An Exploration on the Significance of Data Science in Organisational Development

Faremi O. Oluwatosin
Albany State University,
504 College Dr. Albany Georgia 31705, USA.
tosiinfaremi@gmail.com, olawale_fasuan@yahoo.com

Abstract
The study examines the significance importance of data science to organisational development. The study also described the roles of data scientist in a commercial organization. The study was carried out by digging up issues around the opinions of scholars by the way of conceptual clarifications and content analysis of data. The study was guided by two major objectives and discussions were made based on the clarifications gotten arisen from the content analysis. The study concluded that data science is of great significance important in organization and its values are immeasurable in organizational development. The study recommends among other things that the use of data science should be intensified in any commercial organization.

Keywords: *business values, data science, organisation, scientists.*

Introduction

Data science is a widespread technology that have received an extensive attention in academic, business world, administration and industrial circles over the years and across the globe. Thor Olavsrud (2019) described data science as a method for transforming business data into assets that can help organizations improve revenue, reduce overhead costs, seize business opportunities, improve customer experience, and more. In 2013, Dhar discussed about data science and its various implications from a business and research standpoint. He, according to him, data science helps in drawing out information from a set of statistics" (Dhar, V 2013). Specifically, Dhar explained further that data science includes the technology of data acquisition, computer science, data collection, databases analysis, interpretations and handling. However, data science was introduced with a goal to construct, build and create the means for digging up and established business-focused data vision. Among the requirement for this is by creating an understanding on how data are valued and turned into information that could be used for an effective business opportunities that is comparatively advantage in a business or organization environment.

Data science have created sources of various opportunities and benefits for business organizations', industries and firms to strive and exploits competitive markets by providing comparative advantage across an ever-expanding frontier of domains. Arising from this, data are sourced, analyzed and interpreted as described by Thor Olavsrud (2019) to transform business data into assets that can help organizations improve revenue, reduce costs, seize business opportunities, and improve customer experience etc.
Unfortunately, to explore data, particularly the fast growing available sources of data for gathering, analysis and interpretations is one of the challenges facing business organisation today. In this context, data science applications by organizations and their significance importance together with the efficient and effective of data science to organizational development have largely absent, omitted and unexplored because most are still unaware of the value, the benefits and significance importance of hat data science holds in an organization.

However, the study was carried out to identify the significance importance of data science to organisational development and to categorise data scientist in an organisation.

Background of the Study and Conceptual Clarifications

Accordingly, Data science was said and had been adjudged to be at the rising level used as a competitive advantage in management and administration over the years (Thor Olavsrud 2019). Particularly in business, commercial and industrial sectors. Organizations' have increasingly identify with data science which have provided a development organizational performance, improved organizational structure, monitors the flow of customers and customers satisfactions including the markets chains of supply. Data science has made possible, what was previously impossible or what seems to be impossible. In respect to this, this paper explored the significance importance of data science to organisational development. The result from the study provided information on the importance of data science in an organisation. The study significantly contributed towards the existing body of knowledge and research efforts on data science to validate significant importance of data science applications in an organisation. Based on this, an exploration of significance of data science were made by drawing out the opinions of scholars from the empirical studies.

Data Science as a Science and Tools of Scientific Data

Accordingly, data science is the methods and technologies used to conduct scientific research through management and utilization of scientific data. Data science has been used to better characterize the data-intensive nature of today's science and engineering. Aiden, Michel, and Uncharted (2013) affirmed that many disciplines use data technology to deal with scientific data from their respective areas. From this, X-informatics emerged, including bioinformatics, neuroinformatics, and social informatics. For example, May (2014) said researchers in NuMedii, Inc., a big-data company in Silicon Valley, predicted whether existing drugs could be used to treat ovarian cancer by examining gene expression data from over 2,500 ovarian tumor samples (May, M. 2014). From this point of view, according to (Hey, et al (2009) cited by (Aiden, E., and Michel, J. B. Uncharted, 2013) data mainly refer to data generated and used in scientific studies. This emphasizes that data science is the use of scientific data to support scientific research (Hey, et al 2009).

