Communications of the Association for Information Systems

Volume 42 Article 13

4-2018

Revisiting Ralph Sprague's Framework for Developing Decision Support Systems

Hugh J. Watson
University of Georgia, hwatson@uga.edu

Communications of the Association for Information Systems

Commentary DOI: 10.17705/1CAIS.04213 ISSN: 1529-3181

Revisiting Ralph Sprague's Framework for Developing Decision Support Systems

Hugh J. Watson
Terry College of Business University of Georgia
University of Georgia
hwatson@uga.edu

Abstract:

Ralph H. Sprague Jr. was a leader in the MIS field and helped develop the conceptual foundation for decision support systems (DSS). In this paper, I pay homage to Sprague and his DSS contributions. I take a personal perspective based on my years of working with Sprague. I explore the history of DSS and its evolution. I also present and discuss Sprague's DSS development framework with its dialog, data, and models (DDM) paradigm and characteristics. At its core, the development framework remains valid in today's world of business intelligence and big data analytics. I present and discuss a contemporary reference architecture for business intelligence and analytics (BI/A) in the context of Sprague's DSS development framework. The practice of decision support continues to evolve and can be described by a maturity model with DSS, enterprise data warehousing, real-time data warehousing, big data analytics, and the emerging cognitive as successive generations. I use a DSS perspective to describe and provide examples of what the forthcoming cognitive generation will bring.

Keywords: Sprague, Decision Support Systems, Business Intelligence, Analytics, Maturity Models, Cognitive.

This manuscript underwent editorial review. It was received 08/08/2017 and was with the authors for one month for one revision. Alan Dennis served as Associate Editor.

1 Introduction

Ralph H. Sprague Jr., a decision support systems (DSS) pioneer, passed way in April, 2017. Sprague, who was a friend, colleague, and mentor to me and many others, had a distinguished academic career. He was an important contributor to the emergence of the MIS academic discipline, the founder and first chair of the MIS department at the University of Hawaii, an original supporter of AIS and later an AIS fellow, the long-time conference chair of the Hawaii International Conference on System Sciences (HICSS), and the author of seminal books and papers on DSS (Sprague, 1980; Sprague & Carlson, 1982).

I first met Sprague in 1971 at a Southeastern American Institute for Decision Sciences meeting[1]. This encounter later led to a visiting position at the University of Hawaii for the 1973-74 academic year. During my visit, Sprague shared his thinking on DSS and we published several papers on the conceptual foundation for DSS: its potential, characteristics, architecture, development methodology, and future directions (Sprague & Watson, 1975, 1976). Sprague continued to write about DSS, such as in his seminal paper "A Framework for the Development of Decision Support Systems" that *MIS Quarterly* published in 1980 and the influential book *Building Effective Decision Support Systems* (with Eric Carlson) that he had published in 1992. Sprague and I continued to write papers and books on DSS until the field and our interests moved on to other topics (Sprague & Watson, 1979, 1986, 1989, 1993).

Over the years, Sprague and I continued to have DSS-related conversations. We often observed that the core of the conceptual foundation for DSS remained valid even though the enabling technology changed rapidly. We talked about writing a paper about recent technological developments and how they fit into the conceptual foundations for DSS and were helping realize the early promises of DSS.

This paper is the one that Sprague and I never wrote. In it, I discuss the emergence of DSS as Sprague and I saw and experienced it, summarize Sprague's conceptualizations about DSS, discuss current technological and other developments and how they fit into Sprague's DSS framework, and discuss the future as it applies to DSS[2]. The paper is part homage to Sprague but, more importantly for many people, a way of looking at and understanding current developments in business intelligence and analytics (BI/A).

2 The Emergence and Evolution of DSS

When computers became commercially available in the 1960s, they were first used for electronic data processing (EDP) and scientific applications (e.g., the space program)[3]. The EDP applications are now called operational or transaction processing systems. In addition to processing transactional data, they generated reports that summarized the data processed. These so-called "green bar reports" (because they were printed on paper with alternating green and white rows) provided the first computer-based information to support decision making.

When I arrived at the University of Hawaii in 1973, Sprague was already thinking about and working on DSS. He had a small contract with a bank to conceptualize about how the banking industry could use DSS. We later published a paper that was largely based on that work (Sprague & Watson, 1976). Sprague was eager to talk about DSS, and I was a willing listener. Like many of the early DSS enthusiasts, my background was in operations research/management science (OR/MS). I found the potential to develop systems for supporting decision making to be more interesting than creating additional algorithms. Much of Sprague's thinking about DSS revolved around creating figures and talking with people about his ideas and getting their reactions. We did much of our DSS work on a bench at the Ala Moana Beach Park[4].

Sprague directed me to the early work on DSS[5]. Particularly significant was Scott Morton's (1967) doctoral dissertation in which he describes building, implementing, and testing a system to support planning for laundry equipment. Some of the DSS examples that Sprague often referenced include a media planning-

[1] The organization quickly changed its name to the Decision Sciences Institute (DSI) after the AIDS epidemic began.
[2] Dan Power (2007) maintains the definitive history of the DSS field on his *DSS Resources* website (see http://dssresources.com/history/dsshistory.html). Many people, including myself, have contributed to Power's history of DSS. I have also written about it in with a different focus in Watson (2009).
[3] A very influential paper in the early 1970s was Gipson and Nolan (1974).
[4] The view was great and there was ready access to food and drink. The dedication in the second edition of *Decision Support Systems: Putting Theory into Practice* was "to Hawaii's beachfront benches".
[5] Initially, the DSS field expressed uncertainty about whether to call these new systems management decision systems or decision support systems. In the early to mid-1970s, the field coalesced around DSS.

support system (Little & Lodish, 1969), a portfolio-management system (Gerrity, 1971), and a police beat-allocation system (Carlson & Sutton, 1974) The first use of the term *decision support system* (instead of management decision system) was in Gorry and Scott Morton's (1971) paper. Scott Morton's (1971) and Keen and Scott Morton's (1978) books were also helpful in advancing the nascent field.

Sprague's DSS perspective was application focused. He was interested in building systems that would help users generate information to support decision making for a specific task or problem. In Sprague's mind, a DSS required an easy-to-use interface that accessed specially prepared data that could be analyzed using various models. He later codified this conceptualization in his dialog, data, and models (DDM) paradigm (Sprague & Carlson, 1982). DSS were highly interactive. They allowed users to access data, build and use various models, and explore alternative scenarios. They supported rather than automated decision making.

Steven Alter (1975, 1980) made a significant contribution to the evolution of DSS with his DSS taxonomy. Based on 56 DSS case studies, Alter identified seven kinds of DSS that ranged from simple reporting and data analysis to simulation and optimization[6]. He further classified these seven kinds of DSS as being either data centric or model centric depending on whether they focus on analyzing large volumes of data or using advanced analysis and modeling techniques (e.g., mathematical programming). With this view of DSS, one could classify computer applications as either transaction processing or DSS.

Sprague's DSS development framework and Alter's DSS taxonomy differed yet complemented each other. Sprague's framework focused on how to build DSS, while Alter's focused on how to categorize them. However, the two frameworks also introduced some confusion. The DSS term could refer to either an academic discipline, a decision support application with specific characteristics (Sprague's development framework), or any system that supported decision making (Alter's DSS taxonomy).

Peter Keen and Jerry Wagner also made significant contributions to the growth of DSS. Keen often wrote about DSS (Keen & Scott Morton, 1978) and was a popular conference speaker. In his interesting and thought-provoking presentations, Keen would deliver the "headline news" about what was new in DSS. Wagner was a former University of Texas OR/MS professor who developed a DSS software product called Interactive Financial Planning System (IFPS) and created a highly successful company (EXECUCOM Systems) to further develop and market it. As a university professor, Wagner understood the difficulties that universities had (at that time) gaining access to leading commercial software, and he created a popular university support program that allowed universities to use IFPS free of charge[7]. Wagner was also the key organizer of the first International Conference on Decision Support Systems, and Keen was a featured speaker[8]. This influential annual conference attracted many DSS academic and industry leaders and helped spread the news about current DSS research and cutting-edge practice.

By the mid-1980s, other kinds of systems to support decision making emerged and became the "headline news": executive information systems (Rockart & Treacy, 1982; Houdeshel & Watson, 1987), geographic information systems (Tomlinson, 1968), online analytical processing (with the emergence of relational databases), and group decision support systems (DeSanctis & Gallup, 1987). All of these systems had unique characteristics that warranted special attention, but they all contained Sprague's dialog, data, and model components.

In the early 1990s, Howard Dresner, a Gartner analyst, popularized "business intelligence" as an umbrella term for the applications, technologies, and processes for gathering, storing, accessing, and analyzing data to help business users make better decisions (Watson, 2009)[9]. As the term grew in popularity, especially in industry, it began to replace DSS for describing any system (e.g., dashboards) that supports decision making[10]. Today, the term "analytics" is sometimes used in place of business intelligence.

---

[6] The seven systems were 1) file drawer, 2) data analysis, 3) analysis information, 4) accounting models, 5) representational models, 6) optimization models, and 7) suggestion models. Alter further categorized them as 1) data retrieval, 2) data analysis, 3) simulation, and 4) suggestion and further aggregated into 1) data-centric and 2) model-centric systems.

