

DATA SCIENCE AND ITS RELATIONSHIP TO BIG DATA AND DATA-DRIVEN DECISION MAKING

Foster Provost¹ and Tom Fawcett²



Abstract

Companies have realized they need to hire data scientists, academic institutions are scrambling to put together data-science programs, and publications are touting data science as a hot—even “sexy”—career choice. However, there is confusion about what exactly data science is, and this confusion could lead to disillusionment as the concept diffuses into meaningless buzz. In this article, we argue that there are good reasons why it has been hard to pin down exactly what is data science. One reason is that data science is intricately intertwined with other important concepts also of growing importance, such as big data and data-driven decision making. Another reason is the natural tendency to associate what a practitioner does with the definition of the practitioner’s field; this can result in overlooking the fundamentals of the field. We believe that trying to define the boundaries of data science precisely is not of the utmost importance. We can debate the boundaries of the field in an academic setting, but in order for data science to serve business effectively, it is important (i) to understand its relationships to other important related concepts, and (ii) to begin to identify the fundamental principles underlying data science. Once we embrace (ii), we can much better understand and explain exactly what data science has to offer. Furthermore, only once we embrace (ii) should we be comfortable calling it data science. In this article, we present a perspective that addresses all these concepts. We close by offering, as examples, a partial list of fundamental principles underlying data science.

Introduction

WITH VAST AMOUNTS OF DATA now available, companies in almost every industry are focused on exploiting data for competitive advantage. The volume and variety of data have far outstripped the capacity of manual analysis, and in some cases have exceeded the capacity of conventional databases. At the same time, computers have become far more powerful, networking is ubiquitous, and algorithms have been developed that can connect datasets to enable broader and deeper

analyses than previously possible. The convergence of these phenomena has given rise to the increasingly widespread business application of data science.

Companies across industries have realized that they need to hire more data scientists. Academic institutions are scrambling to put together programs to train data scientists. Publications are touting data science as a hot career choice and even “sexy.”¹ However, there is confusion about what exactly is data science, and this confusion could well lead to

¹Leonard N. Stern School of Business, New York University, New York, New York.

²Data Scientists, LLC, New York, New York and Mountain View, California.

© Foster Provost and Tom Fawcett 2013; Published by Mary Ann Liebert, Inc. This article is available under the Creative Commons License CC-BY-NC (<http://creativecommons.org/licenses/by-nc/4.0>). This license permits non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited. Permission only needs to be obtained for commercial use and can be done via RightsLink.

disillusionment as the concept diffuses into meaningless buzz. In this article, we argue that there are good reasons why it has been hard to pin down what exactly is data science. One reason is that data science is intricately intertwined with other important concepts, like big data and data-driven decision making, which are also growing in importance and attention. Another reason is the natural tendency, in the absence of academic programs to teach one otherwise, to associate what a practitioner actually does with the definition of the practitioner's field; this can result in overlooking the fundamentals of the field.

At the moment, trying to define the boundaries of data science precisely is not of foremost importance. Data-science academic programs are being developed, and in an academic setting we can debate its boundaries. However, in order for data science to serve business effectively, it is important (i) to understand its relationships to these other important and closely related concepts, and (ii) to begin to understand what are the fundamental principles underlying data science. Once we embrace (ii), we can much better understand and explain exactly what data science has to offer. Furthermore, only once we embrace (ii) should we be comfortable calling it data *science*.

In this article, we present a perspective that addresses all these concepts. We first work to disentangle this set of closely interrelated concepts. In the process, we highlight data science as the connective tissue between data-processing technologies (including those for "big data") and data-driven decision making. We discuss the complicated issue of data science as a field versus data science as a profession. Finally, we offer as examples a list of some fundamental principles underlying data science.

Data Science

At a high level, *data science* is a set of fundamental principles that support and guide the principled extraction of information and knowledge from data. Possibly the most closely related concept to data science is *data mining*—the actual extraction of knowledge from data via technologies that incorporate these principles. There are hundreds of different data-mining algorithms, and a great deal of detail to the methods of the field. We argue that underlying all these many details is a much smaller and more concise set of fundamental principles.

These principles and techniques are applied broadly across functional areas in business. Probably the broadest business applications are in marketing for tasks such as targeted marketing, online advertising, and recommendations for cross-selling. Data science also is applied for general customer

relationship management to analyze customer behavior in order to manage attrition and maximize expected customer value. The finance industry uses data science for credit scoring and trading and in operations via fraud detection and workforce management. Major retailers from Wal-Mart to Amazon apply data science throughout their businesses, from marketing to supply-chain management. Many firms have differentiated themselves strategically with data science, sometimes to the point of evolving into data-mining companies.

But data science involves much more than just data-mining algorithms. Successful data scientists must be able to view business problems from a data perspective. There is a fundamental structure to data-analytic thinking, and basic principles that should be understood. Data science draws from many "traditional" fields of study. Fundamental principles of causal analysis must be understood. A large portion of what has traditionally been studied within the field of

statistics is fundamental to data science. Methods and methodology for visualizing data are vital. There are also particular areas where intuition, creativity, common sense, and knowledge of a particular application must be brought to bear. A data-science perspective

provides practitioners with structure and principles, which give the data scientist a framework to systematically treat problems of extracting useful knowledge from data.

Data Science in Action

For concreteness, let's look at two brief case studies of analyzing data to extract predictive patterns. These studies illustrate different sorts of applications of data science. The first was reported in the *New York Times*:

Hurricane Frances was on its way, barreling across the Caribbean, threatening a direct hit on Florida's Atlantic coast. Residents made for higher ground, but far away, in Bentonville, Ark., executives at Wal-Mart Stores decided that the situation offered a great opportunity for one of their newest data-driven weapons...predictive technology.

A week ahead of the storm's landfall, Linda M. Dillman, Wal-Mart's chief information officer, pressed her staff to come up with forecasts based on what had happened when Hurricane Charley struck several weeks earlier. Backed by the trillions of bytes' worth of shopper history that is stored in Wal-Mart's data warehouse, she felt that the company could "start predicting what's going to happen, instead of waiting for it to happen," as she put it.²

Consider *why* data-driven prediction might be useful in this scenario. It might be useful to predict that people in the path

"PUBLICATIONS ARE TOUTING DATA SCIENCE AS A HOT CAREER CHOICE AND EVEN 'SEXY.'"

of the hurricane would buy more bottled water. Maybe, but it seems a bit obvious, and why do we need data science to discover this? It might be useful to project the *amount of increase* in sales due to the hurricane, to ensure that local Wal-Marts are properly stocked. Perhaps mining the data could reveal that a particular DVD sold out in the hurricane's path—but maybe it sold out that week at Wal-Marts across the country, not just where the hurricane landing was imminent. The prediction could be somewhat useful, but probably more general than Ms. Dillman was intending.

It would be more valuable to discover patterns due to the hurricane that were not obvious. To do this, analysts might examine the huge volume of Wal-Mart data from prior, similar situations (such as Hurricane Charley earlier in the same season) to identify unusual local demand for products. From such patterns, the company might be able to anticipate unusual demand for products and rush stock to the stores ahead of the hurricane's landfall.

Indeed, that is what happened. The *New York Times* reported that: "...the experts mined the data and found that the stores would indeed need certain products—and not just the usual flashlights. 'We didn't know in the past that strawberry Pop-Tarts increase in sales, like seven times their normal sales rate, ahead of a hurricane,' Ms. Dillman said in a recent interview.' And the pre-hurricane top-selling item was beer.*"²

Consider a second, more typical business scenario and how it might be treated from a data perspective. Assume you just landed a great analytical job with MegaTelCo, one of the largest telecommunication firms in the United States. They are having a major problem with customer retention in their wireless business. In the mid-Atlantic region, 20% of cell-phone customers leave when their contracts expire, and it is getting increasingly difficult to acquire new customers. Since the cell-phone market is now saturated, the huge growth in the wireless market has tapered off. Communications companies are now engaged in battles to attract each other's customers while retaining their own. Customers switching from one company to another is called *churn*, and it is expensive all around: one company must spend on incentives to attract a customer while another company loses revenue when the customer departs.

You have been called in to help understand the problem and to devise a solution. Attracting new customers is much more expensive than retaining existing ones, so a good deal of marketing budget is allocated to prevent churn. Marketing

has already designed a special retention offer. Your task is to devise a precise, step-by-step plan for how the data science team should use MegaTelCo's vast data resources to decide which customers should be offered the special retention deal prior to the expiration of their contracts. Specifically, how should MegaTelCo decide on the set of customers to target to best reduce churn for a particular incentive budget? Answering this question is much more complicated than it seems initially.

Data Science and Data-Driven Decision Making

Data science involves principles, processes, and techniques for understanding phenomena via the (automated) analysis of data. For the perspective of this article, the ultimate goal of data science is improving decision making, as this generally is of paramount interest to business. Figure 1 places data science in the context of other closely related and data-related processes in the organization. Let's start at the top.

Data-driven decision making (DDD)³ refers to the practice of basing decisions on the analysis

of data rather than purely on intuition. For example, a marketer could select advertisements based purely on her long experience in the field and her eye for what will work. Or, she could base her selection on the analysis of data regarding how consumers react to different ads. She could also use a combination of these approaches. DDD is not an all-or-nothing practice, and different firms engage in DDD to greater or lesser degrees.

The benefits of data-driven decision making have been demonstrated conclusively. Economist Erik Brynjolfsson and his colleagues from MIT and Penn's Wharton School recently conducted a study of how DDD affects firm performance.³ They developed a measure of DDD that rates firms as to how strongly they use data to make decisions across the company. They show statistically that the more data-driven a firm is, the more productive it is—even controlling for a wide range of possible confounding factors. And the differences are not small: one standard deviation higher on the DDD scale is associated with a 4–6% increase in productivity. DDD also is correlated with higher return on assets, return on equity, asset utilization, and market value, and the relationship seems to be causal.

Our two example case studies illustrate two different sorts of decisions: (1) decisions for which "discoveries" need to be

"FROM SUCH PATTERNS, THE COMPANY MIGHT BE ABLE TO ANTICIPATE UNUSUAL DEMAND FOR PRODUCTS AND RUSH STOCK TO THE STORES AHEAD OF THE HURRICANE'S LANDFALL."

*Of course! What goes better with strawberry Pop-Tarts than a nice cold beer?

made within data, and (2) decisions that repeat, especially at massive scale, and so decision making can benefit from even small increases in accuracy based on data analysis. The Wal-Mart example above illustrates a type-1 problem. Linda Dillman would like to discover knowledge that will help Wal-Mart prepare for Hurricane Frances's imminent arrival. Our churn example illustrates a type-2 DDD problem. A large telecommunications company may have hundreds of millions of customers, each a candidate for defection. Tens of millions of customers have contracts expiring each month, so each one of them has an increased likelihood of defection in the near future. If we can improve our ability to estimate, for a given customer, how profitable it would be for us to focus on her, we can potentially reap large benefits by applying this ability to the millions of customers in the population. This same logic applies to many of the areas where we have seen the most intense application of data science and data mining: direct marketing, online advertising, credit scoring, financial trading, help-desk management, fraud detection, search ranking, product recommendation, and so on.

The diagram in Figure 1 shows data science supporting data-driven decision making, but also overlapping with it. This highlights the fact that, increasingly, business decisions are being made automatically by computer systems. Different

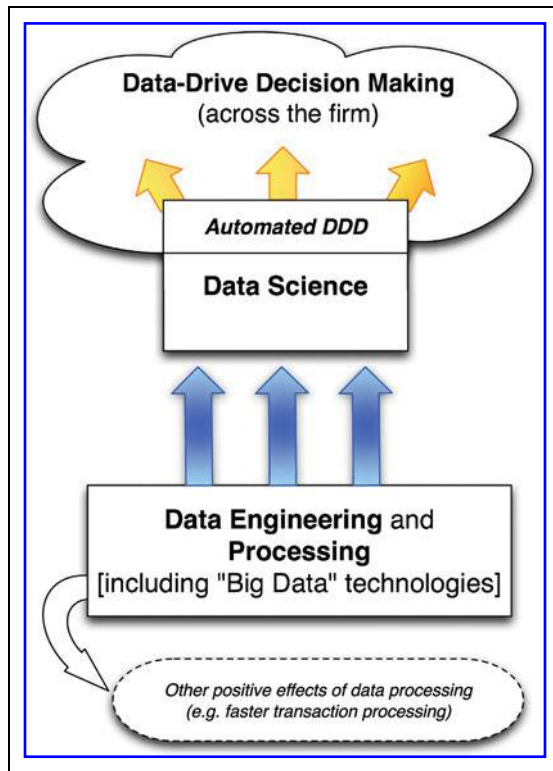


FIG. 1. Data science in the context of closely related processes in the organization.

industries have adopted automatic decision making at different rates. The finance and telecommunications industries were early adopters. In the 1990s, automated decision making changed the banking and consumer-credit industries dramatically. In the 1990s, banks and telecommunications companies also implemented massive-scale systems for

managing data-driven fraud control decisions. As retail systems were increasingly computerized, merchandising decisions were automated. Famous examples include Harrah's casinos' reward programs and the automated recommendations of Amazon and

Netflix. Currently we are seeing a revolution in advertising, due in large part to a huge increase in the amount of time consumers are spending online and the ability online to make (literally) split-second advertising decisions.

Data Processing and "Big Data"

Despite the impression one might get from the media, there is a lot to data processing that is not data science. Data engineering and processing are critical to support data-science activities, as shown in Figure 1, but they are more general and are useful for much more. Data-processing technologies are important for many business tasks that do not involve extracting knowledge or data-driven decision making, such as efficient transaction processing, modern web system processing, online advertising campaign management, and others.

"Big data" technologies, such as Hadoop, Hbase, CouchDB, and others have received considerable media attention recently. For this article, we will simply take *big data* to mean datasets that are too large for traditional data-processing systems and that therefore require new technologies. As with the traditional technologies, big data technologies are used for many tasks, including data engineering. Occasionally, big data technologies are actually used for *implementing* data-mining techniques, but more often the well-known big data technologies are used for data processing *in support of* the data-mining techniques and other data-science activities, as represented in Figure 1.

Economist Prasanna Tambe of New York University's Stern School has examined the extent to which the utilization of big data technologies seems to help firms.⁴ He finds that, after controlling for various possible confounding factors, the use of big data technologies correlates with significant additional productivity growth. Specifically, one standard deviation higher utilization of big data technologies is associated with 1–3% higher productivity than the average firm; one standard deviation lower in terms of big data utilization is associated with 1–3% lower productivity. This leads to potentially very large productivity differences between the firms at the extremes.

From Big Data 1.0 to Big Data 2.0

One way to think about the state of big data technologies is to draw an analogy with the business adoption of internet technologies. In Web 1.0, businesses busied themselves with getting the basic internet technologies in place so that they could establish a web presence, build electronic commerce capability, and improve operating efficiency. We can think of ourselves as being in the era of Big Data 1.0, with firms engaged in building capabilities to process large data. These primarily support their current operations—for example, to make themselves more efficient.

With Web 1.0, once firms had incorporated basic technologies thoroughly (and in the process had driven down prices) they started to look further. They began to ask what the web could do for them, and how it could improve upon what they'd always done. This ushered in the era of Web 2.0, in which new systems and companies started to exploit the interactive nature of the web. The changes brought on by this shift in thinking are extensive and pervasive; the most obvious are the incorporation of social-networking components and the rise of the “voice” of the individual consumer (and citizen).

Similarly, we should expect a Big Data 2.0 phase to follow Big Data 1.0. Once firms have become capable of processing massive data in a flexible fashion, they should begin asking: *What can I now do that I couldn't do before, or do better than I could do before?* This is likely to usher in the golden era of data science. The principles and techniques of data science will be applied far more broadly and far more deeply than they are today.

It is important to note that in the Web-1.0 era, some precocious companies began applying Web-2.0 ideas far ahead of the mainstream. Amazon is a prime example, incorporating the consumer's “voice” early on in the rating of products and product reviews (and deeper, in the rating of reviewers). Similarly, we see some companies already applying Big Data 2.0. Amazon again is a company at the forefront, providing data-driven recommendations from massive data. There are other examples as well. Online advertisers must process extremely large volumes of data (billions of ad impressions per day is not unusual) and maintain a very high throughput (real-time bidding systems make decisions in tens of milliseconds). We should look to these and similar industries for signs of advances in big data and data science that subsequently will be adopted by other industries.

Data-Analytic Thinking

One of the most critical aspects of data science is the support of data-analytic thinking. Skill at thinking data-analytically is important not just for the data scientist but throughout the organization. For example, managers and line employees in other functional areas will only get the best from the company's data-science resources if they have some basic understanding of the fundamental principles. Managers in enterprises without substantial data-science resources should still understand basic principles in order to engage consultants on an informed basis. Investors in data-science ventures

need to understand the fundamental principles in order to assess investment opportunities accurately. More generally, businesses increasingly are driven by data analytics, and there is great professional advantage in being able to interact competently with and within such businesses. Understanding the fundamental concepts, and having frameworks

for organizing data-analytic thinking, not only will allow one to interact competently, but will help to envision opportunities for improving data-driven decision making or to see data-oriented competitive threats.

Firms in many traditional industries are exploiting new and existing data resources for competitive advantage. They employ data-science teams to bring advanced technologies to bear to increase revenue and to decrease costs. In addition, many new companies are being developed with data mining as a key strategic component. Facebook and Twitter, along with many other “Digital 100” companies,⁵ have high valuations due primarily to data assets they are committed to capturing or creating.[†] Increasingly, managers need to manage data-analytics teams and data-analysis projects, marketers have to organize and understand data-driven campaigns, venture capitalists must be able to invest wisely in businesses with substantial data assets, and business strategists must be able to devise plans that exploit data.

As a few examples, if a consultant presents a proposal to exploit a data asset to improve your business, you should be able to assess whether the proposal makes sense. If a competitor announces a new data partnership, you should recognize when it may put you at a strategic disadvantage. Or, let's say you take a position with a venture firm and your first project is to assess the potential for investing in an advertising company. The founders present a convincing argument that they will realize significant value from a unique body of data they will collect, and on that basis, are arguing for a substantially higher valuation. Is this reasonable? With an

“SIMILARLY, WE SHOULD EXPECT A BIG DATA 2.0 PHASE TO FOLLOW BIG DATA 1.0 ... THIS IS LIKELY TO USHER IN THE GOLDEN ERA OF DATA SCIENCE.”

[†]Of course, this is not a new phenomenon. Amazon and Google are well-established companies that obtain tremendous value from their data assets.

understanding of the fundamentals of data science, you should be able to devise a few probing questions to determine whether their valuation arguments are plausible.

On a scale less grand, but probably more common, data-analytics projects reach into all business units. Employees throughout these units must interact with the data-science team. If these employees do not have a fundamental grounding in the principles of data-analytic thinking, they will not really understand what is happening in the business. This lack of understanding is much more damaging in data-science projects than in other technical projects, because the data science supports improved decision making. Data-science projects require close interaction between the scientists and the business people responsible for the decision making. Firms in which the business people do not understand what the data scientists are doing are at a substantial disadvantage, because they waste time and effort or, worse, because they ultimately make wrong decisions. A recent article in Harvard Business Review concludes: “For all the breathless promises about the return on investment in Big Data, however, companies face a challenge. Investments in analytics can be useless, even harmful, unless employees can incorporate that data into complex decision making.”⁶

Some Fundamental Concepts of Data Science

There is a set of well-studied, fundamental concepts underlying the principled extraction of knowledge from data, with both theoretical and empirical backing. These fundamental concepts of data science are drawn from many fields that study data analytics. Some reflect the relationship between data science and the business problems to be solved. Some reflect the sorts of knowledge discoveries that can be made and are the basis for technical solutions. Others are cautionary and prescriptive. We briefly discuss a few here. This list is not intended to be exhaustive; detailed discussions even of the handful below would fill a book.* The important thing is that we understand these fundamental concepts.

Fundamental concept: *Extracting useful knowledge from data to solve business problems can be treated systematically by following a process with reasonably well-defined stages. The Cross-Industry Standard Process for Data Mining⁷ (CRISP-DM) is one codification of this process. Keeping such a process in mind can structure our thinking about data analytics prob-*

lems. For example, in actual practice one repeatedly sees analytical “solutions” that are not based on careful analysis of the problem or are not carefully evaluated. Structured thinking about analytics emphasizes these often underappreciated aspects of supporting decision making with data. Such structured thinking also contrasts critical points at which human intuition and creativity is necessary versus points at which high-powered analytical tools can be brought to bear.

Fundamental concept: *Evaluating data-science results requires careful consideration of the context in which they will be used. Whether knowledge extracted from data will aid in decision*

making depends critically on the application in question. For our churn-management example, how exactly are we going to use the patterns that are extracted from historical data? More generally, does the pattern lead to better decisions than some reasonable alternative? How well would one have done by chance? How well would one do with a smart “default” alternative? Many data science evaluation frameworks are based on

this fundamental concept.

Fundamental concept: *The relationship between the business problem and the analytics solution often can be decomposed into tractable subproblems via the framework of analyzing expected value. Various tools for mining data exist, but business problems rarely come neatly prepared for their application. Breaking the business problem up into components corresponding to estimating probabilities and computing or estimating values, along with a structure for recombining the components, is broadly useful. We have many specific tools for estimating probabilities and values from data. For our churn example, should the value of the customer be taken into account in addition to the likelihood of leaving? It is difficult to realistically assess any customer-targeting solution without phrasing the problem as one of expected value.*

Fundamental concept: *Information technology can be used to find informative data items from within a large body of data. One of the first data-science concepts encountered in business-analytics scenarios is the notion of finding correlations. “Correlation” often is used loosely to mean data items that provide information about other data items—specifically, known quantities that reduce our uncertainty about unknown quantities. In our churn example, a quantity of interest is the likelihood that a particular customer will leave after her contract expires. Before the contract expires, this would be an unknown quantity. However, there may be known data items (usage, service history, how many friends*

“FACEBOOK AND TWITTER, ALONG WITH MANY OTHER ‘DIGITAL 100’ COMPANIES, HAVE HIGH VALUATIONS DUE PRIMARILY TO DATA ASSETS THEY ARE COMMITTED TO CAPTURING OR CREATING.”

*And they do; see <http://data-science-for-biz.com>.

have canceled contracts) that correlate with our quantity of interest. This fundamental concept underlies a vast number of techniques for statistical analysis, predictive modeling, and other data mining.

Fundamental concept: *Entities that are similar with respect to known features or attributes often are similar with respect to unknown features or attributes.* Computing similarity is one of the main tools of data science. There are many ways to compute similarity and more are invented each year.

Fundamental concept: *If you look too hard at a set of data, you will find something—but it might not generalize beyond the data you're observing.* This is referred to as “overfitting” a dataset. Techniques for mining data can be very powerful, and the need to detect and avoid overfitting is one of the most important concepts to grasp when applying data-mining tools to real problems. The concept of overfitting and its avoidance permeates data science processes, algorithms, and evaluation methods.

Fundamental concept: *To draw causal conclusions, one must pay very close attention to the presence of confounding factors, possibly unseen ones.* Often, it is not enough simply to uncover correlations in data; we may want to use our models to guide decisions on how to influence the behavior producing the data. For our churn problem, we want to intervene and *cause* customer retention. All methods for drawing causal conclusions—from interpreting the coefficients of regression models to randomized controlled experiments—incorporate assumptions regarding the presence or absence of confounding factors. In applying such methods, it is important to understand their assumptions clearly in order to understand the scope of any causal claims.

Chemistry Is Not About Test Tubes: Data Science vs. the Work of the Data Scientist

Two additional, related complications combine to make it more difficult to reach a common understanding of just what is data science and how it fits with other related concepts.

First is the dearth of academic programs focusing on data science. Without academic programs defining the field for us, we need to define the field for ourselves. However, each of us sees the field from a different perspective and thereby forms a different conception. The dearth of academic programs is largely due to the inertia associated with academia and the concomitant effort involved in creating new academic programs—especially ones that span traditional dis-

ciplines. Universities clearly see the need for such programs, and it is only a matter of time before this first complication will be resolved. For example, in New York City alone, two top universities are creating degree programs in data science. Columbia University is in the process of creating a master's degree program within its new Institute for Data Sciences and Engineering (and has founded a center focusing on the foundations of data science), and NYU will commence a master's degree program in data science in fall 2013.

“WITHOUT ACADEMIC PROGRAMS DEFINING THE FIELD FOR US, WE NEED TO DEFINE THE FIELD FOR OURSELVES.”

The second complication builds on confusion caused by the first. Workers tend to associate with their field the tasks they spend considerable time on or those they find challenging or rewarding. This is in contrast to the tasks that *differentiate* the field from

other fields. Forsythe described this phenomenon in an ethnographic study of practitioners in artificial intelligence (AI):

The AI specialists I describe view their professional work as science (and in some cases engineering)...The scientists' work and the approach they take to it make sense in relation to a particular view of the world that is taken for granted in the laboratory...Wondering what it means to “do AI,” I have asked many practitioners to describe their own work. Their answers invariably focus on one or more of the following: problem solving, writing code, and building systems.⁸

Forsythe goes on to explain that the AI practitioners focus on these three activities even when it is clear that they spend much time doing other things (even less related specifically to AI). Importantly, *none* of these three tasks differentiates AI from other scientific and engineering fields. Clearly just being very good at these three things does not an AI scientist make. And as Forsythe points out, technically the latter two are not even necessary, as the lab director, a famous AI Scientist, had not written code or built systems for years. Nonetheless, these are the tasks the AI scientists saw as defining their work—they apparently did not explicitly consider the notion of what makes doing AI different from doing other tasks that involve problem solving, writing code, and system building. (This is possibly due to the fact that in AI, there were academic distinctions to call on.)

Taken together, these two complications cause particular confusion in data science, because there are few academic distinctions to fall back on, and moreover, due to the state of the art in data processing, data scientists tend to spend a majority of their problem-solving time on data preparation and processing. The goal of such preparation is either to

subsequently apply data-science methods or to understand the results. However, that does not change the fact that the day-to-day work of a data scientist—especially an entry-level one—may be largely data processing. This is directly analogous to an entry-level chemist spending the majority of her time doing technical lab work. If this were all she were trained to do, she likely would not be rightly called a chemist but rather a lab technician. Important for being a chemist is that this work is in support of the application of the science of chemistry, and hopefully the eventual advancement to jobs involving more chemistry and less technical work. Similarly for data science: a chief scientist in a data-science-oriented company will do much less data processing and more data-analytics design and interpretation.

At the time of this writing, discussions of data science inevitably mention not just the analytical skills but the popular tools used in such analysis. For example, it is common to see job advertisements mentioning data-mining techniques (random forests, support vector machines), specific application areas (recommendation systems, ad placement optimization), alongside popular software tools for processing big data (SQL, Hadoop, MongoDB). This is natural. The particular concerns of data science in business are fairly new, and businesses are still working to figure out how best to address them. Continuing our analogy, the state of data science may be likened to that of chemistry in the mid-19th century, when theories and general principles were being formulated and the field was largely experimental. Every good chemist had to be a competent lab technician. Similarly, it is hard to imagine a working data scientist who is not proficient with certain sorts of software tools. A firm may be well served by requiring that their data scientists have skills to access, prepare, and process data using tools the firm has adopted.

Nevertheless, we emphasize that there is an important reason to focus here on the general principles of data science. In ten years' time, the predominant technologies will likely have changed or advanced enough that today's choices would seem quaint. On the other hand, the general principles of data science are not so different than they were 20 years ago and likely will change little over the coming decades.

