



Big data em saúde e a ciência aberta: um contributo para a gestão do conhecimento em COVID-19

Big data in health and open science: a contribution to the management of knowledge in COVID-19

Big data en salud y ciencia abierta: una contribución a la gestión del conocimiento en COVID-19

Jorge Magalhães¹, Zulmira Hartz², Luc Quoniam³, Gabriela de Paula Pereira⁴, Adelaide Maria de Souza Antunes⁵

Como citar este artigo:

Magalhães J, Hartz Z, Quoniam L, Pereira GP, Antunes AMS. Big data in health and open science: a contribution to the management of knowledge in COVID-19. Rev Pre Infec e Saúde [Internet]. 2020;6:10659. Available from: <https://revistas.ufpi.br/index.php/nupcis/article/view/10659> DOI: <https://doi.org/10.26694/repis.v6i0.10659>

¹ Investigador em Saúde Pública, Núcleo de Inovação Tecnológica de Farmanguinhos-NIT FAR, Fundação Oswaldo Cruz/FIOCRUZ, Ministério da Saúde do Brasil, Rio de Janeiro, Brasil.

² Instituto de Higiene e Medicina Tropical, Universidade NOVA de Lisboa, Lisboa, Portugal.

³ Université du Sud-Toulon Var, Toulon, France, professor visitante da Universidade Federal de Mato Grosso do Sul, Mato Grosso do Sul, Brasil.

⁴ Fundação Oswaldo Cruz/FIOCRUZ, Ministério da Saúde do Brasil, Rio de Janeiro, Brasil.

⁵ Instituto Nacional de Propriedade Industrial/INPI, Universidade Federal do Rio de Janeiro, Departamento de Engenharia de Processos Químicos e Bioquímicos, Rio de Janeiro, Brasil.

ABSTRACT

Introduction: The 21st century is marked by the exponential era of data available daily on the web. This era of knowledge is characterized by the pressing need for new technologies to manage the brutal amount of data. In health, it is no different: scientific and technological information needs to be identified, extracted, treated and made available the essential information for decision making. The objective is to contribute to knowledge management in times of Big Data in health with COVID-19. **Outline:** The work was carried out using a retrospective analysis of the big health data opened for COVID-19. **Results:** They demonstrate the great worldwide effort in science and technology for the topic in question. The Medline database features over 36,000 articles available. In the European Patent Office base there are 2,871 family of patents, with 3 patents being made available and the pharmaceutical company Pfizer leading the technological research. **Implications:** It is concluded that the translation of knowledge has been occurring in an urgent way, since, although there is no vaccine or medicine for the cure so far, never in history has there been so much available volume of science and technology in such a small space of time, generating scientific and technological production for the advancement of science.

DESCRIPTORS

Public Health; Big Data; Knowledge Management; Coronavirus Infections; Coronavirus.

Autor correspondente:

Jorge Magalhães
Endereço: Avenida Comandante Guarany,
447, Jacarepaguá
CEP: 22775-903 – Rio de Janeiro, Rio de
Janeiro, Brasil
Telefone: +55 (21) 3348-5050
E-mail: jorge.magalhaes@far.fiocruz.br

Submetido: 2020-05-20
Aceito: 2020-05-22

INTRODUÇÃO

O século XXI se apresenta como uma era com 40% da população conectada à Internet.¹ Nesse sentido, O'Reilly sugeriu o termo Big Data como um gigantesco banco de dados atualizado em tempo real, que atinge facilmente milhares de *terabytes* de armazenamento em diferentes formatos.² Os sistemas tradicionais de gerenciamento de banco de dados relacional não podem lidar com essas grandes massas de dados.³⁻⁴ O Big Data impulsiona uma nova geração de metodologias desenvolvidas para extrair valor econômico e estratégico de um grande e variado volume de dados (estruturado e não estruturado), permitindo alta velocidade de captura e análise.⁵⁻⁶

Big Data refere-se à terceira geração da era da informação.⁷⁻⁸ Inicialmente, esse volume exponencial de dados atendia aos critérios dos 3 Vs: Volume, Variedade e Velocidade;⁹ posteriormente, foram adicionados 2 Vs: os atributos de Veracidade e Valor. Alguns autores até atribuem os últimos 3 Vs, como Veracidade, Versatilidade e Viabilidade, onde a combinação de todos os "Vs" gera o "V" de Valor.¹⁰ O Big Data se divide em tempestade perfeita de dados, tempestade perfeita de convergência e tempestade perfeita de computação, e esta última é resultante de 4 fenômenos: lei de Moore, computação móvel, redes sociais e computação em nuvem (*cloud computing*). Este acervo de dados deve ser tratado para apresentar informação pesquisada de forma seletiva e objetiva para aumentar a inteligência dos negócios, além de permitir uma melhoria no processo de tomada de decisão.¹¹

Já ao refletirmos sobre saúde, ela é considerada como um bem público global: que não seja excludente, isso é, que ninguém ou nenhuma coletividade seja excluída de sua posse ou de seu consumo; e de que seus benefícios sejam disponíveis a todos. Há, também, o aparente consenso de que a saúde não seja concorrencial, e que não haja rivalidade, isto é, a saúde de uma pessoa não pode se dar às expensas da exclusão de outras pessoas.¹²⁻¹⁵

