Emerald Information Technology & People

Key competencies for big data analytics professions: a multi-method study

Ajax Persaud

University of Ottawa, Canada, ajax@telfer.uottawa.ca; DOI (10.1108/ITP-06-2019-0290)

Abstract

Purpose: This study aims to identify the precise competencies employers are seeking for big data analytics professions and whether higher education big data programs enable students to acquire the competencies.

Design/methodology approach: The study utilizes a multi-method approach involving three data sources: online job postings, executive interviews, and big data programs at universities and colleges. Text mining analysis guided by a holistic competency theoretical framework were used to derive insights into the required competencies.

Findings: We found that employers are seeking workers with strong functional and cognitive competencies in data analytics, computing, and business combined with a range of social competencies and specific personality traits. The exact combination of competencies required varies with job-levels and tasks. Executives clearly indicate that workers rarely possess the competencies and they must provide additional training.

Limitations: A limitation is our inability to capture workers' perspectives to determine the extent to which they think they have the necessary competencies.

Practical implications: The findings can be used by higher educational institutions to design programs to better meet market demand. Job seekers can use it to focus on the types of competencies they need to advance their careers. Policymakers can use it to focus policies and

investments to alleviate skills shortages. Industry and universities can use it to strengthen their collaborations.

Social Implications: Much closer collaborations between public institutions, educational institutions, industry, and community organizations are needed to ensure training programs evolve with the evolving need for skills driven by dynamic technological changes.

Originality: This is the first study on this topic to adopt a multi-method approach incorporating the perspectives of the key stakeholders in the supply and demand of skilled workers. It is the first to employ text mining analysis guided by a holistic competency framework to derive unique insights.

Keywords: Information workers; Multi-method; Competence; Knowledge; Skills; Big Data Analytics, Educational Programs; Text mining; STEM-B.

Introduction

The term 'big data' refers to large and complex datasets requiring advanced analytical techniques and technologies to capture, store, manage, search, share, analyze, and visualize (Chen et al., 2012). These datasets are created through a variety of technologies such as mobile devices, remote sensors, software logs, and social media (Sharma et al., 2014) and are generated from a variety of sources including past and present, structured and unstructured, formal and informal, and social and economic (Bhimani, 2015). The bulk of big data are unstructured (e.g., text, image, audio, video), generated by individuals, and expressed in ambiguous natural language (Müller et al., 2016).

Further, big data analytics (BDA) spans many disciplines including information science, engineering, computer science, mathematics, social science, system science, psychology, management, business, and economics (Jin et al., 2015). It also employs techniques from many fields including probability theory, machine learning, statistical learning, computer programming,

data engineering, pattern recognition, data visualization, data warehousing, and high-performance computing (Jin et al., 2015). It also involves a wide range of new technology hardware, software applications and services.

BDA has the potential to generate insights that can lead to innovation and fundamental shifts in decision-making, productivity, growth, competitiveness, performance, efficiency, commerce, service delivery, cost control, and customer value (Bughin, 2016; Manyika et al., 2011). However, many companies find it challenging to harness the value from BDA (Ransbotham et al., 2016) due to shortages of skills and uncertainties around how best to deploy BDA and transform insights into value (Mazzei and Noble, 2017). Indeed, these difficulties underscore the idea that more data and superior analytical methods do not necessarily equate to better insights or greater value (Boyd and Crawford, 2012) or the era of computational prowess does not obviate the need for intuition and creativity (Hayashi, 2014). In this regard, Bob Muglia, a former Microsoft executive and current CEO of Snowflake Computing, commenting on big data technologies and specifically, Hadoop, the technology that sparked the big data revolution, remarked that, “I can’t find a happy Hadoop customer. It’s sort of as simple as that … the number of customers who have actually successfully tamed Hadoop … might be less than 10 … that’s just nuts given how long that product, that technology has been in the market and how much general industry energy has gone into it.” (Woodie, 2017). But technology is not all. According to a Gartner report, most big data projects fail because of the expertise gap – workers lacking data skills and management lacking big data leadership and capabilities (Axryd, 2019). Faced with this challenge, executives are seeking business-minded data scientists and workers because “successful big data project starts from a deep understanding of the business problems you want to solve and the value you want to gain” (Agarwal, 2017).

These challenges gave rise to a widespread view that BDA requires a different set of technical, managerial, and analytical skills than with predecessor technologies (Comuzzi, 2016; Dubey and Gunasekaran, 2015). The multidisciplinary nature of BDA has also led many to argue that performing analyses on and deriving insights from such voluminous and complex data require not only scaling up of hardware and software platforms (Singh and Reddy, 2015) but also configurations of people with new skills (Ransbotham et al., 2016). Some scholars even argue that BDA will reshape work and the value of skills by deskilling certain workers (Autor, 2015) and requiring workers with new and different skills (Schildt, 2017; Vidgen et al., 2017).

This view strongly suggests that people who are seeking careers in BDA must acquire specific new skills and talents in order to progress in their chosen profession. Additionally, industry analysts have stated that the difficulties in hiring employees with BDA-specific skills have been particularly acute for many organizations (MacDonnel and Castro, 2016; ICTC, 2016). According to Ransbotham et al. (2016), talent is one of the least understood impediments to BDA adoption. Cegielski and Jones-Farmer (2016) found considerable ambiguities and uncertainties among employers as to the precise skills and capabilities needed for BDA careers. De Mauro et al. (2018) went further arguing that the skills and roles of BDA jobs are often nebulous and nearly-mythological citing the role of data scientists as an example. Clearly, organizations are still unclear as to what skills and talents they need for big data professions.

Thus, two very relevant and significant practical questions for which there is very scant academic evidence are: (1) what precise competencies are required for BDA professions? and (2) are higher educational institutions providing students with the required competencies? This study investigates these two questions by analyzing three different qualitative datasets: (1) BDA job

postings, (2) educational programs offered by universities and colleges (higher education dataset), and (3) interviews with senior executives with extensive experience in BDA.

The job posting data was obtained through a combination of web crawling (Olston and Najork, 2010) and manual downloads where web crawling was not possible due to technical limitations. The higher education dataset was obtained by manual download given the lack of consistency in the page structures of each university, college, faculties and programs. Personal interviews with senior executives were conducted via telephone by a professional research organization. The datasets were analyzed via text mining. For the analysis, we employed a competency-based theoretical framework in order to gain insights on the competencies required for big data professions.

The multi-method approach employed in this study combined with text mining analysis is unique since prior studies have relied primarily on a single dataset, mainly job postings, which were analyzed primarily through manual coding. The inclusion of the higher education dataset provide concrete insights into the relevance of the supply-side competencies available, a research-focus requiring further investigation (Miller, 2014; De Mauro et al., 2018; Song and Zhu, 2016). The inclusion of senior executive interviews is intended to corroborate the findings from the job postings data (demand-side) and the higher education data (supply-side). Further, unlike traditional manual content analysis where the researcher develops the themes, text mining generates themes through a computer algorithm and the researchers then provide further explication of the themes based on domain knowledge (Marjanovic and Dinter, 2017). This way, researcher bias are avoided (Cretchley et al., 2010). Moreover, text mining facilitates examination of “very large samples of text” and provides improved “internal, external, construct and statistical conclusion validity” over

traditional methods (McKenny et al., 2018: 2910). Additional details concerning the composition of the datasets and the text mining approach are provided in the methodology section of this article.

This study complements and extends the emerging literature on this topic in several ways. Specifically, our study provides empirical evidence that can help clarify the precise competencies employers are seeking for BDA professions and the extent to which universities and colleges provide opportunities for students to acquire the required competencies. Moreover, insights from senior executives will not only help corroborate any observed supply-demand skills mismatch but may also reveal practical strategies for bridging the skills gap. From a practical perspective, the findings from this study can guide the BDA recruitment, training and retention efforts of organizations. The findings can also point to areas where universities and colleges need to extend or strengthen their existing suite of programs to better meet market demand. Further, the findings can help job seekers understand and focus on the types of competencies they need to acquire in order to secure a job or advance their careers. From a policy perspective, the findings can serve as a guide to policymakers regarding areas of deficiency in the workforce and where they may focus policies and investments to help alleviate skills shortage.

Related work

Many scholars posit that BDA requires specific technical and managerial talents, which are seen as two sides of the same BDA coin (Gupta and George, 2016; Power, 2016). According to Cegielski and Jones-Farmer (2016), technical skills refer to the know-how required to use new forms of technologies, infrastructure, programming paradigms and analytics methods to extract and visualize intelligence from big data. Some of these skills include competencies in machine learning, artificial intelligence, data extraction, data cleaning, statistical analysis, and other

computational methods. Technical analysts focus primarily on capturing and processing data in order to generate insights. Managerial skills are needed to ensure that the insights extracted from data are valued and acted upon (Lycett, 2013; Sharma et al., 2014). On the flip side, technical analysts must be able to understand the questions posed by managers seeking insights to inform decision-making and business strategies (O'Reilly and Paper, 2012). Indeed, it appears that not only managers but workers at all levels are now expected to possess both technical and business competencies in order to work and progress in BDA environments (Gupta and George, 2016; Sharma et al., 2014).