Conclusion

Underlying the extensive collection of techniques for mining data is a much smaller set of fundamental concepts comprising data *science*. In order for data science to flourish as a field, rather than to drown in the flood of popular attention, we must think beyond the algorithms, techniques, and tools in common use. We must think about the core principles and concepts that underlie the techniques, and also the systematic

thinking that fosters success in data-driven decision making. These data science concepts are general and very broadly applicable.

Success in today's data-oriented business environment requires being able to think about how these fundamental concepts apply to particular business problems—to think data-analytically. This is aided by conceptual frameworks that themselves are part of data science. For example, the automated extraction of patterns from data is a process with well-defined stages. Understanding this process and its stages helps structure problem solving, makes it more systematic, and thus less prone to error.

There is strong evidence that business performance can be improved substantially via data-driven decision making,³ big data technologies,⁴ and data-science techniques based on big data.^{9,10} Data science supports data-driven decision making—and sometimes allows making decisions automatically at massive scale—and depends upon technologies for “big data” storage and engineering. However, the principles of data science are its own and should be considered and discussed explicitly in order for data science to realize its potential.

Author Disclosure Statement

F.P. and T.F. are authors of the forthcoming book, *Data Science for Business*.

References

1. Davenport T.H., and Patil D.J. Data scientist: the sexiest job of the 21st century. *Harv Bus Rev*, Oct 2012.
2. Hays C. L. What they know about you. *N Y Times*, Nov. 14, 2004.
3. Brynjolfsson E., Hitt L.M., and Kim H.H. Strength in numbers: How does data-driven decision making affect firm performance? Working paper, 2011. SSRN working paper. Available at SSRN: <http://ssrn.com/abstract=1819486>.
4. Tambe P. Big data know-how and business value. Working paper, NYU Stern School of Business, NY, New York, 2012.
5. Fusfeld A. The digital 100: the world's most valuable startups. *Bus Insider*. Sep. 23, 2010.
6. Shah S., Horne A., and Capellá J. Good data won't guarantee good decisions. *Harv Bus Rev*, Apr 2012.
7. Wirth, R., and Hipp, J. CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, 2000, pp. 29–39.

8. Forsythe, Diana E. The construction of work in artificial intelligence. *Science, Technology & Human Values*, 18(4), 1993, pp. 460–479.
9. Hill, S., Provost, F., and Volinsky, C. Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Science*, 21(2), 2006, pp. 256–276.
10. Martens D. and Provost F. Pseudo-social network targeting from consumer transaction data. Working paper, CEDER-11-05, Stern School of Business, 2011. Available at SSRN: <http://ssrn.com/abstract=1934670>.

Address correspondence to:

F. Provost
Department of Information, Operations,
and Management Sciences
Leonard N. Stern School of Business
New York University
44 W. 4th Street, 8th Floor
New York, NY 10012

E-mail: fprovost@stern.nyu.edu

This article has been cited by:

1. Cuilan Qiao, Yuqing Chen, Qing Guo, Yunwei Yu. 2024. Understanding science data literacy: a conceptual framework and assessment tool for college students majoring in STEM. *International Journal of STEM Education* 11:1. . [[Crossref](#)]
2. Alexander M. Geske, David M. Herold, Sebastian Kummer. 2024. Integrating AI support into a framework for collaborative decision-making (CDM) for airline disruption management. *Journal of the Air Transport Research Society* 3, 100026. [[Crossref](#)]
3. Md. Toriqul Islam, Borhan Uddin Khan. Big Data and Analytics 1-30. [[Crossref](#)]
4. Simon De Vos, Johannes De Smedt, Marijke Verbruggen, Wouter Verbeke. 2024. Data-driven internal mobility: Similarity regularization gets the job done. *Knowledge-Based Systems* 295, 111824. [[Crossref](#)]
5. Pedro Natividade Joergensen, Michael Zaggl. 2024. The role of data science and data analytics for innovation: a literature review. *Journal of Business Analytics* 33, 1-17. [[Crossref](#)]
6. 2024. Optimizing Trade Strategies: The Interactive Trade Decision Making Using Weighted Sum Method. *REST Journal on Banking, Accounting and Business* 3:1, March 2024, 36-44. [[Crossref](#)]
7. 2024. Optimizing Trade Strategies: The Interactive Trade Decision Making Using Weighted Sum Method. *REST Journal on Banking, Accounting and Business* 3:1, March 2024, 36-44. [[Crossref](#)]
8. Herwin Meerveld, Roy Lindelauf. 2024. Data Science in Military Decision-Making: Foci and Gaps. *Global Society* 33, 1-27. [[Crossref](#)]
9. Victor Chang, Karl Hall, Qianwen Xu, Folakemi Amao, Meghana Ganatra, Vladlena Benson. 2024. Prediction of Customer Churn Behavior in the Telecommunication Industry Using Machine Learning Models. *Algorithms* 17:6, 231. [[Crossref](#)]
10. Michelle Novotny, Rachel Dodds, Philip R. Walsh. 2024. Understanding the adoption of data-driven decision-making practices among Canadian DMOs. *Information Technology & Tourism* 26:2, 331-345. [[Crossref](#)]
11. Charlotte Köhler, Ann Melissa Campbell, Jan Fabian Ehmke. 2024. Data-driven customer acceptance for attended home delivery. *OR Spectrum* 46:2, 295-330. [[Crossref](#)]
12. Eleonora Santos, Milena Carvalho, Susana Martins. 2024. Sustainable Enablers of Knowledge Management Strategies in a Higher Education Institution. *Sustainability* 16:12, 5078. [[Crossref](#)]
13. Pengfei Li, Miao Wang. 2024. Navigating the Legal Labyrinth: The Future of Data-Driven Platform Labor in China. *Journal of the Knowledge Economy* 8. . [[Crossref](#)]
14. Sabyasachi Pramanik. The Impact of Data Science and Participated Geographic Metadata on Improving Government Service Deliveries 239-267. [[Crossref](#)]
15. Zheng-Yun Zhuang, Cheng-Kung Chung. 2024. Dissecting the visiting willingness of driving visitors facing a retail market's dual-pricing policy for parking. *Journal of Retailing and Consumer Services* 78, 103655. [[Crossref](#)]
16. L.J. Basile, N. Carbonara, U. Panniello, R. Pellegrino. 2024. The role of big data analytics in improving the quality of healthcare services in the Italian context: The mediating role of risk management. *Technovation* 133, 103010. [[Crossref](#)]
17. Marcela A. D. Di Marzo, Pedro G. Calil, Hossein Nadali Najafabadi, Viviam Lawrence Takase, Carlos H. B. Mourão, Jorge H. Bidinotto. 2024. Computational Tool for Aircraft Fuel System Analysis. *Aerospace* 11:5, 362. [[Crossref](#)]
18. Li Wenting, Wan Mohd Hirwani Wan Hussain, Jia Xinlin, Meng Na, Syed Shah Alam. 2024. Analyzing the Impact on Talent Acquisition and Performance Management. *Journal of Organizational and End User Computing* 36:1, 1-30. [[Crossref](#)]
19. Marvin Niederhaus, Nico Migenda, Julian Weller, Wolfram Schenck, Martin Kohlhas. Technical Readiness of Prescriptive Analytics Platforms: A Survey 509-519. [[Crossref](#)]
20. Anil Meher. Revolution Ethics of Data Science and AI 245-256. [[Crossref](#)]
21. A. S. Anurag. The Evolution of AI and Data Science 295-312. [[Crossref](#)]
22. Vivek Veeraiah, Dharmesh Dhablya, Sukhvinder Singh Dari, Jambi Ratna Raja Kumar, Ritika Dhablya, Sabyasachi Pramanik, Ankur Gupta. The Impact of Data Science and Participated Geographic Metadata on Improving Government Service Deliveries 320-342. [[Crossref](#)]
23. Ki Yong Kwon. 2024. Analyzing fashion item purchase patterns and channel transition patterns using association rules and brand loyalty in big data. *The Research Journal of the Costume Culture* 32:2, 199-214. [[Crossref](#)]
24. Mutaz M. Al-Debei. 2024. The era of business analytics: identifying and ranking the differences between business intelligence and data science from practitioners' perspective using the Delphi method. *Journal of Business Analytics* 7:2, 94-119. [[Crossref](#)]
25. Guanxu Chen, Weiwen Peng, Fangfang Yang. 2024. An LSTM-SA model for SOC estimation of lithium-ion batteries under various temperatures and aging levels. *Journal of Energy Storage* 84, 110906. [[Crossref](#)]

26. Thamir Alaskar. 2024. The impact of organizational capabilities on business analytics use: the moderating role of environmental dynamism. *Information Systems and e-Business Management* 44. . [[Crossref](#)]
27. Olayinka Omole, Somik Ghosh. 2024. Application of Data Science in the US Commercial Construction Industry: Current Trends and Future Opportunities. *International Journal of Construction Education and Research* 76, 1-24. [[Crossref](#)]
28. Temidayo Oluwasola Osunsanmi, Timothy O. Olawumi, Andrew Smith, Suha Jaradat, Clinton Aigbavboa, John Aliu, Ayodeji Oke, Oluwaseyi Ajayi, Opeyemi Oyeyipo. 2024. Modelling the drivers of data science techniques for real estate professionals in the fourth industrial revolution era. *Property Management* 42:2, 310-331. [[Crossref](#)]
29. Faris Mehmedović, Almir Karabegović. Data Science-Driven System Design with LSTM Neural Networks: A COVID-19 Perspective 1-6. [[Crossref](#)]
30. Rajendrakumar Vilas Thorat. 2024. Mathematics in Data Science and Artificial Intelligence. *International Journal of Advanced Research in Science, Communication and Technology* 362-366. [[Crossref](#)]
31. John Metzcar, Catherine R. Jutzeler, Paul Macklin, Alvaro Köhn-Luque, Sarah C. Brüningk. 2024. A review of mechanistic learning in mathematical oncology. *Frontiers in Immunology* 15. . [[Crossref](#)]
32. Daina Bilkštytė-Skanė, Vita Akstinaite. 2024. Strategic organizational changes: Adopting data-driven decisions. *Strategic Change* 33:2, 107-116. [[Crossref](#)]
33. Chao Wu. 2024. Data privacy: From transparency to fairness. *Technology in Society* 76, 102457. [[Crossref](#)]
34. Frank Emmert-Streib, Shailesh Tripathi, Matthias Dehmer. 2024. Human Team Behavior and Predictability in the Massively Multiplayer Online Game WOT Blitz. *ACM Transactions on the Web* 18:1, 1-27. [[Crossref](#)]
35. Seoung-gon Ko. 2024. A study on rethinking EDA in digital transformation era. *Korean Journal of Applied Statistics* 37:1, 87-102. [[Crossref](#)]
36. Ana Paula Ritto, Adriana Ladeira de Araujo, Carlos Roberto Ribeiro de Carvalho, Heraldito Possolo De Souza, Patricia Manga e Silva Favaretto, Vivian Renata Boldrim Saboya, Michelle Louvaes Garcia, Leslie Domenici Kulikowski, Esper Georges Kallás, Antonio José Rodrigues Pereira, Wilson Cobello Junior, Katia Regina Silva, Eidi Raquel Franco Abdalla, Aluisio Augusto Cotrim Segurado, Ester Cerdeira Sabino, Ulysses Ribeiro Junior, Rossana Pulcineli Vieira Francisco, Anna Miethke-Morais, Anna Sara Shafferman Levin, Marcio Valente Yamada Sawamura, Juliana Carvalho Ferreira, Clovis Artur Silva, Thais Mauad, Nelson da Cruz Gouveia, Leila Suemi Harima Letaif, Marco Antonio Bego, Linamara Rizzo Battistella, Alberto José da Silva Duarte, Marília Cerqueira Leite Seelaender, Julio Marchini, Orestes Vicente Forlenza, Vanderson Geraldo Rocha, Maria Cassia Mendes-Correa, Silvia Figueiredo Costa, Giovanni Guido Cerri, Eloisa Silva Dutra de Oliveira Bonfá, Roger Chammas, Tarcisio Eloy Pessoa de Barros Filho, Geraldo Busatto Filho. 2024. Data-driven, cross-disciplinary collaboration: lessons learned at the largest academic health center in Latin America during the COVID-19 pandemic. *Frontiers in Public Health* 12. . [[Crossref](#)]
37. João Farinha, Maria Fatima Pina. Data Science in the Employee Recruitment Process 116-153. [[Crossref](#)]
38. R. Nalini. Transformative Power of Artificial Intelligence in Decision-Making, Automation, and Customer Engagement 189-208. [[Crossref](#)]
39. Abdullah M. Almarashi, Muhammad Daniyal, Farrukh Jamal. 2024. Modelling the GDP of KSA using linear and non-linear NNAR and hybrid stochastic time series models. *PLOS ONE* 19:2, e0297180. [[Crossref](#)]
40. Matti Juhani Haverila, Kai Christian Haverila. 2024. The influence of quality of big data marketing analytics on marketing capabilities: the impact of perceived market performance!. *Marketing Intelligence & Planning* 42:2, 346-372. [[Crossref](#)]
41. Anissa Tanweer, James Steinhoff. 2024. Academic data science: Transdisciplinary and extradisciplinary visions. *Social Studies of Science* 54:1, 133-160. [[Crossref](#)]
42. Diana Korayim, Varun Chotia, Girish Jain, Sharfa Hassan, Francesco Paolone. 2024. How big data analytics can create competitive advantage in high-stake decision forecasting? The mediating role of organizational innovation. *Technological Forecasting and Social Change* 199, 123040. [[Crossref](#)]
43. Aleksandar Radonjić, Henrique Duarte, Nádia Pereira. 2024. Artificial intelligence and HRM: HR managers' perspective on decisiveness and challenges. *European Management Journal* 42:1, 57-66. [[Crossref](#)]
44. A.Mohammed Abubakar, Ahmet Türkmen, Volkan Işık, Patrick Mikalef, Ofir Turel. 2024. Exploring the complementary effects of business analytics capabilities and π -shaped skills on innovation outcomes. *European Journal of Information Systems* 57, 1-18. [[Crossref](#)]
45. Azazkhan I. Pathan, P. G. Agnihotri, Dhruvesh Patel, P. J. Gandhi, Cristina Prieto, Usman Mohseni, Nilesh Patidar. Use of Big Data for Flood Assessment Through HEC-RAS Model: A Study of Purna River of Navsari 87-100. [[Crossref](#)]
46. Sami Seppänen, Minna Saunila, Juhani Ukko. Digital Transformation of Organizational and Management Controls—Review and Recommendations for the Future 1-25. [[Crossref](#)]

47. Dirk Bendlin, Jorge Marx Gómez, H. Kaddoura, A. Kucewicz, M. Werther Häckell. The Data Value Chain Ontology 75-96. [[Crossref](#)]
48. Komalpreet Kaur, Yogesh Kumar, Sukhpreet Kaur. Artificial Intelligence and Machine Learning in Financial Services to Improve the Business System 3-30. [[Crossref](#)]
49. Juan Antonio Torrecilla-García, Agnieszka Grazyna Skotnicka, Juan Carlos Rubio-Romero, Virginia Herrera-Pérez. The Emergent Perspective of Applied Observability in Occupational Health and Safety. The Exploratory Scoping Review for the Future Framework 837-845. [[Crossref](#)]
50. Sahil Sharma. Data Scientist in Tourism: Needs, Skills, and Scope 1-16. [[Crossref](#)]
51. Nguyen Phong Nguyen, Emmanuel Mogaji. Digital Strategies for Aiding Ease of Decision-Making in the Services Sector 75-100. [[Crossref](#)]
52. Dan Zhang, Loo G. Pee, Shan L. Pan, Jingyuan Wang. 2024. Information practices in data analytics for supporting public health surveillance. *Journal of the Association for Information Science and Technology* 75:1, 79-93. [[Crossref](#)]
53. Jaime Govea, Walter Gaibor-Naranjo, Santiago Sanchez-Viteri, William Villegas-Ch. 2024. Integration of Data and Predictive Models for the Evaluation of Air Quality and Noise in Urban Environments. *Sensors* 24:2, 311. [[Crossref](#)]
54. D. Anupama, A. Ravi Kumar, D. Sumathi. Managing Healthcare Data Using ML Algorithms and Society 5.0 71-102. [[Crossref](#)]
55. Christine DiStefano, Geophrey Odero, Angela Starrett. An Introduction to Machine Learning for Educational Researchers 11-30. [[Crossref](#)]
56. Muhammad Ramzan. 2024. The significance and application of data analytics models for strategic management. *International Journal of ADVANCED AND APPLIED SCIENCES* 11:1, 87-94. [[Crossref](#)]
57. Neelima Bhatnagar, Victoria Causer, Michael Lucci, Michael Pry, Dorothy Zilic. 2024. Developing a Data Analytics Practicum Course. *Information Systems Education Journal* 22:2, 47-69. [[Crossref](#)]
58. Juan Liu, Yunpeng Cui, Mengjia Huo, Shumei Zhou. Establishment and Application of Regional Agricultural and Rural Data Asset-Based Management Mode 247-254. [[Crossref](#)]
59. Qinqin Wu, Nur Ajrun Khalid. 2024. Optimization Path for Management Decision-Making of Chinese Public Hospitals Under the Background of Big Data. *Journal of Information Systems Engineering and Management* 9:1, 24423. [[Crossref](#)]
60. Guangmo Tong, Peng Zhao, Mina Samizadeh. Query-Decision Regression Between Shortest Path and Minimum Steiner Tree 314-326. [[Crossref](#)]
61. Ali Vafaei-Zadeh, Jammula Madhuri, Haniruzila Hanifah, Ramayah Thurasamy. 2024. The Interactive Effects of Capabilities and Data-Driven Culture on Sustained Competitive Advantage. *IEEE Transactions on Engineering Management* 71, 8444-8458. [[Crossref](#)]
62. Motahareh Zarefard, Nicola Marsden. The Essential Competencies of Data Scientists: A Framework for Hiring and Training 397-418. [[Crossref](#)]
63. Bernard Yannou, Ghada Bouillass, Michael Saidani, Marija Jankovic. 2024. The Circular Digital Cockpit: Towards an actionable framework for life cycle circularity assessment and decision. *Procedia CIRP* 122, 467-472. [[Crossref](#)]
64. Zsolt Tibor Kosztyán, Tibor Csizmadia, Beáta Sz.G. Pató, Szilárd Berke, Ildiko Neumanné-Virág, Andrea Bencsik. 2023. Measuring organizational happiness in higher education - a data-driven approach. *Cogent Education* 10:1. . [[Crossref](#)]
65. Ebru DEMİREL, Cihan YARALI. 2023. İmalat İşletmelerinin Dijitalleşme Süreçleri Üzerine Nitel Bir Çalışma. *Yönetim ve Ekonomi Dergisi* 21-41. [[Crossref](#)]
66. Danyang Bai. Design of Artificial Intelligence Information System Using NoSQL Database and Ajax Technology 678-683. [[Crossref](#)]
67. Lakshya Sandeep Khanna, Prathamesh Sachin Yadav, Sudhanshu Maurya, Vrince Vimal. Integral Role of Data Science in Startup Evolution 1-7. [[Crossref](#)]
68. Abhishek Singla. The Hybrid Integration of BD & its Analytics in the Area of Corporate Sector: Systematic Review 565-569. [[Crossref](#)]
69. Munir Ahmad. The Role of Data Science and Volunteered Geographic Information in Enhancing Government Service Delivery 254-281. [[Crossref](#)]
70. Aiswarya Raj M, Jan Bosch, Helena Holmström Olsson. Maturity Assessment Model for Industrial Data Pipelines 503-513. [[Crossref](#)]
71. Jürgen Bajorath. 2023. Data science and data analytics in life science research. *Artificial Intelligence in the Life Sciences* 3, 100067. [[Crossref](#)]

72. Jürgen Bajorath. 2023. Specific contributions of artificial intelligence to interdisciplinary life science research – exploring and communicating new opportunities. *Artificial Intelligence in the Life Sciences* 3, 100052. [[Crossref](#)]
73. Paulo Cortez, Albert Bifet. 2023. Editorial: Seventh special issue on Knowledge Discovery and Business Intelligence. *Expert Systems* 40:10. . [[Crossref](#)]
74. Danhuai Guo, Huixuan Chen, Ruoling Wu, Yangang Wang. 2023. AIGC challenges and opportunities related to public safety: A case study of ChatGPT. *Journal of Safety Science and Resilience* 4:4, 329-339. [[Crossref](#)]
75. Parvathy Venkatachalam, Rajhans Mishra. 2023. IMR – 7th INDAM CONFERENCE PAPER: Future of workplace design from a socio-technical perspective. *IIMB Management Review* 35:4, 317-332. [[Crossref](#)]
76. Hamed Taherdoost. 2023. Navigating the Ethical and Privacy Concerns of Big Data and Machine Learning in Decision Making. *Intelligent and Converged Networks* 4:4, 280-295. [[Crossref](#)]
77. Wei Zhang, Xinru Zhong, Fengchun Fan, Xiaoping Jiang. 2023. Unlocking the Creative Potential of Chinese New Liberal Arts: The Role of Interdisciplinary Education, Knowledge Integration, and Metacognitive Awareness. *The Asia-Pacific Education Researcher* 30. . [[Crossref](#)]
78. Moesha Danielle Genola Moore, Kevin Baboolal, Patrick Hosein. On the Prediction of Polycystic Ovarian Syndrome 1-6. [[Crossref](#)]
79. Koshini Muthutantrige, Syed M Hasan, Satya R Shah. Project Management Framework for Digitalization of Logistics Processes 1-6. [[Crossref](#)]
80. Fei Cheng, Jiehui Huang, Huizhen Li, Beate I. Escher, Yujun Tong, Maria König, Dali Wang, Fan Wu, Zhiqiang Yu, Bryan W. Brooks, Jing You. 2023. Text Mining-Based Suspect Screening for Aquatic Risk Assessment in the Big Data Era: Event-Driven Taxonomy Links Chemical Exposures and Hazards. *Environmental Science & Technology Letters* 10:11, 1004-1010. [[Crossref](#)]
81. Dieudonné Tchuenté, Anass El Haddadi. 2023. One decade of big data for firms' competitiveness: insights and a conceptual model from bibliometrics. *Journal of Enterprise Information Management* 36:6, 1421-1453. [[Crossref](#)]
82. Jelita Asian, Oktavianus Tegar N.K Putra, Media Anugerah Ayu, Teddy Mantoro. Sentiment Analysis With N-Gram Preprocessing for Online-Shopping Reviews in Indonesian Language 1-6. [[Crossref](#)]
83. Tugba Karaboga, Cemal Zehir, Ekrem Tatoglu, H. Aykut Karaboga, Abderaouf Bouguerra. 2023. Big data analytics management capability and firm performance: the mediating role of data-driven culture. *Review of Managerial Science* 17:8, 2655-2684. [[Crossref](#)]
84. Catherine Carlson, Susan W. White, Matthew Hudnall, John Lochman, Robert Laird. 2023. Leveraging Data Science to Advance Implementation Science: The Case of School Mental Health. *Journal of School Health* 93:11, 1045-1048. [[Crossref](#)]
85. Neerja Kashive, Vandana Tandon Khanna. 2023. Emerging HR analytics role in a crisis: an analysis of LinkedIn data. *Competitiveness Review: An International Business Journal* 33:6, 1179-1204. [[Crossref](#)]
86. Shiva Tyagi, Riti Rathore. Unveiling the Dynamic Journey from Data Insights to Action in Data Science 1-6. [[Crossref](#)]
87. Samar Sarabhai, Mainak Chakraborty, Mansi Batra, Rajneesh Kler, Supriya Banerjee, Sanjay Mishra. Using AI and Machine Learning to Predict Consumer Buying Behavior: Insights from Behavioral Economics in Case of Alcoholic Beverages 980-986. [[Crossref](#)]
88. Cagla Ozen, Zehra Mine Tas. The Role of Artificial Intelligence (AI) in Improving Product Development Efficiency 313-335. [[Crossref](#)]
89. Aleksandr Gershanok, Iana Lomonosova. 2023. Metamorphosis of the Public Administration Paradigm and the New Qualification Profile of Civil Servants. *Bulletin of Kemerovo State University. Series: Political, Sociological and Economic sciences* 2023:3, 323-333. [[Crossref](#)]
90. Xi Xi, Feifei Ren, Lean Yu, Jing Yang. 2023. Detecting the technology's evolutionary pathway using HiDS-trait-driven tech mining strategy. *Technological Forecasting and Social Change* 195, 122777. [[Crossref](#)]
91. Uta Schirpke, Andrea Ghermandi, Michael Sinclair, Derek Van Berkel, Nathan Fox, Leonardo Vargas, Louise Willemen. 2023. Emerging technologies for assessing ecosystem services: A synthesis of opportunities and challenges. *Ecosystem Services* 63, 101558. [[Crossref](#)]
92. Tamara L. Clegg, Keanna Cleveland, Erianne Weight, Daniel Greene, Niklas Elmqvist. 2023. Data everyday as community-driven science: Athletes' critical data literacy practices in collegiate sports contexts. *Journal of Research in Science Teaching* 60:8, 1786-1816. [[Crossref](#)]
93. Vincent J. Straub, Deborah Morgan, Jonathan Bright, Helen Margetts. 2023. Artificial intelligence in government: Concepts, standards, and a unified framework. *Government Information Quarterly* 40:4, 101881. [[Crossref](#)]
94. Devesh Kumar. The Use of Artificial Intelligence in Data Analysis and Business Intelligence 311-320. [[Crossref](#)]