[7] In 1991, Comshare bought EXECUCOM. Wagner later got the rights to IFPS back, developed a graphical front-end for it, and renamed it Planners Lab. It is once again free for university use and can be accessed through the Teradata University Network at www.teradatauniversitynetwork.com. Applications developed using Planners Lab are excellent examples of model-centric DSS.

[8] Jerry Wagner has noted that the first DSS conference was held in conjunction with the annual EXECUCOM users conference. It quickly became an independent conference and the first International Conference on Decision Support Systems was held in 1981.

[9] Howard Dresner popularized the term business intelligence but did not create it. Wikipedia traces its earliest use to Richard Devens in 1865 (see "Business intelligence", n.d.).

[10] Industry has been quicker than academia to replace the DSS term with BI. However, some companies still have DSS groups that trace their origins to the 1980s.

3 Sprague's DSS Development Framework

Before discussing Sprague's DSS development framework, I reflect on the technology and users in the 1970s. At that time, business people were proudly computer illiterate. Some proclaimed, "I don't use computers; I have people who do that for me". CP/M was the operating system[11], and command-line interfaces (e.g., PRINT Quarterly_Sales_Report.doc) were the norm. Intermediaries often operated systems for executives despite developer's claims that the systems were easy to use. The predominant programming languages were FORTRAN, COBOL, and BASIC, and VisiCalc was the dominant electronic spreadsheet. Processing was batch oriented with limited real-time capabilities.

Organizations used terminals such as the popular Teletype Model 33 to connect to mainframe computers. Laptops, cell phones, and the Internet did not exist.

In addition, relational databases did not exist. The first commercial relational database-management system (RDMS) was Oracle, which Relational Software (now Oracle Corporation) released in 1979. Capturing and storing anything other than transaction data required special effort, and enterprise resource planning (ERP) or customer relationship management (CRM) software did not exist.

Models were plentiful and many had their roots in World War II when they were needed to solve complex logistical problems. Most business school graduates had been exposed at least to inventory-control models, Markov processes, simulation, and linear programming[12]. Artificial intelligence (AI), machine learning, and neural network models existed but were often buried in scholarly journals waiting for today's business need and increased computational power. Models were typically used through a packaged program (e.g., IBM's MPS for linear programming), a callable subroutine (in FORTRAN), or a special-purpose language (e.g., GPSS for simulating queuing systems)[13].

3.1 Framework and Characteristics

Against this technological backdrop, Sprague (and others) envisioned how DSS could support decision makers. The vision was sometimes grander than what contemporary technology would readily allow, but it provided a roadmap for what might be possible once the technology was available[14].

Sprague did not define DSS. Rather, he preferred to use figures and discuss characteristics. Over his roughly ten years of thinking and writing about DSS, his ideas evolved and matured, but the core concepts remained essentially the same (Sprague & Watson, 1975; Sprague, 1980; Sprague & Carlson, 1982).

Figure 1 shows the dialog (i.e., decision maker and user interface), data (i.e., database), and models (i.e., model base) paradigm along with their component parts. The framework forms an integrated system under the decision maker's control.

The most frequently mentioned and salient DSS characteristics include:[15]

- Focus on managers and executives' semi-structured and unstructured decision making tasks
- Support for independent and interdependent decision making (e.g., group) and all phases of the decision making process (i.e., intelligence, design, and choice)
- Use of integrated models with traditional data access and retrieval techniques
- Focus on features that make the system fast and easy to use interactively by non-computer specialists
- Emphasis on flexibility and adaptability to changes in the environment and to users' decision making approach, and
- Built with an evolutionary, iterative development methodology.

---

[11] Only from 1981 after the IBM PC appeared did DOS become the dominant operating system.
[12] Though these topics dropped out of most U.S. business school curricula in the 1980s, many universities are now reintroducing them now that companies increasingly rely on analytics to run their business. Some older faculty are dusting off, updating, and reusing their lecture notes from 30 years ago.
[13] Machine learning research first appeared in computer science journals in the 1950s.
[14] Business and organizational issues (e.g., clearly defined business need, executive support, the difficulty of defining information requirements) have remained the same over the years, but the technology (e.g., data visualization software, Hadoop) has changed dramatically.
[15] Some of these characteristics are referred to as goals but they are also DSS characteristics.

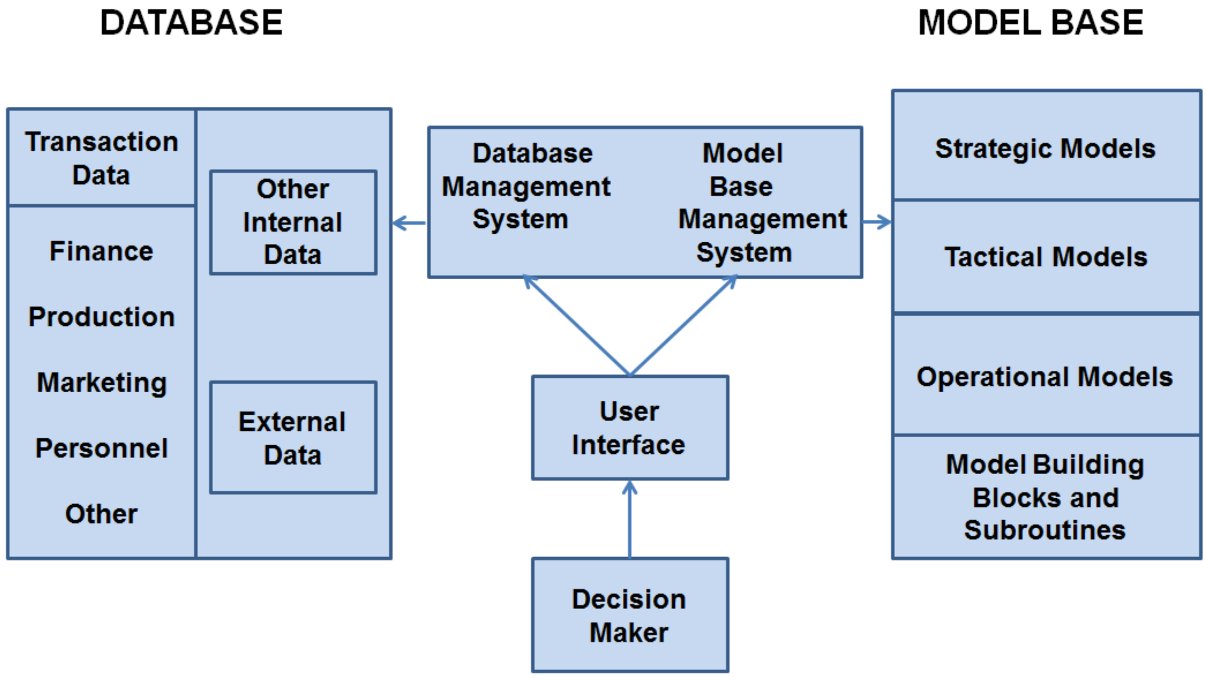

Figure 1. DSS Development Framework (Sprague & Watson, 1979)

3.2 A DSS for Banks

One can better appreciate the various DSS components and characteristics, how they are interrelated and integrated, and the way they can support various related decision making tasks by considering a DSS for banks (Sprague & Watson, 1976).

Decision makers throughout the bank use the system. It serves all managerial levels and cuts across the various functional areas. Some users employ an intermediary to operate the system, while others can operate the system themselves as a result of training, experience, and an easy-to-use interface. A server holds the transaction data (e.g., deposits), other internal data (e.g., subjective estimates of future interest rates), and external data (e.g., industry loan to deposit ratios), which a database-management system manages. Senior managers and executives have access to strategic models that support their strategic planning responsibilities. Middle managers use tactical models for managerial control purposes. Junior managers access operational models for operational control. These managers use modeling building blocks (e.g., multiple regression and linear programming software) to initially build and later update the models, and they use model management software to integrate the various models into an effective system.

Managers use a strategic model for long-range forecasts and to evaluate alternative courses of action. A strategic model is aggregate and econometric in nature and produces 5-10 year forecasts of demand deposits and loans by different categories (e.g., residential, commercial). Managers input the bank's financial records into the model together with forecasts of economic conditions in the area where the bank operates. The long-range planning model also includes the major plans and strategies that management wants to consider.

At the tactical level, financial planning and control models can forecast the bank's profit and loss, cash flow, and balance sheet for the coming twelve months. Managers can visualize the models as a series of accounting-type equations, one for each item on a financial statement. Each equation develops the value of that line or variable for a given time period based on exogenous variables supplied by a source outside the model, other previously defined variables from within the model, and built-in constants and coefficients. Managers also use tactical models to generate forecasts for planning and controlling purposes (e.g., comparing planned versus actual performance).

Banks use operational models for credit scoring and processing loan and credit cards applications. To build the models, banks require their internal data (e.g., checking and savings account balances), data from external sources (e.g., credit scores from a credit bureau), and previous loan and credit card repayment history data. With this data, banks can build multivariate statistical models (e.g., discriminant analysis) that provide the basis for scoring new applications. Based on economic conditions, managers' tolerance for risk, and the benefits and costs associated with good and bad loans, banks can develop and implement credit-granting guidelines.

To avoid suboptimization, banks need to use the strategic, tactical, and operational models in an integrated manner. For example, the forecasts of demand deposits and loans from the strategic models are a key input to the tactical models for financial planning (e.g., profit and loss). Similarly, the availability of funds that the tactical models suggest is an important input to setting the final guidelines for granting credit.

