

DOI: 10.5748/9788599693131-14CONTECSI/RF-4444

## BIBLIOMETRIC STUDY OF ARTICLES WITH REGARD TO DATA AND INFORMATION GOVERNANCES

Patricia Kuzmenko Furlan (Escola Politécnica da Universidade de São Paulo, São Paulo, Brasil) – paty.kf@gmail.com

Fernando José Barbin Laurindo (Escola Politécnica da Universidade de São Paulo, São Paulo, Brasil) – fjblau@usp.br

The disciplines of data and information governances have been developed in the scientific literature, especially in the big data environment. Those governances aim to establish the means by which organizations can use and organize their data and information assets. Thus, it was conducted a bibliometric study in both ISI Web of Science and Scopus platforms to investigate the publication curves of scientific articles with regard to data and information governances. It was concluded that the information governance has more articles and citations on both platforms, but the data governance literature has demonstrated a greatest growth trend in the big data environment. However, both topics can be treated interchangeably, whose main research areas include: medicine, computer science, information science, and business & management.

Keywords: Data governance, Information governance, Big data

## ESTUDO BIBLIOMÉTRICO DOS ARTIGOS EM GOVERNANÇAS DE DADOS E DE INFORMAÇÃO

As disciplinas de governança de dados e de informação tem se desenvolvido na literatura científica, principalmente no ambiente big data. Tais governanças visam instaurar os meios pelos quais as organizações podem utilizar e organizar seus ativos de dados e de informações. Desta forma, foi conduzido um estudo bibliométrico nas plataformas ISI Web of Science e Scopus, de modo a investigar as curvas de publicação dos artigos científicos sobre as governanças de dados e de informação. Conclui-se que a governança de informação possui mais artigos e citações em ambas as plataformas, mas que a literatura de governança de dados tem demonstrado maior crescimento no ambiente big data. No entanto, a maior parte das publicações trata os termos indistintamente, cujas principais áreas de pesquisa são: medicina, ciência da computação, ciência da informação, negócios & gestão.

Palavras-chave: Governança de dados, Governança de informação, Big data

## 1. Introdução

O tópico "governança" corresponde às formas pelas quais uma organização pode assegurar que suas estratégias sejam estipuladas, monitoradas e alcançadas (RAU, 2004; WEBER et al., 2009). A governança se preocupa em *quais* decisões devem ser tomadas, e por *quem* na organização, para assegurar a gestão eficiente, fornecendo uma estrutura para o alcance desses objetivos (KHATRI; BROWN, 2010; FU et al., 2011). Numa escala inferior à governança, o campo da "gestão" se relaciona às tomadas de decisão e aos processos de implementação dessas decisões, atuando na esfera de "fazer e implementar decisões" (KHATRI; BROWN, 2010).

No ambiente com *big data*, a governança de dados e a governança de informação tem demonstrado relevância crescente nas organizações. Elas adquiriram o centro da atenção dos executivos e consistem num tópico de destaque na literatura de gestão dos sistemas da informação (TALLON et al., 2014). Tratam-se de disciplinas emergentes, cuja pesquisa bibliométrica proposta no presente estudo visa investigar sua características, nichos de publicação e principais conteúdos.

Vale ressaltar, no entanto, que a maior preocupação pelas governanças de dados e de informação são resultados da maior quantidade de dados que as organizações tem a disposição. O termo "*big data*" se refere ao alto volume de dados virtuais complexos, diversos, heterogêneos, provenientes de múltiplas e autônomas fontes, com controles distribuídos e descentralizados (McAFEE; BRYNJOLFSSON, 2012; WU et al., 2014; ZHANG, X. et al., 2014; FURLAN, LAURINDO, 2016). Este fenômeno engloba a geração de dados estruturados, semi-estruturados e não estruturados em alto volume, em alta velocidade e em múltiplas fontes (RUSSON, 2011; McAFEE e BRYNJOLFSSON, 2012; CHEN et al., 2012).

Assim, o presente estudo tem como objetivo desenvolver uma pesquisa bibliométrica para investigar como os artigos científicos estão evoluindo no tema, e de que modo eles diferem governança de dados da governança de informação. Tal estudo visa contribuir para a consolidação das pesquisas realizadas, servindo de referencial para as discussões no campo da governança.

## 2. Metodologia de pesquisa

O método de pesquisa utilizado consistiu em pesquisas bibliométricas com a posterior análise dos resultados obtidos nas plataformas ISI *Web of Science* e *Scopus*, de modo a trazer dois parâmetros relevantes nas publicações científicas. As buscas foram realizadas no período de Junho a Setembro de 2016, empregando separadamente as chaves de busca: *<data governance>* e *<information governance>*.

