

EDITORIAL

O compartilhamento de questões e problemas de pesquisa extrapolam as fronteiras de disciplinas científicas e é o cerne da criação de redes e comunidades entre cientistas, especialistas e organizações. Ao se compartilhar as questões e problemas relacionados a determinada temática, também se socializa as soluções e caminhos que estão sendo tomados por estes grupos que atuam de forma colaborativa, possibilitando abordagens interdisciplinares.

A própria área da Ciência da Informação é interdisciplinar por natureza. Sendo assim, é uma área que está em constante contato com outras áreas do conhecimento e, por isso, é capaz de incorporar elementos até então estranhos à área. A interdisciplinaridade se baseia na articulação de diferentes disciplinas e as coloca em inter-relação. Neste número especial, o foco dos trabalhos se dá justamente à emergente discussão sobre a Ciência de Dados no âmbito de diversas disciplinas e à necessidade de pautar essa discussão de forma mais consolidada na área da Ciência da Informação.

A Ciência de Dados tem como matéria prima o dado, que também é uma das matérias primas utilizadas pela Ciência da Informação. Logo, nada mais natural que a colaboração entre estas duas disciplinas. Um levantamento realizado nas tradicionais bases *Information Science & Technology Abstracts* e *Library, Information Science & Technology Abstracts* mostra que essa tendência de interação com a Ciência de Dados já é uma realidade para a área da Ciência da Informação. Ao se buscar pelo termo “*data science*” percebe-se um crescimento exponencial no número de artigos publicados e que incorporaram a temática. O primeiro artigo a abordar o termo remonta ao ano de 1977, mas é a partir do ano de 2016 que se percebe um maior interesse dos pesquisadores da Ciência da Informação sobre o tema. Naquele ano foram mapeados 148 artigos, o dobro que no ano anterior. Em 2020 foram 281 artigos identificados sobre a temática “*data science*”, consolidando um crescimento contínuo que já dura mais de uma década.

No entanto, há de se ressaltar que a Ciência de Dados se refere a um conceito diverso que engloba outros conceitos igualmente abrangentes como *big data*, *machine learning*, *information retrieval*, dentre outros.

Logo, o que se pode esperar da interação entre a Ciência de Dados e a Ciência da Informação é justamente a qualificação desses conceitos a partir da perspectiva própria da nossa área. Os artigos presentes neste número primaram pela diversidade.

A revista aborda em seu artigo de abertura aspectos relacionados à relação entre Ciência de Dados e Ciência da Informação com o título: *Interlocuções bibliográficas e epistemológicas entre a ciência de dados e a ciência da informação*. Depois são abordados temas que tratam de como a Ciência de Dados pode ser utilizada para aprimorar o processo de organização e classificação de dados, metadados e a informação gerida por sistemas informacionais. É o caso dos artigos: *Grau de pertencimento como insumo para classificação automática de textos: uma abordagem sintática*, *Dados e metadados*, *Modelagem de metadados multimídia*, *Recuperação de informação: descoberta e análise de workflows para agregação de dados do patrimônio cultural*, *Dicionário Semântico de Dados: abordagem de anotação de dados aplicada à geração de indicadores de desempenho*, *Modelo de Análise Temporal em Contexto Semântico de Gerenciamento de Emergências e Fusão de dados para análise de imagens registradas por satélites: proposta de modelo de metadados*.

Como não podia deixar de ser, há um grande avanço na discussão acerca dos dados abertos, seus repositórios e ferramentas que favorecem a abertura de dados. Fazem parte desse bloco os seguintes artigos: *Explorando a Reconciliação de Dados Culturais na Wikidata*, *Publicando dados de pesquisa*, *DBacademic: Conectando os dados abertos das instituições de ensino do Brasil*, *GOOGLE DATASET SEARCH: Visão geral e perspectivas para indexação e disponibilização de conjuntos de dados científicos abertos*, *Apliação de Dados Governamentais Abertos à Luz da Ciência da Informação* e *A publicidade de dados abertos pelo tribunal superior eleitoral: o caso do Repositório de Dados Eleitorais*.

Por fim, o último bloco de artigos deste número da revista Ciência da Informação foca nas métricas informacionais e na utilização destas métricas para entender o avanço da pesquisa científica no Brasil.

EDITORIAL

Os artigos que compõem este último bloco são: *Medição da informação científica na Web 2.0, Acervos Culturais Brasileiros no Repositório Wikimedia Commons, Perfil das orientações e produções das mulheres fundamentado em dados da Plataforma Lattes, Um modelo populacional para análise de genealogia acadêmica: Evidências sobre crescimento acadêmico no Brasil e Estimando Futuras Colaborações em Dados sobre Atividades Científicas.*

Após a apresentação deste número é possível perceber que há grande sinergia entre as áreas de Ciência e Dados e Ciência da Informação. É preciso aprofundar a compreensão sobre como essas áreas podem se aprimorar e se alimentar a fim de permitir que a informação seja cada vez mais entendida como matéria prima da ciência. Acreditamos que os artigos aqui apresentados contribuirão significativamente para esse processo. Boa leitura!