Banks initially use the model building blocks to build the strategic, tactical, and operational models and later update them. The updates may involve simply running new data through the models, changing constants or coefficients in the models, or changing some of the variables in the models in order to reflect changes in the decision making environment.

4 The Current DSS Landscape

Having discussed the DSS development framework, characteristics, and a banking example, I now explore whether Sprague's conceptualizations still apply today. The short answer? Yes. Although the users, data, models/applications, and technology significantly differ (and will continue to change), the core concepts do still apply. Most fundamentally, the DDM paradigm remains a useful way to conceptualize the architecture for DSS (and other analytic environments) much like people, process, and technology provide a useful framework for discussing information systems in general. The characteristics are still reasonable goals for decision support applications.

Figure 2 shows a reference architecture for a contemporary BI/analytics environment that is the current equivalent of the DSS development framework presented in Figure 1[16]. It identifies the data, models/applications, and users components and shows typical data flows between the data sources and data stores (i.e., platforms). In Sections 4.1 to 4.4, I look at Figure 2 in more depth.

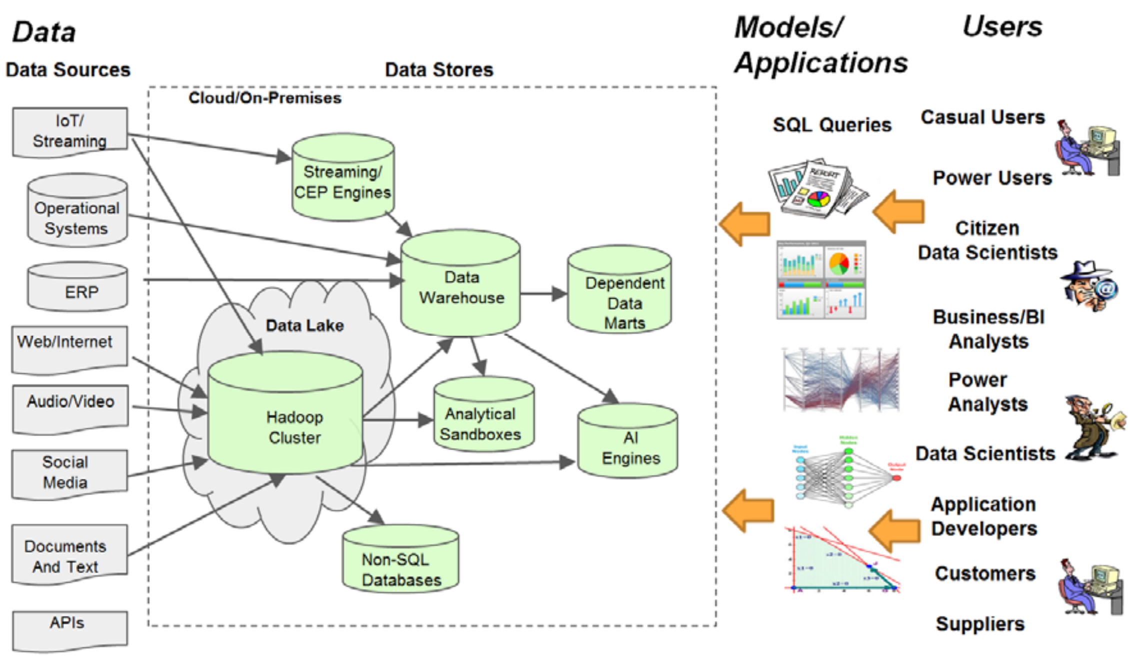

Figure 2. An Analytics/DSS Reference Architecture (Adapted from Watson, 2017b)

[16] When discussing current business intelligence and analytics, the acronym BI/A is used.

4.1 Data

I first consider the traditional and big data that fuel BI/A and the platforms used to store the data.

4.1.1 Traditional Data and Platforms

Sprague recognized the need for DSS data storage that was separate from an organization's operational systems and that would contain all of the data that the DSS needed. The reasons for a separate data store include: preserving the performance speed of operational systems, maintaining historical data, integrating related data (e.g., customer data from all touch points), storing data in a data model that is best for its intended use (e.g., a star schema for OLAP), and delivering faster response times to queries.

Although the terminology did not become popular until the late 1980s, most people now call this data store a data mart or warehouse (Inmon, 1992; Watson, 2002). Many organizations created long-term problems by having separate data marts (called independent data marts) for every DSS application. These marts often have inconsistent data and are costly to maintain because of the hardware, software, and personnel costs. To solve this problem of "analytic silos", many organizations undertook data mart-consolidation projects to integrate the independent data marts into a data warehouse. These projects have experienced varying degrees of success, but one can still find "renegade" independent data marts in almost all organizations today. In contrast, best practices encourage dependent data marts, which source their data from the data warehouse, because they do not create data inconsistency problems and provide faster response times to queries[17].

Today, data warehouses and dependent data marts have become the single version (or source) of the truth and contain the official data needed for decision support purposes: queries, reporting, dashboards/scorecards, and more. The squeaky clean data in the warehouse is especially important for financial, regulatory, and compliance reporting. At one point data warehouses were the focal point for all decision support applications, but, as new storage platforms such as Hadoop emerged, warehouses have become just part of a larger ecosystem.

Data warehouses normally provide most of the data for a real or a virtual sandbox. In a sandbox, analysts and data scientists can "play" with the data and where CPU-analytical intensive applications will not disrupt other warehouse users and applications. Also, one may need to add data to these sandboxes that the warehouse does not contain. With a real sandbox, one downloads data from a warehouse (and, thus, maintains the single version of the truth) to a separate server. With a virtual sandbox, one copies data to a partition in the warehouse where one can use it for analysis purposes.

Sometimes, organizations use an appliance, which are built from the ground up for speed[18], as a real sandbox or for specialized analytical applications (Watson, 2014). Data warehouses and especially appliances are increasingly columnar; that is, the usual columns and rows are switched to achieve greater processing speed. Organizations also now increasingly use in-memory technology (which refers to storing data in RAM rather than on a hard disk) for applications where users require especially fast response times.

4.1.2 Big Data and Platforms

Big data, characterized by the three Vs (high volume, variety, and velocity), has greatly expanded the data that one can capture, store, and analyze (Watson, 2014), Data stores now potentially contain social media data, GPS data, Web log data, RFID data, Internet of things (IoT) data, and image, audio, and video data. These new types of data have spawned additional storage platforms that hold massive amounts of data and do not require a relational model. Most notable is Hadoop, which scales linearly as one adds more data and stores data "as is" rather than in a predefined data model.

Non-relational or NoSQL databases such as Cassandra or Couchbase can store data of any structure and do not rely on SQL to retrieve data, although some do support SQL and are perhaps better called "not only SQL databases". There are also specialized non-SQL databases designed for specific kinds of data such as documents and graphs and use their own storage and retrieval methods.

[17] When one uses a dependent data mart for OLAP, one typically stores the data in a star schema data model. Queries are faster against a star schema data model than the third normal data model that the data warehouse typically uses.
[18] The hardware and software are integrated and optimized for speed.

With the growth of IoT and streaming data, one can capture and analyze vast quantities of data. In some cases, devices, such as a machine, collect and analyze the data themselves. One might use this type of analysis, called edge analytics, to automatically shut off an overheating machine.

Further, specialized devices and platforms such as Apache Kafka ingest, filter, analyze, and either communicate alerts about a condition (e.g., a machine part is predicted to fail) or provide real-time information, such as up-to-the-minute production information to a dashboard. Streaming data refers to data that comes from a single source, while complex event processing (CEP) involves streams from multiple sources.

Because Hadoop can store massive amounts of data at a low cost, a growing number of companies have begun to use the platform as a data lake (Watson, 2015b). They feed data from any source into the data lake, which they then use for a variety of purposes, such as for archiving, for analyzing data so it becomes a data source for a data warehouse, or for analytics as a standalone platform (many of the BI/A tools now connect to Hadoop). Some people even believe that a data lake can eliminate the need for a data warehouse (Smith, 2017).

One must take care with data lakes, however, so that they do not become "data swamps" (i.e., data lakes with poor-quality data, that are weakly controlled and governed, and that are never used). Stressing the need for careful control over the data lake, Gartner and others prefer the term data reservoir. As with a reservoir, one manages the flow of data (i.e., the water) and filters, transforms, and makes it fit for analysis (i.e., drinkable). Bill Inmon, "the father of data warehousing", has also cautioned that one cannot ignore the hard lessons about what one needs to have quality data for analysis purposes in the big data era.

Cloud computing is on the rise. With cloud computing, computing resources (e.g., storage platforms) are virtualized and offered as a service over the Internet (Watson, 2014). Many organizations currently deploy a combination of cloud and on-premises services. For example, an organization may have a data warehouse on its premises but use a BI/A platform in the cloud. The cloud has several potential benefits, such as access to specialized resources, quick deployment, expandable and scalable resources, the ability to discontinue a cloud service when no longer needed, cost savings, and good backup and recovery. These benefits make the cloud especially attractive for big data and analytics.

Both public and private clouds exist. Third party providers such as Amazon and Microsoft offer public clouds, while individual companies implement private clouds behind a firewall. Concerns about data security or government compliance requirements are common reasons to use private rather than public clouds.

Cloud services come in several forms (i.e., software as a service (SaaS), platform as a service (PaaS), or infrastructure as a service (IaaS)) depending on the software provided (Watson, 2014). However, with all cloud services, the cloud stores and analyzes data, and users and applications download the results.