No espectro da saúde, ela carrega desafios e oportunidades no processo de globalização, sendo este o catalisador da evolução do termo "Saúde Global". A saúde global pode ser compreendida ao mesmo tempo como uma condição, uma atividade, uma profissão, uma filosofia, uma disciplina ou um movimento. Todavia, deve-se considerar que não há consenso sobre o que seja Saúde Global, nem uma única definição, e seu campo de ação tem limites imprecisos,¹⁶ contudo é indiscutível que vivemos a Saúde em tempos de globalização.¹⁷

Assim, é mister buscarmos identificar, extrair e tratar o Big Data da Saúde neste mundo globalizado, a fim de focarmos na informação essencial para os tomadores de decisão do presente século. Não obstante, a gestão desse conhecimento não é considerada trivial, uma vez que a abordagem da Ciência Aberta se apresenta iminente. A ciência aberta é um modelo de prática científica que, em consonância com o desenvolvimento da cultura digital, visa a disponibilização das informações em rede de forma oposta à pesquisa fechada dos laboratórios. A expressão também se refere à geração de materiais de pesquisa compartilhados abertamente, sem a necessidade de patentes. Neste âmbito, a Comunidade Europeia tem mostrado maturidade nesta área com a promoção da Ciência Aberta,¹⁸ bem como algumas instituições no Brasil, como, por exemplo, a proposta da Fiocruz para Ciência Aberta em seu Termo de referência: gestão e abertura de dados para pesquisa na Fiocruz.¹⁹

Desta forma, as organizações que lidam com investigação em saúde têm de buscar um melhor manejo da gestão do conhecimento do Big Data em Saúde, cada vez mais numa ciência aberta, a retratar uma inteligência colaborativa e construtiva para a sociedade. Portanto, este trabalho visa contribuir na gestão do conhecimento em tempos de Big Data em saúde e com COVID-19; por meio de um panorama das informações científicas e tecnológicas.

A metodologia utilizada foi identificar o estado da arte em bases de dados indexadas tais como

PubMed, Scielo, Scopus e Web of Science. Da mesma forma, *websites* oficiais como a Organização Mundial da Saúde (OMS) e o Portal Transparência do Governo Brasileiro

(<https://transparencia.registrocivil.org.br/registral-covid>). As palavras-chave utilizadas, no espectro da COVID-19, foram “Coronavir* or MERS-CoV* or SARS-CoV* or COVID* or HCoV or SARS or MERS”.

EVOLUÇÃO INFORMACIONAL DE DADOS E A INTELIGÊNCIA COLABORATIVA PARA SAÚDE GLOBAL

Pode-se considerar que o desenvolvimento de uma investigação não é mais linear. A velocidade dessa evolução tem um ritmo acelerado. O ritmo de mudança de paradigma, de novas ideias tem sido muito acelerado no mundo da era do Conhecimento. Destaca-se que as primeiras mudanças, embora aparentemente rápidas, levaram anos para se desenvolverem. Pode-se exemplificar com a questão do sequenciamento do HIV (*human immunodeficiency virus*) ter levado 15 anos, enquanto o sequenciamento do SARS (*severe acute respiratory syndrome*) foi obtido em somente 21 dias.²⁰⁻²¹ Não obstante, o mesmo ocorre em tempos de COVID-19; onde se observa que em cerca de apenas 06 (seis) meses, o compartilhamento de dados em rede e informações científicas e tecnológicas cresceram exponencialmente, como será discutido no próximo item.

Com o avanço tecnológico em todas as áreas da ciência, é mister o tratamento do grande volume de dados, e ainda mais, os dados relativos à saúde. Destaca-se que 43% de todos os bytes diários na Web são relativos à Saúde e 47% destes são de saúde pública.²² Desta forma, o volume exaustivo de dados requer organização e estruturação para o possível subsídio ao tomador de decisão, o que tem levado organizações a criarem ferramentas de busca em sites específicos na Web, a fim de extrair e tratar os dados para obtenção de informação essencial.²³ Cada vez mais, a ciência da informação se torna mais presente e útil para qualquer outra área da ciência em razão da

quantidade crescente de dados de qualquer natureza. Não obstante, como exemplo, os tratamentos de imunoterapia têm recebido grandes contribuições da inteligência artificial.²⁴

Ferramentas de busca de acesso livre ou pago são cada vez mais comuns e apresentam uma infinidade nessa inteligência competitiva e/ou colaborativa, como os *softwares* de *textmining*, *datamining* etc. O resultado prático dos dados extraídos e processados geram informação essencial e configuram-se como úteis, razoáveis e estratégicos para agilizar o processo de decisão.²⁵ A inteligência colaborativa contribui fortemente para a mudança de conhecimento e do poder do indivíduo para o coletivo.²⁶ O código aberto da inteligência coletiva acabará por gerar resultados superiores aos conhecimentos gerados pelo *software* proprietário desenvolvido dentro das corporações. A educação e a forma como as pessoas estão aprendendo a participar em culturas de conhecimento fora de contextos de aprendizagem formais é determinante no novo contexto global. É crucial a aprendizagem através dos meios de inteligência coletiva, pois é importante para a democratização da ciência, uma vez que está interligada com a cultura baseada no conhecimento e sustentado pela partilha de ideia coletiva, portanto, a contribuir para uma melhor compreensão da diversidade sociedade,²⁷⁻²⁹ com dados abertos à ciência das instituições.³⁰⁻³³ Nesse sentido, novas ferramentas da ciência da informação, proporcionam a facilidade, agilidade e sinergia para o trabalho colaborativo em rede científica e tecnológica, de tal modo, que o conhecimento gerado na área possa ser prospectado e monitorado para uma melhor ação dos tomadores de decisão; como no caso de epidemias e a pandemia COVID-19.

