The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0263-5577.htm

Empirical identification of skills gaps between chief information officer supply and demand: a resource-based view using machine learning

Received 6 January 2021
Revised 21 March 2021
Accepted 23 April 2021

Stuart Barnes
King's Business School, King's College London, London, UK
Richard N. Rutter
School of Business, Australian College of Kuwait, Kuwait City, Kuwait
Ariel I. La Paz
Department of Management Control and Information Systems, Faculty of Economics and Business, University of Chile, Santiago, Chile, and
Eusebio Scornavacca
Merrick School of Business, University of Baltimore, Baltimore, USA

Abstract

Purpose – The role of emerging digital technologies is of growing strategic importance as it provides significant competitive advantage to organisations. The chief information officer (CIO) plays a pivotal role in facilitating the process of digital transformation. Whilst demand continues to increase, the supply of suitably qualified applicants is lacking, with many companies forced to choose information technology (IT) or marketing specialists instead. This research seeks to analyse the organisational capabilities required and the level of fit within the industry between CIO requirements and appointments via the resource-based view.

Design/methodology/approach – Job postings and CIO curriculum vitae were collected and analysed through the lens of organisational capability theory using the machine learning method of Latent Dirichlet Allocation (LDA).

Findings – This research identifies gaps between the capabilities demanded by organisations and supplied by CIOs. In particular, soft, general, non-specific capabilities are over-supplied, while rarer specific skills, qualifications and experience are under-supplied.

Practical implications – The research is useful for practitioners (e.g. potential CIO candidates) to understand current market requirements and for companies aiming to develop internal training that meet present and future skill gaps. It also could be useful for professional organisations (e.g. CIO Forum) to validate the need to develop mentoring schemes that help meet such high demand and relative undersupply of qualified CIOs.

Originality/value – By applying LDA, the paper provides a new research method and process for identifying competence requirements and gaps as well as ascertaining job fit. This approach may be helpful to other domains of research in the process of identifying specific competences required by organisations for particular roles as well as to understand the level of fit between such requirements and a potential pool of applicants. Further, the study provides unique insight into the current supply and demand for the role of CIO through the lens of resource-based view (RBV). This provides a contribution to the stream of information systems (IS) research focused on understanding CIO archetypes and how individual capabilities provide value to companies.

Keywords CIO, Job fit, LDA, Machine learning, Supply, Demand

Paper type Research paper

1. Introduction

In the highly evolving era of digital transformation, the role of the chief information officer (CIO) has become critical to organisational performance (Galbraith, 2014). Their importance

Industrial Management & Data Systems
© Emerald Publishing Limited
0263-5577
DOI 10.1108/IMDS-01-2021-0015

comes hand-in-hand with technological development, and over time, the CIO role has evolved from a technical facilitator to a business strategist (La Paz, 2017; McLean and Smits, 2012). The changes in the CIO role not only reflect the importance of technological innovations, but also reveal an increasing responsibility over human and financial resources, beyond the technological, in order to ensure operational continuity and to lead the development of business strategies as well as new corporate functions (Chun and Mooney, 2009). The changes and evolution of the CIO role also came with challenges for top information technology (IT) professionals to understand and interact with a large array of stakeholders with different profiles, unique technical language, and a variety of objectives and perceptions regarding information systems (Lewis *et al.*, 1995). As a result, the role of the CIO requires, in addition to technical competence, strong managerial and communication skills.

The position of a strategist type of CIO has been prescribed since the early 1990s (Feeny *et al.*, 1992; Grover *et al.*, 1993). This leadership role requires a clear understanding of the corporate vision in order to align the firm's IT strategic direction to gain competitive advantages from information systems (Al-Taie *et al.*, 2014; McAfee *et al.*, 2012). In this modern age of digital business, organisations that are not able to efficiently leverage value from data are potentially losing out on opportunities to gain important business insights from big data (Whitler *et al.*, 2017). In functional areas such as marketing, it is vital to deeply understand customer behaviours and profiles. Also, in operations, big data can significantly enhance multi-phase manufacturing processes and logistics (Chun and Mooney, 2009). In other words, the new CIO has a key role in providing their organisation with strategic analytical tools and skills in order to interpret raw data in meaningful ways (Li and Tan, 2013). A good example of the increasing importance of this functional area is the popularisation of the specialised role of CDO or *chief data officer* (Drechsler *et al.*, 2018; Earley, 2017).

Due to the pervasiveness of digital technologies across the organisation, the contemporary CIO needs to portray a combination of technical and soft skills in order to be able to provide key decision-making analysis in different business areas, from risk management to customer relationships (Carillo *et al.*, 2017) . The CIO must therefore be technically proficient (La Paz *et al.*, 2010) and analytically capable to turn extensive amounts of information into value to the organisation (Feeny *et al.*, 1992; Peppard *et al.*, 2011). In addition, they must be effective communicators and describe complex technological challenges and opportunities in a way that the rest of the board can easily understand (Krotov, 2015). It is also crucial for the CIO to have the ability to lead information security efforts in order to increase organisational resilience. In most cases, CIOs will possess qualifications relevant to at least one technical area of data management and analysis as well as several years of working experience in other areas.

While there is a great demand across the globe for CIOs who are capable to lead organisations' digital transformation initiatives, it has been noted that there is a lack of supply of suitable CIO candidates (Jones *et al.*, 2019; La Paz *et al.*, 2019). As a rapidly evolving and technically demanding senior role, the selection, recruitment and retention of proficient CIOs is highly competitive. Although mentoring and internal training is often carried out by forward-thinking organisations to meet future skill gaps (Boehm *et al.*, 2013), with such high demand and relative undersupply, CIOs are increasingly being recruited from other semi-related roles. As a result, in addition to upskilling information officers, some businesses now recruit CIOs who have previously worked in digital marketing or as IT professionals (Kaarst-Brown, 2008, p. 455). La Paz *et al.* (2019) point out that the skills gap between a CIO's skillset and current organisational requirements shows no sign of changing.

Given the importance of information management within organisations to provide strategic opportunities, and the need to endow companies with the best candidates for their positions, the research questions driving this study are therefore twofold: "What capabilities are required by organisations for chief information officers (demand-side)?" and "What is the level of fit between actual CIO capabilities and required capabilities (supply-side)?" Answering these questions is important in order to ascertain the degree of fit in capabilities between organisational demand and the supply of CIOs. The paper makes three important contributions (discussed in more detail in the conclusions). First, it provides an original test of the resource-based view as a theoretical basis to examine CIO job fit. Second, it uses a novel machine learning approach to examining demand and supply characteristics via text analytics methods. Third, it provides insight into a problem that is of practical importance to organisations.