In an attempt to signal the new competencies required for BDA careers, many new social constructs have been created. These include Chief Data Officer, Data Scientist, Data Architect, Data Engineers, Data Analyst, and Decision-support Designers (Cegielski and Jones-Farmer, 2016). According to Power (2016), the Data Scientists should be able to perform descriptive, diagnostic, predictive, and prescriptive analyses. They should learn many new skills and understand industry- and customer-specific analyses and systems. Additionally, Davenport and Patil (2012) suggest that data scientists should also have the skills of a database designer, software programmer, statistician and storyteller to explore the new data streams for decision-relevance. Moreover, they are seen as someone responsible for managing their organization’s data asset and must be able to provide vision and strategy for data management initiatives.

In assessing the expectations of many BDA jobs, De Mauro et al. (2018) contend that the newness of the many the new job roles have resulted in substantial unrealistic expectations and confusion among employers, current employees and job seekers as to the precise competencies needed for BDA professions. Citing the role of the Data Scientist, De Mauro et al. (2018) conclude that this role is “nearly-mythological” (p. 808) because of the expansive skills and responsibilities

expected of this single role. In addition to these specific job titles and roles, big data professionals have been classified into several categories. For example, Cegielski and Jones-Farmer (2016) suggest three categories, namely, deep knowledge analysts, data-savvy users, and technology support. Deep knowledge analysts must have advanced statistical training to perform analysis on large datasets with new applications; data-savvy users must be capable of interpreting data and making operational and strategic decisions that drive business, and technology support specialists possess technology skills to design, develop, and maintain BDA hardware and software. Table 1 provides further elaboration of BDA skills and capabilities emerging from recent studies.

Insert Table 1 here

Summarizing, it seems that the explosive growth of BDA combined with the power of the Internet, social media and mobility have fundamentally altered the nature of work from the C-suite to the shop floor (O'Reilly and Paper, 2012; Vidgen et al., 2017). The deskilling and reskilling impact of these technological changes (Autor, 2015; Schildt, 2017) have resulted in considerable ambiguity regarding the precise skills needed for BDA professions (De Mauro et al., 2018). The current state of knowledge, which is based primarily on job advertisements, seems like a laundry list of knowledge, skills, and capabilities. Moreover, most prior studies (e.g., Cegielski and Jones-Farmer, 2016; Reaney, 2014; Davenport and Patil, 2012) focus on functional competences, that is, those things that 'a person who works in a given occupational area should be able to do. . . [and] able to demonstrate' (Le Deist and Winterton, 2005: p. 33). While this approach may have been useful in the early phases of BDA, its usefulness can be enhanced by using a more holistic theoretical framework to identifying relevant competencies based on multiple sources of

information. Such an approach could add greater rigor and clarity when identifying relevant competencies. This study adopts such an approach.

Theoretical background

Prior research in information technology (IT), strategy, and management reported strong associations between firm performance and the nature of its resources, capabilities, and competencies. Specifically, the literature on the resource-based view of the firm (Barney, 2001), IT capabilities (Bharadwaj, 2000), and IT competencies (Dehning and Stratopoulos, 2003; Tippins and Sohi, 2003) underscore the importance of human resources (Croteau and Bergeron, 2001) along with technologies, IT infrastructures, and other organizational capabilities in enabling firms to achieve superior performance and sustained competitive advantage (Constantiou and Kallinikos, 2015; Tippins and Sohi, 2003; Coltman et al., 2011; Croteau and Raymond, 2004). Similarly, the core competence approach (Prahalad and Hamel, 1990) shows that firm performance is driven by the complex interactions of people, skills and technologies (Scarbrough, 1998). According to Vidgen et al. (2017), BDA is a business transformation undertaking involving many organizational functions but IT can serve as a key enabler since BDA is multidisciplinary and a substantial component is technical in nature involving hardware, software, technology infrastructure, programming, and a wide array of computational capabilities (Akter et al., 2016; Sharma et al., 2014; Woerner and Wixom, 2015; Wamba and Mishra, 2017).

Since the focus of this study is on knowledge and skills of BDA, we have adopted a competency-based framework to guide our analysis. A competency framework is typically viewed as “a descriptive tool that identifies the skills, knowledge, personal characteristics, and behaviors needed to effectively perform a role in the organization (Lucia and Lepsinger, 1999)” (Le Deist

and Winterton, 2005: p. 33). Further, Le Deist and Winterton (2005) note that competency models have been widely used to align individual capabilities with the core competency of the organization (Rothwell and Lindholm, 1999). Moreover, competency is more than just knowledge and skills; it includes attitudes, behaviours, work habits, abilities and personal characteristics (Gangani et al., 2006; Le Deist and Winterton, 2005; Russ-Eft, 1995). Competency is displayed not in isolation but situated in the kinds of contexts in which they are central to the practice of the profession (Hager and Gonczi, 1996; Russ-Eft, 1995).

A competency approach is employed because it signals labor market needs, helps individuals understand what are need for career mobility (Le Deist and Winterton, 2005), enhances the scope for educational providers to better integrate education and training with labor market needs (Hager and Gonczi, 1996), and shows the synergy between formal education and experiential learning in developing professional competence (Cheetham and Chivers, 1996). In short, a competency approach sheds light on what competencies organizations are seeking, how educational institutions can design programs to meet those needs, and what competencies individuals need to possess to progress in their profession. As noted earlier, there is considerable ambiguity regarding the competencies required in the ever-changing technological environment.

We have adopted the holistic competency model proposed by Le Deist and Winterton (2005), which is depicted in Figure 1 below. This framework provides a comprehensive view of the multidimensionality of competencies that are often required in the context of rapid and sustained technological innovation (Le Deist and Winterton, 2005). The graphic reflects the unity of competence and the practical difficulty of separating the cognitive, functional and social dimensions. Meta-competence is an over-arching input that facilitates the acquisition of one or more of the other competencies. Additionally, the competencies required of an occupation are

invariably described in multidimensional terms (Le Deist and Winterton, 2005). This integrated framework enables us to better classify and makes sense of the long and often conflicting laundry list of knowledge, skills, abilities articulated in job postings.

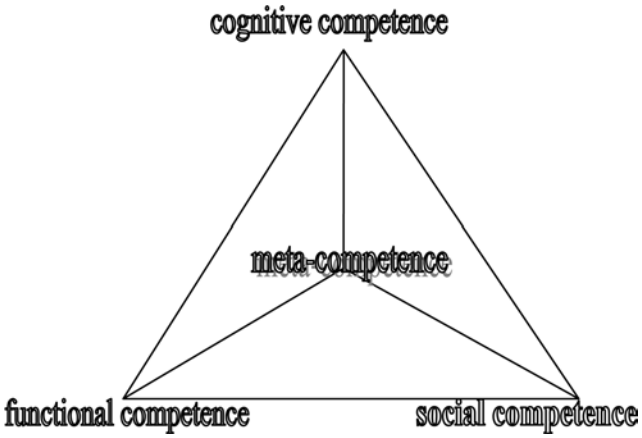

Figure 1. Holistic model of competence.

Source: Le Deist and Winterton (2005)

Each of the four competencies are now described. Functional competence refers to the ability to perform a range of activities in the jobs within an occupation to produce specific outcomes or to demonstrate performance to the standards expected in employment (Cheetham and Chivers, 1996; Le Deist and Winterton, 2005). This requires the possession of discrete skills but the emphasis is on putting these to use to achieve specific outcomes. Essentially, functional competence describes the willingness and ability to employ subject-specific knowledge and skills to carry out tasks and solve problems (Le Deist and Winterton, 2005).

Cognitive competence is defined as knowledge, skills and abilities needed in the specific occupation for mastering tasks and developing appropriate problem-solving strategies. It is the ability to think and act in an insightful and problem-solving way (Le Deist and Winterton, 2005) and being able to apply knowledge in a variety of ways (Cheetham and Chivers, 1996). Cognitive

competence focusses on mastery of the technical and theoretical aspects (principles, theories, etc.) underlying the knowledge base of the profession as well as their application (Cheetham and Chivers, 1996). It also includes tacit, practical and contextual knowledge, which is specific to an organization, industry, or sector (Cheetham and Chivers, 1996).

Social competence is defined as the ability and willingness to cooperate with others in a group, to experience and shape relationships, to identify and understand benefits and tensions, and to interact with others in a rational and conscientious way (Cheetham and Chivers, 1996; Le Deist and Winterton, 2005). Meta-competencies refers to the ability to cope with uncertainty, learning and reflection (Cheetham and Chivers, 1996). It includes personal characteristics such as independence, self-confidence, self-development, creativity, critical thinking, communication, responsibility, reliability, awareness of duty, and professional and ethical values (Le Deist and Winterton, 2005). Further, meta-competence is different from the first three dimensions since it is concerned with facilitating the acquisition of the other substantive competencies (e.g. self-development, creativity) or are capable of enhancing or mediating competence in any or all of the component categories (Cheetham and Chivers, 1996). It is also useful when working in multidisciplinary or team environments.

Methodology

This study employs a multi-method design that considers multiple perspectives, which enhances confidence that the results obtained from multiple methods are valid and not a methodological artefact (Johnson et al., 2007). In this study, we investigate our research questions using three different datasets where each source provides unique insights. The first dataset consists of BDA postings job, the second consists of programs in BDA offered by universities and colleges (higher

education dataset), and the third consists of interviews with senior executives of firms that utilize BDA in their organizations. Each dataset is now described in further detail.

