

(Editorial)**Data Science and Business Review: A necessary outlet to an emerging field****Pedro Jácome de Moura Jr.**Universidade Federal da Paraíba – UFPB – Brasil
pjacome@sti.ufpb.br
ORCID: 0000-0001-6548-4614

Data, a single word with multiple meanings. Meaningless as well, when it comes as mere abstractions from reality. A given thing (from Latin *datum*). Curiosity: currently spoken languages derived from Latin, such as Spanish and Portuguese, use the word "dado" interchangeably, both as data and also as a given thing. One can say "dado é dado" in Portuguese, meaning "data is given", given as the occurrence of any ordinary phenomena which, by it turns, is the most specially attractive thing for scientific inquiring. Data as a phenomena, therefore, resembles to a sparkle, lighting up the DSBR's perspective on data science: a science of data.

A science of data must to acquire knowledge on data (not just through data). Data as a phenomena poses intrinsic properties and behavior to be described, explained, predicted and prescribed. Data properties and behavior express themselves in specific contexts, and organizations and businesses represent our contexts of interest. In the organizational/business context, data emerges also as phenomena subject to processes such as analytics, management, and governance. That's why DSBR topics are directly related to business problems and solutions in Data Science; innovations in the use of data to extract useful business knowledge; machine learning; artificial intelligence; engineering and data-driven decision making; data management (structured, unstructured, distributed, cloud); data quality; data governance; training of data analysts and scientists; human, social and cultural factors and their influence on Data Science processes for business; management of professionals and teams in Data Science projects; and measurement theory, metrics and KPI for business.

In this first issue, we received contributions from eight authors: Francisco José da Costa and Jorge Dias (Management, UFPB), Sílvio Meira (Computer Science, UFPE), and Jorge Viana, Aléssio de Almeida, Antonio Barbosa, Hilton Ramalho e Ignácio Araújo Júnior (Economics, UFPB). The papers they give us have a common subject: data governance within Data Science scope. They have assumed data as the phenomena of Data Science, an assumption in the opposite direction of looking at that science just as a set of analytical techniques with scientific inspiration.

As you, dear reader, will see, Francisco José da Costa analyzes the convergence between management, technology and data analysis. His study contributes by offering "a necessary perspective to the already existing discussions on organizational and training projects in Data Science". With an emphasis on specific academic training for professional performance in the area, Costa proposes that the Chief Data Science Officer (CDSO) masters technical and organizational leadership skills. Among the roles of this new professional, the one that "articulates the internal data policy and interactions with the various agents in the 'data field' (suppliers, regulators)" stands out. In other words, the CDSO will also be involved with data governance aspects.

Jorge Dias contemplates his own experience in academic management in higher education institutions (HEIs), especially his role at the head of the unit responsible for data analysis and decision support in an undergraduate dean's office. Jorge Dias discusses the concept of Academic Analytics, presents the current stage and the challenges faced to consolidate the theme. The lived experience offers the author different perspectives on data management and governance practices. Especially in relation to governance, Dias exposes considerable challenges, expressed in questions such as "who can request the data? From whom should data be requested? What data can be provided? What data should be protected?". The answers to such questions, as the author emphasizes, must be thought in the light of specific rules on data security and privacy, which denotes the leading role of governance when practicing Data Science.

Sílvio Meira, in turn, proposes that organizations focus efforts on data security and privacy, especially when it comes to personal data. Meira compares data to heavy minerals – uranium, for example –, "[d]ata has to be mined from its sources (which includes customers and users), refined to separate what is wanted from what is not, have to reach critical mass to generate energy (i.e. value for the company and customers) and... after that, their storage or disposal can be harmful for the business,

the customer and the ecosystem". Meira's concern with the entire data management cycle also implies a concern with the governance of these same data, as institutions need to demonstrate that they are protecting people from the action of their algorithms.

Finally, Jorge Viana, Aléssio de Almeida, Antonio Barbosa, Hilton Ramalho and Ignácio Araújo Júnior demonstrate how they built indicators for monitoring the Sustainable Development Goals (SDGs) in Paraíba. The study contributes to the identification of bottlenecks, as well as successful cases, in the state's sustainable development agenda, in addition to offering a comparative benchmark with other Brazilian states. As it offers an "instrument for social monitoring of the effectiveness of implemented public policies", the study by Viana and his colleagues is part of the institutional governance aspect, grounded on aspects of data governance.

We hope that this first DSBR issue will be followed by regular publication. As we said in the DSBR website, "Data Science is an essentially interdisciplinary field. DSBR recognizes this interdisciplinarity and accepts original contributions using theoretical, methodological or conceptual contributions from different fields. In particular, studies that mix contributions from more than one field are preferred." Therefore, dear reader, you are welcome to contribute as an author, reviewer or editor as well. Data Science's success depends on the engagement of its community.

Data Science and Business Review: Um meio de divulgação necessário para um campo acadêmico emergente

Dado, uma palavra com múltiplos significados, inclusive incapaz de transmitir significado, quando se trata de mera abstração da realidade. Uma coisa dada (do latim *datum*). Uma curiosidade: algumas línguas faladas atualmente, derivadas do latim, como o espanhol e o português, utilizam a palavra "dado" tanto como uma coisa dada/doada quanto como fato sobre a realidade. Pode-se dizer "dado é dado" em português, significando "dado é doado", doado como a ocorrência de qualquer fenômeno ordinário que, por sua vez, é a coisa mais especial e atraente para qualquer investigação científica. O dado como fenômeno, portanto, se assemelha a uma centelha que acende a perspectiva da DSBR sobre a ciência de dados: uma ciência dos dados.