The paper is structured as follows. In the next section, we examine the changing role of the CIO in the organisation and the rarity of required capabilities. Further, we delineate the gap in knowledge regarding how well the supply of CIOs fit organisational demand. In section three, we describe the research process applied in the study, utilising the machine learning method of Latent Dirichlet Allocation. Subsequently, we present the results of our analysis, fitting and comparing a Latent Dirichlet Allocation (LDA) model to both demand- and supply-side data. Finally, we discuss the results of our study, draw conclusions, and suggest implications for research and practice.

2. Theoretical background and hypotheses development

Historically, there was a generalised perception that a CIO position would only be required in very large organisations that used computers (Fiegener and Coakley, 1995). As such, the CIO used to be much more of a technical role focussed on the particular types of information technology equipment deployed by their organisation (Chun and Mooney, 2009; Ross and Feeny, 1999). At present, we know that not only large corporations need IT executives, since all companies of all sizes, industries or business models already use IT or bear great potential for improvement with its use in reducing costs, assuring operational continuity, exploiting data and information or strategically designing their core processes. The exponential growth of data, security threats in the networks and challenges to keep up with the pace of technological developments, dramatically changed the CIO's role in recent decades (Ross and Feeny, 1999). The Enterprise resource planning (ERP) revolution revealed the strategic importance of the CIO as a business executive who can think technically, in terms of business value, as well as provide support to the development of functional areas across the organisation (Willcocks and Sykes, 2000). Today, companies seek CIOs capable of contributing to the formulation of business strategies that improve business performance through the alignment of information technologies and organisational goals (Andriole, 2009; Carter *et al.*, 2011; Chun and Mooney, 2009; Sobol and Klein, 2009). In the past, a CIO role would frequently also include the responsibilities of chief technology officer (CTO) (Gonzalez *et al.*, 2018), but as the business dynamics and the IT department function became more complex, more positions were related, condensed and coordinated under the CIO umbrella, e.g. to manage risks and security (chief information security officer, CISO), data (chief data officer, CDO) and others that lead digital transformations, innovation or governance (chief innovation officer or chief digital transformation officer). As organisations place greater emphasis on their information services capabilities, many CIOs now have CTOs, CDOs and/or CISOs who report to them (Smith, 2011).

In the literature, there are two streams of studies focussing on CIOs' profiles. One has its emphasis on studying the effects of personal characteristics or attributes. These studies usually look at elements such as age, gender and educational background (Jones *et al.*, 2019; Li and Tan, 2013; Sobol and Klein, 2009) as well as leadership and reporting structure (Chen *et al.*, 2010; Chen and Wu, 2011). The second stream focuses on formulating archetypes based on capabilities and skills to explain how these provide different value to companies (Al-Taie *et al.*, 2014; Carter *et al.*, 2011; La Paz, 2017; La Paz *et al.*, 2010; McLean and Smits, 2012).

The present study is aligned with the second stream of the literature, since the current digital context challenges the CIO not only in the technical aspects, but also in the normative/legal, security, innovation and interpersonal arenas, requiring the acquisition and development of skills according to the emerging corporate functions and responsibilities. Meeting the demands of these challenges is central to the CIO's role and the performance of the entire IT department, as well as providing analytical foresight about future needs.

A widely known theory in management is the resource-based view (RBV), which aims to explain why and how some companies build competitive advantages (Barney, 1991). Within the RBV helm, the concept of IT-related capability captures the relationship between IT assets and resources, aiming to improve organisational performance (Barua *et al.*, 2004; Bharadwaj, 2000; Kearns and Lederer, 2003; Mithas *et al.*, 2012; Nevo and Wade, 2010; Wade and Hulland, 2004). A number of studies on IT-based capabilities associate improvements on company performance with information processing value (Cao *et al.*, 2019), IT infrastructure, human IT skills and IT-enabled intangibles (Bharadwaj, 2000; Wade and Hulland, 2004). On the other hand, just a few argue that the top IT manager role is itself a valuable, rare, inimitable and non-substitutable resource (Spitze and Lee, 2012). Although research in general supports the hypothesis of IT capabilities positively influencing corporate performance, little confirmatory evidence has been found for which CIO skills and abilities are important for different organisational contexts, leaving key questions unanswered that lead to frustration, underperformance, potential business value loss and unnecessary exposition to technological risks and market threats (Nolan and McFarlan, 2005).

Research has shown that managerial skills are not equally distributed among companies (Castanias and Helfat, 2001) and that the most skilled people in certain areas seek employment where their skills and abilities are appreciated and rewarded (Leonard-Barton, 1992). In addition, elements such as market imperfections and deficient recruitment and selection processes make it difficult to find good matches. Spitze and Lee (2012) propose that matching the right person with the defined CIO type can bring extraordinary success for a company, as the IT executive helps develop superior IT capability (Lim *et al.*, 2012). Nevo and Wade (2010) argued that there are levels of acquisition and development of rare capabilities, and in the words of Simon *et al.* (2010), managerial capability is not rare *per se*, but high levels of managerial capability—beyond the skills possessed by competitors—are rare. It is expected that the changes in the requirements and demands for the CIO profile will continue to evolve along with the effects that technological leaps in big data, artificial intelligence, 5G networks and quantum processors will produce on business models (Lasi *et al.*, 2014; Rao and Prasad, 2018; Weill and Woerner, 2013). The rapid evolution of IT and its radical impact on business models and organisations demand agility and flexibility in the IT department. On the other hand, human adaptation to such changes is slower than the evolution of technological disruption, and it takes time and resources to acquire and develop new skills and capabilities, as well as to select and recruit top, scarce and inimitable executives. The exploratory study of La Paz *et al.* (2019) found a relatively small degree of match between the organisational definitions for the CIO position and the individuals hired. About a third of the sample matched the role definition and professional profile and approximately 14% of CIOs were overqualified for the position. In addition, more than a half of the sample was underqualified for the set of responsibilities and expected performance defined in the positions. Hence, we propose the first hypothesis, investigating the extent to which there is a mismatch between required CIO capabilities and the CIO supply:

H1. Required CIO capabilities demanded from organisations do not match CIO supply.