Job postings are considered useful sources of information because they provide insights into what knowledge, skills, and capabilities organizations perceive as being most important for a particular position. Moreover, job postings have a long history of use in IT research to infer job skills (Cappel, 2002; Sodhi and Son, 2010; Todd et al., 1995). For this study, the job titles, descriptions, qualifications, and responsibilities of BDA positions were downloaded from four widely-used job posting websites - Indeed.com, LinkedIn.com, Monster.com, and Procom.com. These websites were selected because they have been used in prior big data research (Cegielski and Jones-Farmer, 2016; De Mauro et al., 2018) and big data vendors and job seekers ranked them as the largest and most popular for big data jobs. The information was gathered through web crawling and manual downloads where web crawling was not feasible due to technical issues. Job titles culled from the BDA literature and the popular press were employed to guide the job search process. Job titles used in the search include big data engineer, data scientist, machine learning expert, data mining expert, data analyst, data architects, big data consultants, applications developer, chief data officer, cloud and artificial intelligence specialists, business analysts, and business intelligence analysts. A total of 3,009 job postings were used in the analysis after removing duplicates. Approximately 43 percent of the jobs were drawn from Indeed.com and 32%, 16%, and 9% were drawn from LinkedIn.com, Monster.com, and Procom.com respectively. We believe that this sample of job postings provide reasonable depth and breadth for assessing the competencies required in the respective jobs. Two graduate students with business and engineering backgrounds assisted with the data collection, cleaning, and pre-processing. The students independently collected and cross-checked the data. Table 2 provides a distribution of the jobs.

The sample list of the job titles shown in Table 2 below indicate that the postings consist of entry-level, mid-level and managerial-level positions.

Insert Table 2 here

The second dataset consists of the undergraduate, graduate, and professional programs in BDA offered at 45 Canadian universities and 16 colleges. The goal was to include all the major universities and colleges in Canada, ensure adequate geographic representation as far as possible, and the universities and colleges must offer identifiable big data analytics programs and courses. Like job postings, educational programs have been used in early IS studies to assess fit between academic preparation and industry expectations (Lee et al., 1995; Nelson, 1991). Our search concentrated on five main discipline areas, commonly referred to as STEMB - Science, Technology, Engineering, Mathematics, and Business since these are currently at the forefront of research and programs in BDA (e.g., De Mauro et al., 2018; Gardiner et al., 2018). The data was collected through manual search and download due to the wide variations in page structures (Olston and Najork, 2010) across universities, colleges, and even faculties. Three graduate students with business and engineering backgrounds assisted with the data collection, cleaning and preprocessing. The students independently collected and cross-checked the data. Appendix 1 provides a list of the universities and colleges used in the study as well as a sample of their programs for illustration purposes, however, the full sample of programs and courses identified were used in our analysis.

The third dataset consists of telephone interviews with senior executives to gain additional insights that are not easily captured in job postings or educational program descriptions and to

corroborate our findings from the other two datasets. This type of data has been used in prior IS research to ascertain the congruency of perceptions regarding skills requirements (Lee et al., 1995; Leitheiser, 1992). Interviews were conducted with 21 senior executives (e.g., vice presidents, senior directors, directors, and senior managers) who are very deeply involved in the BDA initiatives of their organizations. These individuals represent a wide cross-section of industries including banking, insurance, manufacturing, gaming, retailing, transportation, health care, and telecommunications. All the executives have close to a decade of managerial and technical experience with BDA and several have worked in multiple organizations in senior positions. Thus, we consider this group to be highly knowledgeable and experienced.

The executive interviews were conducted by an independent third-party professional research firm. The interviews lasted between 20-40 minutes and were recorded and transcribed. The interviews focused on a range of issues including the ease of recruiting and retaining BDA workers, the extent to which workers and new recruits have the necessary competencies or require additional training, expectations regarding the competencies that recruits should possess or must have in order to be hired, and extent to which recruits have a clear understanding of the competencies required for BDA professions.

The job posting and higher education datasets were analyzed via text mining because text mining facilitates examination of very large samples of text in order to measure constructs, identify semantic structures and discover hidden patterns (McKenny et al., 2018). It also provides improved "internal, external, construct and statistical conclusion validity" over traditional methods (McKenny et al., 2018: p. 2910). The text mining analyses were performed with Leximancer 4.5, a software that has been extensively evaluated for reproducibility, correlative validity and stability of the underlying statistical algorithms (Cretchley et al., 2010). It is widely used in academic and

practitioner-oriented studies including IS (Indulska et al., 2012; Marjanovic and Dinter, 2017). Leximancer is a sophisticated text mining software that analyzes the frequency of co-occurrences of words within blocks of texts to identify semantic concepts in the texts. Based on their derived semantic proximity, the concepts are grouped into themes, which are then visualized in concept maps represented by colored circles. The sizes and brightness of these circles correspond to the occurrences of identified themes within texts (Cretchley et al., 2010). The concept maps establish the relational strengths among different concepts so that it could be used by the researchers to interpret the strength of different associations (Marjanovic and Dinter, 2017). Unlike software that strictly calculate word frequency or suggests significance based on a predetermined dictionary or manual coding, Leximancer's algorithms are designed to intuit significance based on semantic relationships present in the text. Leximancer generates a model that can be used by the researcher to efficiently conduct a sensemaking analysis of the large samples of text (Cretchley et al., 2010; Smith and Humphreys, 2006).

Results

This section presents the results of our analysis. Figure 2, Panel A below depicts the concept map generated by Leximancer for the job description dataset. The colored bubbles represent the most relevant themes from the job descriptions and the dots represent concepts that are related to specific theme. The eight themes along with a sample of the concepts comprising each theme are displayed in Figure 2, Panel B. To explore each theme more closely, the relevant texts for each theme were examined more closely in order to understand how the concepts together constitute a coherent theme.

Insert Figure 2 here

Close examination of the *data* and *computer science* themes unmistakably suggests that employers are seeking individuals who have strong technical knowledge, understanding and experience with big data infrastructure technologies, programming languages, computing and statistical methods, machine learning, artificial intelligence, and cloud computing. They must also demonstrate understanding of the practical applications of these technologies in various industries. This result point clearly to the possession of strong cognitive (technical) and functional competencies for BDA jobs. Detail inspection of the job postings indicate that even for entry-level technical positions, applicants are required to have an undergraduate degree in a quantitative discipline such as computer engineering, computer science, machine learning, mathematics, and statistics coupled with a minimum of between 2 to 3 years industry experience. Many mid- and senior-level technical positions require advanced degrees (Masters or Ph.D.) in a relevant technical discipline with 5 or more years of technical and managerial experience. Basically, applicants for technical positions must be able to demonstrate deep understanding of both the hardware and software side of big data as well as the computing, statistical and analytical processes and results.

Additionally, postings for senior analysts, middle- and senior-managers seek applicants with a demonstrated ability to work effectively with a wide range of technical and business specialists to generate useful insights from data. Not surprisingly, senior analysts and managers are expected to also demonstrate many of the non-technical skills at much higher levels.

The *information* theme emphasizes the ability of workers to sense, gather, and convert data and information into accurate, timely, actionable and impactful intelligence for decision-making. Such intelligence can enable management to configure their firms' resources and capabilities, which can enhance firm competitiveness and performance. Further, they must demonstrate an ability to follow enterprise policies and mandates (regulatory, legal and ethical) concerning data

and information integrity, sharing, and use. The *source* theme emphasizes the ability of workers to recognize the value of open source data and to use open source information for innovation and competitive advantage. This unique finding of our study is consistent with the information management capability literature, which posits that the 'management of information' like the 'management of technology' must be treated as an asset that positively influences competitive advantage and business performance (Mithas et al., 2011; Coltman et al., 2011). These two themes also emphasize cognitive and functional competencies.

With respect to the *business* theme, it seems that companies are seeking individuals who can demonstrate deep knowledge and understanding of the technologies and environment of their industry and who are capable of working in teams to support the firm's operations, customers, suppliers, and other stakeholders. Additionally, they must have the ability to apply business intelligence and insights to develop new projects and business solutions that benefit both the firm and its customers. These competencies are further clarified by the key concepts comprising the *customer* theme, which indicate that job applicants must demonstrate the ability to build strong relationships with customers, develop solutions to customers' challenges, deliver valuable and impactful solutions to customers' needs; and committed to providing the highest level of customer care. The following excepts from the job postings provide insights into the business and customer themes:

⇒ *Provide timely and quality responses to customers*
⇒ *Help our customers achieve their goals*
⇒ *Build resilient partnerships that drive growth*
⇒ *Passionate about helping your customers maximize their benefits*
⇒ *Work with stakeholders across the organization to support data-driven business solutions*
⇒ *Work with internal and external business customers, and subject matter experts to analyze and document needs and requirements in a way that is intuitive for downstream development effort*

The *skills* theme reflects most of the business communication skills that employers are seeking since they provide insights into the ability of job candidates to function effectively within the organization and to work with their clients and partners. Some of these skills include oral and written communications, presentation, problem-solving, organization and time-management. Additionally, the ability and willingness to perform at a high level in a team setting and to interact with others in an honest, reliable and flexible manner are extremely important. The preceding analysis suggests the *business, customer*, and *skills* themes have very strong social emphases, which aligns closely with the social competency of our framework. The following excerpts taken from the job postings provide specific insights into nature of these skills (Table 3).