O CONHECIMENTO GERAL COVID-19

Grandes corporações, utilizam também o conceito de *open innovation* como forma de proporcionar a inovação de seus processos e/ou produtos com contributos de qualquer investigador,

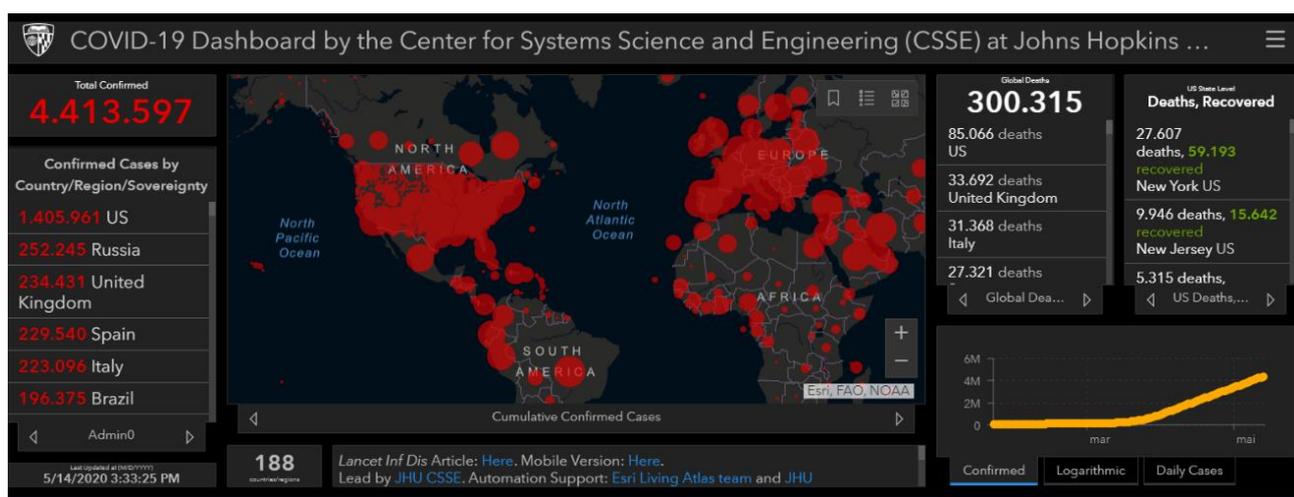
empresa etc.³⁴⁻³⁵ e na área da saúde também se traduz numa oportunidade.³⁶ Os coronavírus são um grupo de vírus grandes, envelopados, com sentido positivo e RNA de cadeia simples, pertencentes à ordem Nidovirales, família Coronaviridae, subfamília Coronavirinae.³⁷ No final de 2019, foi descoberto um novo tipo de coronavírus, chamado provisoriamente de 2019-nCoV e, mais tarde, denominado SARS-CoV-2, devido à sua semelhança com SARS-CoV. A doença causada pelo vírus foi oficialmente denominada Doença de Coronavírus 2019 (COVID-19) pela OMS.³⁸

Dois estudos descrevem que até 2019, apenas seis coronavírus causavam doenças em humanos: HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1, coronavírus com síndrome respiratória aguda grave (SARS-CoV) e coronavírus com vírus respiratório no

Oriente Médio (MERS-CoV).³⁹⁻⁴⁰ Os quatro primeiros são endêmicos localmente e foram associados, principalmente, a doenças leves e limitantes, enquanto os dois últimos podem causar doenças graves. SARS-CoV e MERS-CoV são betacoronavírus e estão entre os patógenos incluídos na Lista Modelo de Doenças Prioritárias da Organização Mundial da Saúde.³⁸

No âmbito de entender o cenário mundial dos casos de COVID-19, a *Johns Hopkins University* disponibilizou as informações tratadas do Big Data em COVID-19, referente ao cenário epidemiológico mundial (Figura 1) com 4.413.597 casos confirmados. Até 14 de maio de 2020, o Brasil apresentava 196.375 casos oficiais confirmados e uma letalidade de 13.555 casos.

Figura 1 – Casos confirmados de COVID-19 global.



Fonte: COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University, 2020.

O Brasil se destaca em quinto lugar com 4,5% dos óbitos frente ao *ranking* mundial com total de 300.315 mortes. Os EUA são o atual epicentro da pandemia com 1.405.961 casos confirmados e 85.066 óbitos.

O CONHECIMENTO CIENTÍFICO EM COVID-19

A base de literatura biomédica Medline, com mais de 30 milhões de citações no PubMed, possui 36.687 produções científicas para o termo de busca “Coronavir* or MERS-CoV* or SARS-CoV* or COVID* or HCoV or SARS or MERS”. Obviamente extrair todos os