However, recently in year 2020, Jian Wang, Jianping Liu, Guomin Zhou, Mo Wang and Lei Shiin in their studies affirmed that scientific data is subcategorized as research data, which according to Engineering and Physical Sciences Research Council (EPSRC), is a factual material that are recorded, interpreted and analyzed. According to Borgman (2013) scientific data is described as an entities that can be used as confirmations of phenomena particularly for a research purpose. Essentially, these definitions have clarified, explained and justified the importance, the functions and the roles of scientific data in an organization. With this, we have demonstrated the concept of scientific data as such a distinctive features that are characterized with different functions. In the study conducted by Jian Wang, Jianping Liu, Guomin Zhou, Mo Wang and Lei Shi in 2020, titled 'How Do People Make Relevance Judgment of Scientific Data?' the scholars affirmed that several efforts have been made to explore user relevance judgment for documents, the scholars mention some of them and listed them to includes images, web pages and music in the field of information retrial. Furthermore, the researcher carried out two-phase research work, the first was carried out using exploratory research design and the second one was an empirical research work to explore the relevance judgment patterns of scientific data users. While conducting the exploratory research study, the researchers engaged and interviewed 23 respondents subject to their knowledge related to scientific data. And, based on the results from the exploratory research study, the study identified three relevant criterial RC, such as

i. Data topicality judgment as the first step or starting point,
ii. Data reliability judgment as the necessary process and
iii. Data utility judgment as final purpose.

Accordingly, the findings from this provided understanding of relevance judgement and behaviours of scientific data users that could however benefit the design for cognitive retrieval systems and algorithms specific to scientific data (Jian Wang, Jianping Liu, Guomin Zhou, Mo Wang and Lei Shi 2020)

Meanwhile, Committee on Data for Science and Technology (CODATA) has also said that, data science task can be done using wide range of tools. Some of the tools were listed to include: SQL, Python, Java, and a cornucopia of open source projects such as Hive, oozie, and TensorFlow. It was said that these tools are used for a variety of data-related tasks, ranging from extracting and cleaning data, to subjecting data to algorithmic analysis via statistical methods or machine learning (CODATA).

The Roles of Data Science in an Organizational Setting?

Today, there are various evidential proof that modern businesses and organizations are awash with data. Enormous value has been attached to data processing and analysis in an organization. Avantika Monnappa (2000) identified some of the benefits of data science in business organizations. Some of the identified advantages as justified by Avantika Monnappa is that, data science helped to mitigate high risk, used as a control mechanism and reduced fraud. Also, it is being used to create statistical, network and path for foretelling fraud tendency therefore creating early warning and alerts that could possibly helped in ensuring timely interventions responses.

Also, in addition to the claims according to Monnappa (2000), data science can be used to determine the delivering of relevant products in an organization particularly it is being used to determine and develop new products, control marketing of the products and take decision on where products can be sold at the right time and according to the request of the customers' needs. As part of what data science does in an organization, Monnappa (2000) added that data science facilitate the understanding of market analysis and therefore creating the best possible customer satisfactions by the sales or marketing teams.

Other areas of importance of data science in an organizations as indicated by Avantika Monnappa (2000) are listed as follows:

i. Data science do help in empowering management and other relevant departments in an organization to make better, timely and relevant decisions through evaluating, measuring, monitoring, tracking, and recording of performance in such a way to bring the best business outcomes.
ii. Data science also help in defining organizational goals through which an action is being directed based on trends, developments and outputs to increase the institutional performance and to ultimately increase profitability in an organization.
iii. In addition, Data science do help business organizations in identifying and utilization of opportunities in such a way to continuously, consecutively and constantly improve upon the value derived from the organization's data to maintain their success growth.

The Significance of Data Science in Organisation

Data science is a major area of concern in any commercial organizations. It assist organizations in the area of taking strategic planning, foresight, trend analysis and, particularly in executive decision-making (Marcus Buckingham and Ashley Goodall 2015). Richard Daft (2009) in his own point of view, said data science applications in an organizations have been assisting commercial organizations in easing their operational efficiency particularly towards improving commercial products and services (Richard Daft 2009).

The application of data science in commercial organization as recorded by James Suroweicki in 2005 was said to have aided the development and the use of prediction markets. Accordingly, prediction markets are any type of distributed computation where independent assessments is being predicted with accuracy (James Suroweicki, 2005). For example, as explained by Jenna Dutcher, (2014) using Hollywood movie studio, Google and Microsoft. Hollywood movie studio was said to have used prediction markets to forecast also, Google and Microsoft have said to have once used prediction markets also to estimate, assessed and evaluate the demand for new products and services and the completion time and release of complex software development projects (Dutcher, 2014).

