

DIMENSIONAMENTO DA FORÇA DE TRABALHO FRENTE A UM GRANDE VOLUME DE DADOS

uma abordagem no âmbito teórico

Lucas Oliveira Gomes Ferreira¹
Universidade de Brasília
lucasoliveira@unb.br

André Luiz Marques Serrano²
Universidade de Brasília
andrelms@unb.br

Nara Cristina Ferreira Mendes³
Universidade de Brasília
naracfm@gmail.com

Resumo

Esta pesquisa emprega técnicas bibliométricas para identificar as características primordiais da produção de artigos científicos sobre Dimensionamento da Força de Trabalho (DFT) e sua relação com *big data*. De natureza aplicada, abordagem qualitativa e caráter exploratório, o estudo faz uso de um levantamento de estudos sobre dimensionamento e *big data*, recuperando um total de 87 artigos científicos publicados entre 1992 a 2020. O objetivo dessa pesquisa é apresentar um *framework* de suporte à produção científica para a avaliação do DFT. A contribuição principal foi a sistematização de um referencial teórico consolidado de DFT para a composição de um ambiente *big data*. A pesquisa gerou perspectiva de continuidade por meios tecnológicos para solucionar problemas de gestão pública e de produtividade da capacidade individuais.

Palavras-chave: Dimensionamento. *Framework*. *Big data*. Produção científica. Setor público.

WORKFORCE SIZING FACE TO LARGE DATA VOLUME

A theoretical approach

Abstract

This research uses bibliometric techniques to identify the main characteristics of the production of scientific articles on Workforce Dimensioning – DFT – and its relationship with big data. With an applied nature, qualitative approach and exploratory character, the study uses a survey of studies on dimensioning and big data, recovering a total of 87 scientific articles published between 1992 to 2020. The objective of this research is to present a support framework for scientific production for the evaluation of the DFT. The main contribution was the systematization of a consolidated theoretical reference of DFT for the composition of a big data environment. The research generated a perspective of continuity by technological means to solve problems of public management and productivity of individual capacity.

Keywords: Sizing. Framework. Big data. Scientific production. Public Sector.

¹ Doutor em Ciências Contábeis pela Universidade de Brasília (UnB). Mestre em Ciências Contábeis pela UnB e Graduado em Ciências Contábeis pela UnB.

² ² Doutor em Economia pela Universidade de Brasília (UnB). Mestre em Economia pela UnB e Graduado em Matemática pela UnB.

³ Doutora em Ciências Contábeis pela Universidade de Brasília (UnB). Mestre em Economia pela UnB, Especialista em Ciências Contábeis com ênfase em IFRS pela UnB e Graduada em Ciências Contábeis pela UnB.



1 INTRODUÇÃO

Mudanças tecnológicas estão ocorrendo a uma velocidade sem precedentes no âmbito das Tecnologias da Informação e Comunicação - TIC, que impuseram conseqüentemente o crescimento das bases e pleitearam o desenvolvimento de novas soluções e tecnologias que auxiliem a gestão pública. Diante desta inevitabilidade, surge o ambiente *big data*, propondo novas interpelações para a geração e manipulação de expressivos volumes de dados, ou seja, o termo *big data* está relacionado a grandes quantidades de dados, providos de diferentes fontes, com controles distribuídos e descentralizados. Por conseguinte, a maior parte dos trabalhos gerados migraram para os computadores, que envolvem tecnologias e processos, armazenamento e disseminação da informação, agora ligados a uma rede, que podem promover decisões que poderão se tornar desatualizadas da "noite para o dia".

Todavia, a incerteza, singularmente em relação ao futuro tecnológico, tem se tornado um fato constante no dia a dia da gestão de empresas privadas e na gestão do poder público. A despeito disso, decisões de toda natureza e grau de importância precisarão ser tomadas cotidianamente. Considerando a lógica keynesiana, expressa por McAfee e Brynjolfsson (2012), “como seres vivos e em movimento, somos forçados a agir... [mesmo quando] nosso conhecimento existente não fornece uma base suficiente para uma expectativa matemática calculada”. Assim sendo, a questão da velocidade, volume e variedade de dados, faz surgir o conceito de ambientes e cenários voltados para o tratamento da informação. Por isso, as decisões de gestão de pessoas têm cada vez mais utilizado modelos quantitativos, estatísticos e lógicos (LONG, 2018).

Nesta perspectiva é necessária a aplicação de novos dispositivos que viabilizem de maneira inteligente as bases de dados disponíveis em informações satisfatórias e úteis para a tomada de decisão, que permitirão a contração sistemática da variabilidade quanto às características das entregas geradas, do refinamento intrínseco da avaliação da produtividade, da confiabilidade e do custo do que está sendo entregue para a sociedade. Sendo assim, a estruturação da força de trabalho, a execução de modelos matemáticos, testes estatísticos, e por fim, em termos de custos e formas inovadoras para o tratamento das informações com vistas a uma melhor percepção e tomada de decisão, tem-se a construção de um *big data*, que poderá trazer vantagens competitivas por disponibilizar e transformar componentes válidos sobre os mais diversos aspectos organizacionais através da análise de banco dados, os quais, (WAXIN *et al.*, 2018), poderão detectar falhas na geração de valor e avançarão para garantir melhorias em processos e elaboração de cartas de controles visando a otimização da produção

(MORSHED *et al.*, 2018; ELLINGTON; WILSON, 2017; IPMA, 2002; HELTON; JACKSON, 2007; JACOBSON 2010; DICKINSON; SULLIVAN; HEAD, 2015).

Diante as demandas governamentais, prevê-se que no serviço público, a força de trabalho será menor e focada principalmente na entrega dos resultados e não exclusivamente na gestão de políticas, o que resulta em maiores desafios para os gestores públicos que planejam a força de trabalho (DICKINSON; SULLIVAN; HEAD, 2015). Dessa forma, o entendimento quanto ao DFT se insere nesse contexto de crescimento exponencial de dados coletados e armazenados. Ao mesmo tempo, a administração pública despende bastante tempo e esforço na construção e manutenção dessas bases de dados que não necessariamente serão analisadas. Algumas dessas bases geram tamanha dimensão que nem os gestores conhecem as informações que poderão ser extraídas ou a relevância que elas poderão ter para a própria máquina pública.

Cabe ressaltar que a utilização de modelos matemáticos, estatísticos e computacionais, através da composição de *big data* poderão ser utilizadas para uma tomada de decisão mais acurada (BUFFA, 1980; LI; CHEN; CAI, 2007; COROMINAS; LUSA; OLIVELLA, 2012). Portanto, o objetivo desse artigo é apresentar um *framework* que dê suporte à produção utilizando aspectos da composição de *big data* para o Dimensionamento da Força de trabalho na administração pública federal. O método de pesquisa foi realizado em três etapas: i) pesquisa bibliográfica sobre DFT; ii) elaboração de um *framework* teórico-conceitual referente a *big data*; e iii) ilustração de aplicação do *framework* proposto.