In order to realise the value of a CIO in building stronger IT capabilities, it is important to precisely describe the role according to organisational needs, and match the skills, knowledge and abilities with the right person. In the literature, there is a wide range of different descriptions of CIOs. These descriptions usually relate the CIO position to corporate performance (Li and Tan, 2013), reporting structure, career paths (Jones *et al.*, 2019; Schambach and Blanton, 2002), personal traits, academic background (Sobol and Klein, 2009), strategic thinking and business vision (Carter *et al.*, 2011). They also relate the role of CIO to managerial competences such as spokesman, innovator, communicator, leader and enabler (Grover *et al.*, 1993). Executive skills and capabilities such as leadership, effective communication, teamwork, persuasion and negotiation are important for the CIO as well as the executives in the C-Suite. The convergence of these elements help shape archetypes of CIOs which are commonly labelled as professional, consultant, executive, activist, collaborator, technician, utilitarian, innovator, enabler and strategist (Carter *et al.*, 2011; McCue, 2007; McLean and Smits, 2012; Sobol and Klein, 2009). These archetypes can be understood as different, unique, rare and inimitable resources that can add value to high level IT capabilities and cannot be substituted as a "commodity asset" according to the industrial sector, business domain, corporate maturity and strategic challenges of a company.

The rapid evolution of technologies and business models requires the CIO to possess domain-specific skills and abilities that fit to the context and reality of the company (Ross and Feeny, 1999), where the ability to adapt to organisational contexts and align to its strategy becomes a critical personal resource. In recent years, the digital economy has put tremendous emphasis on a new set of skills related to data-driven decision-making that are aimed at creating value by capturing, analysing and using information to serve customers (Kane *et al.*, 2018; Hagiu and Wright, 2020). While the softer and more general business skills emphasising the interpersonal and communicational abilities can be found in a large pool of executives (not only IT-specific), those rare capabilities with the potential to create and sustain competitive advantages based on IT and its dynamic alignment to the technological evolution (e.g. data analytics, IT alignment) would be harder to find among applicants with executive experience in the area. As a result, it is expected that in the search process, generic abilities in the profiles of CIOs may abound while candidates with the right balance of rarer technology experience and qualifications as well as experience specific to the CIO position and industry may be scarce. Thus, we propose hypotheses 2 and 3:

H2. Soft, general, non-specific capabilities will be over-supplied.

H3. Hard, rarer specific skills, qualifications and experience will be under-supplied.

The next section describes the methodology used in the study.

3. Methodology

This research is based on the application of LDA, otherwise known as topic modelling, to identify organisational capabilities as well as individual skills mentioned in text documents. Topic modelling is a statistical model for discovering hidden semantic assemblies in textual data. Probabilistic topic models use machine learning algorithms to detect latent groups of words that co-occur in a corpus of documents. The mathematics behind LDA is clearly explained by Blei *et al.* (2003). LDA has become a popular and accepted statistical method in business research for mining detecting latent textual structures in various areas of application with a high degree of scientific rigour (e.g. Bastani *et al.*, 2019; Brown *et al.*, 2020; Guo *et al.*, 2017; Tirunillai and Tellis, 2014).

The methodology developed for this study is illustrated in Figure 1, which outlines the various steps taken in the research process. Let us briefly examine the steps in the research process.

3.1 Source and pre-process text data

All of the data used in the study were retrieved from publicly available sources. The data on CIO job postings and anonymous data on CIO curriculum vitae were scraped using Python

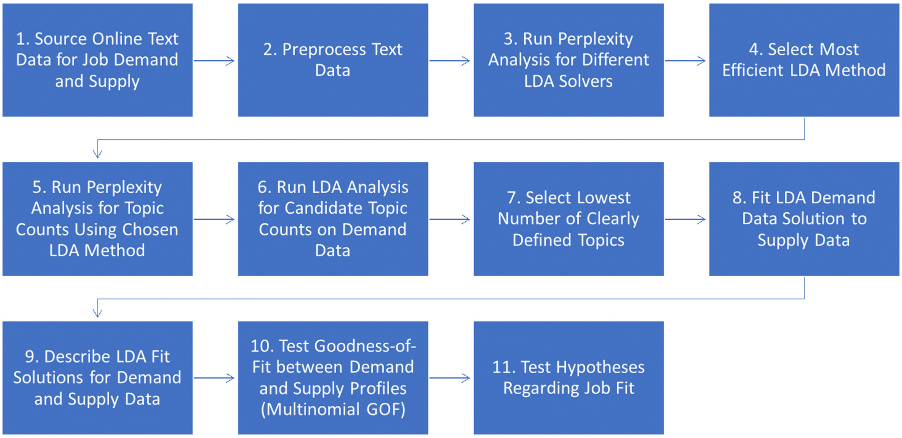

Figure 1.
Research process

code from LinkedIn. Each data set was manually examined by two researchers to remove entries that were not related to CIO positions. This resulted in profiles for 297 CIO job postings and 956 CIO curriculum vitae. The data were dominated by US job postings (85% of jobs) and curriculum vitae (CVs) (100%). The data were further pre-processed for analysis using MATLAB. This included converting the data to lowercase, tokenizing the text, erasing punctuation, removing common stop words, removing short (2 or fewer characters) and long words (15 or more characters) and lemmatising text.

3.2 Select LDA solver method

In order to identify the best LDA solver method for the data set, we ran perplexity analysis in MATLAB to compare the application of three methods: collapsed Gibbs sampling (Griffiths and Steyvers, 2004), zeroth-order collapsed variational Bayes (Schölkopf *et al.*, 2007) and approximate variational Bayes (Asuncion *et al.*, 2009). The results of our analysis are shown in Figure 2. The analysis was run on the jobs data set, with a hold-out sample of 10% for validation.

As we can see from Figure 2, the approximate variational Bayes solver method (AVB) was the fastest. However, the zeroth-order collapsed variational Bayes solver method (CVB0) appeared by far the most efficient in terms of explaining the topics in the data, with a validation perplexity around 600, well below AVB and collapsed Gibbs sampling (CGS). Therefore, we selected CVB0 as the most efficient method for our analysis.

3.3 Select number of topics

Perplexity analysis was run using the zeroth-order collapsed variational Bayes method for numbers of topics from one to forty in MATLAB. The results are shown in Figure 3, which shows potential solutions at seven and nine topics, both of which have the lowest perplexity scores. Subsequently, we ran LDA using CVB0 for seven and nine topics and compared the solutions. We aimed to ascertain whether the seven-topic solution or nine-topic solution was more valid, with clearly defined topics. We found that the nine-topic solution had overlapping topics that could not be clearly distinguished, while the seven-topic solution had much clearer topic areas (shown in section 4). Therefore, we focused on the seven-topic solution for the supply-side data set.