No âmbito das pesquisas em governança de dados, foram encontrados 91 registros no ISI *Web of Science*, dos quais 41 correspondem a artigos (45%). A Figura 1 indica a evolução na publicação de artigos contemplando este tema, bem como evidencia a curva de citações por ano. Por meio das curvas, é possível notar que se trata de um tema recente, possuindo registros de artigos na plataforma após o ano de 2005. A curva de citação tem demonstrado crescimento exponencial, com intensificação após o ano de 2013. A soma dos artigos citados corresponde a 371, com citação média por artigo de 9,05.

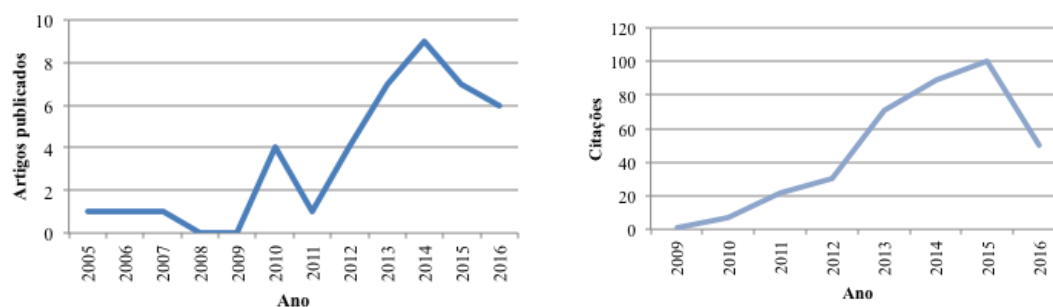


Figura 1 - Evolução nas publicações de artigos sobre governança de dados (à esquerda) e citações por ano (à direita). Fonte: ISI *Web of Science*

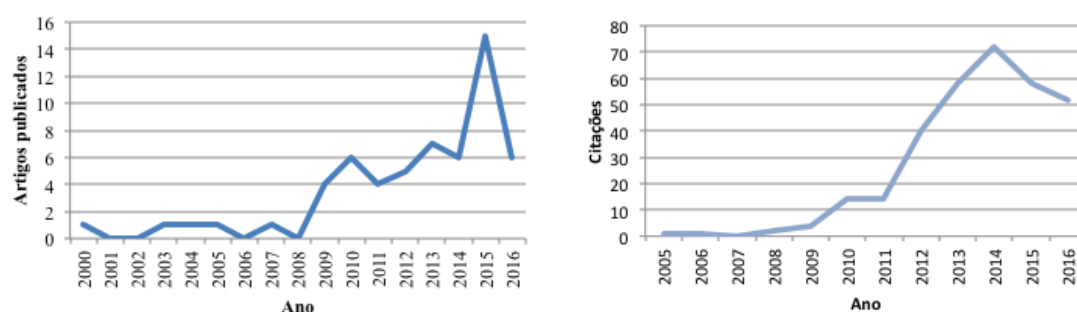


Figura 2 - Evolução nas publicações de artigos sobre governança de informação (à esquerda) e das citações por ano (à direita). Fonte: ISI *Web of Science*

Como pode ser visualizado pela Figura 2, as buscas por artigos sobre governança de informação na ISI *Web of Science* resultaram em itens a partir dos anos 2000, retornando mais artigos que na busca por governança de dados. Foram obtidas 100 publicações, das quais 58 são artigos (58%) e cuja soma das citações corresponde a 320 (média anual de 5.52).

As buscas na *Scopus* retornaram maiores quantidades de publicações de artigos. No âmbito da governança de dados, foram encontrados 213 publicações, das quais 90 são artigos (42,3%). Tais artigos apresentam 496 citações. Já na governança de informação, foram notadas 274 publicações, das quais 171 são artigos científicos (62,4%) e estes apresentam 563 citações.