Este trabalho está estruturado de forma a conter, além dessa seção introdutória (Seção 1), uma seção que trata do referencial teórico alusivo ao tema Dimensionamento da Força de Trabalho e *Big data* (Seção 2), uma seção que descreve os procedimentos metodológicos relativos à realização da revisão sistemática da literatura (Seção 3), uma seção de análise e discussão dos resultados (Seção 4) e, finalmente, uma seção com as considerações finais, limitações do estudo e sugestão para estudos futuros (Seção 5).

2 REFERENCIAL TEÓRICO

Ao longo deste referencial teórico será abordada a relevância do dimensionamento da capacidade produtiva no contexto do planejamento da força de trabalho, bem como os elementos estruturantes de um *framework* diante de um ambiente de *big data*.

2.1 A RELEVÂNCIA DO DIMENSIONAMENTO DA CAPACIDADE PRODUTIVA FRENTE A COMPOSIÇÃO DO PLANEJAMENTO DA FORÇA DE TRABALHO

Analisando a literatura de dimensionamento aplicado como ferramenta de gestão, fica evidente a sua inserção dentro do Planejamento da Força de Trabalho - PFT (MARINHO; VASCONCELLOS, 2007). Em inglês, pode ser chamado de *Workforce Planning* (BARD; MORTON; WANG, 2007), *Manpower Planning* (CAPPELLI; KELLER, 2014), *Strategic Staff Planning* (TORRE; LUSA; MATEO, 2016) ou ainda de *Staff Planning* (LI; CHEN; CAI, 2007). O PFT se refere às ferramentas de gestão de pessoas cujas estratégias buscam sustentar a pluralidade e quantidade de indivíduos com o perfil profissional adequado, alocados nos lugares previstos e ajustados ao longo do tempo para realizar as entregas em curto e longo prazo.

O PFT envolve o uso de estratégias baseadas em evidências para identificação sistemática e análise do que uma organização possui e precisará (STOKKER; HALLAM, 2009). Trata-se de um processo com múltiplas fases cujos componentes principais são a previsão da demanda, a previsão de suprimento, geração de dados e as estratégias para equilibrá-los (CHOWDHURY, 2016). Por exemplo, a literatura sobre alocação de pessoal propõe a estruturação mensal, semanal ou anual para alocação de funcionários considerando uma série de limitações quanto a definição da quantidade de pessoas e tempo disponível (LI; CHEN; CAI, 2007).

Estabelecer a quantidade de pessoas necessárias é uma decisão interdependente da decisão de onde e quando alocá-las. Não obstante, existe uma distinção entre esses dois processos. Os modelos de alocação, optam e se concentram na alocação de uma força de trabalho determinada para turnos ou tarefas (ERNST *et al.*, 2004), enquanto o dimensionamento (staff sizing) se preocupa especificamente com a estimativa quantitativa da força de trabalho. Existe a viabilidade de integrar o dimensionamento com a alocação, mas grande parte dos modelos de alocação de pessoal utilizam quantidades de pessoas predeterminadas para o processo de alocação (LI; CHEN; CAI, 2007).

Embora faça muito mais sentido fazê-los em conjunto, os métodos empregados para cada um, podem ser bastante distintos, e dividi-los auxilia a decompor um problema tão complexo em soluções mais específicas para cada objetivo (SERRANO; MENDES; ABILA, 2019). Isto posto, o DFT se torna uma técnica que se insere no PFT para tomada de decisão, ou seja, faz-se a comparação das necessidades futuras das unidades organizacionais, seja no curto, médio ou longo prazo, com sua situação atual, em termos de pessoal alocado. Tal comparação tem como objetivo fornecer à administração a diferença entre o quantitativo e o perfil de

profissionais adequados ao alcance de seus objetivos estratégicos e seu quadro atual. Essa diferença deverá ser reduzida e suas características nortearão medidas de ajuste, tais como contratação, realocação, qualificação, demissão, entre outras.

No Brasil, apesar de alguns artigos usarem a divisão qualitativa e quantitativa de dimensionamento (TANOS; MASSAROLLO; GAIDZINSKI, 2000), especificamente a literatura de dimensionamento em hospitais (MARINHO; VASCONCELLOS, 2007), não se observa que tal divisão seja internacionalmente replicada (LI; CHEN; CAI, 2007). Essa divisão causa confusão porque, ao contrário do que o nome indica, as propostas qualitativas de dimensionamento não se baseiam em práticas metodológicas qualitativas de pesquisa, como o nome sugere (INOUE; MATSUDA, 2010). As práticas chamadas de dimensionamento que o fazem por meio da identificação de competências ou cargos (MARINHO; VASCONCELLOS, 2007), aproximam-se das técnicas pertencentes ao PFT para determinação do perfil adequado e não de dimensionamento.

O dimensionamento requer técnicas quantitativas para estimar um número (dimensão) e indicadores de força de trabalho para determinada tarefa, portanto, é importante ressaltar que não existe dimensionamento qualitativo ou com o uso de técnicas qualitativas (SERRANO; MENDES; ABILA, 2019). Em seguimento, as metodologias de trabalho devem estar perfeitamente ajustadas à realidade da organização, ou seja, é preciso otimizá-las, eliminando etapas sobrepostas ou desnecessárias (retrabalhos) e “desburocratizando” aquilo que for possível. A utilização de elementos qualitativos, não se baseiam necessariamente em práticas metodológicas de pesquisa e nem sempre possuem um escopo ou objetivo claro, mas sim, em aspectos da qualidade do atendimento ou resultado (INOUE; MATSUDA, 2010).

À vista disso, não há um consenso do que seria o dimensionamento qualitativo. Há, no entanto, evidências de que o tamanho da equipe interfere no desempenho (LATANÉ; WILLIAMS; HARKINS, 1979), além da perda motivacional causada pela dispersão de responsabilidade em grandes equipes (SIMMS; NICHOLS, 2014). A dimensão ideal de uma unidade é resultado do esforço conjunto e influenciado por múltiplos fatores individuais, grupais ou organizacionais, não se constituindo apenas por um aglomerado de competências individuais. Como demonstrado em estudos anteriores, adicionar um trabalhador não conduz necessariamente a um aumento proporcional na eficiência da produção (POWELL; SCHMENNER, 2002; TOHIDI; TAROKH, 2006). Levando à conclusão de que a questão é complexa para ser respondida apenas com mapeamentos, é preciso utilizar a literatura de composição e planejamento de equipes/grupos de trabalho para explicar e identificar outros

determinantes da efetividade de unidades de trabalho (KOZLOWSKI, 2015; MATHIEU; HOLLENBECK; VAN KNIPPENBERG; ILGEN, 2017).