With SaaS, the vendor provides the hardware, application software, operating system, and storage. Users upload data and use the application software to either develop their own application (e.g., reports) or to process the data using the software (e.g., credit scoring). Nearly all BI/analytics vendors offer SaaS versions of their software. SaaS is a particularly attractive option for smaller firms that lack the financial or human resources to implement and maintain the software and develop applications in-house.

With PaaS, the vendor provides only the basic platform, not the software for building or running specific applications. The benefits of this approach include: not having to maintain the computing infrastructure for applications that one develops; access to a dependable, highly scalable infrastructure; greater agility in developing new applications; and potential cost savings. Examples of PaaS include Oracle Cloud Computing, Microsoft Windows Azure, and Google App Engine.

The IaaS option provides only raw computing power and storage; it includes neither the operating system nor the application software. Customers must upload a disk image that includes the operating system and the application. IaaS offerings include Amazon EC2 (part of the Amazon Web Services offerings), Rackspace, and Google Compute Engine.

SaaS and PaaS are the two most important cloud alternatives for decision support. Most BI/A platform vendors offer a SaaS option. Also growing in importance is putting a data warehouse in the cloud or using a cloud-based Hadoop cluster. Amazon has offered its RedShift data warehousing option in Amazon Web Services (AWS) since 2013. Teradata and other data warehousing vendors (such as Oracle) offer data warehousing in the cloud. Once the data warehouse is in place, one can access data through SQL queries and analytic applications.

Application programming interfaces (APIs) provide a fast and easy way to connect systems (MuleSoft, 2017). In some ways, APIs are analogous to plugging into an electrical outlet to access electrical current. Anyone with a working knowledge of another company's API can access whatever data that company shares (or sells). For example, Twitter's API allows other systems to access tweets (e.g., to perform sentiment analytics). In some other cases, the API provides access to analyzed data or to data-analysis software (e.g., BI/A vendors that offer an API to access their SaaS software). APIs are part of the movement toward component-based service architectures and are an increasingly important source of raw and analyzed external data.

4.2 Models and Applications

I now consider different kinds of models, applications, and analytics; the various vendors in the marketplace; how the BI/A stack provides data and model integration; and the need for BI/A metadata.

4.2.1 Various Kinds of Models, Applications, and Analytics

While Sprague focused on the purpose (e.g., strategic planning) and various kinds of models (e.g., linear programming) that one could use when building a DSS, users today focus more on the specific application (e.g., dashboard, data visualization) that they need. Users select software appropriate for building that application. The required models (e.g., OLAP, logistical regression analysis) are integrated into the software and their inclusion depends on the kinds of analyses that the software is designed to perform. The business need (i.e., intended use), application, and models are tightly coupled.

Companies typically progress from descriptive to predictive and prescriptive analytics[19]. In the case of big data, people sometimes use the terms discovery or exploratory analytics instead of predictive analytics. Vendors and consulting firms also recognize this phenomenon but describe it in different ways. For example, Teradata says that companies evolve from asking: "what happened" to "why did it happen" to "what will happen" to "what is happening now" to finally "make it happen."

Vendors' products also reflect this natural progression, and vendors are extending their products' capabilities. For example, users can integrate the popular data-visualization software Tableau and R and RapidMiner to provide predictive-analytics capabilities.

The most significant change in models over the past few years is the growing importance of AI, machine learning, and neural networks. These models rely on the massive amounts of data and computing power available today and allow companies to discover important new insights about their business (especially about customers and their behaviors). Other applications include fraud detection, website recommendation engines, spam filtering, and network-intrusion detection.

4.2.2 BI/A Vendors

A large number and variety of vendors offer decision support products. The four major vendors are IBM, Microsoft, Oracle, and SAP. Each company offers a complete BI/A stack that includes the database, data integration, data warehousing, and data access and analysis tools. These vendors completed their stacks by either internally developing products or, more often, by acquiring smaller companies' products (e.g., IBM acquired Cognos, SAP bought Business Objects)[20]. By having complete stacks, these vendors offer their customers the potential benefits of a "sole source" provider (e.g., a single license agreement).

Some independent vendors have been highly successful over the years, such as MicroStrategy, SAS, and Teradata and newer companies such as Qlik, Tableau, and Looker. These vendors fill a specific need, such as Qlik and Tableau for data visualization and dashboards. Companies that prefer a "best of breed" approach to their software solution typically use independent vendors. Independent companies provide

[19] I experienced this phenomenon early in my career when building expert systems. An example is an expert system built to screen and rank applicants to the University of Georgia's Law School (Watson, Anthony, & Crowder, 1973). The system was performing its job well when the Director of Admissions (the project's sponsor) asked: "What variables should we be considering and how should we be ranking applicants?". He was interested in moving from a descriptive to a predictive or optimization model.

[20] I served as a technical advisor to a start-up company that was developing a data-profiling product. One could see that the founders planned to develop the product, obtain several strong reference accounts, and sell the company to a larger one, which then happened three years later.

most of the innovation in the marketplace because they must continue to offer a unique and compelling product to survive. Unlike the four major vendors, their entire revenue stream depends a single (or a few) products.

ThoughtSpot is a relatively new company that offers users Google-like search capabilities. As users enter a topic into a search box (e.g., "sales in the Western division last week"), ThoughtSpot performs a forward search to provide real-time suggestions. It also displays information (e.g., a chart) that best matches the search phrase. Users can pin, edit, and share this information.

Open-source software is also growing in popularity. It promises lower costs, faster implementation, and greater integration and switching flexibility. Though major open-source vendors have been active for many years, the Apache Foundation has become as important as any of them. Hadoop and its ecosystem (e.g., Hive, Pig) are among the projects/products that a growing number of firms use.

4.2.3 Data and Model Integration through the BI/A Stack

Today, the BI/A stack handles the integration among the various data and models that Sprague envisioned. A user who accesses a dashboard through the Internet represents a good example of integration among key elements. The client software is a Web browser that communicates with a Web server (the dialog) that connects to an application engine/server (the models and analytics) that generates the SQL query to a database (the data). The application engine/server analyzes the results of the query and passes the results to the Web server, which renders the dashboard and sends it back to the browser (see Figure 3).

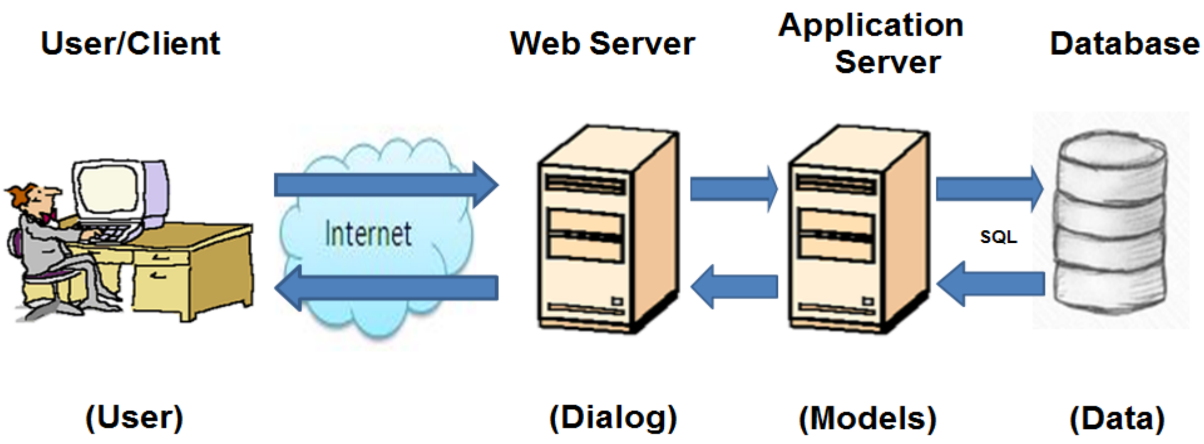

Figure 3. Integration in the BI/A Stack

4.2.4 Metadata for BI/A

As organizations increase their reliance on analytics, they need to maintain documentation (i.e., metadata) about how they develop, use, and maintain models. For every model, the documentation should include: 1) the data used to build (i.e., train) and test (e.g., validate) the model; 2) the programs, algorithms, and software used to create and test the model; 3) the individuals who reviewed the model (preferably data scientists) for accuracy; 4) where and when the model is used; and 5) when and how the model is updated (Watson, 2017b). Without this documentation, one may recreate the "dueling spreadsheets" problem in which people disagree about a problem or decision because they rely on spreadsheets that use different data, formulas, and assumptions.

4.3 Users

I now consider the wide range of BI/A users and how companies are organizing themselves for BI/A.

4.3.1 BI/A Users

In the early 1970s, computers began to appear on desktops, and few workers were computer literate. Computer networks were primitive, email was limited, and people used "sneaker nets" to share information[21]. Not surprisingly, DSS had a small user base, and those who did use the systems often needed help. Many recognized the need for training, human and in-system (especially content-dependent) support, and metadata. Governance was not a concern because of the small user base.

Today the conditions differ significantly. Executives and managers hide computer skill deficiencies due to the fear that others will judge them as being out of date in today's digital business world. New workers who enter the workforce are "Google ready" given that they have used computers and smart devices for their entire (or most of their entire) lifetimes. Analytics has a much broader and more skilled user base than DSS did.