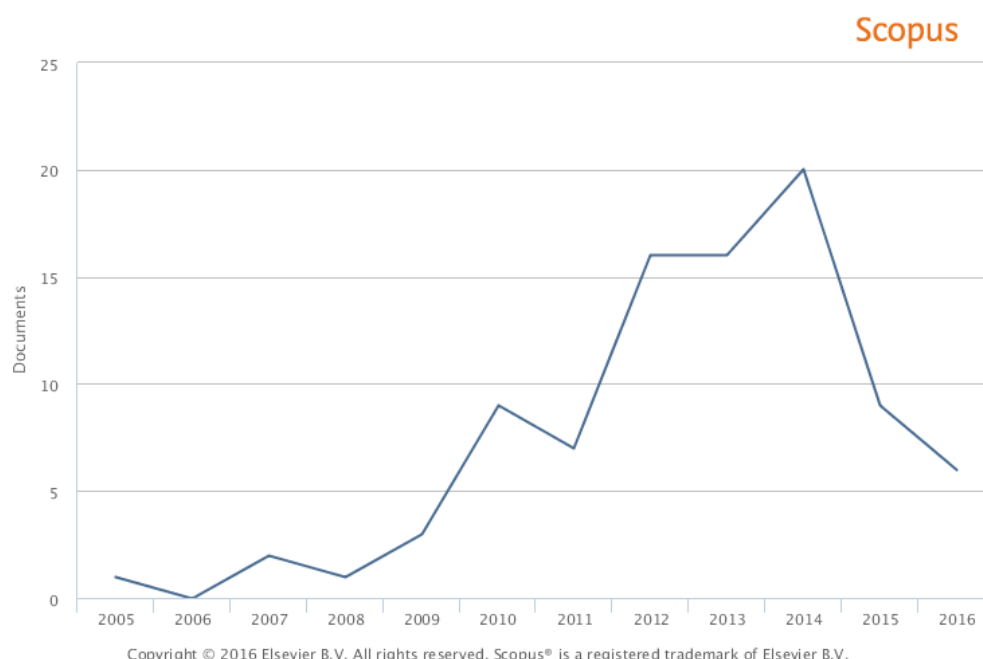


Figura 3 - Curva de evolução dos artigos publicados sobre governança de dados segundo a Scopus. Fonte: Scopus

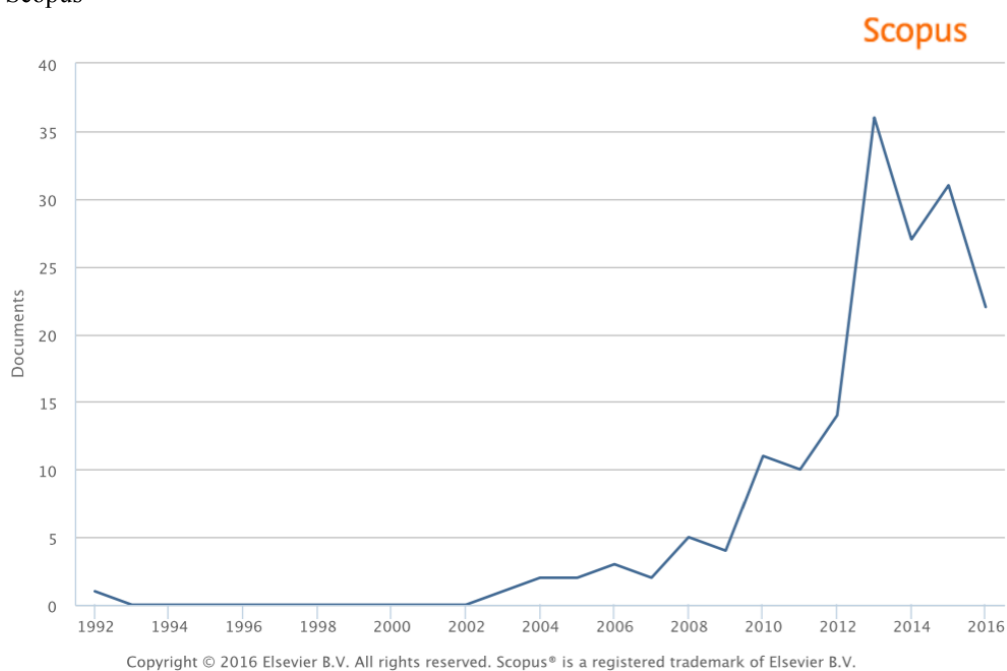


Figura 4 - Curva de evolução dos artigos publicados sobre governança de informação segundo a Scopus. Fonte: Scopus

### 3. Resultados e discussões

Por meio do estudo bibliométrico realizado, foi possível observar que as publicações de artigos sobre governança de informação possuem registros anteriores aos artigos sobre governança de dados. O primeiro registro acerca da governança da informação foi observado na *Scopus* em 1992, cuja evolução se tornou consistente após 2002, coincidindo com a expansão de *big data*. Para este mesmo tema, foram notados registros de pico de mais de 35 artigos no ano de 2012. Na *ISI Web of Science*, o primeiro registro contemplando governança de informação aconteceu no ano 2000, cujo pico nas publicações (de artigos) foi obtido em 2015, com mais de 14

itens. O campo da governança de dados consiste, portanto, em um tema mais recente, cujos primeiros artigos advém do ano de 2005, como retornaram as buscas nas plataformas ISI *Web of Science* e Scopus.