Andre Luiz Appel

Doutor em Ciência da Informação pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e Universidade Federal do Rio de Janeiro (PPGCI-IBICT/UFRJ) – Rio de Janeiro, RJ – Brasil. Assistente de Pesquisa do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) – Brasília, DF – Brasil. Pesquisador do Laboratório Interdisciplinar sobre Informação e Conhecimento (Liinc-UFRJ/IBICT) - Brasil.
E-mail: andreappel@ibict.br

Ricardo Barros Sampaio

Pós-Doutorado pela Universidade de Brasília (UnB) – DF – Brasil. Pós-Doutorado pela Fundação Oswaldo Cruz (FIOCRUZ) - Brasil. Doutor em Ciências da Informação pela Universidade de Brasília (UnB) – DF – Brasil, com período sanduíche em Université de Toulouse - França. Professor e pesquisador do Programa de Pós-Graduação em Ciência da Informação da Universidade de Brasília (UnB) – DF – Brasil. Professor e pesquisador no Mestrado Profissional de Políticas Públicas em Saúde e na especialização em Saúde Coletiva pela Escola Fiocruz de Governo - Brasília, DF – Brasil.
<http://lattes.cnpq.br/3477515781752110>
E-mail: rsampaio.br@gmail.com

Tiago Emmanuel Nunes Braga

Doutor em Ciência da Informação pela Universidade de Brasília (UnB) – Brasília, DF - Brasil. Coordenador Geral de Tecnologias da Informação e Informática (CGTI) e pesquisador do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) - Brasília, DF – Brasil.
E-mail: tiagobraga@ibict.br

EDITORIAL

Sharing research issues and problems crosses the boundaries of scientific fields and is at the heart of creating networks and communities among scientists, experts and organizations. By sharing issues and problems related to a particular theme, the solutions and paths being taken by these groups working collaboratively are also socialized, enabling interdisciplinary approaches.

The field of Information Science is interdisciplinary by nature. As such, it is in constant contact with other fields of knowledge and, therefore, capable of incorporating elements that were previously foreign to the field. Interdisciplinarity is based on the articulation of different disciplines and places them in interrelationship. In this special issue, the focus of the manuscripts is precisely the emerging discussion about Data Science in the scope of several disciplines and the need to guide this discussion in a more consolidated way in the field of Information Science.

Data is Data Science's raw material, which is also one of the raw materials used by Information Science. Therefore, nothing more natural than the collaboration between these two disciplines. A survey carried out in the traditional databases *Information Science & Technology Abstracts* and *Library, Information Science & Technology Abstracts* shows that this trend of interaction with Data Science is already a reality for the field of Information Science. When searching for the term "data science", an exponential growth in the number of articles published that incorporated the theme can be found. The first article to address the term dates back to 1977, but it is from the year of 2016 onwards that Information Science researchers have become more interested in the topic. In that year, 148 articles were mapped, twice as much as in the previous year. In 2020, 281 articles were identified on the theme "data science", consolidating a continuous growth that has lasted more than a decade.

However, it should be noted that Data Science refers to a diverse concept that encompasses other equally comprehensive concepts such as big data, machine learning, information retrieval, among others. Therefore, what can be expected from the interaction between Data Science and Information Science is precisely the qualification of these concepts from the perspective of our field. The articles in this issue are distinguished by their diversity.

In its opening article, the journal addresses aspects related to the relationship between Data Science and Information Science with the title "Bibliographic and epistemological interlocutions between data science and information science". Then, topics that deal with how Data Science can be used to improve the process of organizing and classifying data, metadata and the information managed by informational systems are discussed. This is the case of the articles "Degree of belonging as an input for automatic text classification: a syntactic approach", "Data and metadata", "Multimedia metadata modeling", "Information retrieval: discovery and analysis of workflows for aggregation of cultural heritage data", "Semantic Dictionary of Data: data annotation approach applied to the generation of performance indicators", "Temporal Analysis Model in Semantic Context of Emergency Management" and "Data Fusion for analysis of images recorded by satellites: proposal of a metadata model".