95. Ágnes Szukits, Péter Móricz. 2023. Towards data-driven decision making: the role of analytical culture and centralization efforts. *Review of Managerial Science* 41. . [\[Crossref\]](#)
96. Furkan Çolhak, Bilal Emir Uçar, İpek Saygut, Berkant Düzgün, Ferhat Demirkıran, Hasan Dağ. Garbage in, Garbage Out: A Case Study on Defective Product Prediction in Manufacturing 01-06. [\[Crossref\]](#)
97. Fajar J. Ekaputra, Christoph Fabianek, Gabriel Unterholzer, Eduard Gringinger. The Semantic Overlay Architecture for Data Interoperability and Exchange 232-237. [\[Crossref\]](#)
98. Jiang Xu, Aidong Li, Calvin King Lam Chung, Yang Yue. 2023. Mapping the Unmapped: Investigating Big Data Companies via Online Sources. *The Professional Geographer* 75:5, 816-826. [\[Crossref\]](#)
99. Debolina Dey, Lidia Ghosh, Diptendu Bhattacharya, Amit Konar. 2023. A 2-phase prediction of a non-stationary time-series by Taylor series and reinforcement learning. *Applied Soft Computing* 145, 110565. [\[Crossref\]](#)
100. Jieun Kim, Neal E. Blair. 2023. Biomarker heatmaps: visualization of complex biomarker data to detect storm-induced source changes in fluvial particulate organic carbon. *Earth Science Informatics* 16:3, 2915-2924. [\[Crossref\]](#)
101. Konstantinos Konstas, Panos T. Chountalas, Eleni A. Didaskalou, Dimitrios A. Georgakellos. 2023. A Pragmatic Framework for Data-Driven Decision-Making Process in the Energy Sector: Insights from a Wind Farm Case Study. *Energies* 16:17, 6272. [\[Crossref\]](#)
102. Amjad Ali, Asim Rafiq, Muddassir Hussien, Salman Sarwat, Abdul Raziq. 2023. Exploring Big Data Usage to Predict Supply Chain Effectiveness: A Moderated and Mediated Model Linkage. *Global Business Review* 90. . [\[Crossref\]](#)
103. Yuanju Qu, Jiayun Wang, Bo Jiang, Shenghui Cheng, Yangpeng Wang, Peishan Wu, Xinguo Ming, Xianghua Chu. 2023. Self-decision mechanisms of smart production systems based on improved uncertainty theory and user-CFA. *Computers & Industrial Engineering* 183, 109453. [\[Crossref\]](#)
104. Xileidys Parra, Xavier Tort-Martorell, Fernando Alvarez-Gomez, Carmen Ruiz-Viñals. 2023. Chronological Evolution of the Information-Driven Decision-Making Process (1950–2020). *Journal of the Knowledge Economy* 14:3, 2363-2394. [\[Crossref\]](#)
105. Bridgette J. Befort, Alejandro Garciadiego, Jialu Wang, Ke Wang, Gabriela Franco, Edward J. Maginn, Alexander W. Dowling. 2023. Data science for thermodynamic modeling: Case study for ionic liquid and hydrofluorocarbon refrigerant mixtures. *Fluid Phase Equilibria* 572, 113833. [\[Crossref\]](#)
106. Nna Motlhasedi, Keneilwe Porogo. 2023. The Contribution of Records Management and Big Data Analytics to the Growth of Businesses: An Analysis of Eastern and Southern Africa Regional Branch of the International Council on Archives (ESARBICA) Region. *Mousaion: South African Journal of Information Studies* 41:1. . [\[Crossref\]](#)
107. Nguyen Thi Tuoi, Nguyen Nghi Thanh. 2023. The Impact of Digital Capabilities on the Work Performance of Provincial Civil Servants in Vietnam. *Journal of Law and Sustainable Development* 11:4, e560. [\[Crossref\]](#)
108. Deepika Jhamb, Sukhpreet Kaur, Saurabh Pandey, Amit Mittal. 2023. Pricing model, engagement model and firm performance: evidence from Indian data science industry. *Benchmarking: An International Journal* 22. . [\[Crossref\]](#)
109. David Holger Schmidt, Dirk van Dierendonck, Ulrike Weber. 2023. The data-driven leader: developing a big data analytics leadership competency framework. *Journal of Management Development* 42:4, 297-326. [\[Crossref\]](#)
110. Sarah Pinon, Corentin Burnay, Isabelle Linden. 2023. Business-Driven Data Recommender System: Design and Implementation. *Journal of Computer Information Systems* 69, 1-14. [\[Crossref\]](#)
111. Vlad Krotov, Leigh Johnson. 2023. Big web data: Challenges related to data, technology, legality, and ethics. *Business Horizons* 66:4, 481-491. [\[Crossref\]](#)
112. Bilel Sdiri, Lionel Rigaud, Rym Jemmali, Fatma Abdelhedi. 2023. The Difficult Path to Become Data-Driven. *SN Computer Science* 4:4. . [\[Crossref\]](#)
113. Pamela Tozzo, Arianna Delicati, Beatrice Marcante, Luciana Caenazzo. 2023. Digital Biobanking and Big Data as a New Research Tool: A Position Paper. *Healthcare* 11:13, 1825. [\[Crossref\]](#)
114. Hugh Xuechen Liu, Christian Wagner. 2023. Proxies to the monthly active user number of Geo AR Mobile games – online search volume as a proposal. *Multimedia Tools and Applications* 82:16, 25403-25425. [\[Crossref\]](#)
115. Michael Overton, Stephen Kleinschmit. 2023. Transforming research methods education through data science literacy. *Teaching Public Administration* 41:2, 149-169. [\[Crossref\]](#)
116. Chabi Gupta. Taming the HiPPO (Highest Paid Person's Opinion) With Agile Metrics and Value Management 268-280. [\[Crossref\]](#)
117. Reetika Sarkar, Sithija Manage, Xiaoli Gao. 2023. Stable Variable Selection for High-Dimensional Genomic Data with Strong Correlations. *Annals of Data Science* 1. . [\[Crossref\]](#)

118. Mingjie Fang, Feng Liu, Shufeng (Simon) Xiao, Kwangtae Park. 2023. Hedging the bet on digital transformation in strategic supply chain management: a theoretical integration and an empirical test. *International Journal of Physical Distribution & Logistics Management* **53**:4, 512-531. [[Crossref](#)]
119. Simon Rautenbach, Imke De Kock, Jacomine Grobler. Towards a maturity model for data science in small and medium-sized enterprises from developing countries 1-7. [[Crossref](#)]
120. Alamir Costa Louro, Jurij Jaklič, Marcelo Moll Brandão. 2023. MARKETING ANALYTICS & ORGANIZATIONAL CULTURE FIT. *International Journal of Business Marketing* **8**:1, 45-61. [[Crossref](#)]
121. Dimitris Ampeliotis, Alexandros Gkillas. Subspace Parsimonious Dictionary Learning and its use in Federated Learning 1-5. [[Crossref](#)]
122. Shah Hussain, Muhammad Qasim Khan. 2023. Student-Performulator: Predicting Students' Academic Performance at Secondary and Intermediate Level Using Machine Learning. *Annals of Data Science* **10**:3, 637-655. [[Crossref](#)]
123. Helder Moreira Arruda, Rodrigo Simon Bavaresco, Rafael Kunst, Elvis Fernandes Bugs, Giovanni Cheuiche Pesenti, Jorge Luis Victória Barbosa. 2023. Data Science Methods and Tools for Industry 4.0: A Systematic Literature Review and Taxonomy. *Sensors* **23**:11, 5010. [[Crossref](#)]
124. Krishna Raj Raghavendran, Ahmed Elragal. 2023. Low-Code Machine Learning Platforms: A Fastlane to Digitalization. *Informatics* **10**:2, 50. [[Crossref](#)]
125. Dan Wu, Hao Xu, Yaqi Sun, Siyu Lv. 2023. What should we teach? A human-centered data science graduate curriculum model design for iField schools. *Journal of the Association for Information Science and Technology* **74**:6, 623-640. [[Crossref](#)]
126. Shalini R. Urs, Mohamed Minhaj. 2023. Evolution of data science and its education in iSchools : An impressionistic study using curriculum analysis. *Journal of the Association for Information Science and Technology* **74**:6, 606-622. [[Crossref](#)]
127. You Li, Ye Wang, Yuyung Lee, Huan Chen, Alexis Nicolle Petri, Teryn Cha. 2023. Teaching Data Science through Storytelling: Improving Undergraduate Data Literacy. *Thinking Skills and Creativity* **48**, 101311. [[Crossref](#)]
128. Yu Wang, Zulqurnain Ali. 2023. Exploring big data use to predict supply chain effectiveness in Chinese organizations: a moderated mediated model link. *Asia Pacific Business Review* **29**:3, 632-653. [[Crossref](#)]
129. Aneesh Pradeep, Zulfiya Muytenbaeva. Leveraging Data Science to Advance the United Nations Sustainable Development Goals 1-6. [[Crossref](#)]
130. Sasari Samarasinghe, Sachithra Lokuge. Data Democratization 155-183. [[Crossref](#)]
131. Chao Fu, Zijian Wu, Min Xue, Weiyong Liu. 2023. Cross-domain decision making based on TrAdaBoost for diagnosis of breast lesions. *Artificial Intelligence Review* **56**:5, 3987-4017. [[Crossref](#)]
132. Paul B. Hegedus, Stephanie A. Ewing, Clain Jones, Bruce D. Maxwell. 2023. Using spatially variable nitrogen application and crop responses to evaluate crop nitrogen use efficiency. *Nutrient Cycling in Agroecosystems* **126**:1, 1-20. [[Crossref](#)]
133. Huan Chen, Ye Wang, You Li, Yuyung Lee, Alexis Petri, Teryn Cha. 2023. Computer science and non-computer science faculty members' perception on teaching data science via an experiential learning platform. *Education and Information Technologies* **28**:4, 4093-4108. [[Crossref](#)]
134. Huy Vu Ngoc, J. R. R. Mayer, Elie Bitar-Nehme. 2023. Deep Learning to Directly Predict Compensation Values of Thermally Induced Volumetric Errors. *Machines* **11**:4, 496. [[Crossref](#)]
135. Marcello M. Mariani, Matteo Borghi, Benjamin Laker. 2023. Do submission devices influence online review ratings differently across different types of platforms? A big data analysis. *Technological Forecasting and Social Change* **189**, 122296. [[Crossref](#)]
136. Ragmoun Wided. 2023. IT Capabilities, Strategic Flexibility and Organizational Resilience in SMEs Post-COVID-19: A Mediating and Moderating Role of Big Data Analytics Capabilities. *Global Journal of Flexible Systems Management* **24**:1, 123-142. [[Crossref](#)]
137. Xiaozhen Lian, Liang Hou, Wenbo Zhang, Husehng Yan, Ying Liu. 2023. Identifying risky components of display products for redesign considering user attention and failure causality. *Soft Computing* **27**:6, 2921-2942. [[Crossref](#)]
138. Kristin Wulff, Hanne Finnestrand. 2023. Data-driven information for action. *Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO)* **54**:1, 65-77. [[Crossref](#)]
139. Mohamad Abou-Foul, Jose L. Ruiz-Alba, Pablo J. López-Tenorio. 2023. The impact of artificial intelligence capabilities on servitization: The moderating role of absorptive capacity-A dynamic capabilities perspective. *Journal of Business Research* **157**, 113609. [[Crossref](#)]
140. Hajar Nouinou, Elnaz Asadollahi-Yazdi, Isaline Baret, Nhan Quy Nguyen, Mourad Terzi, Yassine Ouazene, Farouk Yalaoui, Russell Kelly. 2023. Decision-making in the context of Industry 4.0: Evidence from the textile and clothing industry. *Journal of Cleaner Production* **391**, 136184. [[Crossref](#)]

141. Jingran Wang, Yi Liu, Peigong Li, Zhenxing Lin, Stavros Sindakis, Sakshi Aggarwal. 2023. Overview of Data Quality: Examining the Dimensions, Antecedents, and Impacts of Data Quality. *Journal of the Knowledge Economy* 12. . [\[Crossref\]](#)
142. Luigi Jesus Basile, Nunzia Carbonara, Roberta Pellegrino, Umberto Panniello. 2023. Business intelligence in the healthcare industry: The utilization of a data-driven approach to support clinical decision making. *Technovation* 120, 102482. [\[Crossref\]](#)
143. Ozgur Kabadurmus, Yaşanur Kayikci, Sercan Demir, Basar Koc. 2023. A data-driven decision support system with smart packaging in grocery store supply chains during outbreaks. *Socio-Economic Planning Sciences* 85, 101417. [\[Crossref\]](#)
144. Yingen Hu, Xiang Li, Yu Song, Chen Huang. 2023. Data-driven evaluation framework for the effectiveness of rural vitalization in China: an empirical case study of Hubei Province. *Environmental Science and Pollution Research* 30:8, 20235-20254. [\[Crossref\]](#)
145. Yufei Hu, Xin Guan, Benran Hu, Yongnan Liu, Hongyang Chen, Tomoaki Ohtsuki. 2023. Distributed Feature Selection Considering Data Pricing Based on Edge Computing in Electricity Spot Markets. *IEEE Internet of Things Journal* 10:3, 2231-2244. [\[Crossref\]](#)
146. Stefan Sleep, Prachi Gala, Dana E. Harrison. 2023. Removing silos to enable data-driven decisions: The importance of marketing and IT knowledge, cooperation, and information quality. *Journal of Business Research* 156, 113471. [\[Crossref\]](#)
147. Franziska Franke, Martin R.W. Hiebl. 2023. Big data and decision quality: the role of management accountants' data analytics skills. *International Journal of Accounting & Information Management* 31:1, 93-127. [\[Crossref\]](#)
148. Anna Visvizi, Orlando Troisi, Mara Grimaldi. Big data and Decision-making: How Big Data Is Relevant Across Fields and Domains 1-11. [\[Crossref\]](#)
149. Antonietta Megaro. Transparency in AI Systems for Value Co-creation in Healthcare 93-105. [\[Crossref\]](#)
150. Ms. Dalvi Swapnali, Mrs. Akshata Chavan. 2023. Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *International Journal of Advanced Research in Science, Communication and Technology* 326-330. [\[Crossref\]](#)
151. Yu Wang, Zulqurnain Ali, Aqsa Mehreen, Khawar Hussain. 2023. The trickle-down effect of big data use to predict organization innovation: the roles of business strategy alignment and information sharing. *Journal of Enterprise Information Management* 36:1, 323-346. [\[Crossref\]](#)
152. Michela Zambetti, Federico Adrodegari, Giuditta Pezzotta, Roberto Pinto, Mario Rapaccini, Cosimo Barbieri. 2023. From data to value: conceptualising data-driven product service system. *Production Planning & Control* 34:2, 207-223. [\[Crossref\]](#)
153. Murat TAŞDAN, Nuriye KARABULUT, Ziyaeddin Halid İPEK. 2023. Yöneticilerin Bakış Açısından İlköğretim Okullarında Veriye Dayalı Karar Verme. *MANAS Sosyal Araştırmalar Dergisi* 12:1, 139-160. [\[Crossref\]](#)
154. М. В. Карчевський. 2023. ОБЧИСЛЮВАЛЬНЕ КРИМІНОЛОГІЧНЕ АРГУМЕНТУВАННЯ: ПОНЯТТЯ, МОЖЛИВОСТІ ТА ПЕРСПЕКТИВИ ВИКОРИСТАННЯ. *Вісник Луганського державного університету внутрішніх справ імені Е.О. Дідоренка* 4:100, 112-126. [\[Crossref\]](#)
155. Mayk Brendon Almeida Antunes, Matheus de Freitas Issa, Raphael Magalhães Hoed. 2023. TÉCNICAS DE MACHINE LEARNING APLICADA A MINERAÇÃO DE DADOS E ANÁLISE DE SENTIMENTOS PARA PREDIÇÃO DE HOMOFOBIA NO TWITTER. *REVISTA FOCO* 16:1, e853. [\[Crossref\]](#)
156. Louise De Koker, Gabriel Tati, Martin Chanza, Elias Munapo. Re-assessing the Role of the Statistician in the Era of Big Data: A Business Perspective 1176-1191. [\[Crossref\]](#)
157. Wei Li, Shengyu Kang, Yueqiang Sun, Weihua Bai, Yuhe Wang, Hongqing Song. 2023. A Machine Learning Approach for Air-Quality Forecast by Integrating GNSS Radio Occultation Observation and Weather Modeling. *Atmosphere* 14:1, 58. [\[Crossref\]](#)
158. Anca Antoaneta Vărzaru. 2023. Assessing Digital Transformation Acceptance in Public Organizations' Marketing. *Sustainability* 15:1, 265. [\[Crossref\]](#)
159. Febrianti, Hendra Wijaya Salim, Jonathan Buntoro, Oliviana, Ford Lumban Gaol, Tokuro Matsuo, Fonny Hutagalung. The Effect of Big Data Decision-Making Analysis Used in Improving the Effectiveness of the Distance Learning Process at Private University in Jakarta 621-636. [\[Crossref\]](#)
160. Marek Nahotko, Magdalena Zych, Aneta Januszko-Szakiel, Małgorzata Jaskowska. 2023. Big data-driven investigation into the maturity of library research data services (RDS). *The Journal of Academic Librarianship* 49:1, 102646. [\[Crossref\]](#)
161. Shashi, Piera Centobelli, Roberto Cerchione, Myriam Ertz, Eugenio Oropallo. 2023. What we learn is what we earn from sustainable and circular construction. *Journal of Cleaner Production* 382, 135183. [\[Crossref\]](#)
162. Man Li, Ye Zhu, Yuxin Shen, Maia Angelova. 2023. Clustering-enhanced stock price prediction using deep learning. *World Wide Web* 26:1, 207-232. [\[Crossref\]](#)
163. Alessandro Tufano, Rob Zuidwijk, Jan Van Dalen. 2023. The development of data-driven logistic platforms for barge transportation network under incomplete data. *Omega* 114, 102746. [\[Crossref\]](#)

164. Po-Hsin Chou, Jui-Chung John Lin, Tsair-Wei Chien. 2023. Using text mining and forest plots to identify similarities and differences between two spine-related journals based on medical subject headings (MeSH terms) and author-specified keywords in 100 top-cited articles. *Scientometrics* **128**:1, 1-17. [[Crossref](#)]
165. Bernardo Nicoletti. Industrial Revolutions and Supply Network 5.0 43-101. [[Crossref](#)]
166. Oray Kulac, Banu Y. Ekren, A. Ozgur Toy. Digital Twins for Decision Making in Supply Chains 86-96. [[Crossref](#)]
167. Farhad Khoshbakht, S. M. K. Quadri. Big Data Framework for Analytics Business Intelligence 51-69. [[Crossref](#)]
168. Rene Barbie Browne, Jai Narain Vishwakarma, Vedant Vikrom Borah, Raj Kumar Pegu, Jayanti Datta Roy. In silico application of data science, genomics, and bioinformatics in screening drug candidates against COVID-19 107-128. [[Crossref](#)]
169. Prathibha S. Babu, Sangeetha Subhash, K. Ilango. SOC Estimation of Li-Ion Battery Using Hybrid Artificial Neural Network and Adaptive Neuro-Fuzzy Inference System 389-405. [[Crossref](#)]
170. Mohammad Reza Bahrami, Bahareh Bahrami, Farima Behboodi, Samae Pourrafie. Teaching the Future: The Vision of AI/ChatGPT in Education 393-402. [[Crossref](#)]
171. Henry Omietimi, Sunday A. Afolalu, Joseph F. Kayode, Stella I. Monye, Sunday L. Lawal, Moses E. Emeter. 2023. An overview of material removal processes and its importance. *E3S Web of Conferences* **391**, 01077. [[Crossref](#)]
172. Geymerson S. Ramos, Danilo Fernandes, Jorge Artur P. de M. Coelho, Andre L. L. Aquino. Toward Data Lake Technologies for Intelligent Societies and Cities 3-29. [[Crossref](#)]
173. Christoph Schröer, Sven Wittfoth, Jorge Marx Gómez. Multi-metric Approach for Decomposition of Microservice-Based Data Science Workflows 355-369. [[Crossref](#)]
174. Gayathri De Lanerolle, Peter Phiri, Athar Haroon. Data science and clinical informatics 111-152. [[Crossref](#)]
175. Fabiola del Toro Osorio, Victoria Eugenia Ospina Becerra, Elsa Estévez. Designing a Data Strategy for Organizations 143-156. [[Crossref](#)]
176. Parikshit N. Mahalle, Pravin P. Hujare, Gitanjali Rahul Shinde. Introduction to Predictive Analytics 1-10. [[Crossref](#)]
177. Frank Emmert-Streib, Salissou Moutari, Matthias Dehmer. Introduction to Learning from Data 1-14. [[Crossref](#)]
178. Alessandro Ancarani, Carmela Di Mauro. Successful Digital Transformations – Identifying the Role of Leadership 13-26. [[Crossref](#)]
179. Nada Elgendy, Ahmed Elragal, Tero Päiväranta. 2023. Evaluating collaborative rationality-based decisions: a literature review. *Procedia Computer Science* **219**, 647-657. [[Crossref](#)]
180. Frank Acito. Logistic Regression 125-167. [[Crossref](#)]
181. Saeid Pourroostaei Ardakani, Ali Cheshmehzangi. The Role of Big Data Analytics in Urban Systems: Review and Prospect for Smart Transport and Healthcare Systems 1-26. [[Crossref](#)]
182. Yuting Bai, Bin Yan, Chenguang Zhou, Tingli Su, Xuebo Jin. 2023. State of art on state estimation: Kalman filter driven by machine learning. *Annual Reviews in Control* **56**, 100909. [[Crossref](#)]
183. Anna Petrovskis, Betty Bekemeier, Elizabeth Heitkemper, Jenna van Draanen. 2023. The DASH model: Data for addressing social determinants of health in local health departments. *Nursing Inquiry* **30**:1. . [[Crossref](#)]
184. Alejandro Peña-Ayala. Educational Data Science: An “Umbrella Term” or an Emergent Domain? 95-147. [[Crossref](#)]
185. Dominik Lis, Joshua Gelhaar, Boris Otto. Data Strategy and Policies: The Role of Data Governance in Data Ecosystems 27-55. [[Crossref](#)]
186. Destina Ratna Asih Khodijah Kadarsah, Rajesri Govindaraju, Budhi Prihartono. 2023. The Role of Knowledge-Oriented Leadership in Fostering Innovation Capabilities: The Mediating Role of Data Analytics Maturity. *IEEE Access* **11**, 129683-129702. [[Crossref](#)]
187. Tanatip Watthaisong, Khamron Sunat, Nipotept Muangkote. Chaotic Mountain Gazelle Optimizer (CMGO): A Robust Optimization Algorithm for K-Means Clustering of Diverse Data Types 18-34. [[Crossref](#)]
188. Sarah Pinon. Business User-Oriented Recommender System of Data 613-621. [[Crossref](#)]
189. Ronald Buijsse, Martijn Willemsen, Chris Snijders. Data-Driven Decision-Making 239-277. [[Crossref](#)]
190. Willem-Jan van den Heuvel, Werner Liebrechts, Arjan van den Born. Data Science for Entrepreneurship: The Road Ahead 521-532. [[Crossref](#)]
191. Toni Toharudin, Rezzy Eko Caraka, Indah Reski Pratiwi, Yunho Kim, Prana Ugiana Gio, Anjar Dimara Sakti, Maengseok Noh, Farid Azhar Lutfi Nugraha, Resa Septiani Pontoh, Tafia Hasna Putri, Thalita Safa Azzahra, Jessica Jesslyn Cerelia, Gumgum Darmawan, Bens Pardamean. 2023. Boosting Algorithm to Handle Unbalanced Classification of PM 2.5 Concentration Levels by

- Observing Meteorological Parameters in Jakarta-Indonesia Using AdaBoost, XGBoost, CatBoost, and LightGBM. *IEEE Access* **11**, 35680-35696. [[Crossref](#)]
192. Alex D. Rogers, Octavio Aburto-Oropeza, Ward Appeltans, Jorge Assis, Lisa T. Ballance, Philippe Cury, Carlos Duarte, Fabio Favoretto, Joy Kumagai, Catherine Lovelock, Patricia Miloslavich, Aidin Niamir, David Obura, Bethan C. O'Leary, Gabriel Reygondeau, Callum Roberts, Yvonne Sadovy, Tracey Sutton, Derek Tittensor, Enriqueta Velarde. Critical Habitats and Biodiversity: Inventory, Thresholds and Governance 333-392. [[Crossref](#)]
 193. Min-Hooi Chuah, Ramayah Thurusamry. 2022. The relationship between architecture, social, law and market in determine challenges of big data analysis for Malaysia SMEs. *Cogent Business & Management* **9**:1. . [[Crossref](#)]
 194. Sukjin You, Soohyung Joo, Marie Katsurai. 2022. Data mining topics in the discipline of library and information science: analysis of influential terms and Dirichlet multinomial regression topic model. *Aslib Journal of Information Management* **62**. . [[Crossref](#)]
 195. Jagan Mohan Reddy, Abhishek Attuluri, Abhinay Kolli, Nazmus Sakib, Hossain Shahriar, Alfredo Cuzzocrea. A Crowd Source System for YouTube Big Data Analytics: Unpacking Values from Data Sprawl 5451-5457. [[Crossref](#)]
 196. Mavigson Silva, Letícia Azevedo de Farias Pereira, Bruno José Bezerra Silva, Caique Ferreira Borges. 2022. Integração dos sistemas do transporte público intermunicipal do Rio Grande do Norte: Análise da declaração de demanda. *Cadernos Gestão Pública e Cidadania* **28**, e88502. [[Crossref](#)]
 197. Nuria Mollá, Ciara Heavin, Alejandro Rabasa. 2022. Data-driven decision making: new opportunities for DSS in data stream contexts. *Journal of Decision Systems* **31**:sup1, 255-269. [[Crossref](#)]
 198. Wenxia Xu, Lei Feng, Jun Ma. 2022. Understanding the domain of driving distraction with knowledge graphs. *PLOS ONE* **17**:12, e0278822. [[Crossref](#)]
 199. Peter Fröhlich, Alexander G. Mirnig, Damiano Falcioni, Johann Schrammel, Lisa Diamond, Isabel Fischer, Manfred Tscheligi. 2022. Effects of reliability indicators on usage, acceptance and preference of predictive process management decision support systems. *Quality and User Experience* **7**:1. . [[Crossref](#)]
 200. Claudia Eckert, Ola Isaksson, Malin Hane-Hagström, Calandra Eckert. 2022. My Facts Are not Your Facts: Data Wrangling as a Socially Negotiated Process, A Case Study in a Multisite Manufacturing Company. *Journal of Computing and Information Science in Engineering* **22**:6. . [[Crossref](#)]
 201. J. Dingelstad, R. T. Borst, A. Meijer. 2022. Hybrid Data Competencies for Municipal Civil Servants: An Empirical Analysis of the Required Competencies for Data-Driven Decision-Making. *Public Personnel Management* **51**:4, 458-490. [[Crossref](#)]
 202. Kornelia Batko, Andrzej Ślęzak. 2022. The use of Big Data Analytics in healthcare. *Journal of Big Data* **9**:1. . [[Crossref](#)]
 203. Georgios Vranopoulos, Nathan Clarke, Shirley Atkinson. 2022. Addressing big data variety using an automated approach for data characterization. *Journal of Big Data* **9**:1. . [[Crossref](#)]
 204. Ayed Banibaqash, Ziad Hunaiti, Maysam Abbod. 2022. An Analytical Feasibility Study for Solar Panel Installation in Qatar Based on Generated to Consumed Electrical Energy Indicator. *Energies* **15**:24, 9270. [[Crossref](#)]
 205. Yili Chen, Congdong Li, Han Wang. 2022. Big Data and Predictive Analytics for Business Intelligence: A Bibliographic Study (2000-2021). *Forecasting* **4**:4, 767-786. [[Crossref](#)]
 206. Claire Babirye, Chisenga Muyoya, Suvodeep Mazumdar, Andrea Jimenez, Ciira Maina, Jabhera Matogoro, Margaret Nyambura Ndung'u, Dorothea Kleine. 2022. Data science for empowerment: understanding the data science training landscape for women and girls in Africa. *Gender, Technology and Development* **26**:3, 437-462. [[Crossref](#)]
 207. Edwin Arrey Agbor, Yanzhen Qu. Creating a Dataset Used for Applying Machine Learning Techniques to Accurately Forecast the Energy Cost in Home-Based Small Businesses 677-683. [[Crossref](#)]
 208. Marian Pompiliu Cristescu, Raluca Andreea Nerişanu, Dumitru Alexandru Mara. 2022. Using Data Mining in the Sentiment Analysis Process on the Financial Market. *Journal of Social and Economic Statistics* **11**:1-2, 36-58. [[Crossref](#)]
 209. Rossella Pozzi, Violetta Giada Cannas, Maria Pia Ciano. 2022. Linking data science to lean production: a model to support lean practices. *International Journal of Production Research* **60**:22, 6866-6887. [[Crossref](#)]
 210. Eriks Klotins, Tony Gorschek, Katarina Sundelin, Erik Falk. 2022. Towards cost-benefit evaluation for continuous software engineering activities. *Empirical Software Engineering* **27**:6. . [[Crossref](#)]
 211. Muhammad Zahid Sharif, Nayan Di, Fanglin Liu. 2022. Monitoring honeybees (*Apis* spp.) (Hymenoptera: Apidae) in climate-smart agriculture: A review. *Applied Entomology and Zoology* **57**:4, 289-303. [[Crossref](#)]
 212. Serkan Altuntas, Sibel Selim, Fatma Altuntas. 2022. A hierarchical clustering based panel data approach: A case study of regional incentives. *International Journal of Information Management Data Insights* **2**:2, 100098. [[Crossref](#)]
 213. Mohammad Daradkeh, Hamzah Alkhazaleh. Big Data Analytics and Healthcare Performance: A Unified Model of Competencies 1-5. [[Crossref](#)]