Outro ponto observado corresponde na diferença nas citações dos artigos. Com relação aos resultados da *Web of Science*, a literatura de governança de informação caracteriza-se com menor quantidade total de citações, embora tenha maior tempo de existência e maior quantidade de artigos. Segundo o *Web of Science*, a soma dos artigos citados corresponde a 371, com citação média por artigo de 9,05 para a governança de dados, e 320 (média anual de 5.52) para a governança de informação. Por outro lado, os resultados da Scopus indicam que a governança de informação possui maior tempo de desenvolvimento, maior quantidade de artigos publicados e também a maior quantidade de citações. Portanto, não é possível concluir qual literatura é mais relevante com base na análise de citações.

No entanto, apesar das informações acima expostas com relação ao período de desenvolvimento destas literaturas, as publicações de governança de dados e de informação apresentam estes temas de forma indiferente. Ou seja, por meio de Weber et al. (2009) foi constatado que "a maior parte das publicações em governança não diferencia os termos dado e informação", o que denota uma fronteira a ser melhor definida e explorada em futuros trabalhos.

Outra análise corresponde às áreas de pesquisa das buscas realizadas. Na buscas no ISI *Web of Science*, tanto para a governança de dados quanto para a governança de informação, foram identificadas como as áreas de pesquisa mais relevantes:

- Ciência da computação
- Serviços de ciência da saúde
- Ciência da informação
- Informática médica

Já as buscas conduzidas na Scopus, foram notados outros resultados. A Figura 5 ilustra as áreas de pesquisa para a governança de dados e a Figura 6 para a governança de informação, onde podem ser notadas diferenças quanto a participação de cada área no montante das publicações.

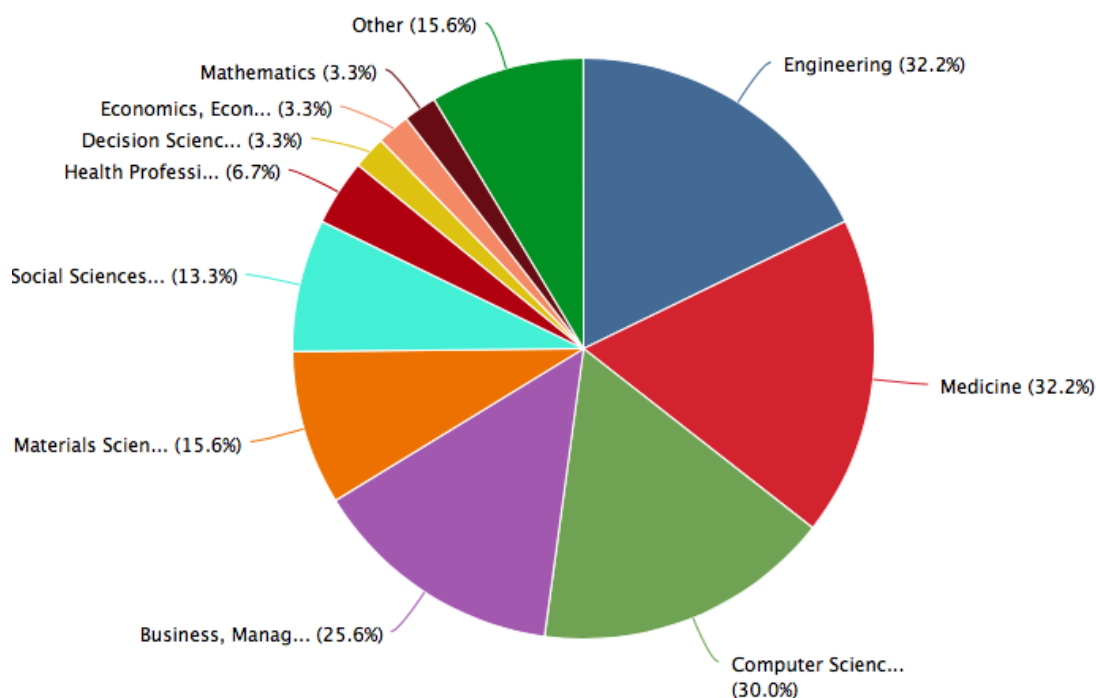


Figura 5 - Áreas de pesquisa dos artigos da Scopus sobre governança de dados. Fonte: Scopus