O dimensionamento, que agora parece redundante apontar sua classificação como quantitativo, possui ampla literatura que abarca a diversidade de métodos aplicados para dar dimensão à força de trabalho, mas poucos trabalhos que sistematizem o desenvolvimento de *big data* aplicados ao DFT. Os artigos que aplicam modelos para dar dimensão à força de trabalho – *staff sizing* – são comumente publicados em revistas de engenharia de produção (McCLEAN, 1991), matemática (NILAKANTAN; RAGHAVENDRA, 2005) e economia (IGHRAVWE; OKE, 2014).

Embora existam diversos objetivos na aplicação do DFT, todos os métodos referentes a estas aplicações são modelos. Modelos desempenham um papel fundamental ao viabilizar a simplificação dos fenômenos com o objetivo de explicar padrões dos dados. Eles são importantes para descrever, prever, explicar fenômenos e, portanto, para tomada de decisão, já que possibilitam compreender uma variedade de questões em torno de um determinado padrão (FUM; MISSIER; STOCCO, 2007; NORRIS, 2005). Entretanto, a construção, utilização e avaliação de modelos quantitativos ou matemáticos dependem também de aspectos práticos (EDWARDS; HAMSON, 2007).

Independente da escolha das variáveis ou do contexto de aplicação, os modelos de dimensionamento, bem como qualquer modelo quantitativo, podem ser divididos conforme sua aplicabilidade ou funcionamento. Cada mecanismo possui formas diferentes de serem avaliados em termos de procedimentos de proposta, foco de análise, dificuldade de implementação, manipulação e avaliação da qualidade de ajuste. Deve-se apontar que o DFT viabiliza a coleta de informações para a proposição de *big data* e assim assumir um papel na formulação de uma tomada de decisão planejada e consciente, mas, como já foi dito anteriormente, trata-se de uma ferramenta que não substitui e, sim, informa gestores de forma válida e fidedigna sobre a sua força de trabalho (OLIVEIRA; BIANCHINI; ABBADE, 2007), estatísticos (AGUIR *et al.*, 2008) e de otimização (ROSA; FILHO, 2008).

Por fim, cabe ressaltar que alguns estudiosos de economia, tais como Greene e Hudzik (1981), apresentaram estudos sobre PFT e DFT em organizações públicas e ressaltam que o planejamento está relacionado com os estados futuros, além de estar intrinsecamente correlacionados a considerações organizacionais mais amplas e econômicas. Os principais motivos para a concentração dos esforços do DFT, talvez esteja na necessidade de otimizar recursos bastante escassos, principalmente em países em desenvolvimento para, ao mesmo tempo, garantir a disponibilidade de categorias de força de trabalho essenciais por meio de

esforços sistemáticos planejados (VERMA, 1984). Portanto, para explicar e identificar outros determinantes da efetividade de unidades de trabalho é preciso ir além, buscando analisar questões igualmente relevantes de qualidade de vida no trabalho, satisfação, bem-estar, fatores socioeconômicos além da utilização de ferramentas de gestão de desempenho e treinamento.

2.2 ELEMENTOS ESTRUTURANTES DE UM *FRAMEWORK*: AMBIENTE *BIG DATA*

Sistemas de *big data* configuram uma nova geração de tecnologias e arquiteturas, delineadas para remover expressivos volumes de uma ampla variedade de dados, permitindo a captura, evidenciação e análise de alta velocidade (CHEN; MAO; LIU, 2014). Outros pesquisadores inferem que o termo *big data* é utilizado de forma deturpada por provedores de soluções para classificar sistemas de análise de dados. Salienta-se que é mais do que isso, pois faculta o início de novos cenários e serviços, contribuindo as empresas responderem às novas demandas de forma mais ágil e mais assertiva, tangendo de forma mais eficiente no mercado, direcionando-se as melhores decisões de negócio a partir de uma diversidade e multidisciplinariedade referentes a fontes de informação combinadas (DAVENPORT; BARTH; BEAN, 2012).

127

Desta forma, *big data* está relacionado a conjuntos de dados, cujo tamanho encontra-se além da capacidade de ferramentas típicas de *software* de banco de dados para capturar, armazenar, gerenciar e analisar. Conforme a tecnologia avança ao longo do tempo, o tamanho dos conjuntos de dados que sejam qualificados, podem estar voltados para analisar questões relacionadas a força de trabalho, demanda essa, que visa aumentar com a necessidade do entendimento da complexidade das entregas (SANTOS, 2016).

Com o entendimento da complexidade das entregas, faz se necessário a transformação e real conhecimento das estruturas organizacionais – transformar a organização no sentido de monitorar o ambiente organizacional, em tempo real, e promover o desenvolvimento de modelos de previsão. Trata-se de um conjunto de soluções que pode aprimorar muito os processos de decisão gerencial, orientando o poder público para análise de dados e com possibilidade de gerar novos modelos de negócios.

As tecnologias atuais de *big data* voltadas a administração pública, sustentam a possibilidade de se trabalhar com um volume expressivo de dados, de diversas origens, fontes e apresentar uma variedade de formatos, além de agregar maior velocidade de processamento e maior confiabilidade em relação à perda de informações, aumentando o potencial das análises

na descoberta de evidências valiosas, como a geração de conhecimento para a tomada de decisão.

Entretanto, o uso efetivo das tecnologias para a composição de *big data* com viés ao dimensionamento da força de trabalho é praticamente nulo na literatura, mas com atributos singulares que já definidos e utilizados em outras áreas do conhecimento. De acordo as publicações levantadas, existe diferentes posições em relação a atributos que compõem um ambiente *big data*, mas nada específico ao DFT, mas representativo para composição de atributos que formularão bases efetivas para pesquisas e análises futuras. Três delas foram preliminarmente identificadas por Doug Laney (2001), descrevendo o impacto do volume, da velocidade e da variedade dos dados em data warehouses corporativos e por fim, outras características foram incorporadas, são elas: veracidade, valor, variabilidade (ZIKOPOULOS *et al.*, 2013; GANDOMI; HAIDER, 2015).

Volume está relacionado a grande quantidade de dados que são constituídas a partir de procedências variadas. É importante observar que a definição de volume pode variar de acordo com o setor, dependendo dos tipos de ferramentas de *software* disponíveis e dos tamanhos de conjuntos de dados usuais do setor. Com essas ressalvas, a definição de *big data* em muitos setores varia de algumas dezenas de terabytes para múltiplos petabytes.