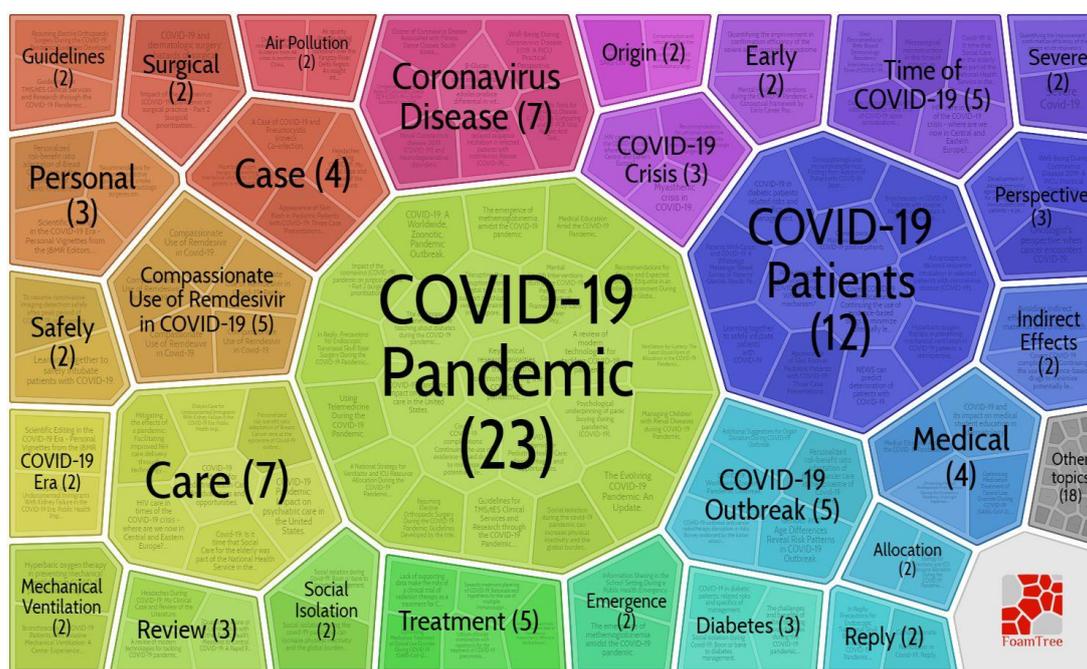
documentos da base e estudá-los pontualmente levaria um tempo grande de análise. Entretanto, usando algoritmos de engenharia de busca e posterior tratamento desse volume, pode-se obter, preliminarmente, informação essencial para auxiliar no gerenciamento do conhecimento. Como por exemplo, a plataforma livre *Carrot Search Lingo4G*.⁴¹ Desta forma, observa-se a contribuição da ciência da informação com a evolução de base algorítmica, onde possibilita cada vez processamentos mais rápidos e

melhores, subsidiando a difusão do conhecimento científico e tecnológico produzido.

Com os mesmos termos dispostos no PubMed, foram extraídos 100 documentos mais relevantes. Destes, foram agregados em subtemas que se correlacionam com novas palavras-chave que mais se caracterizaram (apareceram) dentro de cada produção científica analisada pelo algoritmo. Como se observa na Figura 2, no tema central foram 23 documentos

recuperados (COVID-19 pandemic), já em *COVID-19 Patients* foram separados 12, e para *Coronavirus Disease* 7 etc. Cabe destacar que após a “mineração” dos dados, os termos mais “evidentes” no termo da busca pesquisado, os novos termos que mais se repetem são segregados e geram o cluster dos 100 mais relevantes em relação aos 36.687 iniciais na base PubMed.

Figura 2 – Resultado documentos científicos no MEDLINE sobre COVID-19.



Fonte: Extraído pelos autores em CarrotLingo4g.

No que tange ao número de autores, observou-se que há participação em rede variada, desde opinião de um autor até trabalhos com 11 (onze) autores participando e com instituições diversas internacionalmente.

O CONHECIMENTO TECNOLÓGICO EM COVID-19

O conhecimento tecnológico pode ser medido pelo indicador de patentes depositadas em algum escritório de propriedade industrial e/ou intelectual de algum país. As patentes configuram-se como um forte indicador de inovação em ciência e tecnologia.⁴²⁻⁴⁴ Mesmo com o ambiente de pandemia extremamente

recente de COVID-19, pode-se observar a sinergia dos cientistas e das empresas no mundo, unindo-se da inteligência colaborativa e da ciência aberta, e que tem sido produzida quantidade exponencial de trabalhos científicos, como visto anteriormente, bem como em patentes.

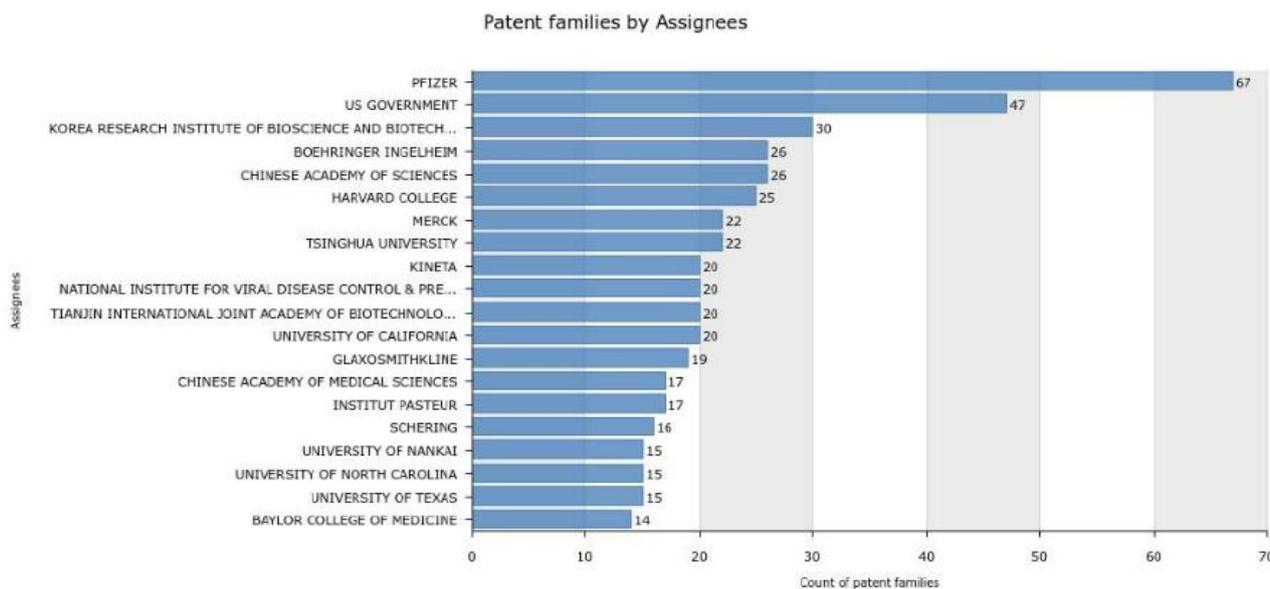
Um desafio na área tecnológica é aliar a socialização do conhecimento patentário de forma igualitária aos cientistas de quaisquer grupos de pesquisa, mesmo que não existam situações emergenciais de calamidade pública nos países. Segundo a Lei de propriedade intelectual (nº 9.279/96), um pedido de patente tem o sigilo mantido durante 18 (dezoito) meses a contar da data de seu