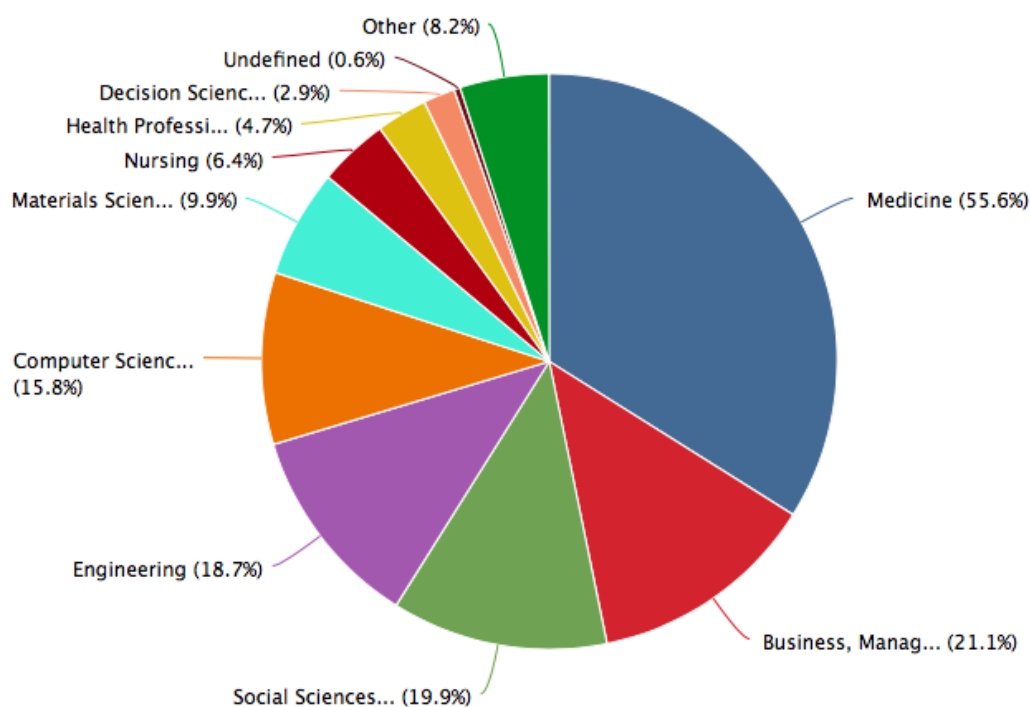


Figura 6 - Áreas de pesquisa dos artigos identificados na Scopus sobre governança de informação. Fonte: Scopus

Comparando os resultados da Scopus e ISI *Web of Science*, foi possível notar diferenças significativas:

- A área da ciência da computação aparece em todos os casos, embora com diferenças quanto a frequência de aparecimento;
- Medicina aparece em primeiro lugar em ambas as buscas na Scopus, de onde se pode inferir que tal área foi contemplada como "serviços de ciência de

saúde" e como "informática médica" pela *Web of Science*;

- As áreas de negócios & gestão são relevantes em ambos os resultados da Scopus; já na *Web of Science* não se tem a definição desta área;
- A *Web of Science* retrata a área Ciência da Informação dentre as áreas de maior destaque na literatura de governança de dados e de informação. Tal campo não aparece nos resultados da Scopus, podendo estar compreendida entre as áreas de negócios e gestão, ou outras.

Através dos resultados obtidos, pode-se notar que a área da computação e da medicina ou saúde compreendem a maior parcela dos resultados. Assim, o maior nicho de desenvolvimento das disciplinas focalizam tais áreas, com um destaque às áreas de negócios & gestão, ciência da informação, economia, entre outras.

Com vista aos resultados obtidos no *Web of Science*, encontram-se listados na sequência os principais artigos identificados em cada pesquisa. Como é de se esperar, uma significativa parcela dos artigos listados nos quadros contemplam domínios da medicina e saúde.

Quadro 1 - Os 10 artigos com maiores citações dentre a pesquisa por governança de dados.

Artigo	Periódico	Citações totais	Média de citação
1- The many faces of ecosystem-based management: Making the process work today in real places (TALLIS et al., 2010)	Marine Policy	100	14.29
2 - Standard Guidelines for the Chromosome-Centric Human Proteome Project (PAIK et al., 2012)	Journal of Proteome Research	78	15.6
3 - The Enterprise Data Trust at Mayo Clinic: a semantically integrated warehouse of biomedical data (CHUTE et al., 2010)	Journal of the American Medical Informatics Association	55	7.86
4 - Designing Data Governance (KHATRI; BROWN, 2010)	Communications of the ACM	39	5.57
5 - Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access (ROSENBAUM, 2010)	Health Services Research	12	1.71
6- Towards a Fuzzy Expert System on Toxicological Data Quality Assessment (YANG et al., 2013 )	Molecular Informatics	10	2.5
7 - Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost (TALLON, 2013)	Computer	9	2.25
8 - Developing a data infrastructure for a learning health system: the PORTAL network (McGLYNN et al.,2014 )	Journal of the American Medical Informatics Association	8	2.67
9 - The Electronic Data Methods (EDM) Forum for Comparative Effectiveness Research (CER) (HOLVE et al, 2012)	Medical Care	8	1.6
10 - Governing Big Data: Principles and practices (MALIK,2013)	IBM Journal of Research and Development	7	1.75

Fonte: ISI *Web of Science*

No Quadro 1, 60% dos artigos retratados são da área da medicina ou saúde. No entanto, o de maior citação contempla o tema da biologia marítima. No que tange a ciência da informação aplicada à gestão, foram encontrados como artigos relevantes o de Khatri e Brown (2010), Rosenbaum (2010), Tallon (2013) e Malik (2013).