Insert Table 3 here

The *people* theme indicates that employers are seeking employees with a wide range of personal characteristics such as independence, critical thinking, creativity, empathy, self-confidence, and self-development. They must also demonstrate intellectual curiosity, a willingness to learn, and being passionate in pursuit of their goals and duties. It is clear that these characteristics are also closely aligned with many of the social and meta-competencies of our competency framework.

In sum, it seems that the four themes – *data, computer science*, *information,* and *sources* – point to skills, abilities, and behaviors that are consistent with the cognitive and functional competencies of our analytical framework. Similarly, the *business, customer,* and *skills* themes fit closely with the social competence dimension while the *people* theme aligns well with the meta-competence dimension. The overlaps among the themes are also consistent with the integrated framework, which underscores that employers always seek multiple competencies in job

applicants even though one or two dimensions may dominate depending on the nature of the job. The fit between the themes and the analytical framework enable a clearer understanding of the rationale behind some of the expansive job requirements and how the different requirements fit together. The competency framework provides a clearer context for understanding the expansive requirement relative to the laundry list approach.

Dataset 2- University and Colleges Programs

Figure 3, Panel A below depicts the text mining concept map generated by Leximancer for dataset 2 comprising descriptions of the BDA programs offered by universities and colleges. Panel A shows the six key themes that emerged from the data and Panel B shows the concepts associated with the themes. The six themes are *students, data, computer, program, research* and *industry*. These themes are discussed next.

Insert Figure 3 here

A content analysis of the relevant texts for the *program* theme indicate that universities and colleges are responding to the demand for BDA professionals by not only offering specific programs in BDA but are also adding various BDA specializations and options into existing curricula. Additionally, the programs and courses are designed to provide students with both core technical knowledge, some soft skills, and practical experience through internships, co-ops, and industry-based research projects. Also, some programs are offered jointly between STEM and Business disciplines and have a distinctive decision-making focus. Excerpts of program descriptions presented in Table 4 provides a glimpse of to the range of BDA competencies being offered. Table 5 in Appendix 1 provides a more comprehensive illustration of the range of program

and courses offered by various educational institutions. Clearly, universities and colleges recognize that employers are looking for analytics professionals with solid technical expertise and excellent business acumen and have tried to design their programs to provide a fair mix of knowledge and skills.

Insert Table 4 here

The *industry* theme shows that many programs are developed and delivered in collaboration with industry in various ways including curriculum development, internships, mentorships, knowledge exchanges with experts, industry-based research projects and industry certification bodies. Several programs offer certification of competency with various business analytics tools and technologies such as SAS™, SPSS, and Tableau. Additionally, some programs offer internships with leading technology firms such as Amazon, Google, and IBM as well as with data-intensive organizations such as hospitals, banks, and retailers.

Finally, the *computer* and *data* themes provide insights regarding the wide range of programs and courses offered. The content of these two themes represent close matches to the *computer* and *data* themes of the job postings. Analysis of this dataset shows that they are designed to give students strong theoretical grounding and as much practical and/or industry exposure as is feasible in a wide variety of areas dealing with tools, technologies, infrastructures, platforms, programming languages, analytics software, and visualization technologies. They are also exposed to many tools that are used in the big data industry including SQL, R Studio, SAS™, and Oracle Warehouse Builder (OWB).

The *research* theme describes the research interests of professors and researchers, the various research centres, institutes, and labs available to support BDA research and programs, and

industry-university research partnerships in the area of BDA and related technologies. The *research* theme generally engages students in advanced degree programs.

Finally, in terms of our competency framework, the results suggest that most BDA programs have an overwhelming bias towards imparting technical and functional competencies. Team work and industry collaborations are the principal ways in which students get opportunities to build their social competence. Little information was observed in the programs in terms of building meta-competence such as personal development, creativity, critical thinking, empathy, communication, responsibility, reliability, awareness of duty, and professional and ethical values.

Dataset 3 - Executive Interviews

Senior executives across a wide variety of industries were asked to share their views and experiences in terms of criteria for hiring BDA workers, the extent to which job applicants demonstrate adequate knowledge, skills, and abilities, and the extent to which they provide additional training beyond familiarizing new employees with the organizations rules, routines, norms, and expectations. The overwhelming sentiment is that all applicants must meet the technical and functional competence of the job. However, this is just the first screening. Beyond this, employers are looking for workers that have demonstrated practical experience within their industry or a related industry and excellent soft skills e.g., social and communication skills coupled with personal qualities such as being passionate about their profession, career, personal development, creativity, willingness to engage in continuous learning, and a strong work ethic. In particular, several executives stated that among equally qualified candidates, they tend to give preference to those who demonstrate interest in expanding their knowledge to other unrelated disciplinary areas, industries, and even countries. According to one executive, this “cognitive

flexibility" tends to make such employees more versatile, innovative, willing to experiment and "go across the grain" or question establish ways of doing things, which can fuel business opportunities.

In terms of readiness, virtually all of the executives indicated that the majority of recruits need additional training, which are offered by their organizations either in-house or externally. They also lament the fact that many applicants who meet or even surpass the technical requirements are not recruited because they lack many of the soft skills (i.e., social- and meta-competencies) and practical business and industry experience. It is clear from the executive interviews that while social- and meta-competencies are highly valued, they are insufficient by themselves to result in hiring. However, cognitive and functional competences combined with strong social- and meta-competence significantly enhances the prospect of being recruited.

Finally, in terms of fit between skills demanded and offered to students by universities and colleges, many executives stated that students are offered the right mix of technical skills for most BDA jobs. However, there is a serious imbalance in business acumen and soft skills (social- and meta-competence). Many stated that they understand that higher educational institutions are hamstrung in their ability to offer the breadth and depth of soft skills needed but they expect students to take the initiative to strengthen or acquire these competencies.

Discussion & Implications

This study endeavours to develop a deeper understanding of the types of competencies employers are seeking for BDA professions and whether universities and colleges offer programs that give students the opportunity to acquire the requisite competencies. The goal was to understand the extent of fit between the demand-side and the supply-side for job talents. Interviews with industry

executives were conducted in order to derive practical corroborative insights on the extent of this fit. The multi-method approach, comprising of data from job postings, educational programs, and company executives, combined with the use of a holistic integrated competency model enabled a clearer understanding of the range of competencies required and the extent of readiness of workers for various BDA professions.

Our text mining analysis generated four broad categories of knowledge, skills, and abilities that employers are seeking for big data professions. These can be summarized as *data analytics, computing, business, and soft skills*. Within the context of our theoretical model, the first two - *data analytics* and *computing* - focus on technical domain knowledge and skills, which correspond to the functional and cognitive competencies of our framework. The latter two - *business, and soft skills* - focus on organizational, business, and management skills and personality traits, which fits into the social- and meta-competencies dimensions of our model. Interviews with executives indicate that they are seeking individuals with a range of solid functional and cognitive competencies complemented with excellent social competencies and personality traits and abilities (meta-competencies).

Additionally, careful examination of the job postings data indicates that the breadth, depth, and combination of competencies required vary by job types and levels. However, there seems to be a recognition that all workers, regardless of job level or job type, are required to demonstrate a basic combination of all four competencies. That is, they must demonstrate proficiency in the technical, analytical, and business aspects of BDA (tools, technologies, platforms, analytic methods) as well as solid communication, relationship-building, and creativity skills. Beyond this basic level, the specific depth, breadth and combination of the four competencies depends on the job tasks and the level at which the individual functions. For instance, workers at the operational

level such as data engineers, data analysts, data architects, and business analysts require very different combinations of functional, cognitive, social, and meta-competencies since their jobs are more technical and less managerial compared to those of chief data scientists and chief data officers, which require higher levels of both technical and managerial sophistication for strategic decision-making.

The above finding, on the surface, seem to suggest that employers are seeking a laundry list of knowledge, skills, and abilities or are creating near-mythical roles (De Mauro et al., 2018), however, interviews with executives indicate the opposite; they are seeking workers with a more coherent and rounded set of mutually reinforcing competencies. Several executives stated that exclusive focus on technical competencies, which emphasize functional and cognitive skills, although necessary, is insufficient. On the flip side, solid social- and meta-competencies that demonstrate effectiveness in working with internal and external organizational stakeholders coupled with key personality traits but without the requisite technical competency will also be inadequate. These executives noted that workers with excellent technical skills but lacking solid social skills and personality traits are often overlooked in hiring while those lacking the requisite technical skills are screened-out very early on in the recruitment process. They also stated that they give greater consideration to applicants with significant practical experience in a technical discipline and industry with evidence of creativity, continuous learning and cognitive flexibility i.e. able to apply their knowledge creatively across multiple domains to identify and drive business value for the organizations and its clients.

The results of our analysis are consistent with those reported in other studies which shows that even for entry-level technical jobs, employers are seeking individuals not only with deep technical knowledge in a wide range of big data technologies, systems, applications, platforms,

and programming languages but also a range of soft skills (e.g., Cegielski and Jones-Farmer, 2016; De Mauro et al., 2018). The corollary is true for non-technical jobs such as business analysts, financial analysts, and business intelligence analysts. Applicants for these jobs, including entry-level positions, must demonstrate strong familiarity with several BDA technologies, tools, and analytics methods which are used to gather data and generate insights. This facilitates effective dialogue and collaboration between technical and on-technical staff.