3.4 Fit LDA demand data solution to supply data set

The seven-topic CVB0 model created on the demand-side data set was saved and then applied to the supply-side data set using MATLAB. This resulted in topic distribution probabilities

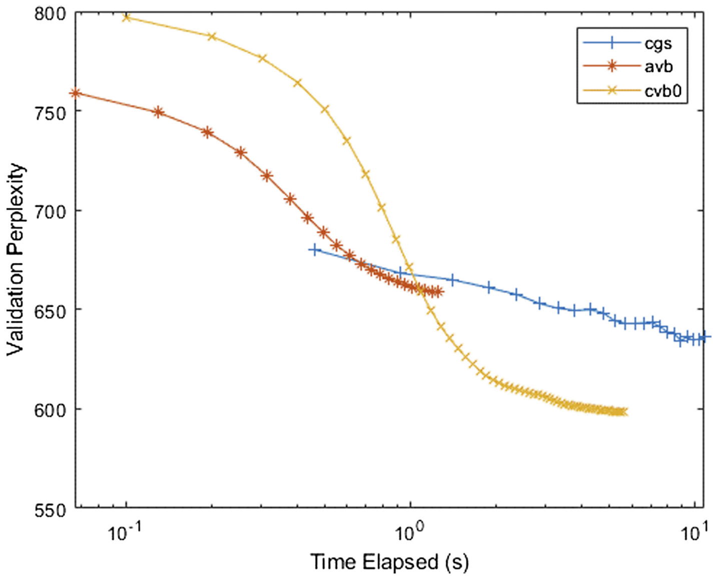

Figure 2.
Comparing solvers for LDA via perplexity analysis

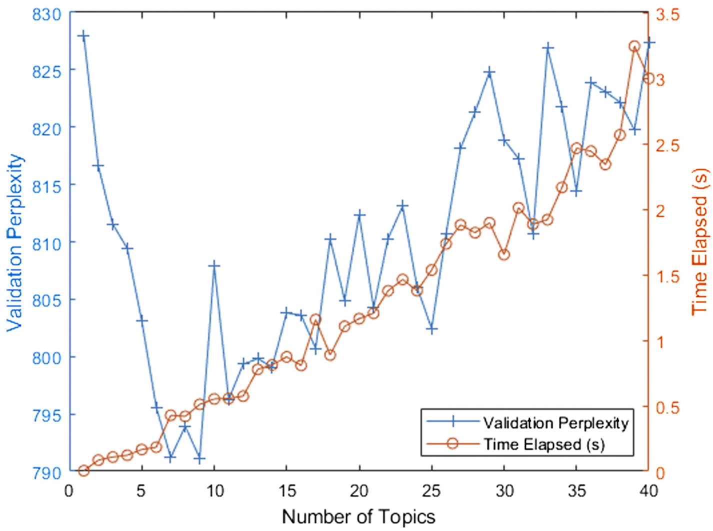

Figure 3.
Perplexity analysis using CVB0

across all documents in both the demand-side and supply-side data sets. The topic probabilities for the seven topics were then used to calculate two distribution profiles in Excel: the frequency profile of the top topic by probability across each data set; and the average topic probability across each data set.

3.5 Descriptive statistics and testing differences in topic profiles

Descriptive statistics and graphical representations for the topics and topic profiles of CIO job advertisements and CVs are then briefly examined. In order to test for differences between the demand and supply profiles for the distribution of topics, we applied multinomial goodness-of-fit tests using the Monte Carlo method and 10,000 simulations (Rao and Scott, 1981; Cressie and Read, 1984). To examine differences for particular topics, we used a test for differences in proportions via the Monte Carlo method and 5000 simulations. The results were used to test the hypotheses for the study.

3.6 Validity and reliability

We used several approaches for ensuring reliability and validity in the model solution (Maier *et al.*, 2018). To ensure reliability in model selection, we examined the data using different model parameters before deciding on the number of topics for investigation. Perplexity analysis was employed to examine all solutions between one and forty topics before finally examining the best candidate models for seven and nine topics. To confirm the validity of the model solution, we examined inter-individual intepretability of the topic model solution (Maier *et al.*, 2018). The process involved deliberation among four researchers. Topic model output was assessed by the first researcher and the topics were identified. Subsequently, topic model output was examined by the second research and their interpretations were compared. This process continued for the third and fourth researchers. Both intra-topic and inter-topic semantic validity were reviewed, and discrepancies discussed and resolved. The final result was a set of agreed topics across the four researchers.

4. Results

The top-100 most frequent words used in our complete bag of words model are shown in Figure 4. As we can see, the most frequent words were “technology”, “information”, “business”, “management”, “experience”, “system” and “service”.

The probabilities of the words for each of the seven topics are shown in the word clouds in Figure 5. While there is some obvious overlap between the common words in the topics, clear themes emerge from each of the seven groupings:

(1) Topic 1 suggests IT-related “Business development experience”, including related words such as business, development, operation, planning, project, company, establish and similar.

(2) Topic 2 surfaces some strategy and leadership words, which we will call “Strategic leadership”.

(3) Topic 3 points to team management, service and support issues, which we term “Team management and support”.

(4) Topic 4 relates to sector-specific experience, such as healthcare, which we label as “Sector-specific experience”.

(5) Topic 5 appears to refer to “IT management”, including words such as information, technology, management, security and system.

(6) Topic 6 examines overall CIO experience, including time-related words. We refer to this topic as “CIO experience”.

(7) Topic 7 specifies qualifications for the position, termed “Qualifications”.

The probability distributions of the seven topics in the demand and supply data sets are shown in Figures 6 and 7. Figure 6 illustrates distributions according to the top topic (highest probability) in each document. This emphasises the most important capability that is demanded in a job ad. Here, we see that IT management experience (−12.5%), CIO experience