depósito. Portanto, considerando que a COVID-19 foi descoberta no final de 2019, há que se esperar não ter nenhum resultado para patente. Desta forma, foi realizada busca com os sinônimos do Coronavírus (família Coronaviridae; Síndrome Respiratória Aguda Grave (SARS); Síndrome Respiratória Aguda do Oriente Médio (MERS)). Desta forma, 03 (três) famílias de patentes foram encontradas: uma chinesa

CN110960532, de inventor independente, uma coreana KR10-2020-0032050 e uma indiana IN202041010021.

Nesse sentido, efetuou-se a busca de forma ampliada com outros coronavírus. Desta forma, no Gráfico 1 observa-se 2.871 famílias de patentes encontradas e distribuídas pela empresa detentora da patente. Destaca-se a empresa farmacêutica Pfizer com 67 patentes.

Gráfico 1 – Principais depositantes de famílias de patentes para Coronavírus.



Fonte: Extraído pelos autores do *European Patent Office* pelo Questel Orbit®. Maio, 2020.

Na Figura 3, nota-se os países onde foram requeridas as patentes pelos depositantes do Gráfico 1 e/ou localização dos inventores delas.

Figura 3 – Países dos depositantes de famílias de patentes para Coronavírus.



Fonte: Extraído pelos autores do *European Patent Office* pelo Patent Inspiration®. Maio, 2020.

No sentido de melhor estratificar o conhecimento gerado e favorecer um rápido entendimento para a gestão deste conhecimento, os dados essenciais plotados em mapa, auxiliam o tomador de decisão, num prático entendimento observacional pela figura plotada. Portanto, na Figura 3, nota-se que quanto mais escuro é a cor disposta no país, representa o país com maior número de conhecimento tecnológico expresso em depósito patentário. Nota-se que a cor mais acentuada para o vermelho (EUA) é onde estão a maioria.

CONCLUSÃO

O século 21 trouxe novos desafios e grandes oportunidades devido ao crescente volume de dados novos adicionados à Web diariamente. Na área da Saúde não é diferente. O conhecimento científico e tecnológico tem avançado de forma exponencial jamais visto outrora. Assim, novas formas de identificar, extrair e tratar o Big Data em saúde se torna fundamental para tomada de decisão.

Políticas de ciência aberta tem sido cada vez mais difundida nesta nova era do século 21, como forma de resposta a este avanço científico e tecnológico. Nesse sentido, para auxiliar a gestão do conhecimento, novas soluções têm sido criadas, a fim de tornar mais célere a árdua tarefa de tomada de

decisão, principalmente na Saúde, especialmente em tempos de pandemia.

Como contributo para gestão do conhecimento adquirido em tempo recorde da pandemia COVID-19, o lastro científico alcançou cerca de 37 mil publicações somente na base PubMed, 03 (três) famílias de patentes e 2.871 família de patentes relacionadas ao Coronavírus.

Os resultados indicam que a ciência aberta proporciona a sinergia dos resultados e o trabalho em rede. Cabe destacar a limitação da “ciência fechada”, por exemplo, o resultado (conhecimento) gerado na plotagem do Gráfico 1, foi obtido pelo sistema de busca e análise de informações Questel Orbit®. Logo, a organização que não tiver a licença não poderá utilizar. Entretanto, pode-se concluir que a investigação individual tem perdido espaço acadêmico nos últimos tempos, proporcionando a inteligência colaborativa.

Pelos dados científicos e tecnológicos identificados e as respectivas informações essenciais disponibilizadas, gerando um arcabouço essencial para tomada de decisão, corrobora-se o papel fundamental da ciência da informação para todas as áreas da Ciência - fundamental para a área da Saúde. Convergência digital e sinergia colaborativa são fatores preponderantes para gestão do conhecimento no Big Data da Saúde.