As it should be, there is a great advance in the discussion about open data, its repositories and tools that favor the opening of data. The following articles are part of this block: Exploring the Reconciliation of Cultural Data on Wikidata, Publishing research data, DBacademic: Connecting open data from educational institutions in Brazil, GOOGLE DATASET SEARCH: Overview and perspectives for indexing and making available sets of Open Scientific Data, Application of Open Government Data in the Light of Information Science and Publicity of Open Data by the Superior Electoral Court: the Electoral Data Repository case.

Finally, the last block of articles in this issue of journal *Ciência da Informação* focuses on informational metrics and the use of these metrics to understand the advancement of scientific research in Brazil. The articles that make up this last block are: Measuring scientific information on Web 2.0, Brazilian Cultural Collections in the Wikimedia Commons Repository, Profile of women's orientations and productions based on data from the Lattes Platform, A population model for academic genealogy analysis: Evidence on academic growth in Brazil and Estimating Future Collaboration on Scientific Activities Data.

EDITORIAL

After the presentation of this issue, it is possible to see that there is great synergy between the fields of Data Science and Information Science. It is necessary to deepen the understanding of how they can improve and feed from each other in order to allow information to be increasingly understood as the raw material of science. We believe that the articles presented here will significantly contribute to this process.

Good reading!

André Luiz Appel

PhD in Information Science from the Brazilian Institute of Information in Science and Technology (Ibict) and Federal University of Rio de Janeiro (PPGCI-IBICT/UFRJ) – Rio de Janeiro, RJ – Brazil. Research Assistant at the Brazilian Institute of Information in Science and Technology (Ibict) – Brasília, DF – Brazil. Researcher at the Interdisciplinary Laboratory on Information and Knowledge (Liinc-UFRJ/IBICT) - Brazil.
E-mail: andreappel@ibict.br

Ricardo Barros Sampaio

Post-Doctorate at the University of Brasília (UnB) – DF – Brazil. Post-Doctorate from the Oswaldo Cruz Foundation (FIOCRUZ) - Brazil. PhD in Information Sciences from the University of Brasília (UnB) – DF – Brazil, with a sandwich period in University of Toulouse - France.
Professor and researcher at the Graduate Program in Information Science at the University of Brasília (UnB) – DF – Brazil. Professor and researcher at the Professional Master's Degree in Public Health Policy and specialization in Public Health at Fiocruz School of Government - Brasília, DF – Brazil.
<http://lattes.cnpq.br/3477515781752110>
E-mail: rsampaio.br@gmail.com

Tiago Emmanuel Nunes Braga

PhD in Information Science from the University of Brasília (UnB) – Brasília, DF - Brazil. General Coordinator of Information Technologies and Informatics (CGTI) and researcher at the Brazilian Institute of Information in Science and Technology (Ibict) - Brasília, DF – Brazil.
E-mail: tiagobraga@ibict.br

EDITORIAL

Compartir cuestiones y problemas de investigación traspasa los límites de las disciplinas científicas y está en el centro de la creación de redes y comunidades entre científicos, expertos y organizaciones. Al compartir las cuestiones y problemas relacionados con un tema en particular, las soluciones y caminos que están tomando estos grupos que trabajan en colaboración también se socializan, lo que permite enfoques interdisciplinarios.

El área de las Ciencias de la Información en sí es de naturaleza interdisciplinaria. De esta forma, es un área que está en constante contacto con otras áreas del conocimiento y, por lo tanto, es capaz de incorporar elementos que antes eran ajenos al área. La interdisciplinariedad se basa en la articulación de diferentes disciplinas y las coloca en interrelación. En este número especial, el foco de los artículos es precisamente la discusión emergente sobre *Data Science* en el ámbito de varias disciplinas y la necesidad de orientar esta discusión de una manera más consolidada en el área de las Ciencias de la Información.

La ciencia de datos tiene como materia prima los datos, que también es una de las materias primas que utiliza la ciencia de la información. Por tanto, nada más natural que la colaboración entre estas dos disciplinas. Una encuesta realizada en las bases de datos tradicionales *Information Science & Technology Abstracts* y *Library, Information Science & Technology Abstracts* muestra que esta tendencia de interacción con Data Science ya es una realidad para el área de Ciencias de la Información. Al buscar el término “data science”, se puede ver un crecimiento exponencial en el número de artículos publicados que incorporaron el tema. El primer artículo que aborda el término remonta a 1977, pero es a partir del año 2016 cuando los investigadores de Ciencias de la Información se han interesado más por el asunto. En ese año se mapearon 148 artículos, el doble que en el año anterior. En 2020 se identificaron 281 artículos sobre el tema “data science”, consolidando un crecimiento continuo que ha durado más de una década.

Sin embargo, cabe destacar que *Data Science* se refiere a un concepto diverso que engloba otros conceptos igualmente completos como *big data*, *machine learning*, recuperación de información, entre otros.