O Quadro 2 retrata a lista do artigos mais citados dentre o universo da governança de informação no ISI *Web of Science*.

Quadro 2 - Os 10 artigos com maiores citações dentre a pesquisa por governança de informação.

Artigo	Periódico	Citações	Citações por ano
The SAIL databank: linking multiple health and social care datasets (LYONS et al., 2009)	BMC Medical Informatics and Decision Making	87	10.88
The SAIL Databank: building a national architecture for e-health research and evaluation (FORD et al., 2009)	BMC Health Services Research	72	9
On the governance of information: Introducing a new concept of governance to support the management of information (KOOOPER et al., 2011)	International Journal of Information Management	24	4
The European Disorder of Sex Development Registry: A Virtual Research Environment (AHMED et al., 2010)	Sexual Development	16	2.29
Bureaucracy stifles medical research in Britain: a tale of three trials (SNOOKS et al., 2012)	BMC Medical Research Methodology	10	2
Information governance in NHS's NPfIT: A case for policy specification (BECKER, 2007)	International Journal of Medical Informatics	9	0.9
Change in alcohol outlet density and alcohol-related harm to population health (CHALICE) (FONE et al., 2012)	BMC Public Health	8	1.6
Scaling up antiretroviral therapy in Uganda: using supply chain management to appraise health systems strengthening (WINDISCH et al., 2011)	Globalization and Health	8	1.33
Information governance - a view from the NHS (DONALDSON; WALKER, 2004)	International Journal of Medical Informatics	8	0.62
The Information Artifact in IT Governance: Toward a Theory of Information Governance (TALLON et al., 2014)	Journal of Management Information Systems SYSTEMS	7	1.75

Fonte: ISI *Web of Science*

Foram listados nos quadros somente os resultados da ISI *Web of Science* com o intuito de ilustrar os principais retornos obtidos. A lista dos artigos com maiores citações da Scopus difere em partes, sem a maior parte dos conteúdos também da medicina ou área da saúde.

#### 4. Conclusões

O Quadro 3 consolida os principais pontos característicos das pesquisas bibliométricas sobre governança de dados e governança de informação.

Quadro 3 - Comparação entre resultados obtidos nas pesquisas na *Web of Science* e Scopus.

	ISI <i>Web of Science</i>	Scopus
Governança de dados	<ul style="list-style-type: none"> <li>• 1º artigo publicado: 2005</li> <li>• 91 publicações, das quais 41 são artigos</li> <li>• Citações: 371</li> <li>• Maior crescimento nas publicações: a partir de 2013</li> <li>• Principais áreas de pesquisa (igual a abaixo): ciência da computação, serviços de ciência da saúde, ciência da informação, informática médica</li> </ul>	<ul style="list-style-type: none"> <li>• 1º artigo publicado: 2005</li> <li>• 213 publicações, das quais 90 são artigos</li> <li>• Citações: 496</li> <li>• Maior crescimento nas publicações: a partir de 2008</li> <li>• Principais áreas de pesquisa: engenharia, medicina, ciência da computação, negócios e gestão, ciência dos materiais</li> </ul>
Governança da informação	<ul style="list-style-type: none"> <li>• 1º artigo publicado: 2000</li> <li>• 100 publicações, das quais 58 são artigos</li> <li>• Citações: 320</li> <li>• Maior crescimento nas publicações: a partir de 2009</li> <li>• Principais áreas de pesquisa: ciência da computação, serviços de ciência da saúde, ciência da informação, informática médica</li> </ul>	<ul style="list-style-type: none"> <li>• 1º artigo publicado: 1992</li> <li>• 274 publicações, das quais 171 são artigos</li> <li>• Citações: 563</li> <li>• Maior crescimento nas publicações: a partir de 2009</li> <li>• Principais áreas de pesquisa: medicina, negócios e gestão, ciências sociais, engenharia, ciência da computação</li> </ul>

Fonte: ISI *Web of Science*; Scopus

A literatura de governança de dados se origina em 2005 no contexto *big data*, conferindo crescimento exponencial no número de artigos e de citações. Ela apresenta picos anuais acima daqueles observados na literatura de governança de informação. No entanto, ao se considerar ambas as plataformas, a governança de informação possui maior quantidade de artigos e de citações.