Variabilidade está relacionado a diversidade das informações. Diante a uma perspectiva analítica, fazer a gestão de dados representa a maior dificuldade para utilização efetiva de grandes volumes de informações. Velocidade se relaciona com comunicações em tempo real a uma maior velocidade para troca de dados. Veracidade infere-se à confiabilidade dos dados, o que pode demandar atividades específicas para identificar os dados que não atendam a essa premissa e assim removê-los dos datasets. Valor infere-se ao retorno do investimento e é o resultado da combinação dos aspectos citados anteriormente. Por fim, variabilidade está relacionada a dados que estão em constante variação, como por exemplo, dados meteorológicos.

Assim, é possível determinar a noção do termo *big data*. Porém, especialistas acreditam que não há a necessidade da existência de todos os fatores ao mesmo tempo para um ambiente seja considerado *big data*, visto que existem casos em que há um maior destaque para uma ou outra característica, e em outros casos elas não são identificadas em sua totalidade

Outrossim, existem grandes adversidades relacionadas à complexidade do uso das tecnologias disponíveis e à escolha do conjunto de ferramentas, que poderão ser corretas para as finalidades pretendidas. Davenport, Barth e Bean (2012) apresentaram 3 dimensões a serem consideradas na estratégia de implantação e utilização de *big data*, que são: focar em um problema com foco nos ambientes internos e externos (dimensão 1); matriz de objetivos e

estágios de aplicação focando baixar custos e tornar mais rápidas as decisões (dimensão 2); e rapidez e agressividade em relação à concorrência para adotar *big data* (dimensão 3).

No que se refere aos trabalhos que buscaram desenvolver estruturas arquiteturais de *big data* para aplicação, pode-se citar, Raghupathi e Raghupathi (2014) que esboçou um *framework* pioneiro no contexto de *big data* para o domínio da saúde, que abarca nas camadas de fonte de dados, transformação, plataforma e aplicações analíticas. Outro trabalho relevante, refere-se a pesquisa de Wang e Hajli (2017), no qual, encontrou na composição de um ambiente *big data*, potenciais benefícios a indústria de medicamentos, que são acionados por um componente de processamento de dados, seguido por uma agregação de informações e um componente de visualização dos mesmos. Teoricamente, os trabalhos mencionados identificaram os distintos e relevantes componentes de uma arquitetura de *big data*.

3 METODOLOGIA

Na bibliometria são utilizados diversos indicadores para estudar a produção científica, as relações e interações entre seus elementos. Em linhas gerais, estes indicadores buscam medir a produtividade, estabelecer relações, identificar colaboração e ocorrência de elementos em uma amostra de documentos. O papel da área acadêmica, em especial os cursos de pós-graduação, é fundamental na produção e disseminação de conhecimento, visando atender às emergentes demandas das organizações e da sociedade. Artigos científicos, dissertações e teses apresentam, além de levantamentos teóricos baseados em outros documentos científicos, padrões rígidos de metodologia e padrões de pesquisa pautados em normas técnicas, sendo consideradas fontes fidedignas para extração de informações a partir de metodologias bibliométricas (MACHADO JUNIOR *et al.*, 2016)

Em termos de classificação de pesquisa, do ponto de vista de sua natureza, este artigo trata-se de uma pesquisa aplicada, do ponto de vista de seus objetivos é descritiva, do ponto de vista da abordagem é qualitativa e, do ponto de vista dos procedimentos técnicos, trata-se de uma revisão da literatura, de acordo com a proposição de Silva e Menezes (2005).

Com a intenção de demonstrar o estado da arte acerca dos temas DFT e *big data*, e o valor agregado aos serviços públicos com criação de um ambiente *big data* para determinação de indicadores, realizou-se a revisão de literatura com a perspectiva sistemática, cuja abordagem se estrutura em um protocolo de seleção e análise das referências, envolvendo as etapas (CRONIN; RYAN; COUGHLAN, 2008): (a) formulação da questão de pesquisa; (b) conjunto de critérios de inclusão e exclusão; (c) seleção e acesso da literatura; (d) avaliação da

qualidade da literatura incluída na revisão; e, (e) análise, síntese e disseminação dos resultados. Assim, baseado neste protocolo, apresenta-se a síntese do procedimento realizado nesta pesquisa:

I. Definição da questão da pesquisa: Quais são e o que abordam os estudos internacionais das bases web of science, da última década, sobre DFT *versus big data* com foco na administração pública?

II. Definição do conjunto de critérios de inclusão e exclusão: Inicialmente foram definidas as seguintes bases científicas para busca dos artigos: *Google Scholar*, *Science Direct* e *Emerald*. Posteriormente, foram definidas as palavras-chave e operadores booleanos: [“*big data*” AND “*workforce planning*” OR “*staff planning*” OR “*manpower planning*” OR “*workforce strategy*” OR “*staff strategy*” OR “*manpower strategy*” OR “*staff system*” OR “*workforce system*” OR “*manpower system*” AND “*public services*” OR “*public administration*”]. Quanto ao tipo, foram considerados artigos completos publicados em periódicos internacionais e nacionais, em língua inglesa ou portuguesa, sem restrição de área de conhecimento, excluindo-se capítulos de livros, livros, citações, patentes, editoriais, anais de eventos.

III) Seleção e acesso da literatura: termos de busca [“*workforce planning*” OR “*manpower planning*” OR “*staff planning*” OR “*workforce strategy*” OR “*staff strategy*” OR “*manpower strategy*” OR “*workforce system*” OR “*staff system*” OR “*manpower system*” AND “*public services*” OR “*public administration*” AND “*big data*”].

Porém, esta configuração inicial não se mostrou totalmente eficaz, uma vez os títulos apresentavam inconsistência entre as bases. Muitos foram os casos onde o título e subtítulo eram separados por hífen em uma base e por dois pontos (:) em outras. Em outros casos, alguns termos foram escritos de maneira diferente como “*data-base*” e “*database*”. Nestes casos, o *software* não considerou o título como sendo o mesmo. Por isso, foi necessário alterar a configuração para que fossem considerados como duplicados artigos com os mesmos autores, anos, periódico, volume e número. A consulta à *Web of Science* foi realizada em 2020 e utilizou-se a *string*: (“*workforce*” e “*big data*”) no campo de pesquisa “Título”. Além das bases internacionais, foi realizada uma busca na base de dados *SciELO* que não retornou nenhum resultado. Esta busca atendeu aos mesmos critérios estabelecidos, todavia com o termo de busca em português (“gestão de dados”). E por fim, quanto ao tipo de documento foram selecionados apenas os artigos e os *proceedings paper* indexados, excluindo os documentos caracterizados por *book review*, *editorial material* e *review* entre outros de menor ocorrência até o ano de 2020, resultando em 87 artigos sobre os termos pesquisados.