214. Santisook Limpeetcharoenchot, Nagul Cooharajanone, Thira Chavarnakul, Nuttirudee Charoenruk, Kanokwan Atchariyachanvanich. 2022. Adaptive big data maturity model using latent class analysis for small and medium businesses in Thailand. *Expert Systems with Applications* **206**, 117965. [[Crossref](#)]
215. Brady Lund, Ting Wang. 2022. What Does Information Science Offer for Data Science Research?: A Review of Data and Information Ethics Literature. *Journal of Data and Information Science* **7**:4, 16-38. [[Crossref](#)]
216. Indraneel Dabhade. Data Science for Industry 4.0 58-71. [[Crossref](#)]
217. Harini Dissanayake, Paul J. Bracewell. Bias in Data-Informed Decision Making 1173-1186. [[Crossref](#)]
218. Matthias Pohl, Christian Haertel, Daniel Staegemann, Klaus Turowski. Data Science Methodology 1201-1214. [[Crossref](#)]
219. Satish Vadlamani, Mayank Modashiya. Improving Transportation Planning Using Machine Learning 3076-3088. [[Crossref](#)]
220. Arko Barman, Leslie S. Beckman, Yasmin Chebaro. Interdisciplinary Computing Education: An Introductory Programming and Data Science Course for Postdoctoral Researchers in the Biosciences 1-8. [[Crossref](#)]
221. Jian-Wei Wu, Yu-Hua Yan, Tsair-Wei Chien, Willy Chou. 2022. Trend and prediction of citations on the topic of neuromuscular junctions in 100 top-cited articles since 2001 using a temporal bar graph: A bibliometric analysis. *Medicine* **101**:40, e30674. [[Crossref](#)]
222. Aleena Shuja, Naveed Yazdani, Aleema Shuja. 2022. Investigating Resilience and Performance of Emergent Financial Technology Startups Endorsed by Knowledge Management. *JISR management and social sciences & economics* **18**:1, 1-13. [[Crossref](#)]
223. Claus Boye Asmussen, Steffen Lundgaard Jørgensen, Charles Møller. 2022. Design and deployment of an analytic artefact – investigating mechanisms for integrating analytics and manufacturing execution system. *Enterprise Information Systems* **16**:10-11, 1587-1616. [[Crossref](#)]
224. Nada Elgendy, Ahmed Elragal, Tero Päivärinta. 2022. DECAS: a modern data-driven decision theory for big data and analytics. *Journal of Decision Systems* **31**:4, 337-373. [[Crossref](#)]
225. Sachin Kumar, Shivam Panwar, Jagvinder Singh, Anuj Kumar Sharma, Zairu Nisha. 2022. iCAD: an intelligent deep learning model to categorise current affairs news article for efficient journalistic process. *International Journal of System Assurance Engineering and Management* **13**:5, 2572-2582. [[Crossref](#)]
226. Francesco Smaldone, Adelaide Ippolito, Jelena Lagger, Marco Pellicano. 2022. Employability skills: Profiling data scientists in the digital labour market. *European Management Journal* **40**:5, 671-684. [[Crossref](#)]
227. Michael Overton, Stephen Kleinschmit. 2022. Data science literacy: Toward a philosophy of accessible and adaptable data science skill development in public administration programs. *Teaching Public Administration* **40**:3, 354-365. [[Crossref](#)]
228. Hidetoshi Togo, Kohei Asanuma, Tatsushi Nishi, Ziang Liu. 2022. Machine Learning and Inverse Optimization for Estimation of Weighting Factors in Multi-Objective Production Scheduling Problems. *Applied Sciences* **12**:19, 9472. [[Crossref](#)]
229. Maryia Zaitsava, Elona Marku, Maria Chiara Di Guardo. 2022. Is data-driven decision-making driven only by data? When cognition meets data. *European Management Journal* **40**:5, 656-670. [[Crossref](#)]
230. Bernard J. Jansen, Joni Salminen, Soon-gyo Jung, Hind Almerexhi. 2022. The illusion of data validity: Why numbers about people are likely wrong. *Data and Information Management* **6**:4, 100020. [[Crossref](#)]
231. Haleh Ayatollahi. Health Data Science and Semantic Technologies 299-321. [[Crossref](#)]
232. Qixin Guo, Liang Yu, Ran Wang, Rui Wang, Weikang Jiang. 2022. The acoustic inverse problem in the inhomogeneous medium by iterative Bayesian focusing algorithm. *Signal Processing* **198**, 108602. [[Crossref](#)]
233. Xuhang Xu, Chunshan Liu, Jianghui Li, Yongchun Miao. 2022. Trajectory clustering for SVR-based Time of Arrival estimation. *Ocean Engineering* **259**, 111930. [[Crossref](#)]
234. Alexander Guzman-Urbina, Kakeru Ouchi, Hajime Ohno, Yasuhiro Fukushima. 2022. FIEMA, a system of fuzzy inference and emission analytics for sustainability-oriented chemical process design. *Applied Soft Computing* **126**, 109295. [[Crossref](#)]
235. М. В. Карчевський. 2022. ПРОТИДІЯ ЗЛОЧИННОСТІ В УКРАЇНІ У ФОРМАТІ DATA SCIENCE. Вісник Луганського державного університету внутрішніх справ імені Е.О. Дідоренка **2**:98, 202-227. [[Crossref](#)]
236. Yaseen Khan, Surendra Thakur. Fake News Detection of South African COVID-19 Related Tweets using Machine Learning 1-5. [[Crossref](#)]
237. Lijia Wang, Xin Tong, Y.X. Rachel Wang. 2022. Statistics in everyone's backyard: An impact study via citation network analysis. *Patterns* **3**:8, 100532. [[Crossref](#)]
238. Ryo Kunimoto, Jürgen Bajorath, Kazumasa Aoki. 2022. From traditional to data-driven medicinal chemistry: A case study. *Drug Discovery Today* **27**:8, 2065-2070. [[Crossref](#)]

239. Austin Pickup. 2022. Toward a historical ontology of the infopolitics of data-driven decision-making (DDDM) in education. *Educational Philosophy and Theory* 54:9, 1476-1487. [[Crossref](#)]
240. Simon Radler, Eugen Rigger, Juergen Mangler, Stefanie Rinderle-Ma. Integration of Machine Learning Task Definition in Model-Based Systems Engineering using SysML 546-551. [[Crossref](#)]
241. Weihua Wang, Jianguo Du, Fakhar Shahzad, Xiangyi Duan, Xiaowen Zhu. 2022. The dynamic evolution of collaborative networks in sustainable development: Untying impact of environmental policy in China using network-based text analysis approach. *Frontiers in Psychology* 13. . [[Crossref](#)]
242. Gianluca Scarano, Barry Colfer. 2022. Linking active labour market policies to digitalisation—a review between remote and automated possibilities. *International Journal of Sociology and Social Policy* 42:13/14, 98-112. [[Crossref](#)]
243. Mohammad Daradkeh, Laith Abualigah, Shadi Atalla, Wathiq Mansoor. 2022. Scientometric Analysis and Classification of Research Using Convolutional Neural Networks: A Case Study in Data Science and Analytics. *Electronics* 11:13, 2066. [[Crossref](#)]
244. Farheen Naz, Rohit Agrawal, Anil Kumar, Angappa Gunasekaran, Abhijit Majumdar, Sunil Luthra. 2022. Reviewing the applications of artificial intelligence in sustainable supply chains: Exploring research propositions for future directions. *Business Strategy and the Environment* 31:5, 2400-2423. [[Crossref](#)]
245. Ying Dai, Lei Dou, Han Song, Lin Zhou, Haiyan Li. 2022. Two-way information sharing of uncertain demand forecasts in a dual-channel supply chain. *Computers & Industrial Engineering* 169, 108162. [[Crossref](#)]
246. Xiaofeng Su, Weipeng Zeng, Manhua Zheng, Xiaoli Jiang, Wenhe Lin, Anxin Xu. 2022. Big data analytics capabilities and organizational performance: the mediating effect of dual innovations. *European Journal of Innovation Management* 25:4, 1142-1160. [[Crossref](#)]
247. Shubham Gujar, Tanishka Shah, Dewen Honawale, Vedant Bhosale, Faizan Khan, Devika Verma, Rakesh Ranjan. GenEthos: A Synthetic Data Generation System With Bias Detection And Mitigation 1-6. [[Crossref](#)]
248. Saeed Hameed Aldulaimi, Basel J. A. Ali, Qahtan M. Yas, Marwan M. Abdeldayem, Abed Rzaij Aswad, Ali Mukhlif Hammad. Application of Big Data Analysis to Foresight the Future: A Review of Opportunities, Approaches, and New Research Directions 346-354. [[Crossref](#)]
249. Syed Ghazi Sarwat, Frank Brückerhoff-Plückelmann, Santiago García-Cuevas Carrillo, Emanuele Gemo, Johannes Feldmann, Harish Bhaskaran, C. David Wright, Wolfram H. P. Pernice, Abu Sebastian. 2022. An integrated photonics engine for unsupervised correlation detection. *Science Advances* 8:22. . [[Crossref](#)]
250. Atif Ismail, Hafiz Muhammad Awais Rashid, Raof Gholami, Arshad Raza. 2022. Characterization based machine learning modeling for the prediction of the rheological properties of water-based drilling mud: an experimental study on grass as an environmental friendly additive. *Journal of Petroleum Exploration and Production Technology* 12:6, 1677-1695. [[Crossref](#)]
251. Henriett Pintér, Franciska Gál, Pál Molnár. 2022. The professional network underlying cerebral palsy intervention research based on systematic reviews and meta-analyses published in international journals: authors' communities, institutional networks, and international collaboration. *Heliyon* 8:6, e09718. [[Crossref](#)]
252. Pedro Jácome de Moura Jr. 2022. Is data science a science? The essence of phenomenon and the role of theory in the emerging field. *Kybernetes* 51:7, 2416-2434. [[Crossref](#)]
253. Arif Mahmud. 2022. Application and Criminalization of the Artificial Intelligence in Business: Recommendation to Counter the Regulatory Challenges. *Journal of Applied Security Research* 7, 1-11. [[Crossref](#)]
254. Prathmesh Vasant Pawar, Ratna Achuta Paluri. 2022. Big Data Analytics in Logistics and Supply Chain Management: A Review of Literature. *Vision: The Journal of Business Perspective* 183, 097226292210916. [[Crossref](#)]
255. Moncef Garouani, Adeel Ahmad, Mourad Bouneffa, Mohamed Hamlich, Gregory Bourguin, Arnaud Lewandowski. 2022. Towards big industrial data mining through explainable automated machine learning. *The International Journal of Advanced Manufacturing Technology* 120:1-2, 1169-1188. [[Crossref](#)]
256. Huy Vu Ngoc, J.R.R. Mayer, Elie Bitar-Nehme. 2022. Deep learning LSTM for predicting thermally induced geometric errors using rotary axes' powers as input parameters. *CIRP Journal of Manufacturing Science and Technology* 37, 70-80. [[Crossref](#)]
257. Wanqing Wu, Wenyu Mao. 2022. An Efficient and Scalable Algorithm to Mine Functional Dependencies from Distributed Big Data. *Sensors* 22:10, 3856. [[Crossref](#)]
258. Ryan P. Kreiser, Aidan K. Wright, Tristan L. McKenzie, J. Alex Albright, Eric D. Mowles, Jared E. Hollows, Scott Limbocker, Melissa Eslinger, Ryan Limbocker, Lam T. Nguyen. 2022. Utilization of Standardized College Entrance Metrics to Predict Undergraduate Student Success in Chemistry. *Journal of Chemical Education* 99:4, 1725-1733. [[Crossref](#)]
259. Orlando Troisi, Mara Grimaldi. 2022. Guest editorial: Data-driven orientation and open innovation: the role of resilience in the (co-)development of social changes. *Transforming Government: People, Process and Policy* 16:2, 165-171. [[Crossref](#)]

260. Ayodeji Folorunsho Ajayi, Emmanuel Tayo Adebayo, Iyanuoluwa Oluwadunsi Adebayo, Olubunmi Simeon Oyekunle, Victor Oluwaseyi Amos, Segun Emmanuel Bamidele, Goodness Olusayo Olatinwo. Application of Data Science Approaches to Investigate Autoimmune Thyroid Disease in Precision Medicine . [[Crossref](#)]
261. Christian Colot, Philippe Baecke, Isabelle Linden. 2022. Toward Decision Support for Telecom External Data Monetization: A Study of the Value of Network- and Personality-Based Metrics for Third-Party Businesses. *Big Data* **10**:2, 115-137. [[Abstract](#)] [[Full Text](#)] [[PDF](#)] [[PDF Plus](#)] [[Supplementary Material](#)]
262. Azazkhan Ibrahimkhan Pathan, Prasit Girish Agnihotri, Dhruv Patel, Cristina Prieto. 2022. Mesh grid stability and its impact on flood inundation through (2D) hydrodynamic HEC-RAS model with special use of Big Data platform—a study on Purna River of Navsari city. *Arabian Journal of Geosciences* **15**:7. . [[Crossref](#)]
263. Ladjel Bellatreche, Carlos Ordonez, Dominique Méry, Matteo Golfarelli, El Hassan Abdelwahed. 2022. The central role of data repositories and data models in Data Science and Advanced Analytics. *Future Generation Computer Systems* **129**, 13-17. [[Crossref](#)]
264. Yasanur Kayikci, Sercan Demir, Sachin K. Mangla, Nachiappan Subramanian, Basar Koc. 2022. Data-driven optimal dynamic pricing strategy for reducing perishable food waste at retailers. *Journal of Cleaner Production* **344**, 131068. [[Crossref](#)]
265. Chen Chen, Zvi Yaari, Elana Apfelbaum, Piotr Grodzinski, Yosi Shamay, Daniel A. Heller. 2022. Merging data curation and machine learning to improve nanomedicines. *Advanced Drug Delivery Reviews* **183**, 114172. [[Crossref](#)]
266. Charuleka Varadharajan, Alison P. Appling, Bhavna Arora, Danielle S. Christianson, Valerie C. Hendrix, Vipin Kumar, Aranildo R. Lima, Juliane Müller, Samantha Oliver, Mohammed Ombadi, Talita Perciano, Jeffrey M Sadler, Helen Weierbach, Jared D. Willard, Zexuan Xu, Jacob Zwart. 2022. Can machine learning accelerate process understanding and decision-relevant predictions of river water quality?. *Hydrological Processes* **36**:4. . [[Crossref](#)]
267. Marwah Ahmed Halwani, S. Yasaman Amirkieae, Nicholas Evangelopoulos, Victor Prybutok. 2022. Job qualifications study for data science and big data professions. *Information Technology & People* **35**:2, 510-525. [[Crossref](#)]
268. S Vijayakumar Bharathi, Dhanya Pramod, Ramakrishnan Raman. Using Statistical Decision Making for a University Examination Scenario 460-464. [[Crossref](#)]
269. Mr. Vasudev Shahapur, Amisha Shetty, Chaitra P Shetty, Gayathri, Prathiksha Shetty. 2022. Data Science. *International Journal of Advanced Research in Science, Communication and Technology* 487-495. [[Crossref](#)]
270. Youetta Kunneman, Mauricy Alves da Motta-Filho, Jasper van der Waa. 2022. Data science for service design: An introductory overview of methods and opportunities. *The Design Journal* **25**:2, 186-204. [[Crossref](#)]
271. Neelam Singh, Devesh Pratap Singh, Bhasker Pant. 2022. Big Data Knowledge Discovery as a Service: Recent Trends and Challenges. *Wireless Personal Communications* **123**:2, 1789-1807. [[Crossref](#)]
272. Ruy Roberto Porto Ascenso Rosa, Ronem Matos Lavareda Filho, José Elislande Breno de Souza Linhares. 2022. Influência das comorbidades para a ocorrência de óbitos por COVID-19 em 2020: razão de chances no estado do Amazonas. *HU Revista* **48**, 1-8. [[Crossref](#)]
273. Vala Ali Rohani, Jahan Ara Peerally, Sedigheh Moghavvemi, Flavio Guerreiro, Tiago Pinho. 2022. Illustrating scholar-practitioner collaboration for data-driven decision-making in the optimization of logistics facility location and implications for increasing the adoption of AR and VR practices. *The TQM Journal* **34**:2, 280-302. [[Crossref](#)]
274. Roberto Salazar-Reyna, Fernando Gonzalez-Aleu, Edgar M.A. Granda-Gutierrez, Jenny Diaz-Ramirez, Jose Arturo Garza-Reyes, Anil Kumar. 2022. A systematic literature review of data science, data analytics and machine learning applied to healthcare engineering systems. *Management Decision* **60**:2, 300-319. [[Crossref](#)]
275. Tobias Kollmann, Lucas Kleine-Stegemann, Katharina de Cruppe, Christina Then-Bergh. 2022. Eras of Digital Entrepreneurship. *Business & Information Systems Engineering* **64**:1, 15-31. [[Crossref](#)]
276. Ahmed M. Shahat Osman, Ahmed A. Elragal, Anna Ståhlbröst. 2022. Data-Driven Decisions in Smart Cities: A Digital Transformation Case Study. *Applied Sciences* **12**:3, 1732. [[Crossref](#)]
277. Iris Figalist, Christoph Elsner, Jan Bosch, Helena Holmström Olsson. 2022. Breaking the vicious circle: A case study on why AI for software analytics and business intelligence does not take off in practice. *Journal of Systems and Software* **184**, 111135. [[Crossref](#)]
278. Ummaporn Pora, Nathasit Gerdri, Natcha Thawesaengskulthai, Sipat Triukose. 2022. Data-Driven Roadmapping (DDRM): Approach and Case Demonstration. *IEEE Transactions on Engineering Management* **69**:1, 209-227. [[Crossref](#)]
279. Christina G. Skarpathiotaki, Konstantinos E. Psannis. 2022. Cross-Industry Process Standardization for Text Analytics. *Big Data Research* **27**, 100274. [[Crossref](#)]
280. Wenjun Li, Kazuki Nagashima, Takuro Hosomi, Chen Wang, Yosuke Hanai, Atsuo Nakao, Atsushi Shunori, Jiangyang Liu, Guozhu Zhang, Tsunaki Takahashi, Wataru Tanaka, Masaki Kanai, Takeshi Yanagida. 2022. Mechanistic Approach for Long-Term Stability of a Polyethylene Glycol–Carbon Black Nanocomposite Sensor. *ACS Sensors* **7**:1, 151-158. [[Crossref](#)]

281. Faten Hamad, Hussam Fakhuri, Sinaria Abdel Jabbar. 2022. Big Data Opportunities and Challenges for Analytics Strategies in Jordanian Academic Libraries. *New Review of Academic Librarianship* **28**:1, 37-60. [[Crossref](#)]
282. Giorgia Brambilla Pisoni, Mariarosaria Taddeo. 2022. Apropos Data Sharing: Abandon the Distrust and Embrace the Opportunity. *DNA and Cell Biology* **41**:1, 11-15. [[Abstract](#)] [[Full Text](#)] [[PDF](#)] [[PDF Plus](#)]
283. Amonida Zadissa, Rolf Apweiler. Data Mining, Quality and Management in the Life Sciences 3-25. [[Crossref](#)]
284. Elena Hernández-Nieves, José A. García-Coria, Sara Rodríguez-González, Ana B. Gil-González. Distributed Decision Blockchain-Secured Support System to Enhance Stock Market Investment Process 48-60. [[Crossref](#)]
285. Lisa Her, Jinzhu Gao, Lewis E. Jacobson, Jonathan M. Saxe, Kathy L. Leslie, Courtney Jensen. Predicting Falls in Older Adults Aged 65 and up Based on Fall Risk Dataset 602-616. [[Crossref](#)]
286. Antonio J. Tallón-Ballesteros, Luís Correia, Rocío Leal-Díaz. Attribute Subset Selection for Image Recognition. Random Forest Under Assessment 821-827. [[Crossref](#)]
287. Ang Liu, Yuchen Wang, Xingzhi Wang. Data-Driven Engineering Design 1-22. [[Crossref](#)]
288. Leila Halawi, Amal Clarke, Kelly George. Introduction to Analytics and Data Science 1-11. [[Crossref](#)]
289. Mojtaba Heidarysafa, Kamran Kowsari, Masoud Bashiri, Donald E. Brown. Towards a Knowledge Discovery Framework for Data Science Job Market in the United States 875-887. [[Crossref](#)]
290. Shen-yang Song, Jing Li, Wei Yan, Jian-xiao Zhang. Comparative Study on the Cleanliness of Ultra-Low Carbon Al-Killed Steel by Different Heating Processes 151-171. [[Crossref](#)]
291. Mario Chiasson, Viktor Freiman. Rethinking the 21st-Century School: New Citizens' Skills for the Digital Era and Their Interaction with Mathematics Teaching and Learning 69-107. [[Crossref](#)]
292. Federico Galli. Algorithmic Manipulation 209-259. [[Crossref](#)]
293. Felix Kruse, René Kessler, Jan-Hendrik Witte. Structured and Targeted Communication as an Enabler for Sustainable Data Science Projects 391-406. [[Crossref](#)]
294. Taha M. Mohamed, Abdulaziz Alharbi, Ibrahim Alhassan, Sherif Kholeif. Applications of Industry 4.0 on Saudi Supply Chain Management: Technologies, Opportunities, and Challenges 189-204. [[Crossref](#)]
295. Lourdes S. Martinez. Data Science 328-331. [[Crossref](#)]
296. Wolfgang Maaß. Smart Services in der Datenökonomie zur Monetarisierung von Fertigungsdaten 417-433. [[Crossref](#)]
297. Andrea Coraddu, Miltiadis Kalikatzarakis, Jake Walker, Davide Ilardi, Luca Oneto. Data science and advanced analytics for shipping energy systems 303-349. [[Crossref](#)]
298. Bart L. MacCarthy, Dmitry Ivanov. The Digital Supply Chain—emergence, concepts, definitions, and technologies 3-24. [[Crossref](#)]
299. Gopi Battineni. Opportunities and challenges in healthcare with the management of big biomedical data 265-275. [[Crossref](#)]
300. Naciye Güliz Uğur, Aykut Hamit Turan. Understanding Big Data 1-21. [[Crossref](#)]
301. Fatma Chiheb, Fatima Boumahdi, Hafida Bouarfa. A Conceptual Model for Describing the Integration of Decision Aspect into Big Data 294-318. [[Crossref](#)]
302. Gabriela Viale Pereira, Gregor Eibl, Constantinos Stylianou, Gilberto Martínez, Haris Neophytou, Peter Parycek. The Role of Smart Technologies to Support Citizen Engagement and Decision Making 687-703. [[Crossref](#)]
303. Cheligeer Cheligeer, Jingwei Huang, Guosong Wu, Nadia Bhuiyan, Yuan Xu, Yong Zeng. 2022. Machine learning in requirements elicitation: a literature review. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* **36**. . [[Crossref](#)]
304. John S. Silva, Marion J. Ball, Mark Polyak, Gabriela Mustata Wilson. The Future of Health Systems: Health Intelligence 461-469. [[Crossref](#)]
305. Abid Haleem, Mohd Javaid, Ravi Pratap Singh, Rajiv Suman. 2022. Medical 4.0 technologies for healthcare: Features, capabilities, and applications. *Internet of Things and Cyber-Physical Systems* **2**, 12-30. [[Crossref](#)]
306. Nabila Mansouri, Khalid Sultan, Aakash Ahmad, Ibrahim Alseadoon, Adal Alkhalil. 2022. A Deep Learning Framework for COVID-19 Diagnosis from Computed Tomography. *Intelligent Automation & Soft Computing* **34**:2, 1247-1264. [[Crossref](#)]
307. Savindu Herath Pathirannehelage, Yash Raj Shrestha, Georg von Krogh. 2022. Design principles for artificial intelligence-augmented decision making in nascent firms: An action design research study. *SSRN Electronic Journal* **31**. . [[Crossref](#)]
308. Md. Ifraham Iqbal, Md. Saddam Hossain Mukta, Ahmed Rafi Hasan, Salekul Islam. 2022. A Dynamic Weighted Tabular Method for Convolutional Neural Networks. *IEEE Access* **10**, 134183-134198. [[Crossref](#)]