When discussing decision support, we can think of an organization's internal users as falling on a continuum from casual users to data scientists (Watson, 2015a). In addition, there are application developers, such as a Web developer who includes a recommendation engine in a company's e-commerce website. Further, there are customers and suppliers external to the organization who access information and systems. Thus, users fall into the following categories.

- End users are business people who access BI/A-related information (most often in the form of reports (often with OLAP functionality) and dashboards/scorecards) when performing their jobs (Eckerson, 2002).
- Power users, also located in business units, who are often experts in Excel and their companies' BI/A tools (e.g., Cognos, Business Objects). They help other users with analyses and build new reports for themselves and their colleagues.
- Business analysts are in the business units (or part of an analytics team) and focus on analytical work such as website optimization in marketing or optimizing supply-chain processes in manufacturing. They typically use Excel, their companies' BI/A tools, and tools that are appropriate for their jobs (e.g., Google Analytics, SAP Supply Chain Analytics).
- BI analysts are on the BI team and focus on supporting enterprise BI/A tools (e.g., MicroStrategy), maintaining the data warehouse (in cases where is their responsibility), and developing applications such as dashboard systems[22].
- Data scientists are important new additions in many organizations. They have what has been called the "sexiest job of the 21st" century" (Davenport & Patil, 2012). Using skills developed through advanced training, they can discover new patterns and relationships in data, such as a previously unknown market segments. Data scientists are highly educated (e.g., PhDs in computer science or statistics) and can work with big data and sophisticated statistical, artificial intelligence, and machine learning models (e.g., neural networks). More than any other category of user, data scientists require highly specialized data-analysis tools.
- The current shortage of data scientists has contributed to the emergence of power analysts and citizen data scientists. Power analysts are BI and business analysts who upgrade their skills through company-sponsored programs, short courses, conferences, and college courses, certificate programs, and degrees. This preparation, along with training on appropriate software (e.g., SAS Enterprise Miner), prepares them to do take on analytical tasks that a data scientist might otherwise perform. Citizen data scientists are analytically inclined and, given the right software (e.g., Watson Analytics), can find interesting and important relationships in data. They are located anywhere in the organization. I discuss these new additions to organizations as analytics users in more detail later.
- With the arrival of the Internet, companies have been able to integrate their business processes with suppliers and customers. Many companies have achieved a high level of integration, but none better than Amazon[23]. A relatively recent development is putting analytics-generated information and analytics tools in the hands of consumers. I also discuss this important trend later.

[21] A "sneaker net" was the term used to describe the practice of having someone (presumably wearing "sneakers) going from PC to PC installing files from a floppy disk.
[22] One can classify analysts as data wizards, Excel/data visualization gurus, stats and algorithm masters, and analyst managers (Data iku, 2017).
[23] With Amazon's customer shopping recommendations, one-click ordering, and next-day delivery rely on BI/A, the company has created a brand as a trusted home supplier that threatens any market it chooses to enter (Little, 2017).

4.3.2 Organizing for BI/A

Companies are organizing themselves for analytics in new ways. Many have assigned someone, possibly the CIO, the responsibility for handling overall company analytics efforts, while others have created a new position (e.g., director of analytics, chief data officer) to take on these responsibilities.

Further, some companies are creating analytics teams that take on specific projects. For example, at Bridgestone Retail Operations (the company's U.S. network of tire and auto-repair stores), an analytics team works on projects across the organization (Ransbothan & Kiron, 2017). Collaborating with the real estate department, the team identifies the best locations for new stores. It also determines the best allocation of 22,000 employees so that Bridgestone stores have the right people on site to deal with peak demand. The staffing of such analytics teams varies, but they often include people with IT, BI, advanced analytics, and business backgrounds. This is the case with the six-person solutions delivery analytics team at Equifax, which also has access to data scientists when it needs their skills (Watson & Obenauf, forthcoming).

Organizations are struggling with how to best organize their data scientists. Some chose a decentralized approach that places their data scientists in the business units where the business problems exist. Others use a centralized approach that places them in an organizational unit (perhaps a center of analytics excellence) from where they are assigned to projects in the business units. A major (and strong_ argument for the centralized approach is that data scientists need ready access to like-minded colleagues who can assist them, help inform and educate them on new analytical techniques, and keep them motivated and happy in their work[24]. Finally, some organizations use a hybrid approach in which data scientists work in both a central unit and in the business units. However, in the centralized and hybrid models, data scientists may not work in the same area long enough (e.g., six months) to become familiar with its people, processes, and applications to develop the necessary domain knowledge.

4.4 Characteristics

I now consider how BI/A satisfies Sprague's DSS characteristics today.

4.4.1 Focus on Semi and Unstructured Decision Making Tasks of Managers and Executives

Today, BI/A has a broader scope than DSS did. It has spread to a broader user base and covers all decisions whether structured, semi-structured, or unstructured.

Sprague recognized the importance of structured decision making, which involves a well-defined (i.e., structured) means-end chain (i.e., analysis)[25], but DSS did not focus on it. At the time, common examples of structured decision making were economic order quantity (EOQ) formulas that identified an optimal order quantity and reorder point. Because the decision was structured, it was a likely candidate for automation.

Such automated decision making has become more important because of the technology for capturing and analyzing real-time data. For example, one can analyze and use IoT/streaming data to automate a decision. Applications include detecting fraud, responding to safety threats, creating customized offers, and preventing machine failures. We can anticipate more automated decision making as companies try to optimize their business processes using real-time data.

BI/A's focus has also shifted from the types of decisions supported to the kinds of analytics (e.g., descriptive, predictive, prescriptive). Currently, in the age of big data analytics, many organizations focus on finding new patterns and relationships in data, such as the leading indicators that a telecommunications customer is going to switch to another company (i.e., churn).

[24] Wayne Eckerson related the story of encountering a young data scientist in a business unit who was struggling because she had no peers with whom to discuss her work.
[25] The idea of a means-end chain as applied to problem solving has existed for many years ("Means-ends analysis", n.d.). It refers to a sequence of actions (i.e., means) that lead to a desirable goal (i.e., end). When one knows the goal and how to achieve, the decision is said to be structured.

4.4.2 Support for Independent and Interdependent Decision Making and All Phases of the Decision Making Process

In the award-winning 1952 western *High Noon*, the sheriff, played by Gary Cooper, faces a gang of killers alone[26]. Many used this metaphor in the early days of DSS to describe a decision maker who faced a difficult decision (i.e., alone and in a tough situation). Sprague correctly argued that a DSS should help support not only independent decision making (the sheriff's situation) but also interdependent decision making, which involves a group of people. The system should also support all phases of the decision making process: intelligence, design, and choice[27].

Today, we have technology that supports all decision making phases and provides group support (often called collaborative, collaboration, or groupware software). For example, one such product, Slack, supports teams' working on a decision, project, topic, or task. Slack allows teams to form and for team members to send real-time messages, have voice or video calls, share screens, store files, search for files, add comments and rate files using "stars", and receive all notifications (e.g., alerts) in one place[28].

BI/A software also includes group or team support capabilities. For example, users commonly use the software to define the team who works on a project, send messages to team members, share any resources (e.g., dashboards) deemed useful, modify the resources, add annotations, rate them, and build presentations.

4.4.3 Use of Integrated Models with Traditional Data Access and Retrieval Techniques

The integrated use of models and data remains important today. One recent development is in-database analytics where one builds data-analysis capabilities into the database software (Watson, 2014). For example, SAS has worked with Teradata and Oracle to integrate SAS' analytical capabilities into Teradata's and Oracle's data warehousing software. In-database analytics has several advantages: one does not need to extract, transform, and load data to a separate server. All (rather than just a sample) of the data remains available for analysis purposes, which improves model accuracy. After one has created a final model, one can easily use it with warehouse data to help close the loop on the process. For example, one might create a propensity-to-buy model, score customers to decide who to target in a marketing campaign, and send messages (with offers) to the targeted customers.

When the data warehouse was the "center of the data universe", model and data integration were relatively easy. All (or nearly all) of the needed data resided in the warehouse. ETL and data integration processes loaded clean data into the data warehouse and the BI/A tools easily accessed it.

The typical data archecture is now more complicated (e.g., Hadoop, appliances). Still, all these technologies should work together in a fast, seemless, and collaborative way (Watson, 2014). Users should not have to know or worry about where and how data is stored—only that it is accurate and easily and quickly accessible. Vendors that have recognized the need to integrate provide software solutions. For example, Teradata offers its Unified Data Architecture, which ties its family of products together, as well as QueryGrid for integration with non-Teradata platforms.

4.4.4 Focus on Features that Make the System Fast and Easy-to-use by Non-computer Specialists in an Interactive Way

In-memory analytics is one way that companies today have met the requirement for fast, easy access to interactive systems (Watson, 2014). With in-memory analytics, one stores data in RAM rather than on disk, which makes data access exceptionally fast[29]. Some applications perform much better with in-

[26] The film won four Academy Awards and four Golden Globe Awards and was selected for preservation in the National Film Registry. In the movie, the sheriff was unable to get help in facing the killers and was told to flee by his new, Quaker, pacifist wife. Instead, he fought the killers and prevailed with the help of his wife who chose her husband's life over her religious beliefs ("High noon", n.d.)

[27] Simon (1960) popularized a decision making process with intelligence, design, and choice phases.

[28] It is interesting how the star rating has migrated to the workplace. Just as we know that a movie with a 4.8 star rating is probably worth watching, a colleague's Excel-based analysis with the same rating is probably worth reviewing. Another interesting example of the use of stars is with sales representatives who need to know the most promising sales leads. Rather than giving them a potentially confusing propensity-to-buy score, some companies convert the scores to 1-5 stars.