Preliminarmente, foi definido no estudo, que a busca contemplaria o campo de pesquisa Tópico, que engloba a consulta ao título, o que resultou em um total de 116 publicações, para as mesmas condições de busca. Porém, após uma análise preliminar nos títulos das publicações, constatou-se que a grande maioria não focava especificamente em de dados em ambientes *big data*, sendo assim, com o objetivo de contemplar somente os estudos relacionados a DFT e *big data*, áreas, optou-se por utilizar apenas o campo Título (SOMASUNDARAM; SHRIVASTAVA, 2009). Para realização do trabalho, foram utilizados os *softwares* Harzing's Publish or Perish (HARZING, 2010) e VOSviewer (VAN ECK; WALTMAN, 2007), que segundo Moreira, Guimarães e Tsunoda (2020), se destacam como *softwares* de análise bibliométrica.

4 ANÁLISE DOS RESULTADOS

Após realização do levantamento dos artigos conforme demonstrado na metodologia, os 87 artigos foram analisados com o auxílio dos *softwares* descritos. A Tabela 1 a seguir evidencia os resultados bibliométricos encontrados:

Tabela 1 – Resultados Bilbiométricos

Quesito	n
Anos das publicações	1992-2020
Anos de citação	29
Artigos	87
Citações	4096
Citações/ano	146,29
Citações/artigo	2,41
h-index	24
g-index	63
hI, norm	14
hI, annual	0,50
Artigos com ACC >=1,2,5,10,20	45,30,11,3,1

Fonte: Elaborado pelos autores (2021), a partir do *software* Harzing's Publish or Perish (HARZING, 2010).

A primeira informação se refere ao tempo que os artigos foram publicados, que totalizaram 29 anos de publicações. Essa informação demonstra que a pesquisa na área não é recente, pois há publicações relevantes desde 1992. Há outras informações anteriores, porém não foram objeto da presente pesquisa porque não entraram nos critérios definidos para a escolha dos trabalhos com base nas bases de dados definidas.

Considerando a totalidade dos 87 artigos analisados, as citações somam 4.096, o que representa 146,29 citações por ano. Tais dados representam que há uma concentração trabalhos na área, porque a citação por artigo totaliza somente 2,41. Quanto ao índice h de 24 da amostra, significa que há 24 artigos na amostra que há pelo menos 24 citações. Uma referência para um pesquisador é um índice h de 20, ou seja, que ele possui 20 artigos publicados que tem pelo menos 20 citações (HIRSCH, 2005). Logo, a amostra está acima de uma referência para o índice.

Tabela 2 – Percentual de citações Ano

Artigos analisados	% Citações/Ano	% Citações ano frente ao h index	% Citações ano frente ao h index
87	3,57%	16,40	43,06%

Fonte: Elaborado pelos autores (2021), a partir do *software* Harzing’s.

O índice g possui métrica semelhante, porém o total de 63 artigos da amostra significa que a amostra possui pelo menos 20 artigos que combinados possuem até 3.969 citações (EGGHE, 2006). O índice é pertinente com o total de 4.096 citações que a amostra possui.

O tema “*big data* como ferramenta do DFT” contém um caráter híbrido em relação à sua classificação ao ponto em que se concentra na aplicação do *big data* como ferramenta, mas seus resultados destacam a sua utilização tanto na gestão organizações quanto no levantamento do DFT. O tema “relações entre dados, pessoas e organizações (DFT)”, por sua vez, debate a interlocução entre essas dimensões sobre uma abordagem teórica, bem como considerando desafios existentes. Por fim, o tema “aplicações do *big data* nas Organizações” se refere à aplicação do *big data* na gestão de organizações correlacionados ao DFT.

132

Tabela 3 – Temas dos trabalhos analisados

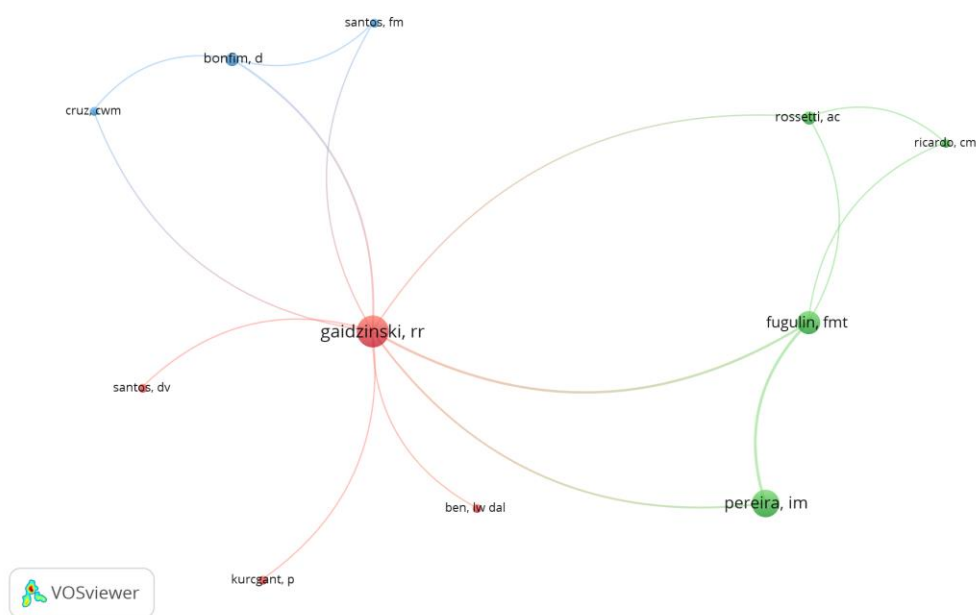
	Tema	Ocorrências por Periódicos	% Ocorrências por tema analisado	Ocorrências por Dissertações
1	Relações entre dados, pessoas e organizações (DFT)	82	94,25%	-
2	Impactos do <i>Big data</i> em áreas do conhecimento	2	2,3%	-
3	<i>Big data</i> como ferramenta de apoio à pesquisa científica	1	1,15%	5
4	Aplicações do <i>Big data</i> nas organizações	1	1,15%	-
5	<i>Big data</i> como ferramenta de Gestão do Conhecimento	1	1,15%	-

Fonte: Elaborado pelos autores (2021), a partir do *software* Harzing’s.

