

Extração da informação e produção de conhecimento por meio da mineração de dados

Extraction of information and knowledge production through Data Mining

Durante a última década, testemunhamos avanços rápidos em Tecnologias da Informação e Comunicação (TICs) e, concomitantemente, o crescimento exponencial de produção de grandes quantidades de dados. Ambos propiciaram uma transformação na forma como lidamos com os dados e informações para um novo formato: o digital. No centro desta transformação digital, segundo Song e Zhu (2017), está a Ciência dos Dados, uma área da computação voltada para o estudo e a análise de dados, estruturados ou não. Paralelamente surgem novos conceitos e formas de tecnologias, tais como *data warehousing*, Internet das Coisas ou *Internet of Things* (IoT), Inteligência Artificial (IA), realidade virtual, *Data Mining*, entre muitas outras.

Essas tecnologias contribuíram para o desenvolvimento de muitas aplicações inovadoras em todos os setores da indústria e da sociedade. Com a evolução das tecnologias, as qualificações profissionais foram modificadas, bem como a forma com que as pessoas vivem, se informam e se comunicam. Mesmo aquelas que atuam nas organizações onde não precisam aprender aspectos técnicos imbricados nessas tecnologias, elas precisam entender os seus usos, para aproveitá-las em seu trabalho e em suas vidas.

Uma dessas tecnologias é a *Data Mining* ou mineração de dados. Vetrivel, Niyoperpasiline e Jeyanthirani (2017) definem mineração de dados como o processo de descoberta de informações a partir de grandes registros porque “é descrito como um método de extração e análise de padrões, relacionamentos e informações de grandes bancos de dados.” (VETRIVEL; NIYOPERPASILINE; JEYANTHIRANI, 2017, p. 80). Grandes quantidades de dados (*Big Date*) não são simplesmente caracterizadas por volume, mas por serem geradas continuamente, exaustivamente, assim são abrangentes e produzidos com condições de serem medidos ou classificados de acordo com uma escala (KITCHIN, 2014).

Seni e Elder (2010 apud KITCHIN, 2014) salientam que se trata de uma abordagem radicalmente diferente da tradicionalmente utilizada, em que o analista seleciona um método apropriado com base em seu conhecimento de técnicas e dados. Em outras palavras, a análise de *Big Data* possibilita uma abordagem epistemológica inteiramente nova para dar sentido ao mundo; em vez de testar uma teoria por meio da análise de dados relevantes, a nova análise de dados busca obter informações ou *insights* dos próprios dados, ou como assinala Kitchin (2014, p. 2, tradução nossa) “nascido dos dados”.

No que diz respeito ao método científico, Kitchin (2014) conclui que o acesso ao *Big Data* pode se configurar em novas práxis na pesquisa científica, porque desafia inclusive o que a ciência estabeleceu como abordagem dedutiva, já que apresenta a possibilidade de um novo paradigma de pesquisa na academia. Como exposto, a mineração ajuda a localizar padrões de informações preditivas a partir da exploração de dados, para pessoas interessadas no seu uso para a solução de suas expectativas. Assim, os dados são o que Fernandez-Molina (1994) chama de informação potencial. O objetivo da mineração de dados é encontrar informações extraídas do conjunto de dados na forma de estruturas compreensíveis pelo ser humano, de modo a possibilitar a criação do conhecimento humano e organizacional. Portanto, o conhecimento, em sentido amplo, não se confunde com a própria informação. Segundo Le Coadic (2004), a informação é o veículo do conhecimento.

Le Coadic (2004, p. 40) considera duas grandes classes de necessidades de informação, ambas derivadas daquelas que ele considera fundamentais: “a necessidade de informação em função do conhecimento e a necessidade de informação em função da ação”. A primeira resulta do desejo de saber, enquanto a outra resulta de necessidades materiais exigidos para a realização de atividades humanas, profissionais e pessoais. De qualquer forma, “a informação permanece sendo o meio de desencadear uma ação com objetivo; é a condição necessária à eficácia dessa ação.” (LE COADIC, 2004, p. 40).

Hjørland (2000) assinala que a informação não é, por si só, a motivação básica ou a meta para a pessoa que está a sua procura. A alimentação, um problema a resolver, ou um lugar seguro para residir podem ser o objetivo. Assim, a informação não é, geralmente, um fim em si mesma, mas algo que ajuda a orientar uma pessoa para um objetivo. É um fenômeno pragmático no sentido de que a informação para uma pessoa não precisa ser informação para outra. Ou seja, ela somente vai fazer sentido no contexto onde ela é produzida, disseminada e utilizada. O conhecimento, fruto da relação pessoal com a informação, é, nesse caso, subjetivo (inerente a cada pessoa), mas ao mesmo tempo social, pois o ser humano interage com o mundo que o circunda, no domínio que ele atua, como defendem Capurro e Hjørland (2007).

Big Data e *Data Mining* poderiam ser vistos, portanto, como fenômenos ou tecnologias da sociedade contemporânea, que permitem uma nova forma de produzir conhecimento. Embora a produção dos dados e sua mineração não sejam por si somente o fazer científico, ainda que por trás existe tanto tecnologia quanto uma ciência que os permite existir, eles alteram profundamente a forma de produção de conhecimento. Estamos, por assim dizer, diante de uma ciência aplicada que permite desmitificar que haja uma dicotomia teoria *versus* prática para um novo fazer teórico-prático ou prático-teórico.

Esperamos, na revista Navus, que tenhamos artigos que desafiem os paradigmas científicos vigentes e tragam à tona questões como estas abordadas neste editorial. Para este volume 1 número 8, trazemos 10 artigos, de temáticas diversas como: o método *Getting Things Done*, qualidade de vida no trabalho, gestão educacional, aprendizagem organizacional e inovação, gestão de Tecnologias da Informação, associativismo em mercado varejista, relacionamento de clientes da telefonia celular, comportamento do consumidor, gestão do conhecimento e seleção de pregoeiros em instituição pública.

Boa Leitura!

Nadi Helena Presser e Eli Lopes da Silva
Editores

REFERÊNCIAS

CAPURRO, R.; HJORLAND, B. O conceito de informação. **Perspectivas em Ciência da Informação**, v. 12, n. 1, p. 148-207, jan./abr. 2007.

FERNANDEZ-MOLINA, J. C. Enfoques objetivo y subjetivo del concepto de información. **Revista Española de Documentación Científica**, v. 17, n. 3, p. 320-331, 1994.

HJORLAND, B. Information seeking behaviour: what should a general theory look like? **The New Review of Information Behaviour Research**, v. 1, Issue December, p. 19-34, 2000.

LE COADIC, Y. A ciência da informação. 2. ed. Brasília, DF: Briquet de Lemos Livros, 2004.

KITCHIN, R. Big Data, new epistemologies and paradigm shifts. **Big Data & Society**, v. 1, n. 1, p. 1-12, Apr./June 2014.

SONG, I. Y.; ZHU, Y. Big Data and Data Science: Opportunities and Challenges of iSchools. **National Science Foundation**, v. 2, n. 3, p. 1-18, 2017.

VETRIVEL, R.S.; NIYOPERPASILINE, I.I.J.; JEYANTHIRANI I.I.I.P. Investigation Study on Vantage Point Indexing Search in Multiple Data Pattern Mining. **International Journal of Advanced Research in Computer Science & Technology (IJARCST)**, v. 5, n. 1, Jan./Mar. 2017.