[29] In-memory analytics exists at the server and the personal computer (PC) levels. SAP's Hana is an example of a server built for speed using in-memory technology. At the PC level, the BI software manages the data in memory. It intelligently decides what data

memory analytics. For example, OLAP users can "slice and dice" data (i.e., perform multidiensional analysis) in order to look at the business from different perspectives much more quickly. Some vendors describe OLAP using in-memory technology as "analysis at the speed of thought".

Sprague recognized the need for flexibility in accommodating different decision making styles and adapting to changing environmental conditions. To a large extent, BI/A software now handles the former. For example, people in finance and accounting tend to prefer tabular presentations of data (such as in spreadsheets), while scientists and engineers prefer graphical presentations[30]. Typical BI/A software allows users to toggle between tabular and graphical presentations of data or include both on the same screen. Users can choose from a wide variety of charts in different colors.

4.4.5 Emphasis on Flexibility and Adaptability to Changes in the Environment and the Decision Approach of the User(s)

Keeping models current when environmental conditions change has always been challenging. For example, structural changes in the marketplace can require new variables in forecasting models or at least changes to their coefficients and constants. In most cases, the analysts have to redevelop the models using new, relevant data. However, keeping models current is becoming easier with AI, machine learning, and neural network-based models because they can more seamlessly process new data and learn.

4.4.6 Built with an Evolutionary, Iterative Development Methodology

Companies now widely recognize agile development methodologies (e.g., scrum, XP, kanban) as the way to build decision support applications. With BI/A, users have a difficult time articulating their information requirements and can better indicate what they need after they receive something to react to. All agile methods include rapid iterations with users until they accept a final version.

Piedmont Healthcare in Atlanta develops some applications with zero latency (Watson & Jackson, 2016). The BI director sits with a user and asks about the user's information needs (most typically a dashboard), develops a prototype using Tableau (the company's choice for an enterprise-wide BI/A tool), gets the user's reactions to the prototype, and iterates until the director has created a final version. The process is agile development to the extreme. Making all this possible is the BI director's knowledge of the hospital, Tableau, and the available data and metrics. Further, this approach has the benefit that it trains the user as the developer develops the application.

Although Sprague called for the use of evolutionary, iterative (the terms in use at the time) development methodologies, the systems development lifecycle (SDLC) prevailed for many years for all applications. While the SDLC is definitely appropriate for building transactional systems, it is not effective when one does not know information requirements well and speed of development is critical. Agile methods also create a sense of "ownership" over the system.

5 DSS in the Future

Technological advances—especially in regards to IT—continue to affect and transform our daily lives at a pace that seems to keep growing. Indeed, these changes have also had a significant impact on decision support. From a maturity or generational perspective, we have seen a new decision support generation approximately every ten years (see Figure 4). After DSS in 1970-80s came enterprise data warehousing in the 1990s. Whereas DSS focused on decisions and had models and data organized around a specific decision making task, enterprise data warehousing focused on data with a centralized data repository that served decision support data needs. Real-time data warehousing emerged after 2000 that, with the availability of real-time data, could better support operational decision making, especially with customer-facing applications (Watson, Wixom, Hoffer, Anderson-Lehman, & Reynolds, 2006). Big data analytics

---

is "hot" (i.e., most likely to be used in an analysis) and stores that data in RAM. When a user needs other ("colder") data, the software accesses it from another storage device and reads into the PC's RAM.

[30] I learned this lesson among others from the Management and Information Decision Support (MIDS) system at Lockheed-Georgia (Houdeshel & Watson, 1987).

became the headline news around 2010 and greatly changed the kinds of information that can support decision making.

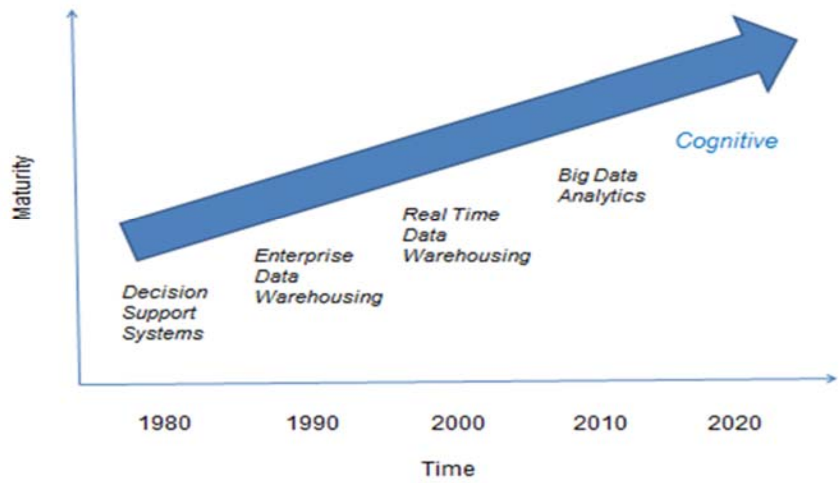

Figure 4. The Decision Support//DSS Generations (Watson, 2017b)

Given this ten-year pattern of generational change motivated, I investigated what might be the headline news in the next generation starting in the 2020s (Watson, 2017a). I analyzed the literature, interviewed 11 leading experts, and drew on my many years of experience in the decision support field. I developed the following themes or salient developments: 1) the widespread use of AI, including natural language processing; 2) the continuing movement of BI/A to the cloud; 3) the greater use of IoT, sensors, and streaming data; 4) more pervasive analytics; 5) different kinds of data scientists; and 6) the greater monetization of data.

Because of the importance of AI in the next generation, I call it the cognitive generation. I now consider some of the most interesting and important examples of this emerging generation. In many cases, the examples include a combination of trends (e.g., natural language interfaces' making BI/A more pervasive).

5.1 Natural Language Interfaces for BI/A

Ease of use has always been an important consideration with decision support applications. For the most part, these applications serve discretionary users; that is, the decision maker can choose whether or not to use the system. An intermediary (also called a "chauffer") sometimes operated a DSS for managers and executives when the system was not sufficiently user friendly. Other times, a DSS was explicitly designed for an intermediary when the necessary system capabilities required an interface beyond the future user's computer skills. One of the reasons that executive information systems became so popular was their relatively simple interfaces for executive users: something akin to a big, red "easy" button.

Sprague spoke about the simplicity/flexibility trade-off in interface design. Systems that are simple to use are not very flexible in terms of what one can do with them and vice versa. This rule once seemed inviolable, but it may not be the case.

While voice response systems have existed for many years, they were overly structured and had a limited vocabulary (e.g., "say or enter one for 'yes'"). As we know from Siri, Google Assistant, and Cortana, natural language (also called "conversational") interfaces have come a long way and will only get better. We can now reasonably think about a voice search capability for BI/A, whether voice or text, where one just asks for the information one wants. Indeed, organizations have already begun to realize this expectation (Eckerson, 2016).

The BI vendor Zoomdata has collaborated with Amazon to provide a conversational interface for accessing BI/A-related information. This collaboration integrates Amazon's Echo and Alexa (Echo's voice assistant) with Zoomdata's Domo BI platform. A user can say: "Alexa, ask Zoomdata who was the top salesperson last month in the Midwest" and receive an answer. Currently, the queries have to be defined in advance, but this requirement will undoubtedly be relaxed over time. Microsoft is on a similar path with integrating Power BI (its BI/A platform) and Cortana (its virtual assistant), and other vendors will also likely follow.

5.2 The Continuing Movement of BI/A to the Cloud

Organizations are continuing to move their data and applications to the cloud. Many CIOs no longer see strategic value in operating their own data centers and would like to free up time and resource for other initiatives. This line of thinking makes moving BI/A to the cloud attractive.

As companies move their enterprise applications, such as ERP and CRM, to the cloud, BI/A will likely follow. The concept of data gravity is a major reason for this anticipated development. Companies tend to keep data where it is rather than moving it. If most of the ERP, CRM, and other enterprise data are in the cloud, it is logical to also maintain warehouse data there.

However, using the cloud for BI/A can create analytic silos (Watson, 2017a). For example, consider a firm that uses cloud-based SAP (for ERP) and Salesforce (for CRM). The firm can use each platform to develop BI/A applications. Silos can occur because the platforms use their own data and analytics and produce output with a different look and feel. These problems are the same ones that enterprise data warehousing largely solved years ago.

5.3 The Greater Use of IoT, Sensors, and Streaming Data

Machines, sensors, smart devices, smartphones, and Web logs are generating a tremendous amount of data. The research firm IDC predicted that over 29 billion "things" will connect to the Internet in some way in 2020 (Norton, 2015). One can use the generated data to fuel dashboards with current conditions, predict future conditions, trigger alerts, direct machines to take automated actions, and more.

The power of this kind of data is even greater when it is aggregated and passed on to higher-order systems. For example, cars currently generate large amounts of machine and sensor data that they use to give out-of-lane warnings, to give alerts that another car is in a "blind spot", and to brake automatically when another car is too close. These and additional capabilities will lead to self-driving cars. However, when all of the self-driving cars become a network of interconnected cars, their movement can be coordinated to optimize safety, road usage, drive time, and fuel usage.

In the future, many applications will combine IoT and sensor data with other data and analytics. Firms will use many of these applications to generate sales and profits and improve the customer experience.