309. Francesco Di Felice, Alessandro Borri, Maria Domenica Di Benedetto. 2022. Deep reinforcement learning for closed-loop blood glucose control: two approaches. *IFAC-PapersOnLine* 55:40, 115-120. [[Crossref](#)]
310. Hélio Castro, Filipe Costa, Luís Ferreira, Paulo Ávila, Goran D. Putnik, Manuela Cruz-Cunha. 2022. Data Science for Industry 4.0: A Literature Review on Open Design Approach. *Procedia Computer Science* 204, 877-884. [[Crossref](#)]
311. Fengsong Zhao, Guangmin Li, Changwei Zhao, Leijiao Ge, Zhicheng Gu, Yuhang Liu. An Power Quality Data Classification Method of Important Power User of Low-voltage Distribution Network based on Twin Networks 191-196. [[Crossref](#)]
312. Prodipto Bishnu Angon, Imrus Salehin, Sujit Mondal, Md. Mahbubur Rahman Khan, Mirza Nazim Uddin, Israt Jahan Lopa. A survey on Healthy Food Demand and Diseases Factors in Urban and Rural Area: Prospective on Bangladesh 316-321. [[Crossref](#)]
313. Shirish Patel, Philip J. Rhodes. Decentralized Storage for Scientific Data 3760-3769. [[Crossref](#)]
314. Carlos Garcia-Alvarado. Coxswain: Guided Data Analytics 3878-3882. [[Crossref](#)]
315. Royston Meriton, Rajinder Bhandal, Gary Graham, Anthony Brown. 2021. An examination of the generative mechanisms of value in big data-enabled supply chain management research. *International Journal of Production Research* 59:23, 7283-7310. [[Crossref](#)]
316. Josef Schwaiger, Timo Hammerl, Johannsen Florian, Susanne Leist. 2021. UR: SMART—A tool for analyzing social media content. *Information Systems and e-Business Management* 19:4, 1275-1320. [[Crossref](#)]
317. Haiko van der Voort, Sabine van Bulderen, Scott Cunningham, Marijn Janssen. 2021. Data science as knowledge creation a framework for synergies between data analysts and domain professionals. *Technological Forecasting and Social Change* 173, 121160. [[Crossref](#)]
318. Mert O. Gökalp, Kerem Kayabay, Ebru Gökalp, Altan Koçyiğit, P. Erhan Eren. 2021. Assessment of process capabilities in transition to a data-driven organisation: A multidisciplinary approach. *IET Software* 15:6, 376-390. [[Crossref](#)]
319. Constantine E. Kontokosta. 2021. Urban Informatics in the Science and Practice of Planning. *Journal of Planning Education and Research* 41:4, 382-395. [[Crossref](#)]
320. Fátima Figueiredo, Maria José Angélico Gonçalves, Sandrina Teixeira. 2021. Information Technology Adoption on Digital Marketing: A Literature Review. *Informatics* 8:4, 74. [[Crossref](#)]
321. Pikee Priya, N. R. Aluru. 2021. Accelerated design and discovery of perovskites with high conductivity for energy applications through machine learning. *npj Computational Materials* 7:1. . [[Crossref](#)]
322. Lingzi Hong, William Moen, Xinchun Yu, Jiangping Chen. 2021. The disciplinary research landscape of data science reflected in data science journals. *Information Discovery and Delivery* 49:4, 287-297. [[Crossref](#)]
323. Alida Cantor, Michael Kiparsky, Susan S. Hubbard, Rónán Kennedy, Lidia Cano Pecharroman, Kamyar Guivetchi, Gary Darling, Christina McCready, Roger Bales. 2021. Making a Water Data System Responsive to Information Needs of Decision Makers. *Frontiers in Climate* 3. . [[Crossref](#)]
324. Xiang Xu. 2021. Research prospect: data factor of production. *Journal of Internet and Digital Economics* 1:1, 64-71. [[Crossref](#)]
325. Aqeel Mohammed, Karesia Ramlal, Patrick Hosein, Nigel Henry. Marketing Channel Recommendations in Banking 60-65. [[Crossref](#)]
326. Buenos Aires de Carvalho, Denis Mayr Lima Martins, Fernando Buarque de Lima Neto. MapView: Exploring Datasets via Unsupervised View Recommendation 1-6. [[Crossref](#)]
327. Tamara KAYNAR, Öyküm Esra YİĞİT. 2021. Öznitelik Mühendisliği ile Makine Öğrenmesi Yöntemleri Kullanılarak BIST 100 Endeksi Değişiminin Tahminine Yönelik Bir Yaklaşım. *Journal of Yaşar University* 16:64, 1741-1762. [[Crossref](#)]
328. Abu Bashar, Mustafa Raza Rabbani, Shahnawaz Khan, Mahmood Asad Moh'd Ali. Data driven finance: A bibliometric review and scientific mapping 161-166. [[Crossref](#)]
329. Giovanni Maccarrone, Giacomo Morelli, Sara Spadaccini. 2021. GDP Forecasting: Machine Learning, Linear or Autoregression?. *Frontiers in Artificial Intelligence* 4. . [[Crossref](#)]
330. Erna Ruijter. 2021. Designing and implementing data collaboratives: A governance perspective. *Government Information Quarterly* 38:4, 101612. [[Crossref](#)]
331. Yi Zhang, Alan L. Porter, Scott Cunningham, Denise Chiavetta, Nils Newman. 2021. Parallel or Intersecting Lines? Intelligent Bibliometrics for Investigating the Involvement of Data Science in Policy Analysis. *IEEE Transactions on Engineering Management* 68:5, 1259-1271. [[Crossref](#)]
332. Jonathan HUKÉ, Patrick SIEGFRIED. 2021. Finance Methods in the Automotive Sector - Business Agility in the Age of Digital Disruption. *International Journal of Automotive Science and Technology* 5:3, 281-288. [[Crossref](#)]
333. Pooya Tabesh. 2021. Who's making the decisions? How managers can harness artificial intelligence and remain in charge. *Journal of Business Strategy* 30. . [[Crossref](#)]

334. Muhammed Çağrı AKSU, Ersin KARAMAN. 2021. Analysis of Turkish Sentiment Expressions About Touristic Sites Using Machine Learning. *Journal of Intelligent Systems: Theory and Applications* 4:2, 103-112. [[Crossref](#)]
335. Bibiana Giudice da Silva Cezar, Antônio Carlos Gastaud Maçada. 2021. Data literacy and the cognitive challenges of a data-rich business environment: an analysis of perceived data overload, technostress and their relationship to individual performance. *Aslib Journal of Information Management* 73:5, 618-638. [[Crossref](#)]
336. Harshal A. Sanghvi, Samvid B. Pandya, Pratyay Chattopadhyay, Riki H. Patel, Abhijit S. Pandya. Data Science for E-Healthcare, Entertainment and Finance 604-611. [[Crossref](#)]
337. John Bang Mathiasen, Henning de Haas. 2021. Digitalization of Shop Floor Management: In Blissful Ignorance of Superfluous Work. *Journal of Industrial Integration and Management* 06:03, 333-352. [[Crossref](#)]
338. Chi-Yo Huang, Chia-Lee Yang, Yi-Hao Hsiao. 2021. A Novel Framework for Mining Social Media Data Based on Text Mining, Topic Modeling, Random Forest, and DANP Methods. *Mathematics* 9:17, 2041. [[Crossref](#)]
339. Ahmad Ibrahim Aljumah, Mohammed T. Nuseir, Md. Mahmudul Alam. 2021. Organizational performance and capabilities to analyze big data: do the ambidexterity and business value of big data analytics matter?. *Business Process Management Journal* 27:4, 1088-1107. [[Crossref](#)]
340. Julian Martinsson, Massimo Panarotto, Michael Kokkolaras, Ola Isaksson. 2021. EXPLORING THE POTENTIAL OF DIGITAL TWIN-DRIVEN DESIGN OF AERO-ENGINE STRUCTURES. *Proceedings of the Design Society* 1, 1521-1528. [[Crossref](#)]
341. Hamza Saleem, Yongjun Li, Zulqurnain Ali, Muhammad Ayyoub, Yu Wang, Aqsa Mehreen. 2021. Big data use and its outcomes in supply chain context: the roles of information sharing and technological innovation. *Journal of Enterprise Information Management* 34:4, 1121-1143. [[Crossref](#)]
342. James Du, Carter Floyd, Amy C. H. Kim, Bradley J. Baker, Mikihiko Sato, Jeffrey D. James, Daniel C. Funk. 2021. To be or not to be: negotiating leisure constraints with technology and data analytics amid the COVID-19 pandemic. *Leisure Studies* 40:4, 561-574. [[Crossref](#)]
343. Georgios Tsoulfas. 2021. Bioinformatics Data Mining: Is the Pipette Mightier than the Scalpel?. *Journal of Investigative Surgery* 34:6, 670-671. [[Crossref](#)]
344. Mert Onuralp Gökalp, Ebru Gökalp, Kerem Kayabay, Altan Koçyiğit, P. Erhan Eren. 2021. Data-driven manufacturing: An assessment model for data science maturity. *Journal of Manufacturing Systems* 60, 527-546. [[Crossref](#)]
345. Vangelis Marinakis, Themistoklis Koutsellis, Alexandros Nikas, Haris Doukas. 2021. AI and Data Democratization for Intelligent Energy Management. *Energies* 14:14, 4341. [[Crossref](#)]
346. Huanmin Jiang, Wenhao Gui. 2021. Bayesian Inference for the Parameters of Kumaraswamy Distribution via Ranked Set Sampling. *Symmetry* 13:7, 1170. [[Crossref](#)]
347. Wenxuan Liu, Junhua Zhao, Dianhui Wang. 2021. Data mining for energy systems: Review and prospect. *WIREs Data Mining and Knowledge Discovery* 11:4. . [[Crossref](#)]
348. Sonia Ferns, Alope Phatak, Susan Benson, Nina Kumagai. 2021. Building employability capabilities in data science students: An interdisciplinary, industry-focused approach. *Teaching Statistics* 43:S1. . [[Crossref](#)]
349. Jürgen Bajorath. 2021. State-of-the-art of artificial intelligence in medicinal chemistry. *Future Science OA* 7:6. . [[Crossref](#)]
350. Luca Urciuoli, Juha Hintsa. 2021. Can digital ecosystems mitigate risks in sea transport operations? Estimating benefits for supply chain stakeholders. *Maritime Economics & Logistics* 23:2, 237-267. [[Crossref](#)]
351. Jacob L Cybulski, Rens Scheepers. 2021. Data science in organizations: Conceptualizing its breakthroughs and blind spots. *Journal of Information Technology* 36:2, 154-175. [[Crossref](#)]
352. Yonghwa Lee, Juhwan Lee, Keumjin Lee. 2021. A Study on the Visualization of an Airline's Fleet State Variation. *Journal of the Korean Society for Aviation and Aeronautics* 29:2, 84-93. [[Crossref](#)]
353. Pulkit Tiwari, P. Vigneswara Ilavarasan, Sushil Punia. 2021. Content analysis of literature on big data in smart cities. *Benchmarking: An International Journal* 28:5, 1837-1857. [[Crossref](#)]
354. Yu Shi, Joe Zhu, Vincent Charles. 2021. Data science and productivity: A bibliometric review of data science applications and approaches in productivity evaluations. *Journal of the Operational Research Society* 72:5, 975-988. [[Crossref](#)]
355. Chanbeom Bak, Abhishek Ghosh Roy, Hungsun Son. 2021. Quality prediction for aluminum diecasting process based on shallow neural network and data feature selection technique. *CIRP Journal of Manufacturing Science and Technology* 33, 327-338. [[Crossref](#)]
356. Mike Tissenbaum, David Weintrop, Nathan Holbert, Tamara Clegg. 2021. The case for alternative endpoints in computing education. *British Journal of Educational Technology* 52:3, 1164-1177. [[Crossref](#)]

357. Bo Wang, Feifan Liu, Lynette Deveaux, Arlene Ash, Samiran Gosh, Xiaoming Li, Elke Rundensteiner, Lesley Cottrell, Richard Adderley, Bonita Stanton. 2021. Adolescent HIV-related behavioural prediction using machine learning: a foundation for precision HIV prevention. *AIDS* 35:Supplement 1, S75-S84. [[Crossref](#)]
358. Godwin-Charles Ogbeide, Yao-Yi Fu, Amanda Kay Cecil. 2021. Are hospitality/tourism curricula ready for big data?. *Journal of Hospitality and Tourism Technology* 12:1, 112-123. [[Crossref](#)]
359. Lili Zhang, Jennifer Priestley, Joseph DeMaio, Sherry Ni, Xiaoguang Tian. 2021. Measuring Customer Similarity and Identifying Cross-Selling Products by Community Detection. *Big Data* 9:2, 132-143. [[Abstract](#)] [[Full Text](#)] [[PDF](#)] [[PDF Plus](#)]
360. Martijn Koot, Martijn R.K. Mes, Maria E. Iacob. 2021. A systematic literature review of supply chain decision making supported by the Internet of Things and Big Data Analytics. *Computers & Industrial Engineering* 154, 107076. [[Crossref](#)]
361. Katarzyna Biesialska, Xavier Franch, Victor Muntés-Mulero. 2021. Big Data analytics in Agile software development: A systematic mapping study. *Information and Software Technology* 132, 106448. [[Crossref](#)]
362. Yashvir Grewal, Sarah Goodwin, Tim Dwyer. Visualising Temporal Uncertainty: A Taxonomy and Call for Systematic Evaluation 41-45. [[Crossref](#)]
363. Yanjun Kong, Yadong Mei, Xianxun Wang, Yue Ben. 2021. Solution Selection from a Pareto Optimal Set of Multi-Objective Reservoir Operation via Clustering Operation Processes and Objective Values. *Water* 13:8, 1046. [[Crossref](#)]
364. Roxanne L. O'Brien, Matt W. O'Brien. 2021. CE: Nursing Orientation to Data Science and Machine Learning. *AJN, American Journal of Nursing* 121:4, 32-39. [[Crossref](#)]
365. Joyce de Souza Zanirato Maia, Ana Paula Arantes Bueno, João Ricardo Sato. 2021. Assessing the educational performance of different Brazilian school cycles using data science methods. *PLOS ONE* 16:3, e0248525. [[Crossref](#)]
366. Sergiy Smetana, Kemal Aganovic, Volker Heinz. 2021. Food Supply Chains as Cyber-Physical Systems: a Path for More Sustainable Personalized Nutrition. *Food Engineering Reviews* 13:1, 92-103. [[Crossref](#)]
367. Rijal Ramdani, Agustiyara, Eko Priyo Purnomo. 2021. Big Data Analysis of COVID-19 Mitigation Policy in Indonesia: Democratic, Elitist, and Artificial Intelligence. *IOP Conference Series: Earth and Environmental Science* 717:1, 012023. [[Crossref](#)]
368. Pritika Bahad, Preeti Saxena, Raj Kamal. 2021. Exploratory and Predictive Analytics of User Preferences from Kaggle LEGO-Toys Datasets Using Spark ML. *IOP Conference Series: Materials Science and Engineering* 1099:1, 012019. [[Crossref](#)]
369. Gilberto Ayala-Bastidas, Hector G. Ceballos, Francisco J. Cantu-Ortiz. 2021. A Systematic Review of Recommendations of Long-Term Strategies for Researchers Using Data Science Techniques. *Computers* 10:3, 35. [[Crossref](#)]
370. Fisnik Doko, Slobodan Kalajdziski, Igor Mishkovski. 2021. Credit Risk Model Based on Central Bank Credit Registry Data. *Journal of Risk and Financial Management* 14:3, 138. [[Crossref](#)]
371. Gang Shao, Jenny P. Quintana, Wei Zakharov, Senay Purzer, Eunhye Kim. 2021. Exploring potential roles of academic libraries in undergraduate data science education curriculum development. *The Journal of Academic Librarianship* 47:2, 102320. [[Crossref](#)]
372. Sanket Bhattarai, Ashish Gupta, Eiman Ali, Moez Ali, Mohamed Riad, Prakash Adhikari, Jihan A Mostafa. 2021. Can Big Data and Machine Learning Improve Our Understanding of Acute Respiratory Distress Syndrome?. *Cureus* . [[Crossref](#)]
373. Thomas Wendt. 2021. Organized Futures. On the Ambiguity of the Digital Absorption of Uncertainty. *Frontiers in Education* 6. . [[Crossref](#)]
374. Jan Jöhnk, Malte Weißert, Katrin Wyrтки. 2021. Ready or Not, AI Comes— An Interview Study of Organizational AI Readiness Factors. *Business & Information Systems Engineering* 63:1, 5-20. [[Crossref](#)]
375. Xi He, Rui Liu, Chimay J. Anumba. 2021. Data-Driven Insights on the Knowledge Gaps of Conceptual Cost Estimation Modeling. *Journal of Construction Engineering and Management* 147:2. . [[Crossref](#)]
376. Anamaria Crisan, Brittany Fiore-Gartland, Melanie Tory. 2021. Passing the Data Baton : A Retrospective Analysis on Data Science Work and Workers. *IEEE Transactions on Visualization and Computer Graphics* 27:2, 1860-1870. [[Crossref](#)]
377. Teresa M. Harrison, Theresa A. Pardo. 2021. Data, Politics and Public Health. *Digital Government: Research and Practice* 2:1, 1-8. [[Crossref](#)]
378. Gerhard Reichmann, Christian Schlögl, Stefan Thalmann. 2021. Das Institut für Informationswissenschaft an der Universität Graz: 1987 – 2020. *Information - Wissenschaft & Praxis* 72:1, 1-9. [[Crossref](#)]
379. Gloria Phillips-Wren, Anna Esposito, Lakhmi C. Jain. Introduction to Big Data and Data Science: Methods and Applications 1-11. [[Crossref](#)]
380. G. Ortegon-Cortazar, J. Samper-Zapater, F. Garcia-Sanchez, O. Garcia-Bedoya, C. O. Diaz. The Application of an Information System in the E-Government of Colombia to Improve Service to Citizens 615-625. [[Crossref](#)]
381. Robert M. Cronin, Holly Jimison, Kevin B. Johnson. Personal Health Informatics 363-389. [[Crossref](#)]

382. Arezou Harraf, Hasan Ghura. Potentials of Artificial Intelligence for Business Performance 99-109. [\[Crossref\]](#)
383. M. Amin Yazdi, Marius Politze. Reverse Engineering: The University Distributed Services 223-238. [\[Crossref\]](#)
384. Ravi Lourdasamy, Xavierlal J. Mattam. Adaptation and Evolution of Decision Support Systems—A Typological Survey 81-94. [\[Crossref\]](#)
385. Veronica Israel-Turim, Josep-Lluís Micó-Sanz. Who Are the Most Followed Following? A Data Analysis Case Study of the Accounts the Top Media Directors from Spain Started Following on Twitter 219-229. [\[Crossref\]](#)
386. Ting Xiao, Mark V. Albert. Big Data in Medical AI: How Larger Data Sets Lead to Robust, Automated Learning for Medicine 11-25. [\[Crossref\]](#)
387. André Ippolito, Augusto Cezar Garcia Lozano. Sammon Mapping-Based Gradient Boosted Trees for Tax Crime Prediction in the City of São Paulo 293-316. [\[Crossref\]](#)
388. Felix Soldner, Leonie Maria Tanczer, Daniel Hammocks, Isabel Lopez-Neira, Shane D. Johnson. Using Machine Learning Methods to Study Technology-Facilitated Abuse: Evidence from the Analysis of UK Crimestoppers' Text Data 481-503. [\[Crossref\]](#)
389. Aiswarya Raj Munappy, Jan Bosch, Helena Homström Olsson. On the Trade-off Between Robustness and Complexity in Data Pipelines 401-415. [\[Crossref\]](#)
390. Stephan Weinert, Klaus P. Stulle. Digitale Kompetenzen im Kontext von Leadership 1-13. [\[Crossref\]](#)
391. Daniel Badura, Alexander Ossa, Michael Schulz. Self-Service Data Science – Vergleich von Plattformen zum Aufbau von Entscheidungsbäumen 215-237. [\[Crossref\]](#)
392. Jinhuan Wang. Study on Evolution of Urban Carrying Capacity of Resources and Environment Under the Background of Big Data 918-930. [\[Crossref\]](#)
393. Arnab Santra, Aniket Mitra, Abhirup Sinha, Amit Kumar Das. Prediction of Most Valuable Bowlers of Indian Premier League (IPL) 211-223. [\[Crossref\]](#)
394. S. Krishnaraj, M. Gomathi Prabha, M. Yuvaraja. Smart Maintenance and Analytics for Indian SMEs 695-706. [\[Crossref\]](#)
395. Hua Song. Intelligent Supply Chain and Tech-Enabled Supply Chain Finance 89-148. [\[Crossref\]](#)
396. Xingxing He, Yang Xu, Jun Liu, Yingfang Li. 2021. α -Paramodulation method for a lattice-valued logic $L_nF(X)$ with equality. *Soft Computing* 25:1, 251-261. [\[Crossref\]](#)
397. Zeinab Arjmandiasl, Narges Manouchehri, Nizar Bouguila, Jamal Bentahar. Variational learning of finite shifted scaled Dirichlet mixture models 175-204. [\[Crossref\]](#)
398. Hyunwoong Ko, Paul Witherell, Yan Lu, Samyeon Kim, David W. Rosen. 2021. Machine learning and knowledge graph based design rule construction for additive manufacturing. *Additive Manufacturing* 37, 101620. [\[Crossref\]](#)
399. Justas Streimikis, Liudmila Kortencko, Marina Panova, Mikhail Voronov. 2021. Development of a smart city information system. *E3S Web of Conferences* 301, 05002. [\[Crossref\]](#)
400. Chen Chen, Yao Zhang, Xiaoyong Shi. 2021. Research on the Construction of Ocean Big Data Platform and Disaster Prevention and Mitigation Application Based on Internet of Things. *IOP Conference Series: Earth and Environmental Science* 632:2, 022009. [\[Crossref\]](#)
401. Luis Bote-Curiel, Sergio Ruiz-Llorente, Sergio Munoz-Romero, Monica Yague-Fernandez, Arantzazu Barquin, Jesus Garcia-Donas, Jose Luis Rojo-Alvarez. 2021. A Resampling Univariate Analysis Approach to Ovarian Cancer from Clinical and Genetic Data. *IEEE Access* 1-1. [\[Crossref\]](#)
402. Jukka Riekkii, Aarne Mammela. 2021. Research and Education Towards Smart and Sustainable World. *IEEE Access* 9, 53156-53177. [\[Crossref\]](#)
403. Alfonso Castro, Victor A. Villagra, Paula Garcia, Diego Rivera, David Toledo. 2021. An Ontological-Based Model to Data Governance for Big Data. *IEEE Access* 9, 109943-109959. [\[Crossref\]](#)
404. Harvey C Turner, David Atkinson. 2021. Strategic Decision Making: The Effects of Big Data. *International Journal of Operations Management* 1:2, 38-45. [\[Crossref\]](#)
405. Haihua Liu, Shan Huang, Peng Wang, Zejun Li. 2021. A review of data mining methods in financial markets. *Data Science in Finance and Economics* 1:4, 362-392. [\[Crossref\]](#)
406. Josiline Phiri Chigwada, Justice Kasiroori. Awareness of Big Data Usage and Applications Among Librarians in Zimbabwe 22-34. [\[Crossref\]](#)
407. Pankaj Pathak, Samaya Pillai Iyengar, Minal Abhyankar. A Survey on Tools for Data Analytics and Data Science 28-49. [\[Crossref\]](#)
408. Omar F. El-Gayar, Martinson Q. Ofori. Disrupting Agriculture 771-812. [\[Crossref\]](#)

409. Miftachul Huda, Ulfatmi, Muhammad Ja'far Luthfi, Kamarul Azmi Jasmi, Bushrah Basiron, Mohd Ismail Mustari, Ajmain Safar, Wan Hassan Wan Embong, Ahmad Marzuki Mohamad, Ahmad Kilani Mohamed. Adaptive Online Learning Technology 639-663. [[Crossref](#)]
410. Raphael de Freitas Saldanha, Diego Ricardo Xavier, Mônica de Avelar Figueiredo Mafra Magalhães, Paulo Roberto Borges de Souza Junior, Marcel de Moraes Pedroso, Christovam Barcellos. 14 - MonitoraCovid-19: informação e disseminação de indicadores em uma pesquisa multidisciplinar 229-249. [[Crossref](#)]
411. Jeffrey C. Oliver, Torbet McNeil. 2021. Undergraduate data science degrees emphasize computer science and statistics but fall short in ethics training and domain-specific context. *PeerJ Computer Science* 7, e441. [[Crossref](#)]
412. Amina Ziegenbein, Joachim Metternich. 2021. Data-based quality analysis in machining production: Influence of data pre-processing on the results of machine learning models. *Procedia CIRP* 104, 869-874. [[Crossref](#)]
413. Alex Singleton, Daniel Arribas-Bel. 2021. Geographic Data Science. *Geographical Analysis* 53:1, 61-75. [[Crossref](#)]
414. Somayeh Dodge. 2021. A Data Science Framework for Movement. *Geographical Analysis* 53:1, 92-112. [[Crossref](#)]
415. Yu-Hua Yan, Tsair-Wei Chien. 2021. The use of forest plot to identify article similarity and differences in characteristics between journals using medical subject headings terms. *Medicine* 100:6, e24610. [[Crossref](#)]
416. Constantine E. Kontokosta, Boyeong Hong. 2021. Bias in smart city governance: How socio-spatial disparities in 311 complaint behavior impact the fairness of data-driven decisions. *Sustainable Cities and Society* 64, 102503. [[Crossref](#)]
417. Nureni Olawale Adeboye, Peter Osuolale Popoola, Oluwatobi Nurudeen Ogunnusi. 2020. Data science skills: Building partnership for efficient school curriculum delivery in Africa. *Statistical Journal of the IAOS* 36, 49-62. [[Crossref](#)]
418. M. D. Cruz, J. J. U. Keh, R. G. Deticio, C. V. T. Tan, J. A. C. Jose, E. Sybingco, E. P. Dadios. Visual-based People Counting and Profiling System for Use in Retail Data Analytics 200-204. [[Crossref](#)]
419. Jeng-Fung Hung, Chun-Yen Tsai. 2020. THE EFFECTS OF A VIRTUAL LABORATORY AND META-COGNITIVE SCAFFOLDING ON STUDENTS' DATA MODELING COMPETENCES. *Journal of Baltic Science Education* 19:6, 923-939. [[Crossref](#)]
420. Riccardo Rialti, Giacomo Marzi, Andrea Caputo, Kayode Abraham Mayah. 2020. Achieving strategic flexibility in the era of big data. *Management Decision* 58:8, 1585-1600. [[Crossref](#)]
421. Jefferson James Keh, Meygen Cruz, Ramiel Deticio, Carl Vincent Tan, John Anthony Jose, Elmer Dadios, Alexis Fillone. Video-Based Gender Profiling on Challenging Camera Viewpoint for Restaurant Data Analytics 1-5. [[Crossref](#)]
422. Piotr Tarka, Elżbieta Jędrych. 2020. On the Unstructured Big Data Analytical Methods in Firms: Conceptual Model, Measurement, and Perception. *Big Data* 8:6, 478-500. [[Abstract](#)] [[Full Text](#)] [[PDF](#)] [[PDF Plus](#)]
423. JingMing Zhang, ShuZhen Zhu, Wei Yan, ZhiPeng Li. 2020. RETRACTED ARTICLE: The construction and simulation of internet financial product diffusion model based on complex network and consumer decision-making mechanism. *Information Systems and e-Business Management* 18:4, 545-555. [[Crossref](#)]
424. Hamed M. Zolbanin, Dursun Delen, Durand Crosby, David Wright. 2020. A Predictive Analytics-Based Decision Support System for Drug Courts. *Information Systems Frontiers* 22:6, 1323-1342. [[Crossref](#)]
425. Saqib Shamim, Jing Zeng, Zaheer Khan, Najam Ul Zia. 2020. Big data analytics capability and decision making performance in emerging market firms: The role of contractual and relational governance mechanisms. *Technological Forecasting and Social Change* 161, 120315. [[Crossref](#)]
426. Andrea Sestino, Maria Irene Prete, Luigi Piper, Gianluigi Guido. 2020. Internet of Things and Big Data as enablers for business digitalization strategies. *Technovation* 98, 102173. [[Crossref](#)]
427. Alexander Stevens, Peter Deruyck, Ziboud Van Veldhoven, Jan Vanthienen. Explainability and Fairness in Machine Learning: Improve Fair End-to-end Lending for Kiva 1241-1248. [[Crossref](#)]
428. Emmanuelle Vaast. 2020. A Seat at the Table and a Room of Their Own: Interconnected processes of social media use at the intersection of gender and occupation. *Organization Studies* 41:12, 1673-1695. [[Crossref](#)]
429. Muhammad Qasim Shabbir, Syed Babar Waheed Gardezi. 2020. Application of big data analytics and organizational performance: the mediating role of knowledge management practices. *Journal of Big Data* 7:1. . [[Crossref](#)]
430. Paul Brous, Marijn Janssen. 2020. Trusted Decision-Making: Data Governance for Creating Trust in Data Science Decision Outcomes. *Administrative Sciences* 10:4, 81. [[Crossref](#)]
431. Ross Towe, Graham Dean, Liz Edwards, Vatsala Nundloll, Gordon Blair, Rob Lamb, Barry Hankin, Susan Manson. 2020. Rethinking data-driven decision support in flood risk management for a big data age. *Journal of Flood Risk Management* 13:4. . [[Crossref](#)]

