang, M., Cheng, X., Yang, Y., Liu, W., Yu, H., Wang, L., & Zhang, J. (2021). Deep Learning-Enabled Sparse Industrial Crowdsensing and Prediction. *IEEE Transactions on Industrial Informatics*, *17*(9). https://doi.org/10.1109/TII.2020.3028616

Wang, G., Gunasekaran, A., Ngai, E. W. T., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. In *International Journal of Production Economics* (Vol. 176). https://doi.org/10.1016/j.ijpe.2016.03.014

Yeboah-Ofori, A., Islam, S., & Brimicombe, A. (2019). Detecting Cyber Supply Chain Attacks on Cyber Physical Systems Using Bayesian Belief Network. *2019 International Conference on Cyber Security and Internet of Things (ICSIoT)*, 37–42. https://doi.org/10.1109/ICSIoT47925.2019.00014

Zhang, S., & Song, H. (2018). Production and distribution planning in Danone waters China division. *Interfaces*, *48*(6). https://doi.org/10.1287/inte.2018.0973

Žulj, I., Salewski, H., Goeke, D., & Schneider, M. (2022). Order batching and batch sequencing in an AMR-assisted picker-to-parts system. *European Journal of Operational Research*, *298*(1). https://doi.org/10.1016/j.ejor.2021.05.033

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author

Weblinks

European Union (2022) https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024_en date accessed 18 Feb 2022

Masters Portal (2022) https://www.mastersportal.com/study-options/268927258/data-science-big-data-united-kingdom.html date accessed 18 Feb 2022

Research and Markets (2022) https://www.researchandmarkets.com/reports/5008078/big-data-market-with-covid-19-impact-analysis-by date accessed: 17 Feb 2022

UK Department for Business, Energy and Industrial Strategy (2022) https://www.gov.uk/government/publications/uk-innovation-strategy-leading-the-future-by-creating-it

US Department of State (2022) https://www.state.gov/innovation-roundtables/ date accessed: 17 Feb 2022

Emel Aktas
Cranfield University, Collage Road, Cranfield, MK43 0AL, UK, e-mail: emel.aktas@cranfield.ac.uk
*Corresponding author