Consider the following example of what we can expect in the future (Watson, 2017a). A family is driving to Disney World for a vacation and hears a restaurant advertisement on SiriusXM radio. The family left Atlanta at 8:00 a.m. and it is now close to noon. The timing of the ad matches the upcoming lunch hour and the restaurant's cuisine corresponds with the family's known dining preferences (from Open Table). The family also receives a text message that offers a discount for the restaurant, which is just off the next exit. Geolocation data, trip plans, social media data, customer preferences, and analytics all drive this pinpoint marketing. The family places an order in advance and the food is ready when they arrive.

5.4 BI/A Becomes More Pervasive

Over the years, the number of BI/A users has grown but not as quickly as one might expect. A 2017 survey of over 2,200 companies found that the percentage of workers who use BI in their work was 26 percent in the U.S, 22 percent in Asia and the Pacific, and 20 percent in Europe (BI-Survey, 2017). Many possible reasons explain this relatively low adoption level. For example, the tools are not sufficiently user-friendly, people are not adequately trained to use them, people are too busy doing their jobs to use the tools, and so on.

An emerging trend that makes BI/A more pervasive is to embed analytics into work processes; that is, make analytics an integrated part of how people do their jobs. It closes "the last mile of BI/A"—putting insight into action—by integrating BI/A into operational systems (Eckerson, 2016).

Some of the ways to embed analytics are straightforward. For example, when an employee processes a new customer's order and the customer requests to open a credit account, the system automatically returns a credit score (from a credit bureau), and the employee either approves or rejects the order based on the score and any other criteria (e.g., size of the order). Another simple example is a sales representative in the field who receives a phone alert that a product is out of stock.

Salesforce provides a more advanced example of embedded analytics. It is striving to enhance its CRM market leadership by acquiring companies and integrating their technologies into its platform. Salesforce used this approach with AI to create Einstein Analytics. A key acquisition for Einstein Analytics was BeyondCore. As an Einstein Analytics capability, BeyondCore automatically analyzes massive amounts of data using machine learning and multiple regression analysis, finds patterns worth examining, and shows why they are statistically relevant. Subsequently, it walks users through the findings, suggests actions that users can take based on the analyses, predicts the impact the actions will have on the business, and generates PowerPoint presentations (Sheridan, 2016). Users can embed these capabilities into analytical applications developed using the Salesforce platform[31].

Self-service is another way to make BI/A more pervasive. With self-service BI/A, users create their own reports, dashboards, visualizations, and analyses rather than relying on IT. Some users find being self-sufficient appealing. They see the potential for getting needed information faster, no mismatches between the BI/A they need and what they get, and greater flexibility and control over BI/A (Eckerson, 2017). For IT departments, the potential advantages include less time spent developing applications for users, a reduced backlog of applications that need to be developed, more time spent focusing on the data and BI/A tools, and happy users.

Although self-service BI/A has been promoted since the early 2000s, several recent developments have sparked renewed interest and optimism. First, many users now recognize the need to be able to quickly create the information they need to effectively perform their jobs. Second, many business units have acquired BI/A tools (e.g., Tableau in the cloud) on their own (relatively independent of IT) and expect their users to employ them[32]. Third, many BI/A tools now include data-wrangling capabilities. Data wrangling includes locating needed data, integrating it when it comes from multiple sources (e.g., data warehouse, Excel spreadsheet), and transforming it into the required format for analysis purposes. The challenges of data wrangling has been a significant barrier to self-service BI/A.

Several factors affect whether self-service BI/A succeeds, including having the right people involved, the right processes in place, appropriate tools, a solid data infrastructure, the right organizational arrangements and structures, the right support, and necessary governance (Eckerson, 2016). The last item is important because of the natural tension between users who want maximum freedom and the IT department who see a need for control. Both parties must achieve a balance that accommodates both of their needs, such as with agreements on what BI/A tools the IT department will support.

5.5 Different Kinds of Data Scientists

An often-cited study by the McKinsey Global Institute has predicted that, by 2018, the United States will face a shortage of 140,000 to 190,000 workers with deep analytical skills (e.g., data scientists) and 1.5 million managers and people to analyze big data and make decisions (Manyika et al., 2011). Follow-up studies and casual observation suggests that these projected shortages are real. The shortage is also significant because inadequate staffing and skills are the leading barriers to success with big data analytics (Russom, 2011).

Universities and companies are responding to the demand for people with advanced analytics skills (Wixom et al., 2014). Schools have quickly rolled out undergraduate degree programs, certificates, MBA concentrations, and graduate degree programs in analytics. However, it will take a while for an adequate number of trained students to graduate.

While comanies can hire people with advanced analytics skills from other companies, many companies have decided to upgrade the skills of the people they already have who posses an analytical mind and a desire to work in analytics. Through tuition-reembursement programs, in-house training, short courses, and conferences, these people can prepare to take on analytics projects.

Gartner has coined the term “citizen data scientist” to describe those people who are not fully trained as data scientists but who, with the right preparation and software tools, can do the work that data scientists

[31] A related development is light or non-programming application development on enterprise platforms such as Salesforce and SAP.
[32] Vendors love to bypass IT and market directly to the business units.

normally do[33]. Gartner predicts that there will be five times as many citizen data scientists as fully trained ones. It also predicts that more than 40 percent of data science tasks will be automated by 2020. However, some experts are concerned about the use of citizen data scientists (Piatetsky, 2016). They argue that, unless people have deep training, they have a high probability of making serious mistakes in building models (e.g., overfitting a model) and interpreting their outputs.

5.6 Consumers as BI/A Users

Because of the technology available at the time, Sprague focused on company employees (managers and executives) as DSS users. Over time, though, the Internet and reliable networks made it possible to include customers and suppliers. B2B and B2C commerce and integrated supply chains became a focal point for companies.

We now need to expand the conceptualization of "users" again. Companies are putting BI/A-generated information and analytics tools directly in the hands of *consumers*, which is consistent with companies' monetizing their data assets by using data and analytics to provide better customer service, increase customer loyalty, generate revenue and profits, and gain a competitive advantage.

Amazon provides a simple, yet powerful example of the value of giving BI/A-generated information to consumers. Its search engine returns results that consumers trust. While the details of the search algorithm are a closely held secret, factors such as search term, profit margins, popularity, and inventory levels of available items probably influences the rank-ordered listing of products (Watson, Hoffer, & Wixom, 2009). Amazon's recommendation engines provide on-site and email recommendations of products that one might be interested in buying. The recommendations are based on market basket analysis of other shoppers' purchases, one's search behavior, one's previous purchases, and items abandoned in one's cart. Customer reviews and answered questions provide helpful information in arriving at a purchasing decision.

Scottrade's client website for online trading provides both information and tools for investment decision making. Users can monitor their accounts in ways that range from balances and pending orders to watch lists and alerts. They can read investment research, market news, analyst reports, and commentary and take advantage of indicators and charts to spot trends. Scottrade SmartText makes technical analysis easy to read and understand. Scottrade's Portfolio Review Tool assists users in selecting the right investment mix for their situation. Pre-defined asset allocation models let users compare current and targeted investments, view investment allocation across market segments, and identify opportunities.

6 Conclusion

Is Sprague's DSS development framework relevant today? The answer is a qualified "yes". The DDM paradigm continues to be useful in thinking about the basic architecture for decision support applications. Users always interact with a system that contains data and models. Sprague's DSS characteristics also remain desirable for today's systems.

Though parts of Sprague's conceptualizations were visionary considering the technology available at the time, anyone would have found it difficult to fully anticipate the changes that were coming. Sprague saw the potential value of text-based data and later researched document-management systems (Sprague, 1995), but he did not foresee Web/Internet, IoT, audio/visual, RFID, and social media data. New storage platforms now allow one to store massive amounts of data with any structure at a relatively low cost. The Internet allows applications to be accessible beyond company walls and BI/A to spread to new user groups. AI-based technology is finding its way into almost everything, including BI/A, and will become increasingly important.

Sprague's DSS conceptualizations concerned a system that supports one or a few related decision making tasks, but BI/A now has an enterprise-wide perspective. Whereas DSS applications provided organizational value, many firms today depend on BI/A to successfully compete in the marketplace (Davenport & Harris, 2007). The various data platforms support a wide variety of users and applications as do data-access and analysis tools (see Figure 2). Governance issues that were not an important issue

[33] Software such as Watson Analytics, SAS Enterprise Miner, and IBM SPSS Modeler have the ability to access data, transform it, split it for model training (i.e., building) and testing (i.e., validation) purposes, process it through multiple models/algorithms, and indicate which model is best.

to Sprague are now very important, especially with business units' investing heavily in their own BI/A capabilities and self-service BI/A. Other issues such as security and BYOD policies for mobile BI are either new or are of heightened concern.

Sprague's DSS development framework could not foresee the applications and societal impacts that analytics would bring about, such as self-driving cars, individualized cancer treatment protocols, enhanced agricultural production, and technology-driven disruptions in labor markets, though nor did he intend it to. However, he provided a starting point and a foundation that helped guide us to where we are today and into the future.

References

Alter, S. L. (1975). *A study of computer aided decision making in organizations* (doctoral dissertation). Massachusetts Institute of Technology, Cambridge, Massachusetts.

Alter, S. L. (1980). *Decision support systems: Current practice and continuing challenge.* Reading, MA: Addison-Wesley.