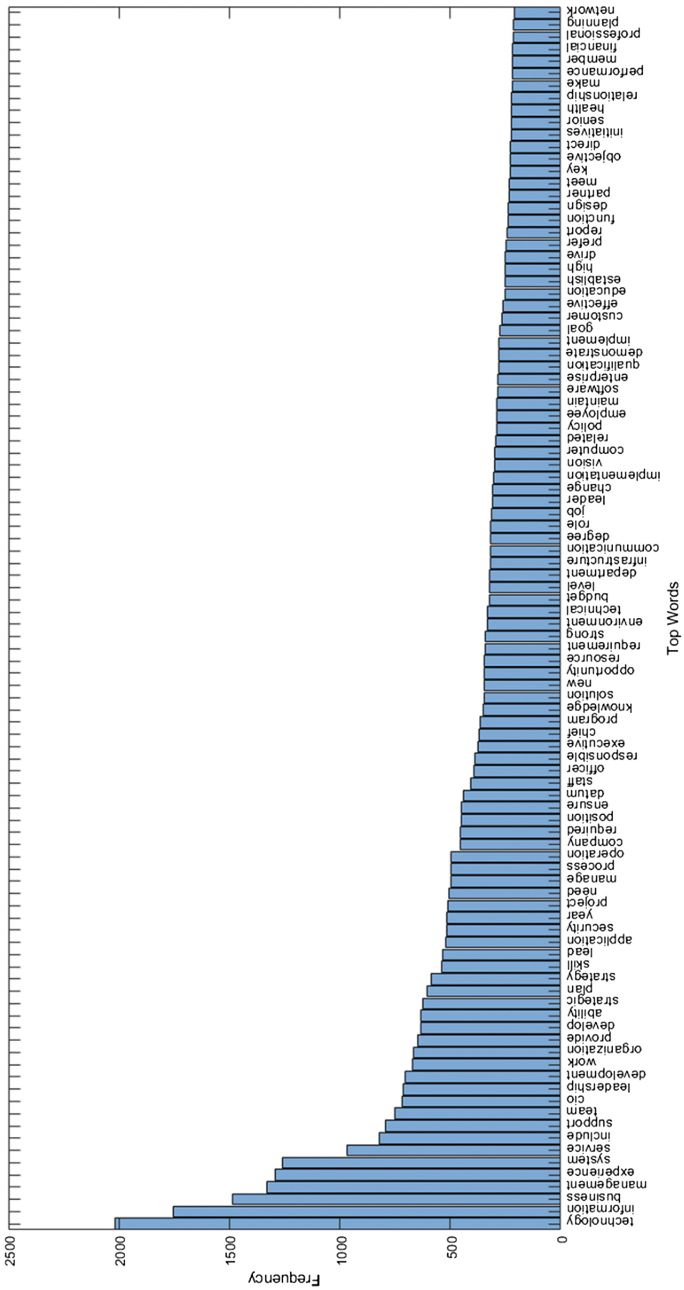

Figure 4.
Top-100 words used in CIO job advertisements

Figure 5.
Word clouds for the 7 topics

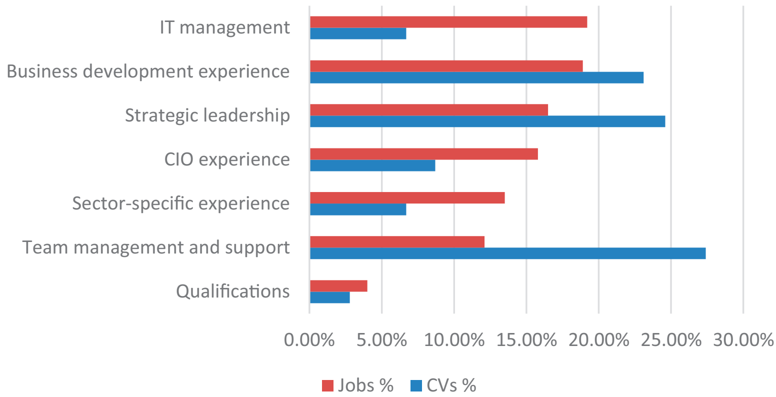

Figure 6.
Demand and supply topic distribution profiles (top topic)

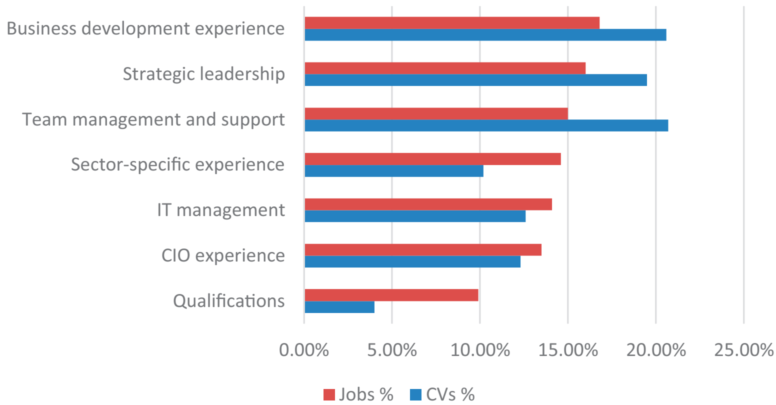

Figure 7.
Demand and supply topic distribution profiles (mean)

(−7.1%) and sector-specific experience (−6.8%), as well as qualifications (−1.2%), are all in scarce supply. This implies that rarer technology experience and qualifications and experience specific to the CIO position and industry are underrepresented in the supply pool. In obverse, softer, more general business skills are overrepresented in the CVs, including team management and support (+15.3%), strategic leadership (+8.1%) and business development experience (+4.2%).

Figure 7 examines the average distributions. This emphasises the general pattern of capabilities demanded and supplied. Here, we see the same pattern of under- or over-representation as in Figure 6. However, the size of the differences is much smaller for all capabilities except for qualifications, which fell to −5.9%.

Tables 1 and 2 formally examine differences between the profiles of the topic distributions for CIO job advertisements and CIO curriculum vitae using the multinomial goodness-of-fit test via the Monte Carlo method and 10,000 simulations. The results from these tables are used to test the hypotheses for the study. Table 1 examines statistical differences between the profiles for the most important topic in documents, while Table 2 examines differences according to the mean distributions. Table 1 confirms that the distributions are statistically different according to the topics (χ^2 = 144.199, DF = 6, p < 0.0001), a finding which is supported by the data for mean distributions (χ^2 = 86.275, DF = 6, p < 0.0.0001). Thus, the

Table 1. Multinomial goodness-of-fit test–Monte Carlo method (top topic)

No	Topic	CVs	CVs % (*C*)	Jobs	Jobs % (*J*)	Diff. (*C-J*)	*Z*-score	*p*-value	Conf. Int.*	*p*-value
1	Business development experience	221	23.1	56	18.9	0.042	1.502	0.133	[−0.013, 0.095]	0.128
2	Strategic leadership	235	24.6	49	16.5	0.081	3.072	0.002	[−0.026, 0.135]	0.003
3	Team management and support	262	27.4	36	12.1	0.153	6.337	<0.0001	[−0.096, 0.209]	<0.0001
4	Sector-specific experience	64	6.7	40	13.5	−0.068	−3.072	0.002	[−0.105, −0.032]	0.001
5	IT management	64	6.7	57	19.2	−0.125	−5.065	<0.0001	[−0.164, −0.087]	<0.001
6	CIO experience	83	8.7	47	15.8	−0.071	−2.985	0.003	[−0.110, −0.032]	<0.001
7	Qualifications	27	2.8	12	4.0	−0.012	−0.78	0.436	[−0.036, 0.010]	0.296
	Total Sample	*956*		*297*						