432. Md. Omar Faruq, A H M Sarowar Sattar. An Approach to Mitigate Composition Attacks on Privacy in Non-coordinated Environments 123-128. [[Crossref](#)]
433. Bright Keswani, Ambarish G. Mohapatra, Poonam Keswani, Ashish Khanna, Deepak Gupta, Joel Rodrigues. 2020. Improving weather dependent zone specific irrigation control scheme in IoT and big data enabled self driven precision agriculture mechanism. *Enterprise Information Systems* 14:9-10, 1494-1515. [[Crossref](#)]
434. Drashti Pathak, Aparna S. Varde, Gerard De Melo, Clement Alo. 2020. Hydroinformatics and the web. *ACM SIGWEB Newsletter* 2020:Autumn, 1-17. [[Crossref](#)]
435. Irma Permata Sari, M. Syafiq Roikhan Maulana, Erna Septiandini, Adnan Rais Rusmantoro. Web-based Educational Administration System for Student Affairs: An Initial Step to Develop Data-Driven Decision Making 142-146. [[Crossref](#)]
436. Lilia Sfaxi, Mohamed Mehdi Ben Aissa. 2020. DECIDE: An Agile event-and-data driven design methodology for decisional Big Data projects. *Data & Knowledge Engineering* 130, 101862. [[Crossref](#)]
437. Lingqiang Kong, Zhifeng Liu, Jianguo Wu. 2020. A systematic review of big data-based urban sustainability research: State-of-the-science and future directions. *Journal of Cleaner Production* 273, 123142. [[Crossref](#)]
438. Ahmet ALBAYRAK. 2020. Doğal Dil İşleme Teknikleri Kullanılarak Disiplinler Arası Lisansüstü Ders İçeriği Hazırlanması. *Bilişim Teknolojileri Dergisi* 13:4, 373-383. [[Crossref](#)]
439. . References 243-256. [[Crossref](#)]
440. Alfred C Ma. 2020. Making Data Reports Useful: From Descriptive to Predictive. *Cureus* . [[Crossref](#)]
441. Malin Song, Qianjiao Xie, Kim Hua Tan, Jianlin Wang. 2020. A fair distribution and transfer mechanism of forest tourism benefits in China. *China Economic Review* 63, 101542. [[Crossref](#)]
442. Orlando Troisi, Gennaro Maione, Mara Grimaldi, Francesca Loia. 2020. Growth hacking: Insights on data-driven decision-making from three firms. *Industrial Marketing Management* 90, 538-557. [[Crossref](#)]
443. Giovanni Cammarota, Gianluca Ianiro, Anna Ahern, Carmine Carbone, Andriy Temko, Marcus J. Claesson, Antonio Gasbarrini, Giampaolo Tortora. 2020. Gut microbiome, big data and machine learning to promote precision medicine for cancer. *Nature Reviews Gastroenterology & Hepatology* 17:10, 635-648. [[Crossref](#)]
444. Anamaria Crisan, Brittany Fiore-Gartland, Melanie Tory. Passing the Data Baton: A Retrospective Analysis on Data Science Work and Workers 1-11. [[Crossref](#)]
445. Doru Rotovei, Daniela Zaharie, Viorel Negru, Daniel Pop. A decision support system for B2B CRM Systems using belief rules 131-138. [[Crossref](#)]
446. Phillip R. Owens, Minerva J. Dorantes, Bryan A. Fuentes, Zamir Libohova, Axel Schmidt. 2020. Taking digital soil mapping to the field: Lessons learned from the Water Smart Agriculture soil mapping project in Central America. *Geoderma Regional* 22, e00285. [[Crossref](#)]
447. Paula Munoz-Lago, Nicola Usula, Emilia Parada-Cabaleiro, Alvaro Torrente. Visualising the Structure of 18 th Century Operas: A Multidisciplinary Data Science Approach 530-536. [[Crossref](#)]
448. David Juárez-Varón, Victoria Tur-Viñes, Alejandro Rabasa-Dolado, Kristina Polotskaya. 2020. An Adaptive Machine Learning Methodology Applied to Neuromarketing Analysis: Prediction of Consumer Behaviour Regarding the Key Elements of the Packaging Design of an Educational Toy. *Social Sciences* 9:9, 162. [[Crossref](#)]
449. Krishna Subramanian, Nur Hamdan, Jan Borchers. Casual Notebooks and Rigid Scripts: Understanding Data Science Programming 1-5. [[Crossref](#)]
450. Paula Pereira, Jacome Cunha, Joao Paulo Fernandes. On Understanding Data Scientists 1-5. [[Crossref](#)]
451. Mei Li, Ying Wu, Yi He, Shuai Huang, Anand Nair. 2020. Sparse Inverse Covariance Estimation: A Data Mining Technique to Unravel Holistic Patterns among Business Practices in Firms. *Decision Sciences* 51:4, 1046-1073. [[Crossref](#)]
452. Rodrigo Hipólito Roza. 2020. Revisitando a Teoria da Criação do Conhecimento Organizacional. *Revista Interamericana de Bibliotecología* 43:3, eRv2. [[Crossref](#)]
453. Ismail Bile Hassan, Jigang Liu. A Comparative Study of the Academic Programs between Informatics/BioInformatics and Data Science in the U.S 165-171. [[Crossref](#)]
454. Alexa Corral-Garza, Isabel Echevarría-Garza, Lorena Flores-Bulnes, Karla Nava-Aguirre. 2020. BIG DATA ANALYTICS COMO RECURSO INTANGIBLE GENERADOR DE VALOR EN LA INDUSTRIA DEL RETAIL: CASO DE ESTUDIO OXXO. *Vinculatégica EFAN* 6:1, 158-173. [[Crossref](#)]
455. Anne L. Washington. 2020. Uncertain risk: assessing open data signals. *Transforming Government: People, Process and Policy* 14:4, 623-637. [[Crossref](#)]

456. Mahyuddin K M Nasution, Opim Salim Sitompul, Erna Budhiarti Nababan. 2020. Data Science. *Journal of Physics: Conference Series* **1566**:1, 012034. [[Crossref](#)]
457. Wolfgang Badewitz, Simon Kloker, Christof Weinhardt. The Data Provision Game: Researching Revenue Sharing in Collaborative Data Networks 191-200. [[Crossref](#)]
458. Songyan Xue, Yi Ma, Rahim Tafazolli. An Orthogonal-SGD Based Learning Approach for MIMO Detection under Multiple Channel Models 1-6. [[Crossref](#)]
459. Catherine Chen, Haoqiang Jiang. 2020. Important Skills for Data Scientists in China: Two Delphi Studies. *Journal of Computer Information Systems* **60**:3, 287-296. [[Crossref](#)]
460. Marva Foster, Zarin Tasnim. 2020. Data Science and Graduate Nursing Education. *Clinical Nurse Specialist* **34**:3, 124-131. [[Crossref](#)]
461. Mauricius Munhoz de Medeiros, Antonio Carlos Gastaud Maçada, José Carlos da Silva Freitas Junior. 2020. The effect of data strategy on competitive advantage. *The Bottom Line* **33**:2, 201-216. [[Crossref](#)]
462. Iván García-Magariño, Carlos Medrano, Jorge Delgado. 2020. Estimation of missing prices in real-estate market agent-based simulations with machine learning and dimensionality reduction methods. *Neural Computing and Applications* **32**:7, 2665-2682. [[Crossref](#)]
463. Shadi A. Aljawarneh. 2020. Reviewing and exploring innovative ubiquitous learning tools in higher education. *Journal of Computing in Higher Education* **32**:1, 57-73. [[Crossref](#)]
464. Andrej Zwitter, Jilles Hazenberg. 2020. Decentralized Network Governance: Blockchain Technology and the Future of Regulation. *Frontiers in Blockchain* **3**. . [[Crossref](#)]
465. Daphne R. Raban, Avishag Gordon. 2020. The evolution of data science and big data research: A bibliometric analysis. *Scientometrics* **122**:3, 1563-1581. [[Crossref](#)]
466. Qian Yi, Shiquan Xiong, Biao Wang, Shuping Yi. 2020. Identification of trusted interactive behavior based on mouse behavior considering web User's emotions. *International Journal of Industrial Ergonomics* **76**, 102903. [[Crossref](#)]
467. Giovanni Ortegon-Cortazar, Olmer Garcia-Bedoya, Jose Samper, Francisco Garcia. A Case of Use in the Management of Administrative Processes, Control of Procedures and Services to Citizens In E-Government 36-40. [[Crossref](#)]
468. Ao Huang, Yanzhu Huo, Juan Yang, Huazhi Gu, Guangqiang Li. 2020. Computational Modeling and Prediction on Viscosity of Slags by Big Data Mining. *Minerals* **10**:3, 257. [[Crossref](#)]
469. G. Rejikumar, A. Aswathy Asokan, V. Raja Sreedharan. 2020. Impact of data-driven decision-making in Lean Six Sigma: an empirical analysis. *Total Quality Management & Business Excellence* **31**:3-4, 279-296. [[Crossref](#)]
470. David Contreras, Maria Salamó. 2020. Data-driven decision making in critique-based recommenders: from a critique to social media data. *Journal of Intelligent Information Systems* **54**:1, 23-44. [[Crossref](#)]
471. Zhiying Wang, Nianxin Wang, Xiang Su, Shilun Ge. 2020. An empirical study on business analytics affordances enhancing the management of cloud computing data security. *International Journal of Information Management* **50**, 387-394. [[Crossref](#)]
472. İnanç KABASAKAL. 2020. Customer Segmentation Based On Recency Frequency Monetary Model: A Case Study in E-Retailing. *Bilişim Teknolojileri Dergisi* **13**:1, 47-56. [[Crossref](#)]
473. Christian Dremel, Emanuel Stoeckli, Jochen Wulf. 2020. Management of analytics-as-a-service - results from an action design research project. *Journal of Business Analytics* **3**:1, 1-16. [[Crossref](#)]
474. Christophe Kolski, Guy André Boy, Guy Melançon, Magalie Ochs, Jean Vanderdonckt. Cross-Fertilisation Between Human-Computer Interaction and Artificial Intelligence 365-388. [[Crossref](#)]
475. Paul Francisco Baldeon Egas, Miguel Alfredo Gaibor Saltos, Renato Toasa. Application of Data Mining and Data Visualization in Strategic Management Data at Israel Technological University of Ecuador 419-431. [[Crossref](#)]
476. Silvana De Gyves Avila, Patricia Ortegon Cano, Ayrton Mondragon Mejia, Ismael Solis Moreno, Arianne Navarro Lepe. A Data Driven Platform for Improving Performance Assessment of Software Defined Storage Solutions 266-275. [[Crossref](#)]
477. Joseph Aversa, Tony Hernandez, Sean Doherty. Spatial Big Data and Business Location Decision-Making: Opportunities and Challenges 205-224. [[Crossref](#)]
478. Riccardo Rialti, Giacomo Marzi. From Big Data to Performance: The Importance of Ambidexterity, Agility and BDA Integration in Business Processes—A Theory-Based Framework 39-67. [[Crossref](#)]
479. Leonard Heilig, Robert Stahlbock, Stefan Voß. From Digitalization to Data-Driven Decision Making in Container Terminals 125-154. [[Crossref](#)]

480. Xuyan Wang, Xi Zhang, Hui Xiong, Patricia Ordóñez de Pablos. KM 3.0: Knowledge Management Computing Under Digital Economy 207-217. [[Crossref](#)]
481. Alex Rabasa, Ciara Heavin. An Introduction to Data Science and Its Applications 57-81. [[Crossref](#)]
482. Sherif Sakr. Introduction 1-16. [[Crossref](#)]
483. Sherif Sakr. General-Purpose Big Data Processing Systems 17-43. [[Crossref](#)]
484. Sherif Sakr. Large-Scale Processing Systems of Structured Data 45-58. [[Crossref](#)]
485. Sherif Sakr. Large-Scale Graph Processing Systems 59-93. [[Crossref](#)]
486. Sherif Sakr. Large-Scale Stream Processing Systems 95-115. [[Crossref](#)]
487. Sherif Sakr. Large-Scale Machine/Deep Learning Frameworks 117-126. [[Crossref](#)]
488. Sherif Sakr. Conclusions and Outlook 127-133. [[Crossref](#)]
489. Kenneth Sawka. The Use and Misuse of Statistics 95-110. [[Crossref](#)]
490. Rongzhen Zhao, Yaochun Wu, Tianjing He. A Kind of Faults Knowledge Discovery Pattern by Means of Rough Set Theory 463-472. [[Crossref](#)]
491. Frederico B. Tiggemann, Bryan G. Pernambuco, Giancarlo Lucca, Eduardo N. Borges, Helida Santos, Graçaliz P. Dimuro, José A. Sanz, Humberto Bustince. An Alternative to Power Measure for Fuzzy Rule-Based Classification Systems 231-244. [[Crossref](#)]
492. Mohamed Mehdi Ben Aissa, Lilia Sfaxi, Riadh Robbana. DECIDE: A New Decisional Big Data Methodology for a Better Data Governance 63-78. [[Crossref](#)]
493. Guanze Peng, Tao Zhang, Quanyan Zhu. A Data-Driven Distributionally Robust Game Using Wasserstein Distance 405-421. [[Crossref](#)]
494. Mario José Diván, María Laura Sánchez Reynoso. A Load-Shedding Technique Based on the Measurement Project Definition 1027-1033. [[Crossref](#)]
495. Yu Li. 2020. Towards fast prototyping of cloud-based environmental decision support systems for environmental scientists using R Shiny and Docker. *Environmental Modelling & Software* **132**, 104797. [[Crossref](#)]
496. Rafael Diaz, Katherine Smith, Rafael Landaeta, Antonio Padovano. 2020. Shipbuilding Supply Chain Framework and Digital Transformation: A Project Portfolios Risk Evaluation. *Procedia Manufacturing* **42**, 173-180. [[Crossref](#)]
497. Andrej Zwitter, Jilles Hazenberg. 2020. Decentralized Network Governance: Blockchain Technology and the Future of Regulation. *SSRN Electronic Journal* . [[Crossref](#)]
498. Omar F. El-Gayar, Martinson Q. Ofori. Disrupting Agriculture 174-215. [[Crossref](#)]
499. Naciye Güliz Uğur, Aykut Hamit Turan. Understanding Big Data 1-29. [[Crossref](#)]
500. Murray Eugene Jennex. A Re-Examination and Re-Specification of the Jennex Olfman Knowledge Management Success Model 1-29. [[Crossref](#)]
501. Boluwaji A. Akinnuwesi, Stephen G. Fashoto, Andile S. Metfula, Adetutu N. Akinnuwesi. Experimental Application of Machine Learning on Financial Inclusion Data for Governance in Eswatini 414-425. [[Crossref](#)]
502. Paul Brous, Marijn Janssen, Rutger Krans. Data Governance as Success Factor for Data Science 431-442. [[Crossref](#)]
503. Vinicius Luiz Ferraz Minatogawa, Matheus Munhoz Vieira Franco, Izabela Simon Rampasso, Rosley Anholon, Ruy Quadros, Orlando Durán, Antonio Batocchio. 2020. Operationalizing Business Model Innovation through Big Data Analytics for Sustainable Organizations. *Sustainability* **12**:1, 277. [[Crossref](#)]
504. . Interpreting Bergson **24**, . [[Crossref](#)]
505. Abhishek Behl, Pankaj Dutta, Stefan Lessmann, Yogesh K. Dwivedi, Samarjit Kar. 2019. A conceptual framework for the adoption of big data analytics by e-commerce startups: a case-based approach. *Information Systems and e-Business Management* **17**:2-4, 285-318. [[Crossref](#)]
506. Riccardo Rialti, Lamberto Zollo, Alberto Ferraris, Ilan Alon. 2019. Big data analytics capabilities and performance: Evidence from a moderated multi-mediation model. *Technological Forecasting and Social Change* **149**, 119781. [[Crossref](#)]
507. Fábio C. P. Navarro, Hussein Mohsen, Chengfei Yan, Shantao Li, Mengting Gu, William Meyerson, Mark Gerstein. 2019. Genomics and data science: an application within an umbrella. *Genome Biology* **20**:1. . [[Crossref](#)]
508. Jose Ramon Saura, Ana Reyes-Menendez, Ferrão Filipe. 2019. Comparing Data-Driven Methods for Extracting Knowledge from User Generated Content. *Journal of Open Innovation: Technology, Market, and Complexity* **5**:4, 74. [[Crossref](#)]
509. Malka N. Halgamuge, Devra Davis. 2019. Lessons learned from the application of machine learning to studies on plant response to radio-frequency. *Environmental Research* **178**, 108634. [[Crossref](#)]

510. Radiyya Dwisaputra, Tricya Esterina Widagdo. Query Handling of Temporal Data with Different Calendar Systems and Time Granularities 1-6. [[Crossref](#)]
511. Michael Kaufmann, Patrick Siegfried, Lukas Huck, Jürg Stettler. 2019. Analysis of Tourism Hotspot Behaviour Based on Geolocated Travel Blog Data: The Case of Qyer. *ISPRS International Journal of Geo-Information* 8:11, 493. [[Crossref](#)]
512. Sirje Virkus, Emmanouel Garoufallou. 2019. Data science from a library and information science perspective. *Data Technologies and Applications* 53:4, 422-441. [[Crossref](#)]
513. Lynn C. Miller, Sonia Jawaid Shaikh, David C. Jeong, Liyuan Wang, Traci K. Gillig, Carlos G. Godoy, Paul R. Appleby, Charisse L. Corsbie-Massay, Stacy Marsella, John L. Christensen, Stephen J. Read. 2019. Causal Inference in Generalizable Environments: Systematic Representative Design. *Psychological Inquiry* 30:4, 173-202. [[Crossref](#)]
514. Daniel Badura, Michael Schulz. 2019. Kleine Barrieren für große Analysen – Eine Untersuchung der Eignung aktueller Plattformen für Self-Service Data Mining. *HMD Praxis der Wirtschaftsinformatik* 56:5, 1053-1067. [[Crossref](#)]
515. Fatma Chiheb, Fatima Boumahdi, Hafida Bouarfa. 2019. A Conceptual Model for Describing the Integration of Decision Aspect into Big Data. *International Journal of Information System Modeling and Design* 10:4, 1-23. [[Crossref](#)]
516. JooSeok Oh, Timothy Paul Connerton, Hyun-Jung Kim. 2019. The Rediscovery of Brand Experience Dimensions with Big Data Analysis: Building for a Sustainable Brand. *Sustainability* 11:19, 5438. [[Crossref](#)]
517. William Villegas-Ch, Jhoann Molina-Enriquez, Carlos Chicaiza-Tamayo, Iván Ortiz-Garcés, Sergio Luján-Mora. 2019. Application of a Big Data Framework for Data Monitoring on a Smart Campus. *Sustainability* 11:20, 5552. [[Crossref](#)]
518. Lorenzo Ardito, Veronica Scutto, Manlio Del Giudice, Antonio Messeni Petruzzelli. 2019. A bibliometric analysis of research on Big Data analytics for business and management. *Management Decision* 57:8, 1993-2009. [[Crossref](#)]
519. Gabriele Santoro, Fabio Fiano, Bernardo Bertoldi, Francesco Ciampi. 2019. Big data for business management in the retail industry. *Management Decision* 57:8, 1980-1992. [[Crossref](#)]
520. Francesco Caputo, Valentina Cillo, Elena Candelo, Yipeng Liu. 2019. Innovating through digital revolution. *Management Decision* 57:8, 2032-2051. [[Crossref](#)]
521. Hui Cai, Yanmin Zhu, Jie Li, Jiadi Yu. 2019. Double Auction for a Data Trading Market with Preferences and Conflicts of Interest. *The Computer Journal* 62:10, 1490-1504. [[Crossref](#)]
522. Kathleen Desveaud, Jean Mallol. 2019. Stakes and challenges in building and analyzing a client database: A case study of a French company. *Management & Data Science* . [[Crossref](#)]
523. Valeriia Boldosova, Severi Luoto. 2019. Storytelling, business analytics and big data interpretation. *Management Research Review* 43:2, 204-222. [[Crossref](#)]
524. Gordon S. Blair, Peter Henrys, Amber Leeson, John Watkins, Emma Eastoe, Susan Jarvis, Paul J. Young. 2019. Data Science of the Natural Environment: A Research Roadmap. *Frontiers in Environmental Science* 7. . [[Crossref](#)]
525. Gisela Fernandes, Filipe Portela, Manuel Filipe Santos. Towards the Development of a Data Science Modular Solution 96-103. [[Crossref](#)]
526. Zhuoxun Li, Xinying Ma, Wenjie Chen, Ningyuan Kuang, Bo Zhang. Neural Network Enhanced Analog Beam Selection Scheme for Terahertz Systems 136-141. [[Crossref](#)]
527. Pawat Tansurat, Nathasit Gerdri. Extended Techniques to Enhance Technology Roadmapping: Research Opportunities and Challenges 1-8. [[Crossref](#)]
528. John R. Walkup, Roger A. Key, Patrick R. M. Talbot, Michael A. Walkup. 2019. Data-driven decision making in an introductory physics lab. *American Journal of Physics* 87:8, 654-659. [[Crossref](#)]
529. José-Antonio Marín-Marín, Jesús López-Belmonte, Juan-Miguel Fernández-Campoy, José-María Romero-Rodríguez. 2019. Big Data in Education. A Bibliometric Review. *Social Sciences* 8:8, 223. [[Crossref](#)]
530. Roberto Espinosa, Diego García-Saiz, Marta Zorrilla, José Jacobo Zubcoff, Jose-Norberto Mazón. 2019. S3Mining: A model-driven engineering approach for supporting novice data miners in selecting suitable classifiers. *Computer Standards & Interfaces* 65, 143-158. [[Crossref](#)]
531. Mikko Hänninen, Lasse Mitronen, Stephen K. Kwan. 2019. Multi-sided marketplaces and the transformation of retail: A service systems perspective. *Journal of Retailing and Consumer Services* 49, 380-388. [[Crossref](#)]
532. Hieu Nguyen, Meeta Kalra, Muhammad Azam, Nizar Bouguila. Data Clustering Using Online Variational Learning of Finite Scaled Dirichlet Mixture Models 267-274. [[Crossref](#)]
533. Jennifer Patterson, George Foxcroft. 2019. Gilt Management for Fertility and Longevity. *Animals* 9:7, 434. [[Crossref](#)]

534. Kevin Daniel André Carillo, Nadine Galy, Cameron Guthrie, Anne Vanhems. 2019. How to turn managers into data-driven decision makers. *Business Process Management Journal* **25**:3, 553-578. [[Crossref](#)]
535. Cristina Orsolin Klingenberg, Marco Antônio Viana Borges, José Antônio Valle Antunes Jr. 2019. Industry 4.0 as a data-driven paradigm: a systematic literature review on technologies. *Journal of Manufacturing Technology Management* **32**:3, 570-592. [[Crossref](#)]
536. Shuyang Li, Guo Chao Peng, Fei Xing. 2019. Barriers of embedding big data solutions in smart factories: insights from SAP consultants. *Industrial Management & Data Systems* **119**:5, 1147-1164. [[Crossref](#)]
537. Andrea Urbinati, Marcel Bogers, Vittorio Chiesa, Federico Frattini. 2019. Creating and capturing value from Big Data: A multiple-case study analysis of provider companies. *Technovation* **84-85**, 21-36. [[Crossref](#)]
538. Hieu Nguyen, Muhammad Azam, Nizar Bouguila. Data Clustering using Variational Learning of Finite Scaled Dirichlet Mixture Models 1391-1396. [[Crossref](#)]
539. Sofia Aparicio, Joao Tiago Aparicio, Carlos J. Costa. Data Science and AI: Trends Analysis 1-6. [[Crossref](#)]
540. Frank Emmert-Streib, Matthias Dehmer. 2019. Large-Scale Simultaneous Inference with Hypothesis Testing: Multiple Testing Procedures in Practice. *Machine Learning and Knowledge Extraction* **1**:2, 653-683. [[Crossref](#)]
541. Emmanouil Perakakis, George Mastorakis, Ioannis Kopanakis. 2019. Social Media Monitoring: An Innovative Intelligent Approach. *Designs* **3**:2, 24. [[Crossref](#)]
542. Luis Bote-Curiel, Sergio Muñoz-Romero, Alicia Gerrero-Curieses, José Luis Rojo-Álvarez. 2019. Deep Learning and Big Data in Healthcare: A Double Review for Critical Beginners. *Applied Sciences* **9**:11, 2331. [[Crossref](#)]
543. Vedika Gupta, Vivek Kumar Singh, Udayan Ghose, Pankaj Mukhija. 2019. A quantitative and text-based characterization of big data research. *Journal of Intelligent & Fuzzy Systems* **36**:5, 4659-4675. [[Crossref](#)]
544. Hsia-Ching Chang, Chen-Ya Wang, Suliman Hawamdeh. 2019. Emerging trends in data analytics and knowledge management job market: extending KSA framework. *Journal of Knowledge Management* **23**:4, 664-686. [[Crossref](#)]
545. Diego Buenaño-Fernández, David Gil, Sergio Luján-Mora. 2019. Application of Machine Learning in Predicting Performance for Computer Engineering Students: A Case Study. *Sustainability* **11**:10, 2833. [[Crossref](#)]
546. William Villegas-Ch, Xavier Palacios-Pacheco, Sergio Luján-Mora. 2019. Application of a Smart City Model to a Traditional University Campus with a Big Data Architecture: A Sustainable Smart Campus. *Sustainability* **11**:10, 2857. [[Crossref](#)]
547. Srikar Velichety, Sudha Ram, Jesse Bockstedt. 2019. Quality Assessment of Peer-Produced Content in Knowledge Repositories using Development and Coordination Activities. *Journal of Management Information Systems* **36**:2, 478-512. [[Crossref](#)]
548. Daniel J. Power, Dale Cyphert, Roberta M. Roth. 2019. Analytics, bias, and evidence: the quest for rational decision making. *Journal of Decision Systems* **28**:2, 120-137. [[Crossref](#)]
549. Patrick Mikalef, John Krogstie. Investigating the Data Science Skill Gap: An Empirical Analysis 1275-1284. [[Crossref](#)]
550. Pervaiz Akhtar, Jędrzej George Frynas, Kamel Mellahi, Subhan Ullah. 2019. Big Data-Savvy Teams' Skills, Big Data-Driven Actions and Business Performance. *British Journal of Management* **30**:2, 252-271. [[Crossref](#)]
551. Patrick Mikalef, Maria Boura, George Lekakos, John Krogstie. 2019. Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. *British Journal of Management* **30**:2, 272-298. [[Crossref](#)]
552. Ao Huang, Yanzhu Huo, Juan Yang, Guangqiang Li. 2019. Computational Simulation and Prediction on Electrical Conductivity of Oxide-Based Melts by Big Data Mining. *Materials* **12**:7, 1059. [[Crossref](#)]
553. A. Ravishankar Rao, Riddhi Dave. Developing hands-on laboratory exercises for teaching STEM students the internet-of-things, cloud computing and blockchain applications 191-198. [[Crossref](#)]
554. Shah Jahan Miah, Huy Quan Vu, John G. Gammack. 2019. A Location Analytics Method for the Utilisation of Geotagged Photos in Travel Marketing Decision-Making. *Journal of Information & Knowledge Management* **18**:01, 1950004. [[Crossref](#)]
555. Frank Emmert-Streib, Matthias Dehmer. 2019. Defining Data Science by a Data-Driven Quantification of the Community. *Machine Learning and Knowledge Extraction* **1**:1, 235-251. [[Crossref](#)]
556. Frank Emmert-Streib, Matthias Dehmer. 2019. Evaluation of Regression Models: Model Assessment, Model Selection and Generalization Error. *Machine Learning and Knowledge Extraction* **1**:1, 521-551. [[Crossref](#)]
557. Bing Wang, Chao Wu, Lang Huang, Liangguo Kang. 2019. Using data-driven safety decision-making to realize smart safety management in the era of big data: A theoretical perspective on basic questions and their answers. *Journal of Cleaner Production* **210**, 1595-1604. [[Crossref](#)]