BI-Survey. (2017). *Pervasive BI—dream or reality?* Retrieved from https://bi-survey.com/bi-deployment

Business intelligence. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Business_intelligence#History

Carlson, E. D., & Sutton. J. A. (1974). *A case study of non-programmer interactive problem solving* (IBM research report RJI382). San Jose, CA.

Data iku. (2017). *The analyst of the future*. Retrieved from https://pages.dataiku.com/data-business-analyst-future-choose-your-own-adventure?utm_campaign=Analyst%20of%20the%20Future&utm_source=paid-eblast-analyst-of-the-future-KdNuggets-Aug-2017&utm_medium=ebook

Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning.* Boston, MA: Harvard Business School Press.

Davenport, T. H., & Patil, D. J. (2012). Data scientist: The Sexiest job of the 21st century. *Harvard Business Review, 90*(10), 70-76.

DeSanctis, G., & Gallupe, R. B. (1987). A foundation for the study of group decision support systems. *Management Science*, *33*(5), 589-609.

Eckerson, H. H. (2016). Spoken queries are now a reality. *Eckerson Group*. Retrieved from https://www.eckerson.com/articles/spoken-queries-now-a-reality

Eckerson, W. (2002). The decision support sweet spot. *Journal of Data Warehousing, 7*(3), 4-9.

Eckerson, W. (2017). *Self-service analytics: Organizational, architectural, and governance success factors* [conference]. Las Vegas, NV: TDWI Conferences.

Gerrity, T. P., Jr. (1971). Design of man-machine decision systems: An application to portfolio management. *Sloan Management Review, 12*(2), 59-75.

Gipson, C. F., & Nolan, R. (1974). Managing the four stages of EDP growth. *Harvard Business Review*. Retrieved from https://hbr.org/1974/01/managing-the-four-stages-of-edp-growth

Gorry, A., & Scott Morton, M. S. (1971). A framework for information systems. *Sloan Management Review, 13*(1), 56-79.

High noon. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/High_Noon

Houdeshel, G., & Watson, H. J. (1987). The management information and decision support (MIDS) system at Lockheed-Georgia. *MIS Quarterly*, *11*(1), 127-140.

Inmon, W. H. (1992). *Building the data warehouse.* New York: Wiley.

Keen, P. G. W., & Scott Morton, M. S. (1978). *Decision support systems: An organizational perspective.* Reading, Mass: Addison-Wesley.

Little, J. D. C., & Lodish, L. M. (1969). A media planning calculus. *Operations Research*, *17*(1), 1-35.

Little, J. R. (2017). Amazon is going to kill your brand. And your job. *LinkedIn*. Retrieved from https://www.linkedin.com/pulse/amazon-going-kill-your-brand-job-jr-little?trk=eml-email_feed_ecosystem_digest_01-hero-0-null&midToken=AQF944vF1u4YDw&fromEmail=fromEmail&ut=1CEQ7eKbHZG7Q1

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier of innovation, competition, and productivity. *McKinsey Global Institute.* Retrieved from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation

Means-ends analysis. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Means-ends_analysis

MuleSoft. (2017). *How to manage and deploy APIs*. Retrieved from http://docs.media.bitpipe.com/io_13x/io_135359/item_1477991/Whitepaper%20-%20How%20to%20Design%20and%20Manage%20APIs.pdf

Norton, S. (2015). Internet of things market to reach $1.7 trillion by 2020: IDC. *Marketwatch.com.* Retrieved from http://www.marketwatch.com/story/internet-of-things-market-to-reach-17-trillion-by-2020-idc-2015-06-02-8103241

Piatetsky, G., (2016). The mirage of the citizen data scientist. *KD Nuggets*. Retrieved from http://www.kdnuggets.com/2016/03/mirage-citizen-data-scientist.html

Power, D. J. (2007). A brief history of decision support systems. *DSSResources.COM*. Retrieved from http://DSSResources.COM/history/dsshistory.html

Ransbothan, S., & Kiron, D. (2017). Analytics as a source of business innovation. *MIT Sloan Management Review.* Retrieved from http://sloanreview.mit.edu/projects/analytics-as-a-source-of-business-innovation/

Rockart, J. F., & Treacy, M. E. (1982). The CEO goes on-line. *Harvard Business Review, 60*(1), 81-93.

Russom, P. (2011). Big data analytics. *The Data Warehousing Institute*. Retrieved from http://tdwi.org/research/2011/09/best-practices-report-q4-big-data-analytics.aspx

Scott Morton, M. S. (1967). *Computer-driven visual display devices—their impact on the management decision-making process* (doctoral dissertation). Harvard Business School, Boston, Massachusetts.

Scott Morton, M. S. (1971). *Management decision systems: Computer-based support for decision making*. Boston, MA: Harvard University.

Sheridan, K. (2016). Salesforce acquiring analytics startup BeyondChoice. *Information Week.* Retrieved from http://www.informationweek.com/software/enterprise-applications/salesforce-acquiring-analytics-startup-beyondcore-/d/d-id/1326639

Simon, H. (1960). *The new science of management decision.* New York: Harper and Row.

Smith, S. J. (2017). The demise of the data warehouse. *Predictive Analytics Trends*. Retrieved from https://www.eckerson.com/articles/the-demise-of-the-data-warehouse

Sprague, R. H., Jr., & Watson, H. J. (1975). MIS concepts part 2. *Journal of Systems Management*, *26*(2), 35-40.

Sprague, R. H., Jr., & Watson, H. J. (1976). A decision support system for banks. *Omega: The International Journal of Management Science, 4*(6), 657-67l.

Sprague, R. H. Jr., & Watson, H. J. (1979). Bit by bit: Toward decision support systems. *California Management Review*, *22*(4), 60-68.

Sprague, R. H., Jr. (1980). A framework for the development of decision support systems. *MIS Quarterly, 4*(4), 1-26.

Sprague, R. H., Jr., & Carlson, E. D. (1982). *Building effective decision support systems.* Englewood Cliffs, NJ: Prentice-Hall.

Sprague, R. H., Jr., & Watson, H. J. (1986). *Decision support systems: Putting theory into practice*. Englewood Cliffs, NJ: Prentice-Hall.

Sprague, R. H., Jr., & Watson, H. J. (1989). *Decision support systems*: *Putting theory into practice* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Sprague, R. H., Jr., & Watson, H. J. (1993). *Decision support systems: Putting theory into practice* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Sprague, R. H, Jr. (1995). Electronic document management: Challenges and opportunities for information systems managers. *MIS Quarterly, 19*(1), 29-49.

Tomlinson, R. (1968). A geographic information system for regional planning. *Department of Forestry and Rural Development.* Retrieved from https://gisandscience.files.wordpress.com/2012/08/1-a-gis-for-regional-planning_ed.pdf

Watson, H. J., Anthony, T. F., & Crowder, W. S. (1973). A Heuristic Model for Law School Admissions Screening," *College and University*, *48*(3), l95-204.

Watson, H. J. (2002). Recent developments in data warehousing. *Communications of the Association for Information Systems*, 8, 1-25.

Watson, H. J., Wixom, B. H., Hoffer, J. A., Anderson-Lehman, R., & Reynolds, A. M. (2006). Real-time business intelligence: Best Practices at Continental Airlines. *Information Systems Management*, *23*(1), 7-18.

Watson, H. J. (2009). Tutorial: Business intelligence—past, present, and future. *Communications of the Association for Information Systems*, *25*, 487-510.

Watson, H. J., Hoffer, J. A., & Wixom, B. H. (2009). BI at RetailStore.com. *Teradata University Network.* Retrieved from www.teradatauniversitynetwork.com

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*, *34*, 1247-1268.

Watson, H. J. (2015a). Should you pursue a career in BI/analytics? *Student Edition of the Business Intelligence Journal.* Retrieved from http://people.terry.uga.edu/csalge/ReadingsBI/Watson%202015z.pdf

Watson, H. J. (2015b). Data lakes, data labs, and sandboxes. *Business Intelligence Journal*, *20*(1), 4-7.

Watson, H. J., & Jackson, M. (2016). Piedmont healthcare: Using dashboards to deliver information. *Business Intelligence Journal*, *21*(3), 5-9.

Watson, H. J. (2017a). The cognitive decision-support generation. *Business Intelligence Journal*, *22*(2), 5-14.

Watson, H. J. (2017b). Preparing for the cognitive generation of decision support. *MISQ Executive, 16*(3),

Watson, H. J., & Obenauf, B. (Forthcoming). Data monetarization at Equifax: Keys to analytics team success. *Business Intelligence Journal.*

Wixom, B., Ariyachandra, T., Douglas, D., Goul, M., Gupta, B., Iyer, L., Kulkarni, U., Mooney, J. G., Phillips-Wren, G., & Turetken, O. (2014). The current state of business intelligence in academia: The arrival of big data. *Communications of the Association for Information Systems*, 34, 1-13.

About the Author

Hugh Watson is a Professor of MIS and holds a C. Herman and Mary Virginia Terry Chair of Business Administration in the Terry College of Business at the University of Georgia. He is a leading scholar and authority on business intelligence and analytics, having authored 24 books and over 200 journal papers. Watson is a fellow of the Association for Information Systems and The Data Warehousing Institute, and is the senior editor of the *Business Intelligence Journal*. For over 20 years, he was the consulting editor for John Wiley & Sons' MIS series. He was the founding director of the Teradata University Network, a free portal for faculty and students interested in BI/analytics, data warehousing and database.