Quanto ao tema dos artigos científicos publicados em periódicos, apesar do baixo número de publicações que correlacionam base de dados e organizações “DFT”, observou-se que os autores têm se concentrado em temas que relacionam o fenômeno *big data* e pessoas e organizações. A tabela 3 acima mostra que estudos sobre os “impactos do *big data* em áreas do conhecimento” (duas ocorrências) e sobre o “*big data* e suas relações com pessoas e organizações” (oitenta e duas ocorrências) somam 96,55% das publicações pesquisadas. Destaca-se que ambos os temas se relacionam ao passo em que o primeiro discute o fenômeno sob a perspectiva de áreas do conhecimento, e o segundo discute a sua utilização prática nas relações organizacionais. Essa constatação tem implicações relevantes; destaca-se que o interesse por estudos bibliométricos foi retomado na década de 1990 em virtude da possibilidade da utilização de computadores. Nesse sentido, esse movimento deve ser incrementado e ampliado com a utilização do *big data*, tanto é que já existem trabalhos que utilizam essa ferramenta, a exemplo de Isasi *et al.* (2015)

Posteriormente, foi realizada análise da concentração gráfica dos autores. A autora Gaidzinski é a que possui maior quantidade de publicação sobre o tema dimensionamento, o total de 13 artigos. Portanto, percebe-se que a autora se posiciona na centralização dos autores que tratam do tema.

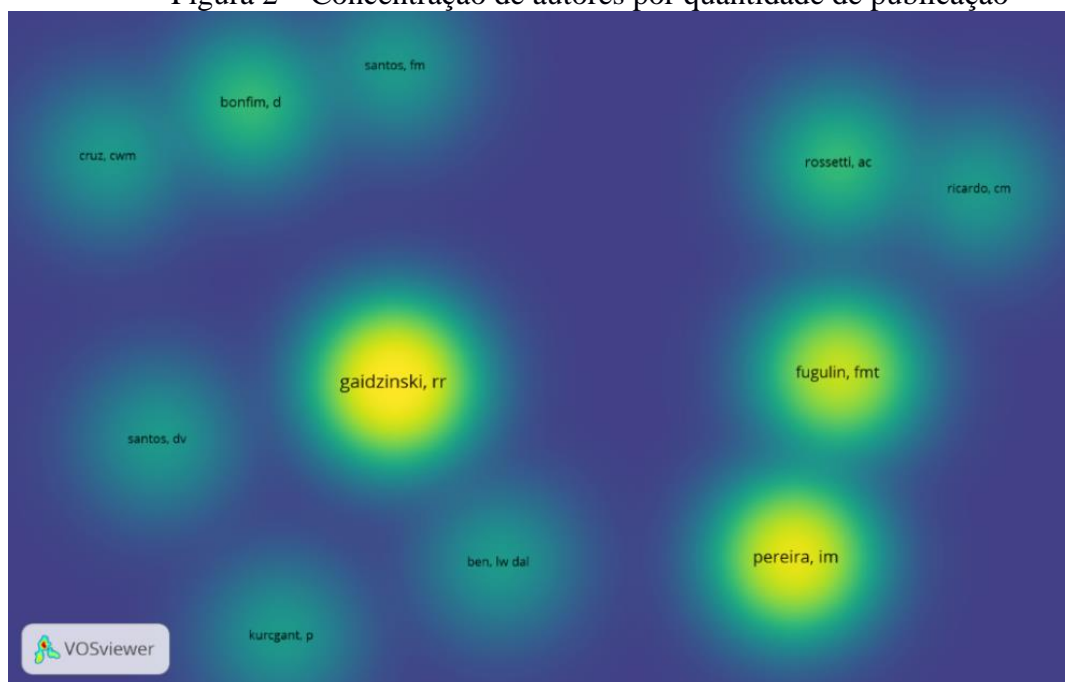
Figura 1 – Concentração de autores por quantidade de publicação



Fonte: autores. Elaborado pelos autores (2021), a partir do software VOSviewer (VAN ECK; WALTMAN, 2007).

Sob a perspectiva da densidade, a distribuição dos autores está representada na figura 2 a seguir, que representa por meio de um gráfico de calor a abrangência e concentração que os autores representam na área.

Figura 2 – Concentração de autores por quantidade de publicação



Fonte: autores. Elaborado pelos autores (2021), a partir do *software* VOSviewer (VAN ECK; WALTMAN, 2007).

Portanto, a pesquisa permite concluir que o tema na área está distribuído em basicamente por 11 autores, sendo que Gaidzinski é a autora com a maior quantidade de trabalhos publicados. Os trabalhos da autora são focados em dimensionamento da força de trabalho na área da saúde, especificamente no setor de enfermagem. Como a utilização de dados é atinente ao tema, *big data* está inserido nesse contexto, mas não é o centro da pesquisa da autora, mas sim uma ferramenta utilizada para realizar os trabalhos empíricos.

Além disso, embora haja um número razoável de publicações (87), o campo de pesquisa precisa evoluir e avançar, principalmente nas outras áreas além da área da saúde.

4 CONSIDERAÇÕES FINAIS

Neste trabalho teve-se como objetivo analisar a produção científica sobre *big data* na área aplicada a Dimensionamento da Força de trabalho, concentrando-se em suas características, produção por periódicos, teses e dissertações, por autores, por instituições, por área de formação, por temas abordados e por conceitos adotados para o termo *big data*. Para cumprir os objetivos foi desenvolvida uma pesquisa descritiva qualitativa e quantitativa, cuja

abordagem adotada foi o estudo bibliométricos. Foram pesquisados periódicos brasileiros e internacionais na área de “Administração pública e de empresas, ciências contábeis e engenharia de produção” com Estratos entre A1 e B2 na Qualis Capes, além de resultados para áreas de Administração no Banco de Teses e Dissertações da Capes.

Nesse sentido, a produção científica sobre *big data versus* DFT observada como resultado do presente estudo pode ser entendida como reflexo da pequena fatia que esta área representa frente a outras áreas do conhecimento, notadamente Ciência da Computação e Engenharia. A mesma observação pode ser comparada em relação à produção científica sobre *big data* em outros países, a exemplo da China, na qual os trabalhos têm se concentrado em algoritmos e técnicas no campo da Tecnologia da Informação, ainda que o tema venha atraindo o interesse crescente dos pesquisadores e embora a produção científica ainda esteja em um estágio exploratório.

Pontua-se, entretanto, que os resultados obtidos foram de publicações recentes, sendo a mais antiga de 1992. Isso pode representar que o assunto tem ganhado espaço na agenda acadêmica recentemente e que, se esta hipótese for verdadeira, o número de publicações pode aumentar em um futuro relativamente próximo. Adicionalmente, os resultados observados na pesquisa feita no banco de teses e dissertações da Capes pode ser considerada como um indicativo de que o tema tem crescido no interesse dos pesquisadores nos últimos anos.