558. Michael Gaies, Jeffrey Anderson, Alaina Kipps, Angela Lorts, Nicolas Madsen, Bradley Marino, John M. Costello, David Brown, Jeffrey P. Jacobs, David Kasnic, Stacey Lihn, Carole Lannon, Peter Margolis, Gail D. Pearson, Jonathan Kaltman, John R. Charpie, Andrew N. Redington, Sara K. Pasquali. 2019. Cardiac Networks United: an integrated paediatric and congenital cardiovascular research and improvement network. *Cardiology in the Young* 29:2, 111-118. [[Crossref](#)]
559. Sonali Vyas, Sai Sathya Jain, Isha Choudhary, Aryaman Chaudhary. Study on Use of AI and Big Data for Commercial System 737-739. [[Crossref](#)]
560. Iztok Fister, Iztok Fister, Dušan Fister. Visualization of Sports Activities Created by Wearable Mobile Devices 223-246. [[Crossref](#)]
561. Thilo Stadelmann, Martin Braschler, Kurt Stockinger. Introduction to Applied Data Science 3-16. [[Crossref](#)]
562. Balázs Bodó, Bob van de Velde. Analyzing Big Data 347-366. [[Crossref](#)]
563. Alex Young Pedersen, Francesco Caviglia, Tom Gislev, Anders Hjortskov Larsen. Learning in Hybrid Protopublic Spaces: Framework and Exemplars 89-110. [[Crossref](#)]
564. Khaled Salah Mohamed. IoT Cloud Computing, Storage, and Data Analytics 71-91. [[Crossref](#)]
565. Pier Francesco De Maria, Leonardo Tomazeli Duarte, Álvaro de Oliveira D'Antona, Cristiano Torezzan. Digital Humanities and Big Microdata: New Approaches for Demographic Research 217-231. [[Crossref](#)]
566. Ivar Hukkelberg, Marthe Berntzen. Exploring the Challenges of Integrating Data Science Roles in Agile Autonomous Teams 37-45. [[Crossref](#)]
567. Hossein Hassani, Xu Huang, Emmanuel Sirimal Silva. Big Data and Blockchain 7-48. [[Crossref](#)]
568. Reza Shahbazian, Francesca Guerriero. Optimal Weight Design for Decentralized Consensus-Based Large-Scale Data Analytics over Internet of Things 279-288. [[Crossref](#)]
569. Sara M. Shaheen, Sawsan Alhalawani, Nuha Alnabet, Dana Alhenaki. Analytical Experiments on the Utilization of Data Visualizations 148-161. [[Crossref](#)]
570. Sirje Virkus, Emmanouel Garoufallou. Data Science from a Perspective of Computer Science 209-219. [[Crossref](#)]
571. Celia Satiko Ishikiriyama, Carlos Francisco Simoes Gomes. Big Data: A Global Overview 35-50. [[Crossref](#)]
572. Ray Cooksey, Gael McDonald. What Data Gathering Strategies Should I Use? 555-687. [[Crossref](#)]
573. Sumira Jan, Parvaiz Ahmad. Ecological Metabolomics: Challenges and Perspectives 293-378. [[Crossref](#)]
574. Stefaan G. Verhulst, Zeynep Engin, Jon Crowcroft. 2019. Data & Policy : A new venue to study and explore policy–data interaction. *Data & Policy* 1. . [[Crossref](#)]
575. Yi-Cheng Tsai, Mu-En Wu, Jia-Hao Syu, Chin-Laung Lei, Chung-Shu Wu, Jan-Ming Ho, Chuan-Ju Wang. 2019. Assessing the Profitability of Timely Opening Range Breakout on Index Futures Markets. *IEEE Access* 7, 32061-32071. [[Crossref](#)]
576. Mikel Canizo, Angel Conde, Santiago Charramendieta, Raul Minon, Raul G. Cid-Fuentes, Enrique Onieva. 2019. Implementation of a Large-Scale Platform for Cyber-Physical System Real-Time Monitoring. *IEEE Access* 7, 52455-52466. [[Crossref](#)]
577. Dickson N. T. How, M. A. Hannan, M. S. Hossain Lipu, Pin Jern Ker. 2019. State of Charge Estimation for Lithium-Ion Batteries Using Model-Based and Data-Driven Methods: A Review. *IEEE Access* 7, 136116-136136. [[Crossref](#)]
578. Huajia Wang, Ruo Hu, Hong Xu. Research on an Improved Algorithm of Professional Information Retrieval System 144-146. [[Crossref](#)]
579. Cindy-Pamela Lopez, Marco Segura, Marco Santórum. Data Analytics and BI Framework based on Collective Intelligence and the Industry 4.0 93-98. [[Crossref](#)]
580. Marié Hattingh, Linda Marshall, Marlene Holmner, Rennie Naidoo. Data Science Competency in Organisations 1-8. [[Crossref](#)]
581. Jens Prufer, Patricia Prufer. 2019. Data Science for Entrepreneurship Research: Studying Demand Dynamics for Entrepreneurial Skills in the Netherlands. *SSRN Electronic Journal* . [[Crossref](#)]
582. James Ming Chen. 2019. Models for Predicting Business Bankruptcies and Their Application to Banking and to Financial Regulation. *SSRN Electronic Journal* 2. . [[Crossref](#)]
583. Murray E. Jennex. Using a Revised Knowledge Pyramid to Redefine Knowledge Management Strategy 1-18. [[Crossref](#)]
584. Miftachul Huda, Ulfatmi, Muhammad Ja'far Luthfi, Kamarul Azmi Jasmi, Bushrah Basiron, Mohd Ismail Mustari, Ajmain Safar, Wan Hassan Wan Embong, Ahmad Marzuki Mohamad, Ahmad Kilani Mohamed. Adaptive Online Learning Technology 163-195. [[Crossref](#)]
585. Pratiyush Guleria, Manu Sood. Big Data Analytics 77-108. [[Crossref](#)]
586. Agata Mardosz-Grabowska. Big Data Myth 223-237. [[Crossref](#)]

587. Brian J. Galli. 2019. Role of Big Data in Continuous Improvement Environments. *International Journal of Applied Logistics* 9:1, 53-72. [[Crossref](#)]
588. Trevor Bogani Shihundla, Khumbulani Mpofo, Olukorede Tijani Adenuga. 2019. Integrating Product-Service Systems into the manufacturing industry: Industry 4.0 perspectives. *Procedia CIRP* 83, 8-13. [[Crossref](#)]
589. Ben Kei Daniel. 2019. Big Data and data science: A critical review of issues for educational research. *British Journal of Educational Technology* 50:1, 101-113. [[Crossref](#)]
590. Richard Berntsson Svensson, Robert Feldt, Richard Torkar. The Unfulfilled Potential of Data-Driven Decision Making in Agile Software Development 69-85. [[Crossref](#)]
591. Vian Ahmed, Zeeshan Aziz, Algan Tezel, Zainab Riaz. 2018. Challenges and drivers for data mining in the AEC sector. *Engineering, Construction and Architectural Management* 25:11, 1436-1453. [[Crossref](#)]
592. Paulo Rita, Nicole Rita, Cristina Oliveira. 2018. Data science for hospitality and tourism. *Worldwide Hospitality and Tourism Themes* 10:6, 717-725. [[Crossref](#)]
593. Muhammad Anwar, Sher Zaman Khan, Syed Zulfiqar Ali Shah. 2018. Big Data Capabilities and Firm's Performance: A Mediating Role of Competitive Advantage. *Journal of Information & Knowledge Management* 17:04, 1850045. [[Crossref](#)]
594. Virginia Ortiz-Repiso, Jane Greenberg, Javier Calzada-Prado. 2018. A cross-institutional analysis of data-related curricula in information science programmes: A focused look at the iSchools. *Journal of Information Science* 44:6, 768-784. [[Crossref](#)]
595. Trang VoPham, Jaime E. Hart, Francine Laden, Yao-Yi Chiang. 2018. Emerging trends in geospatial artificial intelligence (geoAI): potential applications for environmental epidemiology. *Environmental Health* 17:1. . [[Crossref](#)]
596. Changtian Ying, Changyan Ying, Chen Ban. 2018. A performance optimization strategy based on degree of parallelism and allocation fitness. *EURASIP Journal on Wireless Communications and Networking* 2018:1. . [[Crossref](#)]
597. Zhuangli Hu, Tong He, Yihui Zeng, Xiangyuan Luo, Jiawen Wang, Sheng Huang, Jianming Liang, Qinzhang Sun, Hengbin Xu, Bin Lin. 2018. Fast image recognition of transmission tower based on big data. *Protection and Control of Modern Power Systems* 3:1. . [[Crossref](#)]
598. Aleksander Fabijan, Pavel Dmitriev, Colin McFarland, Lukas Vermeer, Helena Holmström Olsson, Jan Bosch. 2018. Experimentation growth: Evolving trustworthy A/B testing capabilities in online software companies. *Journal of Software: Evolution and Process* 30:12. . [[Crossref](#)]
599. Kwangha Lee. 2018. Critical Care Research Using "Big Data": A Reality in the Near Future. *Acute and Critical Care* 33:4, 269-270. [[Crossref](#)]
600. Stefanie Büchner. 2018. Zum Verhältnis von Digitalisierung und Organisation. *Zeitschrift für Soziologie* 47:5, 332-348. [[Crossref](#)]
601. Bas van Raaij, Arco van de Ven. 2018. Betere financiële prognoses in business cases door de toepassing van big data. *Maandblad Voor Accountancy en Bedrijfsconomie* 92:9/10, 255-263. [[Crossref](#)]
602. Shaofu Du, Wenzhi Tang, Jiajia Zhao, Tengfei Nie. 2018. Sell to whom? Firm's green production in competition facing market segmentation. *Annals of Operations Research* 270:1-2, 125-154. [[Crossref](#)]
603. L. Nelson Sanchez-Pinto, Yuan Luo, Matthew M. Churpek. 2018. Big Data and Data Science in Critical Care. *Chest* 154:5, 1239-1248. [[Crossref](#)]
604. P BastinThiyagaraj, A Aloysius. 2018. Distance Based Measurement Approach for Truth Discovery by Resolving the Conflicts in Big Data. *Journal of Physics: Conference Series* 1142, 012013. [[Crossref](#)]
605. Sergio Ribeiro, JingTao Yao, Denis Alcides Rezende. Discovering IMRaD Structure with Different Classifiers 200-204. [[Crossref](#)]
606. Clodomir J. Santana, Pedro Aguiar, Carmelo J. A. Bastos-Filho. Customer Segmentation in a Travel Agency Dataset using Clustering Algorithms 1-6. [[Crossref](#)]
607. Jeppe Hoy Christensen, Michael Kai Petersen, Niels Henrik Pontoppidan, Marco Cremonini. Big Data Analytics in Healthcare: Design and Implementation for a Hearing Aid Case Study 296-303. [[Crossref](#)]
608. Hasna El Alaoui El Abdallaoui, Abdelaziz El Fazziki, Fatima Zohra Ennaji, Mohamed Sadgal. Decision Support System for the Analysis of Traffic Accident Big Data 514-521. [[Crossref](#)]
609. Lin Wang. 2018. Twinning data science with information science in schools of library and information science. *Journal of Documentation* 74:6, 1243-1257. [[Crossref](#)]
610. Forrest J. Bowlick, Dawn J. Wright. 2018. Digital Data-Centric Geography: Implications for Geography's Frontier. *The Professional Geographer* 70:4, 687-694. [[Crossref](#)]
611. Junius Gunaratne, Lior Zalmanson, Oded Nov. 2018. The Persuasive Power of Algorithmic and Crowdsourced Advice. *Journal of Management Information Systems* 35:4, 1092-1120. [[Crossref](#)]

612. Adrian Gardiner, Cheryl Aasheim, Paige Rutner, Susan Williams. 2018. Skill Requirements in Big Data: A Content Analysis of Job Advertisements. *Journal of Computer Information Systems* 58:4, 374-384. [[Crossref](#)]
613. Kyoung-Yun Kim, Fahim Ahmed. 2018. Semantic weldability prediction with RSW quality dataset and knowledge construction. *Advanced Engineering Informatics* 38, 41-53. [[Crossref](#)]
614. Abhilash Acharya, Sanjay Kumar Singh, Vijay Pereira, Poonam Singh. 2018. Big data, knowledge co-creation and decision making in fashion industry. *International Journal of Information Management* 42, 90-101. [[Crossref](#)]
615. Sean Robert Valentine, David Hollingworth, Patrick Schultz. 2018. Data-based ethical decision making, lateral relations, and organizational commitment. *Employee Relations* 40:6, 946-963. [[Crossref](#)]
616. Marian Lamr. Big Data and Its Usage in Systems of Early Warning of Traffic Accident Risks 154-157. [[Crossref](#)]
617. Young Seog Yoon, Hyun-Woo Oh, Kwangroh Park. Mining Twitter to Identify Customers' Requirements and Shoe Market Segmentation 1194-1199. [[Crossref](#)]
618. Xiaoxuan Jiao, Bo Jing, Beiquan Jiao. Research of Prognostic and Health Management for Avionics System Based on Massive Data Mining 350-357. [[Crossref](#)]
619. Gabriela Viale Pereira, Gregor Eibl, Constantinos Stylianou, Gilberto Martínez, Haris Neophytou, Peter Parycek. 2018. The Role of Smart Technologies to Support Citizen Engagement and Decision Making. *International Journal of Electronic Government Research* 14:4, 1-17. [[Crossref](#)]
620. Dorota Kamrowska-Zaluska, Hanna Obracht-Prondzyńska. 2018. The Use of Big Data in Regenerative Planning. *Sustainability* 10:10, 3668. [[Crossref](#)]
621. Changtian Ying, Jiong Yu, JingSha He. 2018. Towards fault tolerance optimization based on checkpoints of in-memory framework spark. *Journal of Ambient Intelligence and Humanized Computing* 275. . [[Crossref](#)]
622. Luca Dezi, Gabriele Santoro, Heger Gabteni, Anna Claudia Pellicelli. 2018. The role of big data in shaping ambidextrous business process management. *Business Process Management Journal* 24:5, 1163-1175. [[Crossref](#)]
623. Andrea De Mauro, Marco Greco, Michele Grimaldi, Paavo Ritala. 2018. Human resources for Big Data professions: A systematic classification of job roles and required skill sets. *Information Processing & Management* 54:5, 807-817. [[Crossref](#)]
624. Sophie Cockcroft, Mark Russell. 2018. Big Data Opportunities for Accounting and Finance Practice and Research. *Australian Accounting Review* 28:3, 323-333. [[Crossref](#)]
625. Jürgen Bajorath. 2018. Foundations of data-driven medicinal chemistry. *Future Science OA* 4:8, FSO320. [[Crossref](#)]
626. Stefan Thalmann. 2018. Data driven decision support. *it - Information Technology* 60:4, 179-181. [[Crossref](#)]
627. Milla Ratia, Jussi Myllärniemi, Nina Helander. 2018. The new era of business intelligence. *Meditari Accountancy Research* 26:3, 531-546. [[Crossref](#)]
628. Patrick Mikalef, Ilias O. Pappas, John Krogstie, Michail Giannakos. 2018. Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management* 16:3, 547-578. [[Crossref](#)]
629. Mario Divan. Applying the Real-Time Monitoring based on Wireless Sensor Networks: The Bajo Giuliani Project 18-24. [[Crossref](#)]
630. Ummaporn Pora, Natcha Thawesaengskulthai, Nathasit Gedsri, Sipat Triukose. Data-driven Roadmapping Turning Challenges into Opportunities 1-11. [[Crossref](#)]
631. Ramon Wenzel, Niels Van Quaquebeke. 2018. The Double-Edged Sword of Big Data in Organizational and Management Research. *Organizational Research Methods* 21:3, 548-591. [[Crossref](#)]
632. Jeeyae Choi, Jeungok Choi, Hee-Tae Jung. 2018. Applying Machine-Learning Techniques to Build Self-reported Depression Prediction Models. *CIN: Computers, Informatics, Nursing* 36:7, 317-321. [[Crossref](#)]
633. Jürgen Bajorath. 2018. Data Analytics and Deep Learning in Medicinal Chemistry. *Future Medicinal Chemistry* 10:13, 1541-1543. [[Crossref](#)]
634. . References 311-314. [[Crossref](#)]
635. Basanta-Val Pablo, Sánchez-Fernández Luis. 2018. Big-BOE: Fusing Spanish Official Gazette with Big Data Technology. *Big Data* 6:2, 124-138. [[Abstract](#)] [[Full Text](#)] [[PDF](#)] [[PDF Plus](#)]
636. Anastasia Griva, Cleopatra Bardaki, Katerina Pramatari, Dimitris Papakiriakopoulos. 2018. Retail business analytics: Customer visit segmentation using market basket data. *Expert Systems with Applications* 100, 1-16. [[Crossref](#)]
637. Wayne Madhlangobe, Ling Wang. Assessment of Factors Influencing Intent-to-Use Big Data Analytics in an Organization: Pilot Study 1710-1715. [[Crossref](#)]

638. Sharon Hewner, Suzanne S. Sullivan, Guan Yu. 2018. Reducing Emergency Room Visits and In-Hospitalizations by Implementing Best Practice for Transitional Care Using Innovative Technology and Big Data. *Worldviews on Evidence-Based Nursing* 15:3, 170-177. [[Crossref](#)]
639. Shirish Jeble, Rameshwar Dubey, Stephen J. Childe, Thanos Papadopoulos, David Roubaud, Anand Prakash. 2018. Impact of big data and predictive analytics capability on supply chain sustainability. *The International Journal of Logistics Management* 29:2, 513-538. [[Crossref](#)]
640. Jenny Farmer, Donald Jacobs. 2018. High throughput nonparametric probability density estimation. *PLOS ONE* 13:5, e0196937. [[Crossref](#)]
641. Kristen L. Fessele. 2018. The Rise of Big Data in Oncology. *Seminars in Oncology Nursing* 34:2, 168-176. [[Crossref](#)]
642. Yin Long, Zhi Chen, Jun Fang, Chintia Tellambura. 2018. Data-Driven-Based Analog Beam Selection for Hybrid Beamforming Under mm-Wave Channels. *IEEE Journal of Selected Topics in Signal Processing* 12:2, 340-352. [[Crossref](#)]
643. Aleksej Heinze, Marie Griffiths, Alex Fenton, Gordon Fletcher. 2018. Knowledge exchange partnership leads to digital transformation at Hydro-X Water Treatment, Ltd. *Global Business and Organizational Excellence* 37:4, 6-13. [[Crossref](#)]
644. John Storm Pedersen, Adrian Wilkinson. 2018. The digital society and provision of welfare services. *International Journal of Sociology and Social Policy* 38:3/4, 194-209. [[Crossref](#)]
645. Patrick Mikalef, Michail N. Giannakos, Ilias O. Pappas, John Krogstie. The human side of big data: Understanding the skills of the data scientist in education and industry 503-512. [[Crossref](#)]
646. Judson Costa, Daniel Araujo, Diego R. C. Silva, Marcelo B. Nogueira, Marconi C. Rodrigues. Home Automation Architecture Based on IOT Technologies 63-67. [[Crossref](#)]
647. Luke Stark. 2018. Algorithmic psychometrics and the scalable subject. *Social Studies of Science* 48:2, 204-231. [[Crossref](#)]
648. Ehi E. Aimiwu. 2018. Using Social Media to Target Customers for Green Technology Use. *International Journal of Virtual Communities and Social Networking* 10:2, 41-61. [[Crossref](#)]
649. Alon Friedman. 2018. Measuring the promise of Big Data syllabi. *Technology, Pedagogy and Education* 27:2, 135-148. [[Crossref](#)]
650. Marynia Kolak. 2018. Ian Foster, Rayid Ghani, Ron S Jarmin, et al. (eds), Big data and social science: A practical guide to methods and tools. *Environment and Planning B: Urban Analytics and City Science* 45:2, 388-389. [[Crossref](#)]
651. Kyungtae Kim, Sungjoo Lee. 2018. How Can Big Data Complement Expert Analysis? A Value Chain Case Study. *Sustainability* 10:3, 709. [[Crossref](#)]
652. Qingqi Long. 2018. Data-driven decision making for supply chain networks with agent-based computational experiment. *Knowledge-Based Systems* 141, 55-66. [[Crossref](#)]
653. Pierre-Étienne Heudel, Thierry Durand, Jean-Yves Blay. 2018. Projets d'intelligence artificielle à l'échelle d'un établissement de santé : l'exemple du centre Léon Bérard. *Revue française des affaires sociales* :4, 133-140. [[Crossref](#)]
654. Cecilia Fredriksson. 2018. Big data creating new knowledge as support in decision-making: practical examples of big data use and consequences of using big data as decision support. *Journal of Decision Systems* 27:1, 1-18. [[Crossref](#)]
655. Boyeong Hong, Awais Malik, Jack Lundquist, Ira Bellach, Constantine E. Kontokosta. 2018. Applications of Machine Learning Methods to Predict Readmission and Length-of-Stay for Homeless Families: The Case of Win Shelters in New York City. *Journal of Technology in Human Services* 36:1, 89-104. [[Crossref](#)]
656. Houcine Dammak, Mickaël Gardoni. Improving the Innovation Process by Harnessing the Usage of Content Management Tools Coupled with Visualization Tools 642-655. [[Crossref](#)]
657. Kamaluddeen Usman Danyaro, M. S. Liew. A Proposed Methodology for Integrating Oil and Gas Data Using Semantic Big Data Technology 30-38. [[Crossref](#)]
658. Ettore Bolisani, Constantin Bratianu. Strategic Performance and Knowledge Measurement 175-198. [[Crossref](#)]
659. Konstantinos Vassakis, Emmanuel Petrakis, Ioannis Kopanakis. Big Data Analytics: Applications, Prospects and Challenges 3-20. [[Crossref](#)]
660. Stephen Jia Wang, Patrick Moriarty. The Potential for Big data for Urban Sustainability 45-63. [[Crossref](#)]
661. Riccardo Fini, Monica Bartolini, Stefano Benigni, Paolo Ciancarini, Angelo Di Iorio, Alan Johnson, Marcello Maria Mariani, Silvio Peroni, Francesco Poggi, Einar Rasmussen, Riccardo Silvi, Maurizio Sobrero, Laura Toschi. Collaborative Practices and Multidisciplinary Research: The Dialogue Between Entrepreneurship, Management, and Data Science 129-152. [[Crossref](#)]
662. Longbing Cao. What Is Data Science 29-58. [[Crossref](#)]
663. Alessandro Ancarani, Carmela Di Mauro. Successful digital transformations need a focus on the individual 11-26. [[Crossref](#)]

664. Arnold Picot, Yvonne Berchtold, Rahild Neuburger. Big Data aus ökonomischer Sicht: Potenziale und Handlungsbedarf 309-416. [[Crossref](#)]
665. X. Chen, W. S. Lu. Scenarios for Applying Big Data in Boosting Construction: A Review 1299-1306. [[Crossref](#)]
666. Weisheng Lu, Yi Peng, Fan Xue, Ke Chen, Yuhan Niu, Xi Chen. The Fusion of GIS and Building Information Modeling for Big Data Analytics in Managing Development Sites 345-359. [[Crossref](#)]
667. Rafael de Oliveira Werneck, Waldir Rodrigues de Almeida, Bernardo Vecchia Stein, Daniel Vatanabe Pazinato, Pedro Ribeiro Mendes Júnior, Otávio Augusto Bizetto Penatti, Anderson Rocha, Ricardo da Silva Torres. 2018. Kuaa: A unified framework for design, deployment, execution, and recommendation of machine learning experiments. *Future Generation Computer Systems* **78**, 59-76. [[Crossref](#)]
668. Luca Urciuoli. 2018. An algorithm for improved ETAs estimations and potential impacts on supply chain decision making. *Procedia Manufacturing* **25**, 185-193. [[Crossref](#)]
669. Russell Tatenda Munodawafa, Satirenjit Kaur Johl. 2018. Eco-Innovation and Industry 4.0: A Big Data Usage conceptual model. *SHS Web of Conferences* **56**, 05003. [[Crossref](#)]
670. Maqbool Ali, Rahman Ali, Wajahat Ali Khan, Soyeon Caren Han, Jaehun Bang, Taeho Hur, Dohyeong Kim, Sungyoung Lee, Byeong Ho Kang. 2018. A Data-Driven Knowledge Acquisition System: An End-to-End Knowledge Engineering Process for Generating Production Rules. *IEEE Access* **6**, 15587-15607. [[Crossref](#)]
671. Shahid Shayaa, Noor Ismawati Jaafar, Shamshul Bahri, Ainin Sulaiman, Phoong Seuk Wai, Yeong Wai Chung, Arsalan Zahid Piprani, Mohammed Ali Al-Garadi. 2018. Sentiment Analysis of Big Data: Methods, Applications, and Open Challenges. *IEEE Access* **6**, 37807-37827. [[Crossref](#)]
672. W. Velasquez, A. Munoz-Arcentales, T. M. Chalen, Joaquin Salvachua. Survival analysis of people with cardiac problems in a simulated earthquake environment 702-706. [[Crossref](#)]
673. Gabriela Viale Pereira, Gregor Eibl, Peter Parycek. The Role of Digital Technologies in Promoting Smart City Governance 911-914. [[Crossref](#)]
674. Vukosi Marivate, Nyalleng Moorosi. Exploring data science for public good in South Africa 1-6. [[Crossref](#)]
675. Casandra Grundstrom, Maria Karampela. A Transforming Insurance Company and the 4 Types of Health Data Challenges that Arise 310-317. [[Crossref](#)]
676. # #. 2018. Research and Design of Hydrological Big Data Sharing Platform. *Journal of Water Resources Research* **07**:01, 10-18. [[Crossref](#)]
677. Małgorzata Przybyła-Kasperek. Application of the Pairwise Comparison Matrices into a Dispersed Decision-Making System With Pawlak's Conflict Model 392-404. [[Crossref](#)]
678. Prasanna Tambe. 2018. The Growing Importance of Algorithmic Literacy. *SSRN Electronic Journal* **37**. . [[Crossref](#)]
679. Alberto Fernández, Sara del Río, Abdullah Bawakid, Francisco Herrera. 2017. Fuzzy rule based classification systems for big data with MapReduce: granularity analysis. *Advances in Data Analysis and Classification* **11**:4, 711-730. [[Crossref](#)]
680. Thomas M. Kreuzer, Martina Wilde, Birgit Terhorst, Bodo Damm. 2017. A landslide inventory system as a base for automated process and risk analyses. *Earth Science Informatics* **10**:4, 507-515. [[Crossref](#)]
681. Carlos Costa, Maribel Yasmina Santos. 2017. The data scientist profile and its representativeness in the European e-Competence framework and the skills framework for the information age. *International Journal of Information Management* **37**:6, 726-734. [[Crossref](#)]
682. Jinhai Zhang. Research on Big Data Storage Structure and Query Optimization 1508-1511. [[Crossref](#)]
683. Mario Jose Divan. Data-driven decision making 50-56. [[Crossref](#)]
684. Mansi Khurana, Deepak Kumar. The study of data analytics in inventory management 140-144. [[Crossref](#)]
685. Ming-Chi Liu, Yueh-Min Huang. 2017. The use of data science for education: The case of social-emotional learning. *Smart Learning Environments* **4**:1. . [[Crossref](#)]
686. ###, Kidae Kim, Kwangyong Kim, Chungsik Jeong, DOHYEON KIM, ###. 2017. A Development on a Predictive Model for Buying Unemployment Insurance Program Based on Public Data. *The Korea Journal of BigData* **2**:2, 17-31. [[Crossref](#)]
687. Jun Li, Ming Lu, Guowei Dou, Shanyong Wang. 2017. Big data application framework and its feasibility analysis in library. *Information Discovery and Delivery* **45**:4, 161-168. [[Crossref](#)]
688. Murray E. Jennex. 2017. Big Data, the Internet of Things, and the Revised Knowledge Pyramid. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* **48**:4, 69-79. [[Crossref](#)]