We also note that our findings are closely aligned with and lend support to the emerging literature on T-shaped and Pi-shaped knowledge and skills (Demirkan and Spohrer, 2015; Demirkan and Spohrer, 2018; Donofrio et al., 2010; Karjalainen et al., 2009). In the T- and Pi-shaped literature, professionals have deep knowledge or are experts in one or two disciplines but have good knowledge and experience of several other disciplines (Demirkan and Spohrer, 2015; Iansiti, 1999; Karjalainen et al., 2009). They are life-long learners, deeply engaged critical thinkers, seek challenges, excellent communicators, open-minded, and can collaborate easily (Demirkan and Spohrer, 2015). It is clear from the job postings and executive interviews that organizations are looking for workers with both depth and breadth of functional (technical and business), social, and meta-competencies since these "allow for enhanced adaptive and innovation capacity … collaboration and teamwork [that] support the kind of close customer engagement required" for BDA professions (Demirkan and Spohrer, 2015: p. 13). Indeed, multidisciplinary teams with strong technical competencies combined with excellent social- and meta-competencies can generate creative abrasion that leads to increased effectiveness, efficiency, and innovativeness (Karjalainen et al., 2009). These qualities and the integrative competencies found in our study are consistent with the view that analytics-driven workers and managers have to combine expertise in computing, statistics, experiment design, interpretation and analytics with fundamental business

knowledge and acumen in order to pose the right questions and get data to tell them the things that matter (Hopkins et al., 2010). It is also in line with the contention of Lycett (2013) who posits that a deep understanding of the complex relationships between data, analytical tools and human sense making is key to unlocking the value of big data.

In sum, it seems clear from the job posting data and executive interviews that virtually all BDA professions require workers at all levels to posses an integrated set of technical, social and meta-competencies that enables them to cultivate decision making that blends analytical insights with intuition (Ransbotham et al., 2016).

Our analysis of the program and courses indicates that universities and colleges clearly recognize the need for professionals with strong technical, analytical and business knowledge and skills coupled with industry experience i.e. cognitive and functional competencies. Moreover, virtually all the BDA programs were designed in collaboration with industry partners and many offer internships, co-ops, mentorships, and practical research projects based on real datasets and use software and hardware tools that are commonly used in industry. A few even offer certification for certain analytics tools such as SAS™ (Statistical Analysis Software). Further, most of the programs and courses have strong theoretical and technical biases that are oriented towards cognitive and functional competencies. In the context of STEMB, Business disciplines offer greater opportunities than STEM disciplines for students to develop their social- and meta-competencies. Also, most programs continue to be very discipline-oriented with minimal overlap among key disciplines. We do however observe an emerging trend to give students, particularly at the Master's and Ph.D. levels, practical exposure to industry and integrating business courses aimed at developing students' social- and meta-competencies. However, multidisciplinary teams and projects involving STEMB students are rare, except for specially designed certificate or

professional programs where both technical and business experience are required entry conditions. However, enrolment in these programs tend to be small by design. The extent to which the current approach by universities and colleges to modify existing programs to incorporate BDA or develop small, boutique BDA programs can meet market demand for talents is an open question, which is beyond the scope of this study. Against this backdrop, it is probably worthwhile to recognize that industry tends to be more dynamic and fast-changing than academia and there are too much for students to know and too little time to teach (Donofrio et al., 2010).

Overall, it seems fairly clear that universities and colleges recognize the multi-faceted competencies that industry requires for various BDA professions and have and continue to design programs to equip students with them. Unfortunately, it seems that their efforts are often limited due to a variety of structural constraints such as academic requirements, curricula development and approval process, limited resources and time, and other capabilities. It seems that some of the challenges can be alleviated with greater collaboration among academia, industry, and policymakers. Such collaborations could produce programs that provide students with both theoretical and practical knowledge and experience (multi-faceted competencies) that are needed for BDA professions. Closer industry-academic collaboration could also strengthen the BDA knowledge and capabilities of faculties.

Conclusion

This study attempts to identify the precise competencies that are required for big data analytics jobs and the extent to which academic programs at higher educational institutions are preparing students with the relevant competencies. The study employs a multimethod approach consisting of three unique datasets – BDA job postings, higher education academic programs, and interviews

with senior executives who are knowledgeable and experienced in the BDA. The data was analyzed via text mining and assessed using an integrated competency theoretical model.

We found that organizations are expecting workers for virtually every job role to possess all four types of competencies of our model framework to different degrees based on the nature of the job role. We observed that cognitive and functional competencies are necessary for every job role but insufficient to guarantee hiring. Workers must also demonstrate both social competence (people skills) and meta-competencies (positive personality traits and abilities). On the flip side, while social- and meta-competencies are important, they will not be sufficient for hiring without the necessary cognitive and functional competence.

Essentially, we note that employers are particularly interested in hiring workers that are very knowledgeable in one or two discipline areas (depth of knowledge and experience) and have good knowledge in a variety of other disciplines or domains (breadth of knowledge and experience). For instance, workers from the STEM disciplines (science, technology, engineering, and mathematics) must demonstrate good business acumen and knowledge on non-STEM domains. Similarly, workers with a business discipline background must demonstrate proficiency on the technical aspects of big data analytics (e.g., tools, techniques, programming languages, infrastructure). Social competence (people skills) are given additional emphasis given that big data analytics is a multidisciplinary undertaking involving various knowledge sources and methods. Excellent personality traits such as creativity, empathy, communications, and work ethic are deemed extremely important given that big data analytics employees have to work very closely on a routine basis with multiple internal (e.g., many teams, mangers) and external stakeholders (e.g., customers, suppliers).

We also observe that most higher education programs in big data analytics provide students the opportunity to develop deep cognitive and functional competence in their relevant disciplines but are quite limited in their efforts to target social- and meta-competencies. This seems to be due primarily due to the disciplinary focus of most programs although many programs are using innovative mechanisms to improve in these areas. Although there is increasing collaboration with industry in BDA-specific programs, more collaboration is needed in all disciplines and programs.

Finally, despite our findings, this study has certain limitations. First, we only used a sample of just over 3,000 job postings, while considered reasonable, could be enlarged in future studies. Second, we did not include students and employees in our study to get their perspectives and experience on their readiness. This aspect could be undertaken in a future study. Another topic that we did not address but could enhance our understanding pertains to how the demand and supply of big data skills are managed in cities and regions with vibrant, dynamic BDA ecosystems.

References

Agarwal S. (2017) *Why big data projects fail and how to make 2017 different*. Available at: https://www.networkworld.com/article/3170137/why-big-data-projects-fail-and-how-to-make-2017-different.html.

Akter S, Wamba SF, Gunasekaran A, et al. (2016) How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics* 182: 113-131.

Autor DH. (2015) Why are there still so many jobs? The history and future of workplace automation. *The Journal of Economic Perspectives* 29: 3-30.

Axryd S. (2019) *Why 85% of Big Data Projects Fail*. Available at: https://www.digitalnewsasia.com/insights/why-85-big-data-projects-fail

Barney JB. (2001) Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of management* 27: 643-650.

Bharadwaj AS. (2000) A resource-based perspective on information technology capability and firm performance: an empirical investigation. *MIS quarterly*: 169-196.

Bhimani A. (2015) Exploring big data's strategic consequences. *Journal of Information Technology* 30: 66-69.

Boyd D and Crawford K. (2012) Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15: 662-679.

Bughin J. (2016) Big data, Big bang? *Journal of Big Data* 3: 1.

Cappel JJ. (2002) Entry-level IS job skills: A survey of employers. *Journal of Computer Information Systems* 42: 76-82.

Cegielski CG and Jones-Farmer LA. (2016) Knowledge, skills, and abilities for entry-level business analytics positions: A multi-method study. *Decision Sciences Journal of Innovative Education* 14: 91-118.