Note(s): **Confidence interval for differences based on 5000 Monte Carlo simulations
Multinomial GOF, Monte Carlo Method, $\chi^2 = 144.199$, DF = 6, $p < 0.0001$, Simulations = 10,000

No	Topic	CVs	CVs % (C)	Jobs	Jobs % (J)	Diff. (C-J)	Z-score	p-value	Conf. Int.*	p-value
1	Business development experience	221	20.6	56	16.8	0.038	1.413	0.158	[−0.014, 0.092]	0.165
2	Strategic leadership	235	19.5	49	16.0	0.035	1.321	0.187	[−0.015, 0.085]	0.165
3	Team management and support	262	20.7	36	15.0	0.057	2.235	0.025	[0.004, 0.108]	0.027
4	Sector–specific experience	64	10.2	40	14.6	−0.044	−1.841	0.066	[−0.087, −0.002]	0.045
5	IT management	64	12.6	57	14.1	−0.015	−0.559	0.576	[−0.059, 0.026]	0.497
6	CIO experience	83	12.3	47	13.5	−0.012	−0.435	0.663	[−0.057, 0.030]	0.594
7	Qualifications	27	4.0	12	9.9	−0.059	−3.078	0.002	[−0.088, −0.031]	0.001
	Total Sample	*956*		*297*	*297*					

Note(s): *Confidence interval for differences based on 5000 Monte Carlo simulations
Multinomial GOF, Monte Carlo Method, $\chi^2 = 86.275$, DF = 6, $p < 0.0.0001$, Simulations = 10,000

Table 2. Multinomial goodness-of-fit test–Monte Carlo method (topic distribution)

results support H1: Required CIO capabilities demanded from organisations do not match CIO supply.

We further examined differences between the importance of particular topics in the profiles by testing for differences in the proportions of the topics between the job advertisements and the CVs. The results in Table 1 show a statistically significant oversupply of strategic leadership ($p = 0.003$) and team management and support ($p < 0.0001$) for top topics. However, in the mean distributions, only team management and support is found to be in oversupply ($p = 0.027$). Thus, we are able to offer partial support for H2: Soft, general, non-specific capabilities will be over-supplied.

Turning to capabilities with an undersupply, we find that sector-specific experience ($p = 0.001$), IT management ($p < 0.001$) and CIO experience ($p < 0.001$) are in statistically significant deficit in Table 1 for top topics. Of these capabilities, only sector-specific experience is in statistically significant deficit in Table 2 ($p = 0.045$). However, qualifications is added as an area with a very significant undersupply ($p = 0.001$). Overall, our results provide partial support for H3: Hard, rarer specific skills, qualifications and experience will be under-supplied.

5. Discussion and conclusions

Over the years, the widespread adoption and strategic importance of digital technologies in organisations has required significant changes to the fundamental role of the CIO (La Paz, 2017; McLean and Smits, 2012). In a world driven by digital business, the CIO plays a pivotal role in facilitating the process of digital transformation throughout the organisation. This requires not only technical competence but the ability to also understand the needs of a large number of stakeholders from different areas of the business (Al-Taie *et al.*, 2014).

While the demand for CIOs who are capable to lead organisations' digital transformation efforts continues to increase, suitable qualified applicants for the role are in short supply (Jones *et al.*, 2019; La Paz *et al.*, 2019). As a result, in order to help address some of these challenges, this study aimed to unveil the current competences required by organisations for the role of CIO as well as to explore the level of fit between these requirements and CIOs' curriculum vitae. In order to achieve this goal, using organisational capability as a theoretical foundation, state-of-the art machine learning techniques were used to collect and analyse job postings and curriculum vitae of CIOs on LinkedIn.

Seven clear themes emerged from the data set. These themes confirmed that the current organisational requirements for CIOs represent a mix of general managerial capabilities such as leadership and team management, business development experience and strategic leadership (also traits for other C-Suite executives), and elements specific to the role, functional area and sector of the business – e.g. IT management skills, IT-related CIO experience and sector-specific experience. As a result, it appears that the "ideal" CIO must possess a combination of soft and hard skills in order to be able to create value for the organisation from vast amounts of data and increase organisational resilience through information security (Feeny *et al.*, 1992; Peppard *et al.*, 2011; La Paz *et al.*, 2010).

The analysis also identified a gap between the capabilities demanded by organisations and supplied by CIOs–supporting the first hypothesis. This is in line with La Paz *et al.* (2019), who also found a clear gap between the organisational definitions for the CIO position and the individuals hired. As the effects of technological advances produce disruptive change in businesses, it is expected that the requirements and demands for the CIO profile evolve at a pace that the workforce pool is not able to supply (Rao and Prasad, 2018).

The most salient gap relates to the lack of specific skills, qualifications and experience such as IT management experience, CIO experience and sector-specific experience. This underrepresentation in the supply pool is not surprizing as an experienced CIO can bring extraordinary value to organisations by leveraging the company's IT capability (Lim *et al.*, 2012). An experienced CIO in a particular sector or industry is, without a doubt, a very scarce resource (Nevo and Wade, 2010; La Paz *et al.*, 2019).

On the other hand, as expected, general business skills were overrepresented in the supply pool. C-suite executive skills and capabilities such team management and strategic leadership were the most common capability in oversupply. Perhaps this is because these softer, general and industry non-specific managerial skills are also available from executives originating from other areas of the business such as digital marketing (Kaarst-Brown, 2008, p. 455).

This study makes three distinctive contributions. First, by applying LDA, it provides a new research method and process for identifying competence requirements and gaps as well as ascertaining job fit. This approach may be helpful to other domains of research (e.g. human resource management, public policy, and economic and workforce development) in the process of identifying specific competences required by organisations for particular roles as well as to understand the level of fit between such requirements and a potential pool of applicants. Second, it provides unique insight into the current supply and demand for the role of CIO through the lens of RBV. This provides a contribution to the stream of information systems (IS) research focused on understanding CIO archetypes and how individual capabilities provide value to companies. Third, this can be useful for practitioners (e.g. potential CIO candidates) to understand current market requirements and for companies aiming to develop internal training that meet present and future skill gaps. It also could be useful for professional organisations (e.g. CIO Forum) to validate the need to develop mentoring schemes that help meet such high demand and relative undersupply of qualified CIOs.