Nos trabalhos analisados foi possível observar uma predominância da utilização do conceito de *big data* como um fenômeno holístico, o qual não se restringe apenas a um grande volume de dados armazenados, mas refere-se a um conjunto complexo de tecnologias capazes produzir informações aplicáveis, contribuindo para a gestão de organizações públicas e privadas. Nesse sentido, os temas predominantes relacionaram-se à aplicação do *big data* na produção científica e na sua aplicação em organizações, o que pode representar um interesse da Academia em entender e possibilitar a aplicação do fenômeno na gestão de organizações.

Ante o exposto, considera-se que os objetivos propostos para este trabalho foram alcançados, muito embora a baixa produção científica observada tenha impossibilitado a aplicação de fórmulas matemáticas referentes as leis de Lotka ou de Bradford e, portanto, o aprofundamento do estudo sob o ponto de vista bibliométrico mais tradicional. Contudo, a análise sobre as características dos trabalhos, bem como sobre parte de seu conteúdo se mostrou satisfatória e foi capaz de apresentar não só um retrato da situação existente, mas também uma perspectiva do desenvolvimento deste campo de estudos.

Adicionalmente, acrescenta-se que, como demonstrado, os resultados têm aderência em relação ao cenário internacional, ainda que signifique que o campo de estudos ainda esteja em

desenvolvimento. Diante das oportunidades apresentadas neste trabalho para a aplicação do fenômeno do *big data* e de suas potencialidades para a gestão pública e considerando que as organizações, como sistemas permeáveis que são, se utilizam de conhecimentos de outros campos científicos para sua gestão, considera-se relevante como agenda de pesquisa a realização de um estudo mais amplo sobre a produção científica no país, observando e analisando trabalhos de todas as áreas do conhecimento.

Este estudo mostrou que a produção científica é escassa, mas a verificação desta produção de maneira mais abrangente, incluindo todas as áreas científicas ou, no mínimo, as de Administração Pública, Ciência da Informação e Engenharia, poderão contribuir para um conhecimento maior sobre a produção no tema, bem como sobre a contribuição que o país tem para dar no debate quanto ao verdadeiro conhecimento da utilização de *big data* para definição da quantidade ótima de servidores em uma organização.

REFERÊNCIAS

- AGUIR, M. S.; AKŞIN, O. Z.; KARAESMEN, F.; DALLERY, Y. On the interaction between retrials and sizing of call centers. **European Journal of Operational Research**, v. 191, n. 2, p. 398-408, 2008.
- BARD, J. F.; MORTON, D. P.; WANG, Y. M. Workforce planning at USPS mail processing and distribution centers using stochastic optimization. **Annals of Operations Research**, v. 155, n. 1, p. 51-78, 2007.
- BUFFA, E. S. Research in operations management. **Journal of Operations Management**, v. 1, n. 1, p. 1-14, 1980.
- CAPPELLI, P.; KELLER, J. R. Talent management: conceptual approaches and practical challenges. **Annual Review of Organizational Psychology and Organizational Behavior**, v. 1, n. 1, p. 305-331, 2014.
- CHEN, M.; MAO, S.; LIU, Y. *Big data: A Survey*. **Mobile Networks and Applications**, v. 19, n. 2, p. 171-209, 2014.
- CHOWDHURY, S. Optimization and business improvement studies in upstream oil and gas industry. **John Wiley & Sons**: New Jersey, 2016.
- COROMINAS, A.; LUSA, A.; OLIVELLA, J. A manufacturing and remanufacturing aggregate planning model considering a non-linear supply function of recovered products, **Production Planning & Control: The Management of Operations**, v. 23, n. 2-3, p. 194-204, 2012.
- CRONIN, P.; RYAN, F.; COUGHLAN, M. Undertaking a literature review: a step-by-step approach. **British Journal of Nursing**, v. 17, n. 1, p. 38-43, 2008.
- DAVENPORT, T. H.; BART, P.; BEAN, R. How *Big data* is Different. **MIT Sloan Management Review**, July 30, pp. 43-6, 2012.
- DICKINSON, H.; SULLIVAN, H.; HEAD, G. The future of the public service workforce: a dialogue. **Australian Journal of Public Administration**, v. 74, n. 1, p. 23-32, 2015.
- EDWARDS, D.; HAMSON, M. **Guide to mathematical modelling**. Industrial Press: New York, 2007.
- ELLINGTON, J. K.; WILSON, M. A. The performance appraisal milieu: a multilevel analysis of context effects in performance ratings. **Journal of Business and Psychology**, v. 32, n. 1, p. 87-100, 2017.
- ERNST, A.; JIANG, H.; KRISHNAMOORTHY, M.; SIER, D. An annotated bibliography of personnel scheduling and rostering. **Annals of Operations Research, Kluwer Academic Publishers**, v. 127, n. 1-4, p. 21-144, 2004.
- FUM, D.; MISSIER, F. D.; STOCCO, A. The cognitive modeling of human behavior: why a model is (sometimes) better than 10,000 words. **Cognitive Systems Research**, v. 8, p. 135-142, 2007.

GANDOMI, A.; HAIDER, M. Beyond the hype: *Big data* concepts, methods, and analytics. **International Journal of Information Management**, v. 35, n. 2, p. 137–144, 2015.

GREENE, J. R.; HUDZIK, J. K. Present trends and future prospects for criminal justice manpower planning. **Computers, Environment and Urban Systems**, v. 6, n.4, p. 193-210, 1981.

HARZING, A. W. **The Publish or Perish book**. Tarma *Software* Research Pty Limited, 2010.

HELTON, K. A.; JACKSON, R. D. Navigating Pennsylvania's dynamic workforce: succession planning in a complex environment. **Public Personnel Management**, v. 36, n. 4, 335-347, 2007.

HIRSCH, J. E. An index to quantify an individual's scientific research output. **Proceedings of the National Academy of Sciences**, v. 102, n. 46, p. 16569–16572, 2005.

IGHRAVWE, D. E.; OKE, S. A. A non-zero integer non-linear programming model for maintenance workforce sizing. **International Journal of Production Economics**, v. 150, p. 204-214, 2014.

INOUE, K. C.; MATSUDA, L. M. Dimensionamento de pessoal de enfermagem em Unidade de Terapia Intensiva para adultos. **Acta Paulista de Enfermagem**, v. 23, n. 3, p. 379-384, 2010.

INTERNATIONAL PERSONNEL MANAGEMENT ASSOCIATION – IPMA. Workforce Planning Resource Guide for Public Sector Human Resource Professionals. **IPMA Workforce Planning Resource Guide**, 2002.