Cheetham G and Chivers G. (1996) Towards a holistic model of professional competence. *Journal of European Industrial Training* 20: 20-30.
Chen H, Chiang RH and Storey VC. (2012) Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly* 36: 1165-1188.
Coltman T, Devinney TM and Midgley DF. (2011) Customer relationship management and firm performance. *Journal of Information Technology* 26: 205-219.
Comuzzi MP, Anit. (2016) How organisations leverage Big Data: a maturity model. *Industrial Management & Data Systems* 116: 1468-1492.
Constantiou ID and Kallinikos J. (2015) New games, new rules: big data and the changing context of strategy. *Journal of Information Technology* 30: 44-57.
Cretchley J, Rooney D and Gallois C. (2010) Mapping a 40-year history with Leximancer: Themes and concepts in the Journal of Cross-Cultural Psychology. *Journal of Cross-Cultural Psychology* 41: 318-328.
Croteau A-M and Bergeron F. (2001) An information technology trilogy: business strategy, technological deployment and organizational performance. *The journal of strategic information systems* 10: 77-99.
Croteau A-M and Raymond L. (2004) Performance Outcomes of Strategic and IT Competencies Alignment1. *Journal of Information Technology* 19: 178-190.
Davenport TH and Patil DJ. (2012) Data scientist. *Harvard Business Review* 90: 70-76.
De Mauro A, Greco M, Grimaldi M, et al. (2018) Human resources for Big Data professions: A systematic classification of job roles and required skill sets. *Information Processing & Management* 54: 807-817.
Dehning B and Stratopoulos T. (2003) Determinants of a sustainable competitive advantage due to an IT-enabled strategy. *The journal of strategic information systems* 12: 7-28.
Demirkan H and Spohrer J. (2015) T-shaped innovators: Identifying the right talent to support service innovation. *Research-Technology Management* 58: 12-15.
Demirkan H and Spohrer JC. (2018) Commentary—cultivating T-shaped professionals in the era of digital transformation. *Service Science* 10: 98-109.
Donofrio N, Spohrer J, Zadeh HS, et al. (2010) Driven medical education and practice: A case for T-shaped professionals.
Dubey R and Gunasekaran A. (2015) Education and training for successful career in Big Data and Business Analytics. *Industrial and Commercial Training* 47: 174-181.
Gangani N, McLean GN and Braden RA. (2006) A competency-based human resource development strategy. *Performance Improvement Quarterly* 19: 127-139.
Gardiner A, Aasheim C, Rutner P, et al. (2018) Skill requirements in big data: a content analysis of job advertisements. *Journal of Computer Information Systems* 58: 374-384.
Gupta M and George JF. (2016) Toward the development of a big data analytics capability. *Information & Management* 53: 1049-1064.
Hager P and Gonczi A. (1996) What is competence? *Medical teacher* 18: 15-18.
Hayashi AM. (2014) Thriving in a big data world. *MIT Sloan Management Review* 55: 35.
Hopkins MS, LaValle S and Balboni F. (2010) 10 Insights: A First Look at The New Intelligent Enterprise Survey. *IT Management Select*: 14.
Iansiti M. (1999) Real-World R&D: Jumping the product generation gap. *Harvard Business Review on managing high-tech industries*. Harvard Business School Press, 91-116.
ICTC. (2016) Digital Talent: Road to 2020 and Beyond.
Indulska M, Hovorka DS and Recker J. (2012) Quantitative approaches to content analysis: identifying conceptual drift across publication outlets. *European Journal of Information Systems* 21: 49-69.
Jin X, Wah BW, Cheng X, et al. (2015) Significance and challenges of big data research. *Big Data Research* 2: 59-64.
Johnson RB, Onwuegbuzie AJ and Turner LA. (2007) Toward a definition of mixed methods research. *Journal of Mixed Methods Research* 1: 112-133.
Karjalainen T, Koria M and Salimäki M. (2009) Educating T-shaped design, business and engineering professionals. *Proceedings of the 19th CIRP Design Conference–Competitive Design*. Cranfield University Press.
Le Deist FD and Winterton J. (2005) What is competence? *Human Resource Development International* 8: 27-46.
Lee DM, Trauth EM and Farwell D. (1995) Critical skills and knowledge requirements of IS professionals: a joint academic/industry investigation. *MIS quarterly*: 313-340.
Leitheiser RL. (1992) MIS skills for the 1990s: A survey of MIS managers' perceptions. *Journal of Management Information Systems* 9: 69-91.

Lycett M. (2013) 'Datafication': making sense of (big) data in a complex world. *European Journal of Information Systems* 22: 381-386.
MacDonnel P and Castro D. (2016) Europe Should Embrace the Data Revolution. Centre for Data innovation.
Manyika J, Chui M, Brown B, et al. (2011) Big data: The next frontier for innovation, competition, and productivity.
Marjanovic O and Dinter B. (2017) 25+ Years of Business Intelligence and Analytics Minitrack at HICSS: A Text Mining Analysis. *Proceedings of the 50th Hawaii International Conference on System Sciences.*
Mazzei MJ and Noble D. (2017) Big data dreams: A framework for corporate strategy. *Business horizons* 60: 405-414.
McKenny AF, Aguinis H, Short JC, et al. (2018) What doesn't get measured does exist: Improving the accuracy of computer-aided text analysis. *Journal of management* 44: 2909-2933.
Miller S. (2014) Collaborative approaches needed to close the big data skills gap. *Journal of Organization design* 3: 26-30.
Mithas S, Ramasubbu N and Sambamurthy V. (2011) How information management capability influences firm performance. *MIS quarterly* 35: 237.
Müller O, Junglas I, vom Brocke J, et al. (2016) Utilizing big data analytics for information systems research: challenges, promises and guidelines. *European Journal of Information Systems.*
Nelson RR. (1991) Educational needs as perceived by IS and end-user personnel: A survey of knowledge and skill requirements. *MIS quarterly*: 503-525.
O'Reilly K and Paper D. (2012) Want value from big data? Close the gap between the C-Suite and the server room. *Journal of Information Technology Case and Application Research* 14: 3-10.
Olston C and Najork M. (2010) Web crawling. *Foundations and Trends® in Information Retrieval* 4: 175-246.
Power DJ. (2016) Data science: supporting decision-making. *Journal of Decision Systems* 25: 345-356.
Prahalad C and Hamel G. (1990) The core competence of the corporation. *Boston (MA)*.
Ransbotham S, Kiron D and Prentice PK. (2016) Beyond the Hype: The Hard Work Behind Analytics Success. *MIT Sloan Management Review* 57.
Reaney M. (2014) The 22 skills of a data scientist.
Rothwell WJ and Lindholm JE. (1999) Competency identification, modelling and assessment in the USA. *International journal of training and development* 3: 90-105.
Russ-Eft D. (1995) Defining competencies: A critique. *Human Resource Development Quarterly* 6: 329.
Scarbrough H. (1998) Path (ological) dependency? Core competencies from an organizational perspective. *British Journal of Management* 9: 219-232.
Schildt H. (2017) Big data and organizational design–the brave new world of algorithmic management and computer augmented transparency. *Innovation: Organization & MAnagement* 19: 23-30.
Sharma R, Mithas S and Kankanhalli A. (2014) Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems* 23: 433-441.
Singh D and Reddy CK. (2015) A survey on platforms for big data analytics. *Journal of Big Data* 2: 1-20.
Smith AE and Humphreys MS. (2006) Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior research methods* 38: 262-279.
Sodhi M and Son B-G. (2010) Content analysis of OR job advertisements to infer required skills. *Journal of the Operational Research Society* 61: 1315-1327.
Song IY and Zhu Y. (2016) Big data and data science: what should we teach? *Expert Systems* 33: 364-373.
Tippins MJ and Sohi RS. (2003) IT competency and firm performance: is organizational learning a missing link? *Strategic Management Journal* 24: 745-761.
Todd PA, McKeen JD and Gallupe RB. (1995) The evolution of IS job skills: a content analysis of IS job advertisements from 1970 to 1990. *MIS quarterly*: 1-27.
Vidgen R, Shaw S and Grant DB. (2017) Management challenges in creating value from business analytics. *European Journal of Operational Research* 261: 626-639.
Wamba SF and Mishra D. (2017) Big data integration with business processes: a literature review. *Business Process Management Journal* 23: 477-492.
Woerner S and Wixom BH. (2015) Big data: extending the business strategy toolbox. *Journal of Information Technology* 30: 60-62.
Woodie A. (2017) *Hadoop Has Failed Us, Tech Experts Say*. Available at: https://www.datanami.com/2017/03/13/hadoop-failed-us-tech-experts-say/

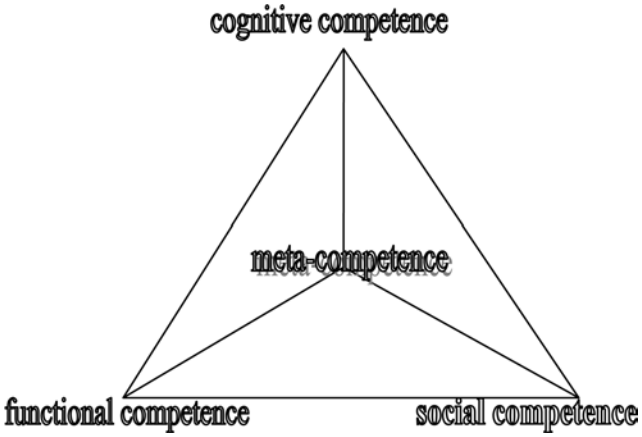

Figure 1. Holistic model of competence.