While this paper makes positive contributions, there are some limitations. The sample involved is limited in size and to a specific point in time and the information available on LinkedIn. For instance, there is considerable scope for future research assessing the changes in the job-fit gap through a longitudinal study with a larger, broader sample from other sources that would help to explain better the effects of market trends and technological disruption on the required profile for the CIO role. Moreover, there are strong opportunities to apply and extend the methodology used in this study to other roles and contexts.

References

Al-Taie, M.Z., Lane, M. and Cater-Steel, A. (2014), "The relationship between organisational strategic IT vision and CIO roles: one size does not fit all", *Australasian Journal of Information Systems*, Vol. 18 No. 2, pp. 59-89.

Andriole, S.J. (2009), "Boards of directors and technology governance: the surprising state of the practice", *Communications of the Association for Information Systems*, Vol. 24, pp. 373-394.

Asuncion, A., Welling, M., Smyth, P. and Teh, Y.W. (2009), "On smoothing and inference for topic models", *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*, June, Quebec.

Barney, J. (1991), "Firm resources and sustained competitive advantage", *Journal of management*, Vol. 17 No. 1, pp. 99-120.

Barua, A., Konana, P., Whinston, A.B. and Yin, F. (2004), "An empirical investigation of net-enabled business value", *MIS Quarterly*, Vol. 28 No. 4, pp. 585-620.

Bastani, K., Namavari, H. and Shaffer, J. (2019), "Latent Dirichlet Allocation (LDA) for topic modeling of the CFPB consumer complaints", *Expert Systems with Applications*, Vol. 127, pp. 256-271.

Bharadwaj, A.S. (2000), "A resource-based perspective on information technology capability and firm performance: an empirical investigation", *MIS Quarterly*, Vol. 24 No. 1, pp. 169-196.

Blei, D.M., Andrew, Y.N.G. and Michael, I.J. (2003), "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, Vol. 3, pp. 993-1022.

Boehm, M., Stolze, C. and Thomas, O. (2013), "Teaching the chief information officers: an assessment of the interrelations within their skill set", *Wirtschaftsinformatik Proceedings*, pp. 1573-1587, Article 98, available at: https://aisel.aisnet.org/wi2013/98/ (accessed 8 May 2021).

Brown, N.C., Crowley, R.M. and Elliott, W.B. (2020), "What Are You Saying? Using topic to detect financial misreporting", *Journal of Accounting Research*, Vol. 58 No. 1, pp. 237-291.

Cao, G., Duan, Y. and Cadden, T. (2019), "The link between information processing capability and competitive advantage mediated through decision-making effectiveness", *International Journal of Information Management*, Vol. 44, pp. 121-131.

Carillo, K., Scornavacca, E. and Za, S. (2017), "The role of media dependency in predicting continuance intention to use ubiquitous media systems", *Information and Management*, Vol. 54 No. 3, pp. 317-335.

Carter, M., Grover, V. and Thatcher, J.B. (2011), "The emerging CIO role of business technology strategist", *MIS Quarterly Executive*, Vol. 10, pp. 19-29.

Castanias, R.P. and Helfat, C.E. (2001), "The managerial rents model: theory and empirical analysis", *Journal of Management*, Vol. 27 No. 6, pp. 661-678, doi: 10.1016/S0149-2063(01)00117-9.

Chen, Y.-C. and Wu, J.-H. (2011), "IT management capability and its impact on the performance of a CIO", *Information & Management*, Vol. 48 No. 4, pp. 145-156.

Chen, D.Q., Preston, D.S. and Weidong, X.I.A. (2010), "Antecedents and effects of CIO supply-side and demand-side leadership: a staged maturity model", *Journal of Management Information Systems*, Vol. 27 No. 1, pp. 231-271, doi: 10.2753/MIS0742-1222270110.

Chun, M. and Mooney, J. (2009), "CIO roles and responsibilities: twenty-five years of evolution and change", *Information and Management*, Vol. 46 No. 6, pp. 323-334.

Cressie, N. and Read, T.R.C. (1984), "Multinomial goodnet-of-fit tests", *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 46 No. 3, pp. 440-464.

Drechsler, K.L., Hund, A. and Wagner, H.-T. (2018), *Championing Digital Innovation Success: The Role of CDOs*, *Paper Presented at the PACIS*.

Earley, S. (2017), "The evolving role of the CDO", *IT Professional*, Vol. 19 No. 1, pp. 64-69.

Feeny, D., Edwards, B.R. and Simpson, K.M. (1992), "Understanding the CEO/CIO relationship", *MIS Quarterly*, Vol. 16, pp. 435-448.

Fiegener, M. and Coakley, J. (1995), "CIO problems and practices: 'impression management'", *Journal of Systems Management*, Vol. 46 No. 6, p. 56.

Galbraith, J. (2014), "Organizational design challenges resulting from big data", *Journal of Organization Design*, Vol. 3 No. 1, pp. 2-13.

Gonzalez, P.A., Ashworth, L. and McKeen, J. (2018), "The CIO stereotype: content, bias, and impact", *The Journal of Strategic Information Systems*, Vol. 28 No. 1, pp. 83-99.

Griffiths, T.L. and Steyvers, M. (2004), "Finding scientific topics", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 101 Supplement 1, pp. 5228-5235.

Grover, V., Seung-Ryul, J., Kettinger, W.J. and Lee, C.C. (1993), "The Chief Information Officer: a study of managerial roles", *Journal of Management Information Systems*, Vol. 10, pp. 107-130.

Guo, Y., Barnes, S.J. and Jia, Q. (2017), "Understanding quality using online reviews: tourist satisfaction analysis using Latent Dirichlet Allocation", *Tourism Management*, Vol. 59, pp. 467-483.

Hagiu, A. and Wright, J. (2020), "When data creates competitive advantage", *Harvard Business Review*, Vol. 98 No. 1, pp. 94-101.

Jones, M.C., Kappelman, L., Pavur, R., Nguyen, Q.N. and Johnson, V.L. (2019), "Pathways to being CIO: the role of background revisited", *Information and Management*, Vol. 57 No. 5, 103234, doi: 10.1016/j.im.2019.103234.

Kaarst-Brown, M.L. (2008), "Understanding an organization's view of the CIO: the role of assumptions about IT", *MIS Quarterly Executive*, Vol. 4 No. 2, Article 4, available at: https://aisel.aisnet.org/misqe/vol4/iss2/4.