RESUMO

Introdução: O século 21 é marcado pela era exponencial de dados disponíveis diariamente na web. Esta era do conhecimento caracteriza-se pela necessidade premente de novas tecnologias para gerenciar a quantidade brutal de dados. Na área da saúde não é diferente: informações científicas e tecnológicas precisam ser identificadas, extraídas, tratadas, e disponibilizadas as informações essenciais para tomada de decisão. Objetiva-se contribuir com a gestão do conhecimento em tempos de Big Data em saúde com COVID-19. **Delineamento:** O trabalho foi realizado utilizando análise retrospectiva dos grandes dados de saúde abertos para COVID-19. **Resultados:** Demonstram o grande esforço mundial na área da ciência e tecnologia para o tema em questão. A base Medline apresenta mais de 36.000 artigos disponíveis. Na base *European Patent Office* há 2.871 família de patentes, sendo 03 patentes disponibilizadas, e a empresa farmacêutica Pfizer liderando a pesquisa tecnológica. **Implicações:** Conclui-se que a translação do conhecimento vem ocorrendo de maneira urgente, uma vez que, embora, não haja vacina ou medicamento para a cura até o momento, jamais na história houve tanto volume disponível de ciência e tecnologia num espaço tão pequeno de lapso temporal, gerando produção científica e tecnológica para o avanço da ciência.

DESCRITORES

Saúde Pública; Big Data; Gestão do Conhecimento; Infecções por Coronavírus; Coronavirus.

RESUMEN

Introducción: El siglo XXI está marcado por la era exponencial de los datos disponibles a diario en la web. Esta era del conocimiento se caracteriza por la urgente necesidad de nuevas tecnologías para gestionar la brutal cantidad de datos. En el área de la salud no es diferente: es necesario identificar, extraer, tratar la información científica y tecnológica y disponer de la información esencial para la toma de decisiones. El objetivo es contribuir a la gestión del conocimiento en tiempos de Big Data en salud con COVID-19. **Delineación:** El trabajo se llevó a cabo utilizando un análisis retrospectivo de los grandes datos de salud abiertos para COVID-19. **Resultados:** Demostrar el gran esfuerzo mundial en el área de ciencia y tecnología para el tema en cuestión. La base de datos de Medline cuenta con más de 36.000 artículos disponibles. En la base de la Oficina Europea de Patentes hay 2.871 familias de patentes, de las cuales 03 patentes están disponibles, y la compañía farmacéutica Pfizer lidera la investigación tecnológica. **Implicaciones:** Se concluye que la traducción del conocimiento se viene dando de forma urgente, ya que, aunque hasta el momento no existe una vacuna o medicamento para la cura, nunca en la historia se ha contado con tanto volumen de ciencia y tecnología disponible en tan poco tiempo, generando producción científica y tecnológica para el avance de la ciencia.

DESCRIPTORES

Salud Pública; Macrodatos; Gestión del Conocimiento; Infecciones por Coronavirus; Coronavirus.

REFERÊNCIAS

1. McKinsey Global Institute. Big data: The next frontier for innovation, competition, and productivity. 2011. Available from: http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation
2. O'Reilly T. What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Rochester, NY: Social Science Research Network; 2007. Available from: <http://papers.ssrn.com/abstract=1008839>
3. Magalhães JL, Quoniam L. Perception of the Information Value for Public Health: A Case Study for Neglected Diseases: Library and Information Science Book Chapter. In: Rethinkin the Conceptual Base for New Practical Applications in Information Value and Quality. IGI Global; 2013. p. 345. Available from: <http://www.igi-global.com/chapter/perception-information-value-public-health/84218>
4. Quoniam L, Lucien A. Intelligence compétitive 2.0: organisation, innovation et territoire. France: Librairie Lavoisier; 2010. Available from: <http://www.lavoisier.fr/livre/notice.asp?ouvrage=2139418&pos=8>
5. Gray J, Chambers L, Bounegru L. The Data Journalism Handbook, O'Reilly Media, 2012 - InfoVis: Wiki; 2012. Available from: http://www.infovis-wiki.net/index.php?title=Gray,_J._and_Chambers,_L._and_Bounegru,_L._The_Data_Journalism_Handbook,_O%27Reilly_Media,_2012
6. Sivarajah U, Kamal MM, Irani Z, Weerakkody V. Critical analysis of Big Data challenges and analytical methods. J Busin Resear [Internet]. 2017; Jan [cited 31 Mar 2020]; 70:263–286. Available from: <https://doi.org/10.1016/j.jbusres.2016.08.001>
7. Magalhaes JL, Quoniam L. Percepção do valor da informação por meio da inteligência competitiva 2.0 e do Big Data na saúde. In: Tarapanoff, K. Análise da Informação para Tomada de Decisão: desafios e soluções. Brasil: InterSaber; 2015. p. 365.
8. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst [Internet]. 2020 Feb [cited 31 Mar 2020];2. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4341817/>
9. Laney D. 3D Data Management: Controlling Data Volume, Velocity, and Variety. 2001 Available from: <http://blogs.gartner.com/douglaney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>
10. Aleixo JA, Duarte P. Big data opportunities in healthcare. How can medical affairs contribute? Rev Port Farmacoter [Internet]. 2015 Feb [cited 31 Mar 2020]; 7:230–6. Available from: <https://pdfs.semanticscholar.org/2066/744d3646ad986ec33c2f582c9ef05773213c.pdf>
11. Minelli M, Chambers M, Dhiraj A. Big Data, Big Analytics. EUA: John Wiley & Sons, Inc.; 2013.
12. Buse K, Waxman A. Public-private health partnerships: a strategy for WHO. Bull World Health Organ [Internet]. 2001 Jan [cited 31 Mar 2020]; 79(8):748–54. Available from: <https://pubmed.ncbi.nlm.nih.gov/11545332/>
13. Haines A, McMichael AJ, Smith KR, Roberts I, Woodcock J, Markandya A, et al. Public health benefits of strategies to reduce greenhouse-gas emissions: overview and implications for policy makers. Lancet [Internet]. 2009 Dec [cited 31 Mar 2020]; 374(9707):2104–14. Available from: [https://doi.org/10.1016/S0140-6736\(09\)61759-1](https://doi.org/10.1016/S0140-6736(09)61759-1)
14. Hartz ZMA. Meta-evaluation of health management: Challenges for “new public health”. Cienc Saude Coletiva [Internet]. 2012 Feb [cited 31 Mar 2020]; 17(4):832–4. Available from: <https://doi.org/10.1590/S1413-81232012000400004>
15. Vance K, Howe W, Dellavalle RP. Social Internet Sites as a Source of Public Health Information. Dermatol Clin [Internet]. 2009 Feb [cited 31 Mar 2020]; 27(2):133–6. Available from: <https://doi.org/10.1016/j.det.2008.11.010>
16. Fortes PA de C, Ribeiro H. Saúde Global em tempos de globalização. Saúde Soc [Internet]. 2014 Jun [cited 31 Mar 2020]; 23(2):366–75. Available from: <https://doi.org/10.1590/S0104-12902014000200002>