Source: Le Deist and Winterton (2005)

Figure 2 Panel A Concept Map of Themes and Concepts of Big Data Jobs

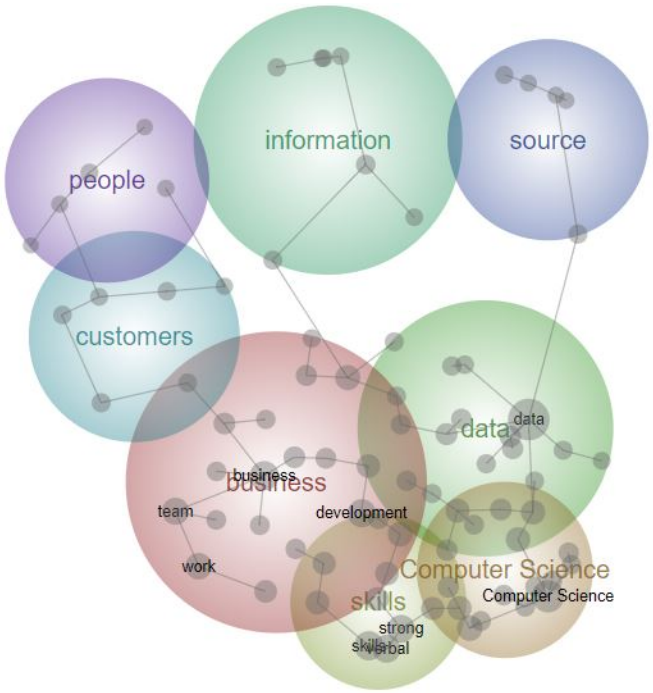

Figure 2 Panel B Sample list of Concepts Associated with each Theme

Theme	Associated Concepts
Data	analytics, learning, machine, applications, modeling, and statistics
Computer Science	degree, experience, tools, technologies, software, programming, and related systems including Hadoop, SQL, Spark, Python, etc.
Information	management, sharing
Source	Open, value
Business	environment, technology, development, projects, support services, team, work
Customers	relationship, building, focused, solutions
Skills	strong, communication, verbal, presentation, writing, and management
People	ideas, trust, passionate, connections, goals, learning, curious, self-directed

Figure 3 Panel A Concept Map of University and College Dataset

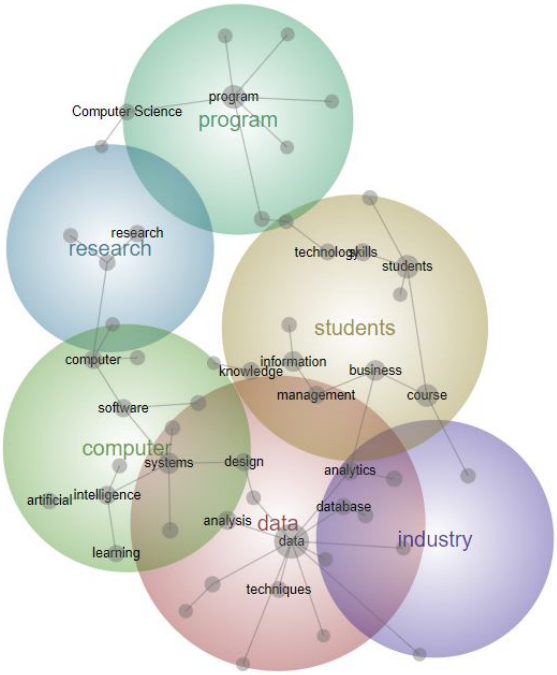

Figure 3 Panel B Sample list of Concepts Associated with each Theme

Theme	Concepts
Students	students, course, business, information, knowledge, management, skills, technology
Data	data, systems, analytics, analysis, design, database, techniques, tools, mining, insight
Computer	computer, intelligence, software, learning, artificial, include
Program	computer science, business, statistics, decision-making
Research	research, science
Industry	industry, partners, experience

Table 1 Recent Research on BDA Capabilities

Authors	Details	Key Conclusions
Rialti et al. (2019)	Survey of 259 managers from the European Union (EU). Structural Equation Modeling (SEM) Analysis	Top managers to empower people who have strong problem-solving skills about big data processes so they can exploit its potentialities.
Mikalef et al. (2019)	Mixed method - survey data from 175 chief information officers and 3 case studies from Greece. Fuzzy set qualitative comparison analysis (fsQCA) & SEM with Smart PLS	Strong technical skills are necessary to convert data into actionable insights for firms operating in contexts of high dynamism and an absence of heterogeneity (e.g., consumer goods, media, transport, and industrials). Managerial skills were not core contributors to firm performance in large firms but were core contributors in SMEs in these industries. For firms operating under conditions of high dynamism and hostility (e.g., oil & gas, banking, and financial sector), both technical and managerial skills are core contributors to performance.
Mikalef et al. (2019)	Mixed method - survey data from 175 chief information officers and 3 case studies from Greece. fsQCA & SEM with Smart PLS	Both technical and managerial skills are core components of BDA capabilities, and big data capabilities enhance firms' dynamic capabilities, which in turn influence both radical and incremental innovation capabilities
Wang et al. (2019)	Survey of 600 firms from 6 countries: Australia, Canada, China, Korea, Malaysia, and the USA. Method - PLS.	Business analytics competency positively impacts organizational absorptive capacity and business analytics assimilation, which in turn influences competitive advantage.
Mandal (2018)	Online survey of travel and tour operators in India. 212 responses. SEM with Smart PLS (Partial Least Squares)	Big data analytics (BDA) management capabilities are key to sustainable tourism SC performance. BDA planning, coordination, and control capabilities are prominent enablers of sustainable tourism SC performance.
Mauro et al. (2018)	2786 Job postings from various online websites. Basic text mining techniques and expert judgment to classify jobs into 4 categories. Then used Latent Dirichlet Allocation, LDA to match skills to job types	**Business Analysts** need a mix of industry-specific skills and broader management competencies, such as effective communication, business process transformation and financial acumen. **Data scientists** need to understand the business context in which they operate and use project management techniques in order to interact effectively with the rest of the organization. **Big Data Developers** need coding skills and solid expertise in systems management, cloud computing and distributed technologies, database management, corporate data architecture, and how analytics are used in their company. **Big Data Engineers** need competencies to construct and manage the corporate big data ecosystem in a sustainable manner.
Vidgen et al. (2017)	Two rounds of Delphi Study & 3 case studies. Delphi round 1 involved 72 responses from 36 practitioners, 23 consultants, and 13 academics and round 2 produced 42 responses. Qualitative analysis involves coding data	Data scientists must have a problem-solving orientation, capable of working independent, and be able to work with the business to co-create plausible and convincing stories based on data and actionable insights. Data scientists need strong statistical and mathematical skills as well as IT skills, notably an ability to program (e.g., R) and an ability to manipulate data (e.g., SQL). They need visualization skills because it is part of the storytelling.

Authors	Details	Key Conclusions
Lugmayr et al. (2016)	Literature review and focus group. Qualitative analysis	The job of a business analyst or a data analyst inside a corporation will need to be re-centered to become the role of a “perception data designer” who understands ways to make humans understand principles, knowledge, and wisdom that emerges from manifold corporate data sources.
Cegielski & ones-Farmer (2016)	Mixed method research design with three separate data collection methods: a qualitative Delphi study, a survey, and archival content analysis. The Delphi study and survey used data analytics professionals. Round 1 had 47 participants and round 2 had 27. The archival content analysis used 186 job postings for entry-level business analytics professionals. Qualitative analysis using manual coding	Postings were classified into 3 categories - Technical, Business, and Analytical - and in terms of Knowledge, Skills, and Abilities (KSAs). In terms of **Analytical KSAs**, the top ranked items according to the *survey* were the ability to integrate analyses, data visualization, and problem framing ability, whereas the *content analysis* suggested that problem solving, decision making, and integrative analysis were the highest ranked skills. Regarding **Technical KSAs**, Excel, SAS, SQL, and R skills were ranked as the top four skills necessary for an entry level position in business analytics. In the category of **Business KSAs,** the top ranked skills from the *survey* were independent learner, organizational skills, and industry specific knowledge whereas the *content analysis* ranked communication skills, project management, and organizational ability as the top ranked skills.
Dubey & Gunasekaran (2014)	review of extant literature and interviews with 10 heads of business analytics for companies situated in India. Qualitative analysis.	Both hard and soft skills needed. Hard skills are domain knowledge or technical skills, which are important for successful execution of the task (e.g., statistics, forecasting, and optimization). IT skills supported by communication and analytical skills are very important. Soft skills represent individuals’ attitudes and communication skills, leadership ability and passion for excellence

Table 2 Sample Job Titles

Position Title	Number
Senior-level Positions	
Vice Presidents	22
Chief Data Officer, Chief Data Scientist	123
Senior (Manager, Director, Lead, Developer, Architect, etc.)	294
Chief (Marketing Officer, Technology Officer, Architect, etc.)	105
Mid-level Positions	
Managers	187
Consultants	103
Specialists (Cloud, software, support, big data, etc.)	191
Artificial Intelligence (Experts, Lead, Director, etc.)	186
Entry-Level	
Engineer	502
Developer	288
Architect (Data, Analytics, Applications, etc.)	202
Machine Learning (Expert, Researcher, Engineer, Developer, etc.)	152
Analyst (BI, Big Data, Marketing, Supply Chain, Insight Analyst, Strategy, Systems, etc.)	203
Analyst (Bi Data, Data Science, Big Data Engineering, etc.)	307
Application (developer, architect, consultant, etc.)	144
Total	**3009**