Accordingly, one of the major probable prospects of data science in any business organization is to:

i. Monitor the popularity of the organization,
ii. Measure the market value of the products, and
iii. Evaluate the acceptability of the services in organization.

To achieve all the listed in above, Jenna Dutcher (2014) said, sentiment analysis data science application was designed for commercial organizations to monitor and to follow up the popularity of their product, evaluate the acceptability of the services, and to measure the market value of the products in a given organization. These were said to have assisted in tracking the strength of their products, and monitoring market opportunities (Jenna Dutcher 2014).

Data Scientist and Their Roles in Commercial Organizations.

Data scientists are professionals. They are data science professionals with various specializations that run through various potential job categories and career service in an organization. According to RAND researchers, there are four major identified distinct specialties of data science professionals. These are:

i. The data communications specialist (data visualization)
ii. The Data engineer (computer science),
iii. Computational Social Data Scientist, and
iv. The data analyst (statistics).

The various position descriptions according to the distinct specialties of data science professionals is as follows:

The Data Communications Specialist (Data Visualization): The data communications specialist is a categories of job specialization in an organization that responsible and dealing in the communication and presentation of data in an organization in either structured or unstructured form in a format like visual, text-based, written, oral, and graphical communication interactive formats. The data communications specialist work to collect and manipulate the control of data in support of analytic, collection, and managerial activities in such a way that viewers can interact, engaged and make use of. The study according to RAND researchers indicated that the position of data communications specialist requires a strong, solid and vast technical knowledge. RAND researchers (Bradley M. Knopp, Sina Beaghley, Aaron Frank, Rebeca Orrie, Michael Watson 2016).

Data Engineer: Engineer everywhere are always regarded as a technical adviser to any organization. Data Engineer in this respect is a technical adviser in respect to organization databases and every other matters related to data science in an organization, such as data capturing, data processing, streaming, and distribution of data. Data Engineer according to RAND researchers is a specialist that designs, links, develops customized data collection, and

install a retrieval systems to support, protect and back up the data collection, data processing, data exploitation, data analysis, and data dissemination from a very complex datasets.
Accordingly, data engineer are specialist that works to develop software systems required to promote operational effectiveness in an organization. Also, Data Engineer are responsible in giving technical advice in acquisition and development computational resources in an organization and to articulate requirements for specialized hardware and other infrastructure needs of the organization RAND researchers (Bradley M. Knopp, Sina Beaghley, Aaron Frank, Rebeca Orrie, Michael Watson 2016).

Computational Social Scientist*:* This categories of data scientists as recorded by RAND researchers is also known as data computational technical adviser. RAND researchers affirmed that computational social data scientist is a data specialist working to translate social science theory into computational algorithms and formulate a prototype models used to simulate social behavior and processes and analyzing data on social structures, configurations and interactions. Computational social data scientist was said to have specialized in provision of methodological support that could assist in the development, formulation and application of quantitative and computational analytic methods, produce finished analysis through the generation and evaluation of tailored data sets RAND researchers (Bradley M. Knopp, Sina Beaghley, Aaron Frank, Rebeca Orrie, Michael Watson 2016).

Data Analyst: The data analyst is a categories of job specialization in an organization that works on raw data. The data analyst processed a raw data to provide descriptive statistics, probability models, and other quantitative assessments of generated data. This specialist, as RAND researchers put, uses a combination of traditional statistical and machine learning/artificial intelligence techniques to analyze complex data sets in support of analytic, collection, and managerial activities. Accordingly, data analysts works with multiple quantitative and computational methodologies for data collection, sorting, data analysis data interpretation using variety of databases and types. RAND researchers (Bradley M. Knopp, Sina Beaghley, Aaron Frank, Rebeca Orrie, Michael Watson 2016).

Today, research has enable us to identify how data science focuses more on common issues in the analysis of various business data and a significant portion of the work performed by many data scientists. To effect this, Hunter Heyck (2015) said a large number of practitioners have transitioned into data scientists. Amazon, Google, LinkedIn, Facebook, and other internet companies opened job positions for data scientists and established data science teams. These data scientists study and analyze business data to provide services for management decision making. For example, Amazon uses collaborative filtering to generate high-quality product recommendations, and Facebook uses a “People you may know” feature to recommend friend connections. From this point of view, the acquisition of knowledge from business data in order to make decisions is one aspect of data science.

