

Garantindo imparcialidade, precisão, confidencialidade e transparência aos dados na perspectiva da Ciência dos Dados Responsáveis

Morgana Andrade¹, Paula Amorim², Caio Saraiva Coneglian³, Decio Wey Berti Junior⁴
Ana Alice Baptista⁵

¹ <https://orcid.org/0000-0002-4055-5882>

Universidade Federal do Espírito Santo, Vitória, Espírito Santo
morganaandrade@hotmail.com

² <https://orcid.org/0000-0002-5480-4106>

Universidade Federal do Espírito Santo, Vitória, Espírito Santo.
paulaventuram2orim@gmail.com

³ <https://orcid.org/0000-0002-6126-9113>

Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP, Marília, Brasil
caio.coneglian@gmail.com

⁴ <http://orcid.org/0000-0003-4313-2727>

Universidade Estadual de Londrina-UEL, Londrina, Paraná, Brasil.
deciowbj@gmail.com

⁵ <http://orcid.org/0000-0003-3525-0619>

Centro Algoritmi – Universidade do Minho, Guimarães, Portugal
analice@dsi.uminho.pt

Resumo

A demanda por acesso e análise de dados, sejam privados ou públicos, que promovem a tomada de decisão comercial e a ciência, tem impulsionado a Economia e os avanços científicos, de forma a atrair a atenção de vários segmentos da sociedade. No contexto Big Data, surge, como necessidade urgente, a aplicação de direitos individuais e empresariais e de normas regulatórias que resguardem a privacidade, a imparcialidade, a precisão e a transparência. Este artigo aborda possíveis alternativas que podem assegurar a aplicação da ética, bem como a regulamentação para acesso e reuso de dados, a Responsible Data Science (Ciência de Dados Responsáveis). Nesse cenário, a Responsible Data Science desponta como uma iniciativa que tem como base as diretrizes FACT, que correspondem à

adoção de quatro princípios: imparcialidade, precisão, confidencialidade e transparência. Para a implementação dessas diretrizes, deve-se considerar o uso de técnicas e abordagens que estão sendo desenvolvidas pela Green Data Science, ou seja, que preconiza o uso de mecanismos desde a fase inicial da análise à disseminação dos dados. Nesse sentido, esse estudo explora os conceitos que norteiam as diretrizes FACT, suas orientações e suas iniciativas para a efetivação de sua adoção. Para tanto, foi desenvolvida pesquisa exploratória e descritiva com abordagem qualitativa. Na busca por literatura sobre os temas, foram realizadas buscas nas bases de dados bibliográficas Web of Science, Scopus e pelo motor de busca Scholar.Google. Foram adotados os termos “Responsible Data Science”, “Fairness, accuracy, confidentiality, transparency + Data science”, FACT e FAT relacionados com Data Science. Após a análise dos documentos selecionados, concluiu-se que a Green Data Science e as diretrizes FACT contribuem significativamente para a salvaguarda dos direitos individuais, não sendo necessário recorrer a medidas que impeçam o acesso e a reutilização de dados. Tanto a iniciativa Responsible Data Science é vista pelos autores como boas práticas no contexto do Big Data, principalmente pela possibilidade de reuso dos dados, um dos fundamentos das diretrizes FAIR - Findable, Accessible, Interoperable e Reusable, que têm se destacado nas comunidades acadêmicas e científicas em relação aos dados disponibilizados na Web. No entanto, os desafios para implementar as diretrizes FACT requerem estudos, condições *sine qua non* para que as ferramentas para análise e disseminação dos dados sejam desenvolvidas ainda na fase de design, ou seja, na concepção de metodologias que garantam a imparcialidade, a precisão, a transparência e a confidencialidade. Caso contrário, a sua não efetivação poderá resultar em consequências indesejáveis para a sociedade.