17. Koplan JP, Bond TC, Merson MH, Reddy KS, Rodriguez MH, Sewankambo NK, et al. Towards a common definition of global health. *Lancet Lond Engl* [Internet]. 2009 Jun [cited 31 Mar 2020]; 373(9679):1993–5. Available from: [https://doi.org/10.1016/S0140-6736\(09\)60332-9](https://doi.org/10.1016/S0140-6736(09)60332-9)
18. Resolução Conselho de Ministros. Resolução do Conselho de Ministros n.º 20/2016. Diário da República n.º 70/2016, Série I de 2016-04-11. Série I de 2016-04-11, 70/2016 abr 11, 2016. Available from: <http://legislacaoportuguesa.com/resolucao-do-conselho-de-ministros-n-o-202016-diario-da-republica-n-o-702016-serie-i-de-2016-04-11/>
19. FIOCRUZ. Grupo de Trabalho em Ciência Aberta. Termo de referência: gestão e abertura de dados para pesquisa na Fiocruz. Rio de Janeiro: FIOCRUZ/Presidência; 2018. Available from: <https://www.arca.fiocruz.br/handle/icict/26803>
20. Goes ACS, Oliveira BVX. Projeto Genoma Humano: um retrato da construção do conhecimento científico sob a ótica da revista Ciência Hoje. *Ciênc. educ. (Bauru)* [Internet]. 2014 Sep [cited 31 Mar 2020]; 20(3):561–577. Available from: <https://doi.org/10.1590/1516-73132014000300004>
21. The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East by John Gantz and David Reinsel sponsored by EMC [Internet]. 2012 [cited 20 Apr 2020]. Available from: <https://www.emc.com/leadership/digital-universe/2012iview/index.htm>
22. Magalhães JL, Quoniam L. Perception of the Information Value for Public Health: A Case Study for Neglected Diseases. In: Magalhães JL, Quoniam L. Rethinkin the Conceptual Base for New Practical Applications in Information Value and Quality. Pennsylvania, USA: Information Science Reference; 2014. p. 345.
23. Tarapanoff, K. *Análise da Informação para Tomada de Decisão Desafios e Soluções*. vol. 1. Brasil: Editora InterSaberes; 2015.
24. Jabbari P, Rezaei N. Artificial intelligence and immunotherapy. *Expert Rev Clin Immunol* [Internet]. 2019 Jul [cited 31 Mar 2020]; 15(7):689–91. Available from: <https://doi.org/10.1080/1744666X.2019.1623670>
25. Magalhães JL, Oliveira DA, Costa JCS, Hartz Z. Strategic Vision on the Chain of Decision in Modern Competitive Scenarios: A Case Study in Material Planning in a Public Company. Pennsylvania, USA: IGI Global; 2020.
26. Bonabeau, E. Decisions 2.0: The Power of Collective Intelligence. *MIT Sloan Manag Rev* [Internet]. 2009 Jul [cited 31 Mar 2020]; 45–52. Available from: <https://sloanreview.mit.edu/article/decisions-20-the-power-of-collective-intelligence/>
27. Burke RR. Reasoning with empirical marketing knowledge. *Int J Res Mark* [Internet]. 1991 Apr [cited 31 Mar 2020]; 8(1):75–90. Available from: [https://doi.org/10.1016/0167-8116\(91\)90008-U](https://doi.org/10.1016/0167-8116(91)90008-U)
28. Collective intelligence. In: Wikipedia, the free encyclopedia. 2016. Available from: https://en.wikipedia.org/w/index.php?title=Collective_intelligence&oldid=732642532
29. Trigo MR, Gouveia LB, Quoniam L, Riccio EL. Using competitive intelligence as a strategic tool in a higher education context. *Nr Reading: Academic Conferences Ltd*; 2007.
30. Pordes R, Petravick D, Kramer B, Olson D, Livny M, Roy A, et al. The open science grid. *J Phys Conf Ser* [Internet]. 2007 Jul [cited 31 Mar 2020]; 78(1):1–12. Available from: <https://doi.org/10.1088/1742-6596/78/1/012057>
31. Molloy JC. The Open Knowledge Foundation: Open Data Means Better Science. *PLOS Biol* [Internet]. 2011 Dec [cited 31 Mar 2020]; 9(12):1–12. Available from: <https://doi.org/10.1371/journal.pbio.1001195>
32. David PA. Understanding the emergence of ‘open science’ institutions: functionalist economics in historical context. *Ind Corp Change* [Internet]. 2004 Jan [cited 31 Mar 2020]; 13(4):571–89. Available from: <https://doi.org/10.1093/icc/dth023>
33. Errington TM, Iorns E, Gunn W, Tan FE, Lomax J, Nosek BA. An open investigation of the reproducibility of cancer biology research. *eLife* [Internet]. 2014 Dec [cited 31 Mar 2020]; 3:1–12. Available from: <https://doi.org/10.7554/eLife.04333>
34. Celadon KL. Knowledge Integration and Open Innovation in the Brazilian Cosmetics Industry. *J Technol Manag Innov* [Internet]. 2014 Sep [cited 31 Mar 2020]; 9(3):34–50. Available from: <https://www.jotmi.org/index.php/GT/article/view/1612>
35. Michelino F, Cammarano A, Lamberti E, Caputo M. Knowledge Domains, Technological Strategies and Open Innovation. *J Technol Manag Innov* [Internet]. 2015 Jul [cited 31 Mar 2020]; 10(2):50–78. Available from: https://www.academia.edu/29792874/Knowledge_Domains_Technological_Strategies_and_Open_Innovation
36. Chaifetz S, Chokshi DA, Rajkumar R, Scales D, Benkler Y. Closing the access gap for health innovations: an open licensing proposal for universities. *Glob Health* [Internet]. 2007 Feb [cited 31 Mar 2020]; 3(1):1–10. Available from: <https://globalizationandhealth.biomedcentral.com/articles/10.1186/1744-8603-3-1>
37. Bastos LFCS. <https://www.facebook.com/pahowho>. OPAS/OMS Brasil. Folha informativa: COVID-19 (doença causada pelo novo coronavírus). / OPAS/OMS. Pan American Health Organization. Brasília: WHO; 2020. Available from: https://www.paho.org/bra/index.php?option=com_content&view=article&id=6101:covid19&Itemid=875
38. Coronavírus disease 2019. Genebra: WHO; 2019. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019?gclid=Cj0KCQjw2PP1BRCiARIsAEqv-pTucWCIRp6LtBdqlo321_Dzgbm-9Kky2CPJqXWBncwyzKp29ueLpc0aAgarEALw_wcB
39. Saxena SK. Coronavírus Disease 2019 (COVID-19): Epidemiology, Pathogenesis, Diagnosis, and Therapeutics. Springer Nature [Internet]. 2020 Jul [cited 31 Jul 2020]. Available from: <https://link.springer.com/book/10.1007/978-981-15-4814-7>