Observations

From this section the study were discussed to make major observations from the content analysis of the secondary data. The submissions here however reflected findings from the extant literature and previous research empirical studies carried out by scholars through relevant studies on the exploration on the significance of data science in organisation. The observations made here provided a clear understanding of the significance of data science in business organisation.
However, as a procedure of this research work, conceptual clarification were made to establish the real understanding of data science. By the concept of data science, the study observed that

data science is a method for transforming business data into assets that can help organizations improve revenue, reduce costs, seize business opportunities, improve customer experience, and more.

Arising from the above, some observations were made to justify the significance of data science in business organisation as follows:

First, the study observed that data science is the methods and technologies used to conduct scientific research through management and utilization of scientific data. Data science has been used to better characterize the data-intensive nature of today's science that is of significant important to organisational development. Arising from the studies conducted by scholars, it was affirmed that many disciplines use data technology to deal with scientific data from their respective areas, particularly in today's organizations.

Accordingly, the study has provided a clear understanding of relevance data science to the scientific data users and, in particular, on how this could benefit the design for cognitive retrieval systems and algorithms specific to scientific data.

Second, the study observed that enormous value was attached to data processing and analysis in an organization. This has helped to determine how to deliver the right products at the right time and that can help companies develop new products to meet their customers' needs. In addition, the study observed that personalized customer experiences according researchers is one of the most buzzworthy benefits and power of data science for it helps to create the ability for sales and marketing teams to understand their audience on a very granular level. With this knowledge, the study noted that an organization can create the best possible customer experiences. Also, from the studies according to previous researchers, the study observed some other roles of data science to include empowering management and officers to make better decisions and to identify opportunities. And, based on several interview conducted by scholars, it was revealed that data science applications has helped to analyze day-to-day activity, both at the program and enterprise levels and particularly with the basis to improve performance.

In addition, analysis from the studies previously carried out by scholars identified four distinct specialties that are required for a complete data science capabilities. Accordingly, the study noted the following data science specialist:

i. Data science engineer (computer science),
ii. data science analyst (statistics),
iii. data science communications specialist (data visualization), and
iv. Computational data social scientist.

The study also noted the roles according to each of the data scientist in organization as affirmed by the scholars as follows:

Data Engineer. The study noted that data engineer are technical adviser in respect to organizational databases and every other matters related to data science in an organization such as data capturing, data processing, streaming, and distribution of data. The study further noted that data Engineer is a specialist that designs, links, develops customized data collection, and install a retrieval systems to support, protect and back up the data collection, data processing, data exploitation, data analysis, and data dissemination from a very complex datasets.

Data Analyst. The data analyst according to the study was said to be a categories of job specialization in an organization that works on raw data. The study noted that data analyst processed a raw data to provide descriptive statistics, probability models, and other quantitative assessments of generated data in an organization.

Data Communications Specialist. According to the studies carried out by the scholars, it was revealed that data communications specialist is a categories of job specialization in an organization that responsible and dealing in the communication and presentation data in an

organization in either structured or unstructured form in a format like visual, text-based, written, oral, and graphical communication interactive formats.

Computational Social Scientist. This categories of data scientist as indicated by researchers is known as Computational Social Scientist. The study noted that the computational social scientist also known as data computational technical adviser. The study further noted that computational social data scientist is a data specialist working to translate social science theory into computational algorithms and formulate a prototype models used to simulate social behavior and processes and analyzing data on social structures, configurations and interactions.

Conclusion/Recommendations

Based on the forgoing, and arising from the number of studies carried out by scholars, a conclusion is therefore drawn here to effect and established that data science is of great significance important in organization which can add value to any organizational setting. Particularly, it was adjudged from the study that one of the major area where data science can be explored in an organization is the area of data science applications which focuses on providing unique, exploitable benefits about their corporate strategies, assisting commercial, public or corporate organizations with strategic foresight, trend analysis, and executive decision-making, organizational behavior and structure, customers, markets, supply chains, and more. Data science has indeed made what was previously impossible or what seems to be impossible possible in today organisation. However, the following recommendations are made:

i. Commercial data science applications that focuses on assisting commercial organizations with strategic foresight, trend analysis, and executive decision-making should be adopted by all commercial organizations for effective and efficient organizational performance.

ii. Internet Networking should be made readily available in any commercial organization with various connected devices that could offer data scientists' opportunities to identify, control, determine, predict, and address any foreseeable anomalies or failures that organization may be subjected to. This is to enable data scientists have an automate system preemptive and predict the likelihood of specific failure modes and to prevent failures or minimize repair times and costs.

iii. Data scientist and data science specialist such as Data science engineer (computer science), data science analyst (statistics), data science communications specialist (data visualization), and Computational data social scientist should be give adequate training and re-training at a regular interval in any given organization to promote organizational development.