Uma ciência dos dados deve adquirir conhecimento sobre dados (não apenas a partir de dados). Dado como fenômeno apresenta propriedades e comportamentos intrínsecos, que devem ser descritos, explicados, previstos e prescritos. As propriedades e o comportamento do dado se expressam em contextos específicos, e as organizações representam nosso contexto de interesse. No contexto organizacional o dado surge também como fenômeno sujeito a processos de análises, gestão e governança. É por isso que os tópicos da DSBR estão diretamente relacionados aos problemas e soluções de negócios em Ciência de Dados; inovações no uso de dados para extração de conhecimento útil para negócios; aprendizado de máquina; inteligência artificial; engenharia e tomada de decisão baseada em dados; gerenciamento de dados (estruturados, não estruturados, distribuídos, na nuvem); qualidade de dados; gestão de dados; treinamento de analistas e cientistas de dados; fatores humanos, sociais e culturais e sua influência nos processos de ciência de dados para negócios; gestão de profissionais e equipes em projetos de ciência de dados; e teoria da medição, métricas e indicadores para negócios.

Neste primeiro número, recebemos contribuições de oito autores: Francisco José da Costa e Jorge Dias (Administração, UFPB), Sílvio Meira (Ciência da Computação, UFPE), e Jorge Viana, Aléssio de Almeida, Antonio Barbosa, Hilton Ramalho e Ignácio Araújo Júnior (Economia, UFPB). Os artigos que eles nos trazem têm um tema central em comum: governança de dados no âmbito da ciência de dados. Esses autores assumiram o dado como fenômeno da ciência de dados, uma suposição que vai na direção oposta de olhar para esta ciência apenas como um conjunto de técnicas analíticas com inspiração científica.

Como verá, prezado leitor, Francisco José da Costa analisa a convergência entre gestão, tecnologia e análise de dados gerenciais. Seu estudo contribui ao oferecer "perspectiva necessária às discussões já existentes sobre projetos organizacionais e formativos em Ciência de Dados". Com ênfase na formação acadêmica específica para atuação profissional na área, Costa propõe que o *Chief Data Science Officer* (CDSO) possua competências técnicas e de liderança organizacional. Destaca-se dentre os papéis deste novo profissional aquele que "articula a política interna de dados e as interações com os diversos agentes do 'campo de dados' (fornecedores, reguladores, demandantes)". Ou seja, o CDSO estará também envolvido com aspectos de governança de dados.

Jorge Dias contempla a sua própria experiência na gestão acadêmica em instituições de ensino superior (IES), especialmente a sua atuação à frente de unidade responsável pela análise de dados e

suporte à decisão em uma pró-reitoria de graduação. Jorge Dias discute o conceito de Academic Analytics, apresenta o estágio atual e os desafios enfrentados para consolidação do tema. A experiência vivenciada oferece ao autor perspectivas diferenciadas das práticas de gestão e governança de dados. Especialmente em relação à governança, Jorge Dias expõe desafios consideráveis, expressos em perguntas do tipo “quem pode solicitar os dados? A quem os dados devem ser solicitados? Quais dados podem ser fornecidos? Quais dados devem ser protegidos?”. As respostas a tais questionamentos, conforme enfatiza o autor, devem ser elaboradas à luz de regramentos específicos sobre segurança e privacidade de dados, o que denota o protagonismo do tema governança quando se pratica ciência de dados.

Sílvio Meira, por sua vez, propõe às organizações que concentrem esforços em segurança e privacidade de dados, especialmente quando se trata de dados pessoais. Meira compara dados a minerais pesados – urânio, por exemplo –, “[d]ata has to be mined from its sources (which includes customers and users), refined to separate what is wanted from what is not, have to reach critical mass to generate energy (i.e. value for the company and customers) and... after that, their storage or disposal can be harmful for the business, the customer and the ecosystem”. A preocupação de Meira com todo o ciclo de gestão de dados implica também uma preocupação com governança destes mesmos dados, já que as instituições precisam demonstrar que estão protegendo as pessoas da ação dos seus algoritmos.

Por fim, Jorge Viana, Aléssio de Almeida, Antonio Barbosa, Hilton Ramalho e Ignácio Araújo Júnior demonstram como construíram indicadores para monitoramento dos Objetivos do Desenvolvimento Sustentável (ODS) na Paraíba. O estudo contribui para a identificação de gargalos, bem como de casos de sucesso, na agenda de desenvolvimento sustentável do estado, além de oferecer *benchmark* comparativo com outros estados brasileiros. Por oferecer um “instrumento de monitoramento social da efetividade das políticas públicas implementadas”, o estudo de Viana e seus colegas se insere na vertente de governança institucional, fundamentado em aspectos de governança de dados.

Esperamos que esta edição inaugural da DSBR seja seguida por muitas outras publicações regulares. Como dissemos no website DSBR, “Ciência de dados é um campo essencialmente interdisciplinar. A DSBR reconhece essa interdisciplinaridade e aceita contribuições originais teóricas, metodológicas ou conceituais de diferentes campos. Em particular, os estudos que combinam contribuições de mais de um campo são preferidos”. Portanto, caro leitor, você está convidado a contribuir com a DSBR como autor, revisor ou editor também. O sucesso da ciência de dados depende do pleno envolvimento de sua comunidade.