ISASI, N. K. G.; FRAZZON, E. M.; URIONA, M. *Big data* and Business Analytics in the supply chain: a review of the literature. **IEEE Latin America Transactions**, v. 13, n. 10, p. 3382-3391, 2015.

JACOBSON, W. S. Preparing for tomorrow: a case study of workforce planning in North Carolina municipal governments. **Public Personnel Management**, v. 39, n. 4, p. 353-377, 2010.

KOZLOWSKI, S. W. Advancing research on team process dynamics: theoretical, methodological, and measurement considerations. **Organizational Psychology Review**, v. 5, n. 4, p. 270-299, 2015.

LANEY, D. 3-D data management: Controlling data volume, velocity and variety. **META group research note**, v. 6, n. 70, p. 1, 2001.

LATANÉ, B.; WILLIAMS, K.; HARKINS, S. Many hands make light the work: the causes and consequences of social loafing. **Journal of Personality and Social Psychology**, v. 37, n. 6, p. 822-832, 1979.

LI, Y.; CHEN, J.; CAI, X. An integrated staff-sizing approach considering feasibility of scheduling decision. **Annals of Operations Research**, Springer: 2007.

LONG, Q. Data-driven decision making for supply chain networks with agent-based computational experiment. **Knowledge-Based Systems**, v. 141, p. 55-66, 2018.

MACHADO JUNIOR, C.; SOUZA, M. T. S.; PARASOTTO, I. R. S.; PALMISANO, A. As leis da bibliometria em diferentes bases de dados científicos. **Revista de Ciências da Administração**, v. 18, n. 44, p. 111-123, 2016.

MARINHO, B. L.; VASCONCELLOS, E. P. G. Dimensionamento de recursos humanos: desenvolvimento de um modelo conceitual e sua aplicação. **REGE Revista de Gestão**, v. 14, n. 2, p. 61-76, 2007.

MATHIEU, J. E.; HOLLENBECK, J. R.; VAN KNIPPENBERG, D.; ILGEN, D. R. 2017. A century of work teams in the Journal of Applied Psychology. **Journal of Applied Psychology**, v. 102, n. 3, p. 452-467, 2017.

MCAFEE, A.; BRYNJOLFSSON, E. *Big data*: The management revolution. **Harvard Business Review**, v. 90, n. 10, p. 60-68, 2012.

MOREIRA, P. S. C.; GUIMARÃES, A. J. R.; TSUNODA, D. F. Qual ferramenta bibliométrica escolher? um estudo comparativo entre *softwares*. **Revista P2P e INOVAÇÃO**, v. 6, n. 2, p. 140-158, 2020.

NILAKANTAN, K.; RAGHAVENDRA, B. G. Control aspects in proportionality Markov manpower systems. **Applied Mathematical Modelling**, v. 29, n. 1, p. 85-116, 2005.

NORRIS, D. How do computational models help us develop better theories. **Twentyfirst century psycholinguistics: Four cornerstones**, p. 331-346, 2005.

OLIVEIRA, A. G.; BIANCHINI, D.; ABBADE, M. L. F. Métricas para dimensionar recursos humanos nos Centros de Operações de Redes. **Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos**, Belém do Pará, 2007.

POWELL, P. T.; SCHMENNER, R. W. Economics and operations management: towards a theory of endogenous production speed. **Managerial and Decision Economics**, v. 23, p. 331-342, 2002.

RAGHUPATHI, W.; RAGHUPATHI, V. *Big data* analytics in healthcare: promise and potential. **Health Information Science and Systems**, v. 2, n. 1, p. 3, 2014.

ROSA, B. A.; FILHO, E. M. S. Dimensionamento de recursos humanos em uma empresa fabricante de materiais sanitários: uma abordagem via programação inteira. **Simpósio de Pesquisa Operacional e Logística da Marinha**, v. 11, p. 5-6, 2008.

SANTOS, M. S. F. **Análise do dimensionamento do pessoal de enfermagem da unidade de terapia intensiva de um hospital de ensino em Cascavel-PR**. 2016.

SERRANO, A. L. M.; MENDES, N. C. F.; ABILA, N. **Dimensionamento na Administração Pública Federal**: uma ferramenta de gestão da força de trabalho. Enap: Brasília, 2019.

SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**, 4ª ed. Florianópolis: UFSC, 2005.

SIMMS, A.; NICHOLS, T. Social loafing: a review of the literature. **Journal of Management Policy and Practice**, v. 15, n. 1, p. 58-67, 2014.

SOMASUNDARAM, G.; SHRIVASTAVA, A. **Armazenamento e gerenciamento de informações: como armazenar, gerenciar e proteger informações digitais**. Bookman Editora, 2009.

STOKKER, J. H.; HALLAM, G. The right person, in the right job, with the right skills, at the right time: a workforce-planning model that goes beyond metrics. **Library Management**, v. 30, n. 8/9, p. 561-571, 2009.

TANOS, M. A. A.; MASSAROLLO, M. C. K. B.; GAIDZINSKI, R. R. Dimensionamento de pessoal de enfermagem em uma unidade especializada em transplante de fígado: comparação do real com o preconizado. **Revista da Escola de Enfermagem da USP**, v. 34, n. 4, p. 376-382, 2000.

TOHIDI, H.; TAROKH, M. J. Productivity outcomes of teamwork as an effect of information technology and team size. **International Journal of Production Economics**, v. 103, p. 610-615, 2006.

TORRE, R.; LUSA, A.; MATEO, M. A MILP model for the long term academic staff size and composition planning in public universities. **Omega**, v. 63, p. 1-11, 2016.

VAN ECK, N. J.; WALTMAN, L. VOS: A new method for visualizing similarities between objects. In: **Advances in data analysis**. Springer: Berlim, p. 299-306, 2007.

VERMA, M. C. Review of skilled manpower forecasts and changes in occupational structures of India. **International Journal of Educational Development**, v. 4, n. 3, p. 173-221, 1984.

WANG, Y.; HAJLI, N. Exploring the path to *big data* analytics success in healthcare, **Journal of Business Research**, v. 70, p. 287-299, 2017.

WAXIN, Marie-France.; LINDSAY, V.; BELKHODJA, O.; ZHAO, F. Workforce localization in the UAE: Recruitment And Selection Challenges And Practices In Private And Public Organizations. **The Journal of Developing Areas**, v. 52, n. 4, p. 99-113, 2018.

ZIKOPOULOS, P. C.; DEROOS, D.; PARASURAMAN, K. **Harness the power of big data - The IBM Big data Platform**. McGraw-Hill, 2013.