De esa forma, lo que se puede esperar de la interacción entre Data Science y Ciencias de la Información es precisamente la calificación de estos conceptos desde la perspectiva de nuestro campo. Los artículos de este número se distinguen por su diversidad.

En su artículo de apertura, la revista aborda aspectos relacionados con la relación entre la ciencia de datos y la ciencia de la información con el título: Interlocuciones bibliográficas y epistemológicas entre ciencia de datos y ciencia de la información. Luego, se discuten temas que tratan sobre cómo la ciencia de datos se puede utilizar para mejorar el proceso de organización y clasificación de datos, metadatos e información administrada por sistemas de información. Este es el caso de los artículos: Grado de pertenencia como insumo para la clasificación automática de textos: un enfoque sintáctico, Datos y metadatos, Modelado de metadatos multimedia, Recuperación de información: descubrimiento y análisis de flujos de trabajo para la agregación de datos del patrimonio cultural, Diccionario semántico de datos : enfoque de anotación de datos aplicado a la generación de indicadores de desempeño, Modelo de Análisis Temporal en Contexto Semántico de Manejo de Emergencias y Fusión de Datos para análisis de imágenes registradas por satélites: propuesta de un modelo de metadatos.

Como debe ser, hay un gran avance en la discusión sobre datos abiertos, sus repositorios y herramientas que favorecen la apertura de datos. Los siguientes artículos forman parte de este bloque: Explorando la reconciliación de datos culturales en Wikidata, Publicación de datos de investigación, DBacademic: Conectando datos abiertos de instituciones educativas en Brasil, GOOGLE DATASET SEARCH: Visión general y perspectivas para indexar y poner a disposición conjuntos de Datos científicos abiertos , Aplicación de Datos de Gobierno Abierto a la Luz de la Ciencia de la Información y Publicidad de Datos Abiertos por el Tribunal Superior Electoral: el caso del Repositorio de Datos Electorales.

Finalmente, el último bloque de artículos de este número de la revista *Ciência da Informação* se centra en las métricas de información y el uso de estas métricas para comprender el avance de la investigación científica en Brasil.

EDITORIAL

Los artículos que componen este último bloque son: Medición de información científica en Web 2.0, Colecciones Culturales Brasileñas en el Repositorio de Wikimedia Commons, Perfil de orientaciones y producciones de mujeres con base en datos de la Plataforma Lattes, Un modelo poblacional para análisis de genealogía académica: Evidencia sobre crecimiento académico en Brasil y Estimación de la colaboración futura en datos de actividades científicas.

Tras la presentación de este número, se puede constatar que existe una gran sinergia entre las áreas de Data Science y Information Science. Es necesario profundizar en la comprensión de cómo estas áreas pueden mejorar y alimentarse para permitir que la información se entienda cada vez más como la materia prima de la ciencia. Creemos que los artículos aquí presentados contribuirán significativamente a este proceso.

¡Buena lectura!

André Luiz Appel

Doctor en Ciencias de la Información por el Instituto Brasileño de Información en Ciencia y Tecnología (Ibict) y Universidad Federal de Rio de Janeiro (PPGCI-IBICT / UFRJ) - Rio de Janeiro, RJ - Brasil. Asistente de investigación en el Instituto Brasileño de Información en Ciencia y Tecnología (Ibict) - Brasilia, DF - Brasil. Investigador del Laboratorio Interdisciplinario de Información y Conocimiento (Liinc-UFRJ / IBICT) - Brasil. Correo electrónico: andreappel@ibict.br

Ricardo Barros Sampaio

Postdoctorado en la Universidad de Brasilia (UnB) - DF - Brasil. Postdoctorado de la Fundación Oswaldo Cruz (FIOCRUZ) - Brasil. Doctorado en Ciencias de la Información por la Universidad de Brasilia (UnB) - DF - Brasil, con un período sándwich en Universidad de Toulouse - Francia. Profesor e investigador del Programa de Posgrado en Ciencias de la Información de la Universidad de Brasilia (UnB) - DF - Brasil. Profesor e investigador de la Maestría Profesional en Políticas de Salud Pública y especialización en Salud Pública de la Escuela de Gobierno Fiocruz - Brasilia, DF - Brasil. <http://lattes.cnpq.br/3477515781752110> Correo electrónico: rsampaio.br@gmail.com

Tiago Emmanuel Nunes Braga

Doctorado en Ciencias de la Información por la Universidad de Brasilia (UnB) - Brasilia, DF - Brasil. Coordinador General de Tecnologías de la Información e Informática (CGTI) e investigador del Instituto Brasileño de Información en Ciencia y Tecnología (Ibict) - Brasilia, DF - Brasil. Correo electrónico: tiagobraga@ibict.br