689. Dominikus Kleindienst. 2017. The data quality improvement plan: deciding on choice and sequence of data quality improvements. *Electronic Markets* 27:4, 387-398. [Crossref]
690. Ivana Semanjski, Sidharta Gautama, Rein Ahas, Frank Witlox. 2017. Spatial context mining approach for transport mode recognition from mobile sensed big data. *Computers, Environment and Urban Systems* 66, 38-52. [Crossref]
691. Kasper Welbers, Wouter Van Atteveldt, Kenneth Benoit. 2017. Text Analysis in R. *Communication Methods and Measures* 11:4, 245-265. [Crossref]
692. Sarah Cheah, Shenghui Wang. 2017. Big data-driven business model innovation by traditional industries in the Chinese economy. *Journal of Chinese Economic and Foreign Trade Studies* 10:3, 229-251. [Crossref]
693. Ju Yeon Lee, Joo Seong Yoon, Bo-Hyun Kim. 2017. A big data analytics platform for smart factories in small and medium-sized manufacturing enterprises: An empirical case study of a die casting factory. *International Journal of Precision Engineering and Manufacturing* 18:10, 1353-1361. [Crossref]
694. Sarah Giest. 2017. Big data for policymaking: fad or fasttrack?. *Policy Sciences* 50:3, 367-382. [Crossref]
695. S. Vijayakumar Bharathi. 2017. Prioritizing and Ranking the Big Data Information Security Risk Spectrum. *Global Journal of Flexible Systems Management* 18:3, 183-201. [Crossref]
696. Jie Sheng, Joseph Amankwah-Amoah, Xiaojun Wang. 2017. A multidisciplinary perspective of big data in management research. *International Journal of Production Economics* 191, 97-112. [Crossref]
697. Mohammed Zuhair Al-Taie, Naomie Salim, Adekunle Isiaka Obasa. 2017. Successful Data Science Projects: Lessons Learned from Kaggle Competition. *Kurdistan Journal of Applied Research* 2:3, 40-49. [Crossref]
698. Il-Yeol Song, Yongjun Zhu. 2017. Big Data and Data Science: Opportunities and Challenges of iSchools. *Journal of Data and Information Science* 2:3, 1-18. [Crossref]
699. Marko Bohanec, Marko Robnik-Šikonja, Mirjana Kljajić Borštnar. 2017. Decision-making framework with double-loop learning through interpretable black-box machine learning models. *Industrial Management & Data Systems* 117:7, 1389-1406. [Crossref]
700. Marko Bohanec, Marko Robnik-Šikonja, Mirjana Kljajić Borštnar. 2017. Organizational Learning Supported by Machine Learning Models Coupled with General Explanation Methods: A Case of B2B Sales Forecasting. *Organizacija* 50:3, 217-233. [Crossref]
701. Dominik Krimpmann, Anna Stühmeier. 2017. Big Data and Analytics. *International Journal of Service Science, Management, Engineering, and Technology* 8:3, 79-92. [Crossref]
702. Matthias Murawski, Markus Bick. 2017. Digital competences of the workforce – a research topic?. *Business Process Management Journal* 23:3, 721-734. [Crossref]
703. Kevin Daniel André Carillo. 2017. Let's stop trying to be "sexy" – preparing managers for the (big) data-driven business era. *Business Process Management Journal* 23:3, 598-622. [Crossref]
704. Samuel T. McAbee, Ronald S. Landis, Maura I. Burke. 2017. Inductive reasoning: The promise of big data. *Human Resource Management Review* 27:2, 277-290. [Crossref]
705. Ruben Costa, Paulo Figueiras, Ricardo Jardim-Goncalves, Jose Ramos-Filho, Celson Lima. Semantic enrichment of product data supported by machine learning techniques 1472-1479. [Crossref]
706. Mikel Canizo, Enrique Onieva, Angel Conde, Santiago Charramendieta, Salvador Trujillo. Real-time predictive maintenance for wind turbines using Big Data frameworks 70-77. [Crossref]
707. Ling Cen, Dymitr Ruta. A Map-Based Gender Prediction Model for Big E-Commerce Data 1025-1029. [Crossref]
708. Diana Heredia Vizcaino, Wilson Nieto Bernal. Proposal for a methodology for the adoption of the big data 1-6. [Crossref]
709. Hansol Lee, Eunkyung Kweon, Minkyun Kim, Sangmi Chai. 2017. Does Implementation of Big Data Analytics Improve Firms' Market Value? Investors' Reaction in Stock Market. *Sustainability* 9:6, 978. [Crossref]
710. Guillaume Coqueret. 2017. Approximate NORTA simulations for virtual sample generation. *Expert Systems with Applications* 73, 69-81. [Crossref]
711. Allen D. Allen. 2017. Algorithms that extract knowledge from fuzzy big data: Conserving traditional science1. *Journal of Intelligent & Fuzzy Systems* 32:5, 3689-3694. [Crossref]
712. Saša Baškarađa, Andy Koronios. 2017. Unicorn data scientist: the rarest of breeds. *Program* 51:1, 65-74. [Crossref]
713. Marko Bohanec, Mirjana Kljajić Borštnar, Marko Robnik-Šikonja. 2017. Explaining machine learning models in sales predictions. *Expert Systems with Applications* 71, 416-428. [Crossref]
714. Andoni Elola, Javier Del Ser, Miren Nekane Bilbao, Cristina Perfecto, Enrique Alexandre, Sancho Salcedo-Sanz. 2017. Hybridizing Cartesian Genetic Programming and Harmony Search for adaptive feature construction in supervised learning problems. *Applied Soft Computing* 52, 760-770. [Crossref]

715. Ali Intezari, Simone Gressel. 2017. Information and reformation in KM systems: big data and strategic decision-making. *Journal of Knowledge Management* 21:1, 71-91. [[Crossref](#)]
716. Elias G. Carayannis, Evangelos Grigoroudis, Manlio Del Giudice, Maria Rosaria Della Peruta, Stavros Sindakis. 2017. An exploration of contemporary organizational artifacts and routines in a sustainable excellence context. *Journal of Knowledge Management* 21:1, 35-56. [[Crossref](#)]
717. Lourdes S. Martinez. Data Science 1-4. [[Crossref](#)]
718. Florin Gheorghe Filip, Constantin-Bălă Zamfirescu, Cristian Ciurea. Essential Enabling Technologies 121-176. [[Crossref](#)]
719. Muhammad Fahim, Thar Baker. Knowledge-Based Decision Support Systems for Personalized u-lifecare Big Data Services 187-203. [[Crossref](#)]
720. Lynda R. Hardy, Philip E. Bourne. Data Science: Transformation of Research and Scholarship 183-209. [[Crossref](#)]
721. Youakim Badr, Soumya Banerjee. Developing Modified Classifier for Big Data Paradigm: An Approach Through Bio-Inspired Soft Computing 109-122. [[Crossref](#)]
722. Carlos Costa, Maribel Yasmina Santos. A Conceptual Model for the Professional Profile of a Data Scientist 453-463. [[Crossref](#)]
723. Chhaya S. Dule, H. A. Girijamma. Guaging the Effectivity of Existing Security Measures for Big Data in Cloud Environment 209-219. [[Crossref](#)]
724. Hiroshi Mamiya, Arash Shaban-Nejad, David L. Buckeridge. Online Public Health Intelligence: Ethical Considerations at the Big Data Era 129-148. [[Crossref](#)]
725. Eduan Kotzé. Augmenting a Data Warehousing Curriculum with Emerging Big Data Technologies 128-143. [[Crossref](#)]
726. Abeer Ahmed Abdullah AL-Hakimi. Big Data Skills Required for Successful Application Implementation in the Banking Sector 381-392. [[Crossref](#)]
727. Xiaohong Liu, Junfei Chu, Pengzhen Yin, Jiasen Sun. 2017. DEA cross-efficiency evaluation considering undesirable output and ranking priority: a case study of eco-efficiency analysis of coal-fired power plants. *Journal of Cleaner Production* 142, 877-885. [[Crossref](#)]
728. Jaime Campos, Pankaj Sharma, Unai Gorostegui Gabiria, Erkki Jantunen, David Baglee. 2017. A Big Data Analytical Architecture for the Asset Management. *Procedia CIRP* 64, 369-374. [[Crossref](#)]
729. Jiang Zeyu, Yu Shuiping, Zhou Mingduan, Chen Yongqiang, Liu Yi. 2017. Model Study for Intelligent Transportation System with Big Data. *Procedia Computer Science* 107, 418-426. [[Crossref](#)]
730. Fahim Ahmed, Kyoung-Yun Kim. 2017. Data-driven Weld Nugget Width Prediction with Decision Tree Algorithm. *Procedia Manufacturing* 10, 1009-1019. [[Crossref](#)]
731. Md Tarique Hasan Khan, Fahim Ahmed, Kyoung-Yun Kim. 2017. Weldability Knowledge Visualization of Resistance Spot Welded Assembly Design. *Procedia Manufacturing* 11, 1609-1616. [[Crossref](#)]
732. Dorota Jelonek. 2017. Big Data Analytics in the Management of Business. *MATEC Web of Conferences* 125, 04021. [[Crossref](#)]
733. Stavros Sindakis. Applying Data Analytics for Innovation and Sustainable Enterprise Excellence 271-275. [[Crossref](#)]
734. P. Y. Zhao, Y. M. Shi. Predicting the likelihood of purchase by big data 040002. [[Crossref](#)]
735. Fernando Rosas, Jen-Hao Hsiao, Kwang-Cheng Chen. 2017. A Technological Perspective on Information Cascades via Social Learning. *IEEE Access* 5, 22605-22633. [[Crossref](#)]
736. Joohee Choi, Yla Tausczik. Characteristics of Collaboration in the Emerging Practice of Open Data Analysis 835-846. [[Crossref](#)]
737. Samantha Hautea, Sayamindu Dasgupta, Benjamin Mako Hill. Youth Perspectives on Critical Data Literacies 919-930. [[Crossref](#)]
738. Pan Xiang. Decision Support Assistant Management in Intelligent Logistics System 254-258. [[Crossref](#)]
739. Andrzej Szwabe, Pawel Misiorek, Michal Ciesielczyk. Logistic Regression Setup for RTB CTR Estimation 61-70. [[Crossref](#)]
740. Eduardo Nigri, Ognjen Arandjelovic. Light Curve Analysis From Kepler Spacecraft Collected Data 93-98. [[Crossref](#)]
741. Megan K. Sutherland, Meghan E. Cook. Data-Driven Smart Cities 471-476. [[Crossref](#)]
742. I. Tikito, N. Souissi. Data Collect Requirements Model 1-7. [[Crossref](#)]
743. Qi Wang, Xia Zhao, Jincui Huang, Yanghe Feng, Jiahao Su, Yehan Zou. Back-End 318-324. [[Crossref](#)]
744. Magdalena Kalińska-Kula. 2017. The Use of Big Data in Corporate Decision-Making. *Marketing i Zarządzanie* 48, 141-153. [[Crossref](#)]
745. Magdalena Kalińska-Kula. 2017. Role of Marketing Research in Acquiring Knowledge for Data-Driven Management. *Marketing i Zarządzanie* 50, 15-23. [[Crossref](#)]

746. Elena Falletti. 2017. Could Wearable Technology Transform the Traditional Concept of Habeas Corpus?. *SSRN Electronic Journal* . [\[Crossref\]](#)
747. Wardah Zainal Abidin, Nur Amie Ismail, Nurazean Maarop, Rose Alinda Alias. Skills Sets Towards Becoming Effective Data Scientists 97-106. [\[Crossref\]](#)
748. Andrzej Szwabe, Pawel Misiorek, Michal Ciesielczyk. Tensor-Based Modeling of Temporal Features for Big Data CTR Estimation 16-27. [\[Crossref\]](#)
749. Andrzej Szwabe, Paweł Misiorek, Michał Ciesielczyk. Evaluation of Tensor-Based Algorithms for Real-Time Bidding Optimization 160-169. [\[Crossref\]](#)
750. Asma Al-Drees, Reem Bin-Hezam, Ruba Al-Muwayshir, Wafa' Haddoush. Unified Retrieval Model of Big Data 323-332. [\[Crossref\]](#)
751. Rudolph Pienaar, Ata Turk, Jorge Bernal-Rusiel, Nicolas Rannou, Daniel Haehn, P. Ellen Grant, Orran Krieger. CHIPS – A Service for Collecting, Organizing, Processing, and Sharing Medical Image Data in the Cloud 29-35. [\[Crossref\]](#)
752. Babak Yadranshaghdam, Nathan Pool, Nasseh Tabrizi. A Survey on Real-Time Big Data Analytics: Applications and Tools 404-409. [\[Crossref\]](#)
753. Leonardo M. Millefiori, Dimitrios Zissis, Luca Cazzanti, Gianfranco Arcieri. Scalable and Distributed Sea Port Operational Areas Estimation from AIS Data 374-381. [\[Crossref\]](#)
754. Yan Li, Manoj Thomas, Kweku-Muata Osei-Bryson, Jason Levy. 2016. Problem Formulation in Knowledge Discovery via Data Analytics (KDDA) for Environmental Risk Management. *International Journal of Environmental Research and Public Health* **13**:12, 1245. [\[Crossref\]](#)
755. Yan Li, Manoj A. Thomas, Kweku-Muata Osei-Bryson. 2016. A snail shell process model for knowledge discovery via data analytics. *Decision Support Systems* **91**, 1-12. [\[Crossref\]](#)
756. Ivana Semanjski, Rik Bellens, Sidharta Gautama, Frank Witlox. 2016. Integrating Big Data into a Sustainable Mobility Policy 2.0 Planning Support System. *Sustainability* **8**:11, 1142. [\[Crossref\]](#)
757. Kayode Ayankoya, Andre P. Calitz, Jean H. Greyling. 2016. Real-Time Grain Commodities Price Predictions in South Africa: A Big Data and Neural Networks Approach. *Agrekon* **55**:4, 483-508. [\[Crossref\]](#)
758. Shirley Coleman, Rainer Göb, Giuseppe Manco, Antonio Pievatolo, Xavier Tort-Martorell, Marco Seabra Reis. 2016. How Can SMEs Benefit from Big Data? Challenges and a Path Forward. *Quality and Reliability Engineering International* **32**:6, 2151-2164. [\[Crossref\]](#)
759. Ajaya K. Swain. 2016. Mining big data to support decision making in healthcare. *Journal of Information Technology Case and Application Research* **18**:3, 141-154. [\[Crossref\]](#)
760. Shanliang Yang, Mei Yang, Song Wang, Kedi Huang. 2016. Adaptive immune genetic algorithm for weapon system portfolio optimization in military big data environment. *Cluster Computing* **19**:3, 1359-1372. [\[Crossref\]](#)
761. Satyanarayana V Nandury, Beneyaz A Begum. Strategies to handle big data for traffic management in smart cities 356-364. [\[Crossref\]](#)
762. Xiangzheng Deng. 2016. Urgent need for a data sharing platform to promote ecological research in china. *Ecosystem Health and Sustainability* **2**:9. . [\[Crossref\]](#)
763. Lisbeth Rodríguez-Mazahua, Cristian-Aarón Rodríguez-Enríquez, José Luis Sánchez-Cervantes, Jair Cervantes, Jorge Luis García-Alcaraz, Giner Alor-Hernández. 2016. A general perspective of Big Data: applications, tools, challenges and trends. *The Journal of Supercomputing* **72**:8, 3073-3113. [\[Crossref\]](#)
764. Russell Newman, Victor Chang, Robert John Walters, Gary Brian Wills. 2016. Model and experimental development for Business Data Science. *International Journal of Information Management* **36**:4, 607-617. [\[Crossref\]](#)
765. Il-Yeol Song, Yongjun Zhu. 2016. Big data and data science: what should we teach?. *Expert Systems* **33**:4, 364-373. [\[Crossref\]](#)
766. Rong Tang, Watinee Sae-Lim. 2016. Data science programs in U.S. higher education: An exploratory content analysis of program description, curriculum structure, and course focus. *Education for Information* **32**:3, 269-290. [\[Crossref\]](#)
767. Hao Peifeng, Cui Yuzhe, Song Jingping, Hu Zhaomu. Smart wardrobe system based on Android platform 279-285. [\[Crossref\]](#)
768. Jianwei Yin, Yan Tang, Wei Lo, Zhaohui Wu. From Big Data to Great Services 165-172. [\[Crossref\]](#)
769. Zhenning Xu, Gary L. Frankwick, Edward Ramirez. 2016. Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research* **69**:5, 1562-1566. [\[Crossref\]](#)
770. Carolina Lagos, Sebastian Gutierrez, Felisa Cordova, Guillermo Fuertes, Raul Carrasco. Data analysis methods related to energetic consumption in copper mining — A Test case in Chile 244-249. [\[Crossref\]](#)

771. Jean-Luc Giannelloni, Emmanuelle Le Nagard. 2016. Big Data et marketing prédictif. Vers un « consommateur nu » ?. *Décisions Marketing* N° 82:2, 5-15. [[Crossref](#)]
772. Myongho Yi. 2016. A Study on the Curriculums of Data Science. *Journal of the Korean BIBLIA Society for library and Information Science* 27:1, 263-290. [[Crossref](#)]
773. Thomas Hart, Lei Xie. 2016. Providing data science support for systems pharmacology and its implications to drug discovery. *Expert Opinion on Drug Discovery* 11:3, 241-256. [[Crossref](#)]
774. Meiyu Fan, Jian Sun, Bin Zhou, Min Chen. 2016. The Smart Health Initiative in China: The Case of Wuhan, Hubei Province. *Journal of Medical Systems* 40:3. . [[Crossref](#)]
775. 2016. A Project-Based Case Study of Data Science Education. *Data Science Journal* 15. . [[Crossref](#)]
776. David M. Steinberg. 2016. Industrial statistics: The challenges and the research. *Quality Engineering* 28:1, 45-59. [[Crossref](#)]
777. Christos Alexakos, Konstantinos Arvanitis, Andreas Papalambrou, Thomas Amorgianiotis, George Raptis, Nikolaos Zervos. ERMIS: Extracting Knowledge from Unstructured Big Data for Supporting Business Decision Making 611-622. [[Crossref](#)]
778. Yuriko Yano, Yukari Shiota. SVD and Text Mining Integrated Approach to Measure Effects of Disasters on Japanese Economics 20-29. [[Crossref](#)]
779. Harald Foidl, Michael Felderer. Data Science Challenges to Improve Quality Assurance of Internet of Things Applications 707-726. [[Crossref](#)]
780. Kurt Stockinger, Thilo Stadelmann, Andreas Ruckstuhl. Data Scientist als Beruf 59-81. [[Crossref](#)]
781. Go Muan Sang, Lai Xu, Paul de Vrieze. A reference architecture for big data systems 370-375. [[Crossref](#)]
782. Gerald Schermann, Dominik Schöni, Philipp Leitner, Harald C. Gall. Bifrost 1-14. [[Crossref](#)]
783. Jenna K. Simandl, Andrew J. Graettinger, Randy K. Smith, Steven Jones, Timothy E. Barnett. 2016. Making Use of Big Data to Evaluate the Effectiveness of Selective Law Enforcement in Reducing Crashes. *Transportation Research Record: Journal of the Transportation Research Board* 2584:1, 8-15. [[Crossref](#)]
784. Siddhartha Duggirala. Big Data Architecture 315-344. [[Crossref](#)]
785. Saurabh Brajesh. Big Data Analytics in Retail Supply Chain 1473-1494. [[Crossref](#)]
786. Saurabh Brajesh. Big Data Analytics in Retail Supply Chain 269-289. [[Crossref](#)]
787. Tien-Chi Huang, Chieh Hsu, Zih-Jin Ciou. 2015. Systematic Methodology for Excavating Sleeping Beauty Publications and Their Princes from Medical and Biological Engineering Studies. *Journal of Medical and Biological Engineering* 35:6, 749-758. [[Crossref](#)]
788. Pedro Quelhas Brito, Carlos Soares, Sérgio Almeida, Ana Monte, Michel Byvoet. 2015. Customer segmentation in a large database of an online customized fashion business. *Robotics and Computer-Integrated Manufacturing* 36, 93-100. [[Crossref](#)]
789. Hamid Afshari, Qingjin Peng. 2015. Modeling and quantifying uncertainty in the product design phase for effects of user preference changes. *Industrial Management & Data Systems* 115:9, 1637-1665. [[Crossref](#)]
790. Rameshwar Dubey, Angappa Gunasekaran. 2015. Education and training for successful career in Big Data and Business Analytics. *Industrial and Commercial Training* 47:4, 174-181. [[Crossref](#)]
791. Banage T.G.S. Kumara, Incheon Paik, Jia Zhang, T.H.A.S. Siriweera, Koswatte R.C. Koswatte. Ontology-Based Workflow Generation for Intelligent Big Data Analytics 495-502. [[Crossref](#)]
792. Amy Sliva, Joe Gorman, Christopher Bowman, Martin Voshell. Dual node decision wheels: an architecture for interconnected information fusion and decision making 94640F. [[Crossref](#)]
793. Seth C. Lewis. 2015. Journalism In An Era Of Big Data. *Digital Journalism* 3:3, 321-330. [[Crossref](#)]
794. Giulio Aliberti, Alessandro Colantonio, Roberto Di Pietro, Riccardo Mariani. 2015. EXPEDITE: EXPress closED ITemset Enumeration. *Expert Systems with Applications* 42:8, 3933-3944. [[Crossref](#)]
795. Ashok Kumar Wahi, Yajulu Medury, Rajnish Kumar Misra. 2015. Big Data. *International Journal of Service Science, Management, Engineering, and Technology* 6:2, 1-17. [[Crossref](#)]
796. Li Kung, Hsiao-Fan Wang. A recommender system for the optimal combination of energy resources with cost-benefit analysis 1-10. [[Crossref](#)]
797. Vincenzo Morabito. Managing Change for Big Data Driven Innovation 125-153. [[Crossref](#)]
798. Conrad Boton, Gilles Halin, Sylvain Kubicki, Daniel Forgues. Challenges of Big Data in the Age of Building Information Modeling: A High-Level Conceptual Pipeline 48-56. [[Crossref](#)]

799. Jussi Ronkainen, Antti Iivari. 2015. Designing a Data Management Pipeline for Pervasive Sensor Communication Systems. *Procedia Computer Science* **56**, 183-188. [[Crossref](#)]
800. Zhecheng Zhu, Heng Bee Hoon, Kiok-Liang Teow. Interactive Data Visualization Techniques Applied to Healthcare Decision Making 46-59. [[Crossref](#)]
801. Vincenzo Morabito. Big Data and Analytics for Competitive Advantage 3-22. [[Crossref](#)]
802. Laura Azzimonti, Marzia A. Cremona, Andrea Ghiglietti, Francesca Ieva, Alessandra Menafoglio, Alessia Pini, Paolo Zanini. BarCamp: Technology Foresight and Statistics for the Future 53-67. [[Crossref](#)]
803. Daniel J. Power. 'Big Data' Decision Making Use Cases 1-9. [[Crossref](#)]
804. Brunella Caroleo, Andrea Tosatto, Michele Osella. Making Sense of Governmental Activities Over Social Media: A Data-Driven Approach 34-45. [[Crossref](#)]
805. Jiuyong Li, Sarowar A. Sattar, Muzammil M. Baig, Jixue Liu, Raymond Heatherly, Qiang Tang, Bradley Malin. Methods to Mitigate Risk of Composition Attack in Independent Data Publications 179-200. [[Crossref](#)]
806. Yi Shen. 2015. Strategic planning for a data-driven, shared-access research enterprise: virginia tech research data assessment and landscape study. *Proceedings of the Association for Information Science and Technology* **52**:1, 1-4. [[Crossref](#)]
807. Shiming Yang, Mary Njoku, Colin F Mackenzie. 2014. 'Big data' approaches to trauma outcome prediction and autonomous resuscitation. *British Journal of Hospital Medicine* **75**:11, 637-641. [[Crossref](#)]
808. Stefan Debortoli, Oliver Müller, Jan vom Brocke. 2014. Vergleich von Kompetenzenanforderungen an Business-Intelligence- und Big-Data-Spezialisten. *WIRTSCHAFTSINFORMATIK* **56**:5, 315-328. [[Crossref](#)]
809. Stefan Debortoli, Oliver Müller, Jan vom Brocke. 2014. Comparing Business Intelligence and Big Data Skills. *Business & Information Systems Engineering* **6**:5, 289-300. [[Crossref](#)]
810. A.H.M. Sarowar Sattar, Jiuyong Li, Jixue Liu, Raymond Heatherly, Bradley Malin. 2014. A probabilistic approach to mitigate composition attacks on privacy in non-coordinated environments. *Knowledge-Based Systems* **67**, 361-372. [[Crossref](#)]
811. Alberto Fernández, Sara del Río, Victoria López, Abdullah Bawakid, María J. del Jesus, José M. Benítez, Francisco Herrera. 2014. Big Data with Cloud Computing: an insight on the computing environment, MapReduce, and programming frameworks. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **4**:5, 380-409. [[Crossref](#)]
812. Solon Barocas, Helen Nissenbaum. Big Data's End Run around Anonymity and Consent 44-75. [[Crossref](#)]
813. Provost Foster. 2014. ACM SIGKDD 2014 to be Held August 24-27 in Manhattan. *Big Data* **2**:2, 71-72. [[Citation](#)] [[Full Text](#)] [[PDF](#)] [[PDF Plus](#)]
814. Prasanna Tambe. 2014. Big Data Investment, Skills, and Firm Value. *Management Science* **60**:6, 1452-1469. [[Crossref](#)]
815. Shujaat Hussain, Byeong Ho Kang, Sungyoung Lee. A Wearable Device-Based Personalized Big Data Analysis Model 236-242. [[Crossref](#)]
816. Tiago Cunha, Carlos Soares, Eduarda Mendes Rodrigues. TweepProfiles: Detection of Spatio-temporal Patterns on Twitter 123-136. [[Crossref](#)]
817. Jay Lee, Hung-An Kao, Shanhu Yang. 2014. Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. *Procedia CIRP* **16**, 3-8. [[Crossref](#)]
818. Dylan Walker, Lev Muchnik. 2014. Design of Randomized Experiments in Networks. *SSRN Electronic Journal* . [[Crossref](#)]
819. Xiuli He, Satyajit Saravane, Qiannong Gu. Supply Chain Analytics 2364-2375. [[Crossref](#)]
820. Siddhartha Duggirala. Big Data Architecture 129-156. [[Crossref](#)]
821. Matthew A. Waller, Stanley E. Fawcett. 2013. Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics* **34**:2, 77-84. [[Crossref](#)]
822. Prasanna Tambe. 2013. Big Data Investment, Skills, and Firm Value. *SSRN Electronic Journal* . [[Crossref](#)]
823. Zhecheng Zhu, Heng Bee Hoon, Kiok-Liang Teow. Interactive Data Visualization Techniques Applied to Healthcare Decision Making 1157-1171. [[Crossref](#)]