Kane, G.C., Palmer, D., Phillips, A.N., Kiron, D. and Buckley, N. (2018), "Coming of age digitally", *MIT Sloan Management Review and Deloitte Insights*, June, available at: https://sloanreview.mit.edu/projects/coming-of-age-digitally/ (accessed 8 May 2021).

Kearns, G.S. and Lederer, A.L. (2003), "A resource-based view of strategic IT alignment: how knowledge sharing creates competitive advantage", *Decision Sciences*, Vol. 34 No. 1, pp. 1-29.

Krotov, V. (2015), "Bridging the CIO-CEO gap: it takes two to tango", *Business Horizons*, Vol. 58 No. 3, pp. 275-283.

La Paz, A. (2017), "How to become a strategist CIO", *IT Professional*, Vol. 19 No. 1, pp. 48-55.

La Paz, A., Laengle, S. and Cancino, C. (2010), "The core skills of the CIO we need", *Dyna*, Vol. 77 No. 163, pp. 172-180.

La Paz, A., Vasquez, J. and Miranda, J. (2019), "The CIO gap and mismatch", *IT Professional*, Vol. 21 No. 2, pp. 66-72, doi: 10.1109/MITP.2018.2883089.

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T. and Hoffmann, M. (2014), "Industry 4.0", *Business and Information Systems Engineering*, Vol. 6 No. 4, pp. 239-242.

Leonard-Barton, D. (1992), "Core capabilities and core rigidities: a paradox in managing new product development", *Strategic Management Journal*, Vol. 13 No. S1, pp. 111-125.

Lewis, B.R., Snyder, C.A. and Rainer, J.R.K. (1995), "An empirical assessment of the information resource management construct", *Journal of Management Information Systems*, Vol. 12, pp. 199-223.

Li, Y. and Tan, C.-H. (2013), "Matching business strategy and CIO characteristics: the impact on organizational performance", *Journal of Business Research*, Vol. 66 No. 2, pp. 248-259.

Lim, J.-H., Stratopoulos, T.C. and Wirjanto, T.S. (2012), "Role of IT executives in the firm's ability to achieve competitive advantage through IT capability", *International Journal of Accounting Information Systems*, Vol. 13 No. 1, pp. 21-40.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H. and Adam, S. (2018), "Applying LDA topic modeling in communication research: toward a valid and reliable methodology", *Communication Methods and Measures*, Vol. 12 Nos 2-3, pp. 93-118.

McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D. and Barton, D. (2012), "Big data: the management revolution", *Harvard Business Review*, Vol. 90 No. 10, pp. 60-68.

McCue, A. (2007), "In need of an upgrade", *Director (00123242)*, Vol. 60 No. 8, pp. 60-65, available at: http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=24415701&lang=es&site=ehost-live.

McLean, E. and Smits, S. (2012), "Management, leadership, and the roles of the CIO", *Robinson College of Business*, Georgia State Univ., available at: http://bit.ly/2hscfqL.

Mithas, S., Tafti, A., Bardhan, I. and Goh, J.M. (2012), "Information technology and firm profitability: mechanisms and empirical evidence", *MIS Quarterly*, Vol. 36 No. 1, pp. 205-224.

Nevo, S. and Wade, M.R. (2010), "The formation and value of IT-enabled resources: antecedents and consequences of synergistic relationships", *MIS Quarterly*, Vol. 34 No. 1, pp. 163-183.

Nolan, R. and McFarlan, F.W. (2005), "Information technology and the board of directors", *Harvard Business Review*, Vol. 83, pp. 96-106.

Peppard, J., Edwards, C. and Lambert, R. (2011), "Clarifying the ambiguous role of the CIO", *MIS Quarterly Executive*, Vol. 10 No. 1, pp. 31-44.

Rao, J.N.K. and Scott, A.J. (1981), "The analysis of categorical data from complex sample surveys: chi-squared tests for goodness-of-fit and independence in two-way tables", *Journal of the American Statistical Association*, Vol. 76, pp. 221-230.

Rao, S.K. and Prasad, R. (2018), "Impact of 5G technologies on industry 4.0", *Wireless Personal Communications*, Vol. 100 No. 1, pp. 145-159.

Ross, J.W. and Feeny, D.F. (1999), "The evolving role of the CIO", CISR Working Paper No. 308, Massachusetts Institute of Technology, Cambridge, MA.

Schambach, T. and Blanton, J.E. (2002), "The professional development challenge for IT professionals", *Communications of the ACM*, Vol. 45 No. 4, pp. 83-87.

Schölkopf, B., Platt, J. and Hofmann, T. (Eds). (2007), *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, Vol. 19.

Sirmon, D.G., Hitt, M.A., Arregle, J.L. and Campbell, J.T. (2010), "The dynamic interplay of capability strengths and weaknesses: investigating the bases of temporary competitive advantage", *Strategic Management Journal*, Vol. 31 No. 13, pp. 1386-1409.

Smith, G.S. (2011), *Straight to the Top: Becoming a World-Class CIO*, John Wiley & Sons, Hoboken, NJ.

Sobol, M.G. and Klein, G. (2009), "Relation of CIO background, IT infrastructure, and economic performance", *Information and Management*, Vol. 46 No. 5, pp. 271-278.

Spitze, J.M. and Lee, J.J. (2012), "The renaissance CIO project: the invisible factors of extraordinary success", *California Management Review*, Vol. 54 No. 2, pp. 72-91.

Tirunillai, S. and Tellis, G.J. (2014), "Mining marketing meaning from online chatter: analysis of big data using Latent Dirichlet Allocation", *Journal of Marketing Research*, Vol. 51 No. 4, pp. 463-479.

Wade, M. and Hulland, J. (2004), "The resource-based view and information systems research: review, extension, and suggestions for future research", *MIS Quarterly*, Vol. 28 No. 1, pp. 107-142.

Weill, P. and Woerner, S.L. (2013), "The future of the CIO in a digital economy", *MIS Quarterly Executive*, Vol. 12 No. 2, Article 3, available at: https://aisel.aisnet.org/misqe/vol12/iss2/3.

Whitler, K.A., Boyd, D.E. and Morgan, N.A. (2017), "The criticality of CMO-CIO alignment", *Business Horizons*, Vol. 60 No. 3, pp. 313-324.

Willcocks, L.P. and Sykes, R. (2000), "The role of the CIO and IT function in ERP", *Communications of the ACM*, Vol. 43 No. 4, pp. 32-38.

Corresponding author
Stuart Barnes can be contacted at: stuart.barnes@kcl.ac.uk

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**