Table 3 Sample of Social- and Meta-competencies Required for Big Data Jobs

Skill	Excerpts from Job Postings
Communication - oral & written	- Ability to communicate complex analysis in a clear and actionable manner; - Excellent skills to communicate with technical and non-technical audiences - Clearly communicate the value of new capabilities and technologies - Ability to clearly present analyses, with the help of visualization tools, to guide decision; - Ability to communicate with data-driven stories; - Excellent storytelling and presentation skills
Team/Collaboration	- Takes a collaborative approach to assignments and works well with others - Foster an environment of teamwork; - Reliable and flexible team player - Ability to work in a dynamic, high performing and rewarding team
Problem-solving	- Must have experience working and solving highly complex problems, - Ability to solve ambiguous problems; - Unrelenting focus on closing the gap between a brilliant idea and actual real-world problem
Organization	- Attention to detail and organization/documentation skills; - Ability to prioritize and triage deadline-driven tasks in a high-pressure environment; - Attention to detail and in-depth planning, organizational and time management skills, strong project management skills; - Ability to multi-task and prioritize
Self-directed	- Highly self-directed with ability to manage ambiguity; - Ability to work independently with high-level direction
Learning	- Demonstrated willingness to learn new technologies and evaluating fit for a client solution delivery environment; - Capable of learning new technologies; - Strong desire to learn new skills and techniques, intellectually curious; - Ability to quickly adapt to new technologies

Table 4 Sample BDA Program Highlights

The PhD program aims is to produce computer science graduates with a *broad background in information technology along with project management and people skills*. Graduates will work in interdisciplinary teams to tackle problems that require both technical and non-technical solutions. This program concentrates on *both applied research and the development of professional skills*.
This one-year MSc program enables students to *develop interdisciplinary* skills and gain a deep understanding of *technical and applied knowledge in data science and analytics*. It is delivered in both lecture-based and *hands-on lab learning* environments and uses complex, real-world datasets and data science and analytics problems.
The Graduate Diploma offer graduate-level professional education in data analytics to professionals at all levels of organisations and across different industries. Students will learn how to conduct *descriptive, predictive, and prescriptive analytics*.
The Master of Data Analytics program provides students with the *technical skill-sets* required in data analytics and the opportunity to gain practical experience by applying those skills in an organizational setting. The program culminates with an *internship where students will get practical, hands-on experience using the analytics skills in a workplace environment*.
The capstone of the Master of Business Analytics program is an Analytics Consulting Project in which *students will spend two terms working with real data* in service to the community. Students will complete a hands-on, problem-driven analytics project and develop applicable business solutions. They will *directly interface with industry leaders* and develop both *technical and organizational expertise*. Students are awarded *SAS™ (Statistical Analysis Software) certification* upon completion of the program.
The Master of Management Analytics program give students *advanced data management, analytic, managerial and communication skills*. We *consulted with analytic professionals in industry* who identified the need for graduates who excel at the *technical aspects required AND who possess business acumen*.
The Digital Business Management program provides the essential components of a *traditional business degree plus a broad selection of cutting-edge technology and advanced digital business management and marketing tools*. Students get *real-world experience* by attending industry events, participating in competitions and learning first-hand from guest speakers.
The Marketing Research and Analytics business graduate certificate program combines advanced courses in marketing research and BDA with training *in leading business intelligence and marketing research technologies and tools used in the field*. They include SAS Enterprise Guide and SAS Enterprise Miner, Environics Analytics Envision, SPSS, Tableau, Excel, and XL Miner.

Appendix 1

Table 5 Sample list of Programs and Courses Specific to Big Data Analytics

Institution	Discipline - STEMB	Program (P)/ Course (C)	Title of Programs & Courses	Degree Level: G/UG/O)
1. Univ. of Toronto	T	P	Master of Management Analytics	G
2. Queen's University	B	P	Master of Management Analytics	G
3. Queen's University	B	P	Strategic Analytics Program	O
4. Univ. of Ottawa	E	P	Data Science Option	U
5. Univ. of Waterloo	B	P	Data Analytics	G
6. Univ. of Waterloo	M	C	Big Data Analytics	U
7. Univ. of Western Ontario	S	P	Master of Data Analytics	G
8. Univ. of Western Ontario	B	P	MSc in Management - Business Analytics	G
9. York University	CS (Comp. Sci. Dept.)	P	Certificate in Big Data Analytics	O
10. York University	CS	C	Advanced Data Science & Predictive Analytics	O
11. McMaster University	B	C	Masters Certificate in Analytic for Leaders	G
12. Univ. of British Columbia	CS	P	Master of Data Science	G
13. Simon Fraser University	S	P	M.Sc. in Big Data	G
14. Univ. of Victoria	B	P	Master's in Business Analytics	G
15. Univ. of Alberta	CS/M	P	M.Sc. & Ph.D. in Statistical Machine Learning	G
16. Univ. of Calgary	M/B	P	Minor in Data Science & Certificate in Business Intelligence and Analytics	U/O
17. Univ. of Manitoba	S/B	C	AI, Machine Learning, Big Data, Topics in Data Analytics	U
18. Univ. of Winnipeg	S	P	Applied Computer Science	U
19. McGill University	B	P	Master of Science in Analytics	G
20. Concordia University	E	P	Diploma in Big Data Analytics & Infrastructure; Certificate in Big Data Analytics	U
21. HEC Montréal	B	P	Master of Science in Business Analytics	G
22. Dalhousie University	CS	P	Big Data Analytics, AI, & Machine Learning	U
23. St. Mary's University	CS	P	M.Sc. in Computing & Data Analytics	P
24. Univ. of New Brunswick	CS	C	Computer Science – Data Science, AI, Big Data Analytics, Machine Learning, etc.	G/U
25. Memorial Univ. of Newfoundland	S	P	Scientific Computing	G
26. Univ. of Saskatchewan	S	C	AI, Big Data Analytics, Data Science Bootcamp (data science, machine learning, visualization, statistical methods)	G/O
27. Univ. of Regina	E/B	C	AI, Machine Learning, Analytics & Decision-making	G
28. Ontario Tech University	E	P	Data Science specialization	U
29. Carleton University	B	P	Business Analytics	G
30. Univ. of Guelph	CS	C	Data Science	O
31. Univ. of Guelph	CS	C	Big Data Analytics	O

Institution	Discipline - STEMB	Program (P)/ Course (C)	Title of Programs & Courses	Degree Level: G/UG/O)
32. Wilfrid Laurier University	S	P	B.Sc. Data Science	U
33. Brock University	B	P/C	MBA Business Analytics Stream & various courses/workshops (e.g., Datathon, Data Science through Centre of Business Analytics & Digital Scholarship Lab)	P/O
34. Univ. of Windsor	E	C	Big Data Analytics, Artificial Intelligence	O
35. Royal Military College	B	C	Business Analytics	U
36. Ryerson University	CS	C	Data Science Bootcamp	O
37. Ryerson University	CS	C	Data Science & Analytics, Big Data, & Predictive Analytics	O
38. Trent University	S	P	Master of Science Big Data Analytics	G
39. Univ. of Northern BC	B/S	C	Business Intelligence	
40. Thompson Rivers University	S	C	Artificial Intelligence	U
41. Lakehead University	E	C	Introduction to Data Science	O
42. Laurentian University	MS	C	Artificial Intelligence	U
43. OCAD University	IS	C	Web Analytics; Text analysis & machine learning	O
44. Algoma University	E	C	Artificial Intelligence	U
45. Univ. of PEI	S	C	Artificial Intelligence & Automated Reasoning, Machine Learning, Data Visualization	U
46. St. Francis Xavier	CS	C	Data science concentration & courses in coding	U
47. Univ. of Lethbridge	S	C	Studies in Computational Intelligence	U
48. Athabasca University	S	C	Statistical Language Processing for Text Analytics	G
49. Univ. of Fraser Valley	S	P/C	Data Analysis Postbaccalaureate Certificate, Artificial Intelligence, Machine Learning, Data Mining,	U/G
50. MacEwan University	CE	C	Business Analysis Professional Development Certificate	O
51. Nipissing University	B	P	Business Analytics and Technology	U

Appendix 2

Sample list of College Programs and Courses Specific to Big Data Analytics

Institution	Discipline - STEMB	Program (P)/ Course (C)	Title
1. Algonquin College	S	C	Critical Data Analysis in Regulation
2. Algonquin College	S	C	Computer Forensics & Data Mining
3. Centennial College	B	P	Marketing Research & Analytics
4. Conestoga College	CS	P	Business Analysis/SQL & Data Analysis
5. Conestoga College	CS	C	Introduction to Analytics & Big Data
6. Durham College	B	P	Data Analytics for Business Decision Making
7. George Brown College	B	P	Digital Analytics Certificate
8. Georgian College	CS	C	Data Analytics Project/Data Visualization
9. Humber Inst. of Tech.	B	C	CRM & Data Mining
10. Humber Inst. of Tech.	B	C	Data Analysis & Knowledge Management
11. Mohawk College	B	P	Analytics for Business Decision Making
12. Niagara College	E	C	Big Data Programming & Reporting/Architecture & Administration
13. Sheridan College	CS	C	Big Data Tools/Data Science
14. St. Lawrence College	B	C	Data Acquisition, Analysis & Modelling
15. Douglas College	S	C	Data Analytics Fundamentals, Data Analytics for Managers, Data Visualization
16. BCIT	S	C	Applied Data Analytics, Data Analytics Fundamentals, Introduction to Big Data & Hadoop
17. Saskatchewan Polytechnic	S	C	Advanced Data Analytics, Data Analysis Certificate,
18. Northern Alberta Institute of Technology	T	C	Privacy and Security in Big Data
19. Northern Alberta Institute of Technology	B	P	Business Intelligence and Data Management Certificate