No entanto, tratam-se de campos novos, relevantes e multidisciplinares, com destaque em diferentes áreas do campo científico. Com vistas aos resultados obtidos em ambas as plataformas, pode-se sugerir que tais literaturas contemplam principalmente aspectos da medicina, ciência da computação, ciência da informação e negócios & gestão. Apesar das peculiaridades de cada governança, a maior parte das publicações retratam de forma indistinta a governança de dados e a de informação.

Com base nas buscas realizadas, conclui ser relevante o uso de ambas as plataformas ISI *Web of Science* e Scopus para as pesquisas com relação a governança de dados e de informação, pois os resultados são bastante divergentes (quantidade de artigos indexados, número de citações).

Deve-se ressaltar que o crescimento das aplicações *big data* não somente potencializaram, a capacidade analítica das organizações como também aumentaram muito a quantidade de dados e informações consideradas relevantes por elas, tornando mais importante a questão de suas respectivas governanças.

Futuros estudos deverão explorar como estas governanças são de fato implementadas e usadas por diferentes tipos de organizações.

### Referências bibliográficas

AHMED, S. F.; RODIE, M.; JIANG, J.; SINNOTT, R. O.. The European Disorder of Sex

Development Registry: A Virtual Research Environment. **Sexual Development**, v. 4, 2010.

BECKER, Moritz Y.. Information governance in NHS's NPfIT: A case for policy specification. **International Journal of Medical Informatics**, v. 76, n. 5, 2007.

CHEN, Hsinchun; CHIANG, Roger H. L.; STOREY, Veda C. Business Intelligence and Analytics: From big data to big impact. **MIS Quarterly**. Vol. 36, edição 4, p. 1165 - 1188. 2012.

CHUTE, Christopher G.; BECK, Scott A.; FISK, Thomas B.; MOHR, David N. The Enterprise Data Trust at Mayo Clinic: a semantically integrated warehouse of biomedical data. **Journal of the American Medical Informatics Association**, v. 17, n. 2, p. 131 - 135, 2010.

DONALDSON, Alistair; WALKER, Phil. Information governance - a view from the NHS. **International Journal of Medical Informatics**, v. 73, p. 281 - 284, 2004.

FONE, David; DUNSTAN, Frank; WHITE, James; WEBSTER, Chris; RODGERS, Sarah; LEE, Shin; SHIODE, Narushige; ORFORD, Scott; WEIGHTMAN, Alison; BRENNAN, Iain; SIVARAJASINGAM, Vas; MORGAN, Jennifer; FRY, Richard; LYONS, Ronan. Change in alcohol outlet density and alcohol-related harm to population health (CHALICE). **BMC Public Health**, v. 12, 2012.

FORD, David V.; JONES, Kerina H.; VERPLANCKE, Jean-Philippe; LYONS, Ronan A.; JOHN, Gareth; BROWN, Ginevra; BROOKS, Caroline J.; THOMPSON, Simon; BODGER, Owen; COUCH, Tony; LEAKE, Ken. The SAIL Databank: building a national architecture for e-health research and evaluation. **BMC Health Services Research**, v. 9, 2009.

FU, Xin; WOJAK, Anna; RIDLEY, Mick; TRAVIS, Kim. Data governance in predictive toxicology: A review. **Journal of Cheminformatics**, p. 3 - 24, 2011.

FURLAN, Patricia Kuzmenko; LAURINDO, Fernando José Barbin. Um estudo bibliométrico sobre big data na literatura de gestão de projetos **13th INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS & TECHNOLOGY MANAGEMENT - CONTECSI**, São Paulo: 2016.

HOLVE, Holve, Erin; SEGAL, Courtney; LOPEZ, Marianne Hamilton; REIN, Alison; JOHNSON, Beth H. The Electronic Data Methods (EDM) Forum for Comparative Effectiveness Research (CER). **Medical Care**, v. 50, n. 7, 2012.

KHATRI, Vijay; BROWN, Carol V. Designing Data Governance. **Communications of the ACM**, v. 53, n. 1, 2010.

KOOPER, M. N; MAES, R.; LINDGREEN, E. E. O. Roos. On the governance of information: Introducing a new concept of governance to support the management of information. **International Journal of Information Management**, v. 31, p. 195 - 200, 2011.

LYONS, Ronan A.; JONES, Kerina H.; JOHN, Gareth; BROOKS, Caroline J.; VERPLANCKE, Jean-Philippe; FORD, David V.; BROWN, Ginevra; LEAKE, Ken. The SAIL databank: linking multiple health and social care datasets. **BMC Medical Informatics and Decision Making**, v. 9, 2009.