Reference.

Aiden, E., and Michel, J. B. Uncharted (2013): Big data as a lens on human culture. Penguin Group (2013).

Avantika Monnappa (2000): Why Data Science Matters And How It Powers Business Value. Simplilearn.com |

Bob Violino in (2018) Contributing Writer, CIO | 22 January 2018 13:00 SAST

Borgman, C. (2013). Scholarship in the networked world: Big data, little data, no data. Oxford: Oliver Smithies Lecture, Balliol College.

Daft, Richard,(2009) *Management*, Boston: Cengage Learning, 2009

Dhar, Vasant, (2013) "Data Science and Prediction," *Communications of the ACM*, Vol. 56, No. 12, December 2013, pp. 64–73.

Dutcher, Jenna,(2014) "Sentiment Analysis Symposium: Uncovering

Human Motivations," blog post, *datascience@berkeley*, March 7, 2014. http://datascience.berkeley.edu/sentiment-analysis-symposium-uncovering-human-motivations

Hal R. Varian, Justin Wolfers, and Eric Zitzewitz (2008), "The Promise of Prediction Markets," *Science*, Vol. 320, No. 5878, May 16, 2008, pp. 877–878;

Hey, T. et al.(2009) The fourth paradigm: Data intensive scientific discovery. Microsoft Research. http://research.microsoft.com/en-us/collaboration/fourthparadigm/.

Heyck, Hunter, "A. M. Turing Award: Herbert ('Herb') Alexander Simon (2016)," web page, Association of Computing Machinery, n.d. As of January 22, 2016: http://amturing.acm.org/award_winners/simon_1031467.cfm

Jian Wang, Jianping Liu, Guomin Zhou, Mo Wang and Lei Shi (2020): How Do People Make Relevance Judgment of Scientific Data? Data Science Journal, 19(1), p.9. DOI: http://doi.org/10.5334/dsj-2020-009

Kelly-Detwiler, Peter, (2013) "Machine to Machine Connections—the Internet of Things—and Energy," *Forbes*, August 6, 2013.

Kenneth J., et al (2008), "The Promise of Prediction Markets," *Science*, Vol. 320, No. 5878, May 16, 2008, pp. 877–878.

Los Alamos National Laboratory, (2016) "GENetic Imagery Exploitation," web page, 2011. As of June 22, 2016: http://www.genie.lanl.gov/ *Forecasting*, Vol. 8, 2011, pp. 169–184.

Marcus Buckingham and Ashley Goodall (2015), "Reinventing Performance Management," *Harvard Business Review*, April 2015, pp. 40–50

May, M. (2014) Life science technologies: Big biological impacts from big data. *Science (Science Products, Product Articles*), (13 June 2014); http://www.sciencemag.org/site/products/ lst_20140613.xhtml.

Murray, Adrienne (2013), "What Is the Value of Healthcare Dashboards?" web page, HealthCatalyst, August 8, 2013. ttps://www.healthcatalyst.com/value-of-healthcare-dashboards

O'Leary, Daniel, (2011)"Prediction Market as a Forecasting Tool," *Advances in Business and Management*

RAND researchers: Bradley M. Knopp, Sina Beaghley, Aaron Frank, Rebeca Orrie, Michael Watson (2016). RAND Corporation, "Robust Decision Making," web page, n.d. As of May 4, 2015: http://www.rand.org/topics/robust-decision-making.htmlRAND Corporation, Santa Monica, Calif. © Copyright 2016 RAND Corporation

Richard Daft, *Management*, Boston: Cengage Learning, (2009), p. 218; Fred Seibel and Chuck Thomas, "Manifest Destiny: Adaptive Cargo Routing at Southwest Airlines," Uncluttered.com, 2000;

Suroweicki, James,(2005) *The Wisdom of C2rowds*, New York: Anchor, 2005.

Thor Olavsrud (2019). What is data science? Transforming data into value: CIO | 3 July 2019 12:00 SAST