40. Bonilla-Aldana DK, Holguin-Rivera Y, Cortes-Bonilla I, Cardona-Trujillo MC, García-Barco A, Bedoya-Arias HA, et al. Coronavirus infections reported by ProMED, February 2000–January 2020. *Travel Med Infect Dis* [Internet]. 2020 Feb [cited 31 Mar 2020]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7102564/>
41. Magalhães J, Bastos AC, Barroso W. Cenário Global e Glocal das Tendências Científicas e Tecnológicas em Diabetes: Uma Abordagem do Big Data em Saúde no Século 21. *Rev Gest Em Sist Saúde* [Internet]. 2016 Jun [cited 31 Mar 2020]; 5(1):1–14. Available from: <http://www.revistargss.org.br/ojs/index.php/rgss/article/view/191>
42. Freitas CCG, Segatto AP. Science, technology and society from the perspective of social technology: a study from the Critical Theory of Technology. *Cad EBAPEBR* [Internet]. 2014 Jun [cited 31 Mar 2020]; 12(2):302–20. Available from: <https://doi.org/10.1590/1679-39517420>
43. Leitão DM. O conhecimento tecnológico e sua importância. Possibilidades de sua transferência internacional. *Ciênc Informação* [Internet]. 1981 May [cited 31 Mar 2020]; 10(2):1–12. Available from: <http://revista.ibict.br/ciinf/article/view/149>
44. Antunes A, Mercado A. A aprendizagem tecnológica no Brasil: a experiência da indústria química e petroquímica. Editora E-papers; 2000.

COLABORAÇÕES

Jorge Magalhães (JM) e Zulmira Hartz (ZH) foram idealizadores do artigo. JM e Gabriela de Paula Pereira desenvolveram a coleta de dados juntamente com Adelaide Antunes (AA) e a redação do presente trabalho. Luc Quoniam e ZH analisaram, discutiram a revisão final do trabalho com JM e AA.

AGRADECIMENTOS

Os autores agradecem a infraestrutura de Farmanguinhos/FIOCRUZ e da Gestec/FIOCRUZ pela disponibilidade das bases Questel Orbit, do Instituto Nacional de Propriedade Industrial (INPI) e a do Instituto de Higiene e Medicina Tropical (IHMT) da Universidade NOVA de Lisboa.

DISPONIBILIDADE DOS DADOS

Não se aplica.

FONTE DE FINANCIAMENTO

Próprio financiamento institucional.

CONFLITOS DE INTERESSE

Não há conflitos de interesses a declarar.