MALIK, Piyush. Governing Big Data: Principles and practices. **IBM Journal of Research and Development**, v. 57, n. 4, 2013.

McAFEE, Andrew; BRYNJOLFSSON, Erik. Big Data: The Management Revolution. **Harvard Business Review**. Vol. 90, edição 10, p. 60. 2012.

McGLYNN, Elizabeth A.; LIEU, Tracy A.; DURHAM, Mary L.; BAUCK, Alan; LAWS, Reesa; GO, Alan S.; CHEN, Jersey; FEIGELSON, Heather Spencer; CORLEY, Douglas A.; YOUNG, Deborah. Developing a data infrastructure for a learning health system: the PORTAL network. **Journal of the American Medical Informatics Association**, v. 21, n. 4, p. 596- 601, 2014.

PAIK, Young-Ki; OMENN, Gilbert S.; UHLEN, Mathias; HANASH, Samir; MARKO-VARGA, Gyorgy; AEBERSOLD, Ruedi; BAIROCH, Amos; YAMAMOTO, Tadashi; LEGRAIN, Pierre; LEE, Hyung-Joo; NA, Keun; JEONG, Seul-Ki; HE, Fuchu; BINZ, Pierre-Alain; NISHIMURA, Toshihide; KEOWN, Paul; BAKER, Mark S.; YOO, Jong Shin; GARIN, Jerome; ARCHAKOV, Alexander; BERGERON, John; SALEKDEH, Ghasem Hosseini; HANCOCK, William S. Standard Guidelines for the Chromosome-Centric Human Proteome Project. **Journal of Proteome Research**, v. 11, n. 4, p. 2005- 2013, 2012.

RAU, Kenneth G. Effective Governance of IT: Design objectives, roles and relationships. **Information Systems Management**. 2004.

ROSENBAUM, Sara. Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access. **Health Services Research**, v. 45, n. 5, 2010.

RUSSON, Philip. Big Data Analytics - Executive Summary. **TDWI Best Practices Reports**. 4th Quarter. TDWI Research. 2011.

SNOOKS, Helen; HUTCHINGS, Hayley; SEAGROVE, Anne; STEWART-BROWN, Sarah; WILLIAMS, John; RUSSEL, Ian. Bureaucracy stifles medical research in Britain: a tale of three trials. **BMC Medical Research Methodology**, v. 12, 2012.

TALLIS, Heather; LEVIN, Phillip S.; RUCKELSHAUS, Mary; LESTER, Sarah E.; McLEOD, Karen L.; FLUHARTY, David L.; HALPERN, Benjamin S. The many faces of ecosystem-based management: Making the process work today in real places. **Marine Policy**, v. 34, n. 2, p. 340 - 348, 2010.

TALLON, Paul P. Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost. **IEEE Computer**, v. 46, n. 6, p. 32–38, 2013.

TALLON, Paul P; RAMIREZ, Ronald V.; SHORT, James E. The Information Artifact in IT Governance: Toward a Theory of Information Governance. **Journal of Management Information Systems**, v. 30, ed. 3, p. 141 - 177, 2014.

WEBER, Kristin; OTTO, Boris; ÖSTERLE, Hubert. One Size Does Not Fit All—A Contingency Approach to Data Governance. **ACM Journal of Data and Information Quality**, vol. 1, n. 1, 2009.

WINDISCH, Ricarda; WAISWA, Peter; NEUHANN, Florian; SCHEIBE, Florian; DE SAVIGNY, Don. Scaling up antiretroviral therapy in Uganda: using supply chain management to appraise health systems strengthening. **Globalization and Health**, v. 7, 2011.

WU, Xindong; ZHU, Xingquan; WU, Gong-Qing; Ding, Wei. Data Mining with Big Data. **IEEE Transactions on Knowledge and Data Engineering**. Vol. 26, edição 1, p. 97 - 107. 2014.

YANG, Longzhi; NEAGU, Daniel; CRONIN, Mark T. D.; HEWITT, Mark; ENOCH, Steven J.; MADDEN, Judith C.; PRZYBYLAK, Katarzyna. Towards a Fuzzy Expert System on Toxicological Data Quality Assessment. **Molecular Informatics**, v. 32, n. 1, p. 65- 78, 2013.

ZHANG, Xuyun; YANG, Laurence T.; LIU, Chang; CHEN, Jinjun. A Scalable Two-Phase Top-Down Specialization Approach for Data Anonymization Using MapReduce on Cloud. **IEEE Transactions on Parallel and Distributed Systems**. Vol. 25, edição 2, p. 363 - 373. 2014.