Título do Tema: Gerência, Análise e Qualidade de Dados em Larga Escala

Orientador: Fábio José Coutinho da Silva

Resumo:
O avanço das tecnologias de informação e comunicação tem impulsionado um crescente aumento do volume de dados processados por empresas, governos, organizações e instituições de pesquisa. A geração contínua e acelerada de grandes volumes de dados provenientes de redes sociais, dispositivos IoT, sistemas geográficos, serviços de saúde e redes de sensores tem se consolidado como uma característica marcante da sociedade contemporânea. Nesse sentido, esta pesquisa volta-se ao desafio de lidar eficientemente com volumes massivos de dados heterogêneos produzidos por pessoas, dispositivos e sistemas em diversos domínios de aplicação. Esse desafio inclui manter, gerenciar, integrar, analisar e compartilhar dados de forma ágil e eficiente.

Para tanto, busca-se investigar estratégias, métodos e frameworks voltados ao processamento, integração e análise de dados em larga escala, com ênfase na avaliação, monitoramento e garantia da qualidade dos dados ao longo de todo o seu ciclo de vida. Destaca-se, nesse contexto, a investigação qualidade de dados espaciais, considerando aspectos como completude, consistência, atualidade e precisão posicional, bem como à qualidade de dados empregados em pipelines de Inteligência Artificial, abrangendo atividades de preparação, integração, validação e monitoramento de dados utilizados no treinamento, avaliação e operação de modelos. Esse processo envolve múltiplas facetas e visões, abrangendo diferentes abordagens, técnicas e tecnologias aplicadas em diversos domínios.

Referências:

- Wang, J., Xu, C., Zhang, J., & Zhong, R. (2022). Big data analytics for intelligent manufacturing systems: A review. *Journal of Manufacturing Systems*, *62*, 738-752. https://doi.org/10.1016/j.jmsy.2021.03.005
- Clarindo, J. P., Silva, R. L. L., Alves, J. M. S., Castro, J. P. C., Coutinho, F. J. and Aguiar, C. D. (2026). Empowering Large-Scale Spatial Analytics in Smart Cities with a Fog-Cloud Lakehouse Architecture. In Proceedings of the *28th International Conference on Enterprise Information Systems* - Volume 1: ICEIS; ISBN 978-989-758-834-1; ISSN 2184-4992, SciTePress, pages 121-132. DOI: 10.5220/0014971500004018
- Naumann, F., Ehrlinger, L., Harmouch, H., Mohammed, S., Srivastava, D. (2026). Data Quality in the Age of AI. In: Chrysanthis, P.K., Nørvåg, K., Stefanidis, K., Zhang, Z. (eds) Advances in Databases and Information Systems. ADBIS 2025. Lecture Notes in Computer Science, vol 16043. Springer, Cham. https://doi.org/10.1007/978-3-032-05281-0_1
- SILVA, Rafael Luciano L.; CORREIA, Diêgo de A.; MENDES, Letícia A.; DE MELO, Ruan T.; COUTINHO, Fabio J.. QualiBus: implementando métricas de qualidade para dados de transporte coletivo. In: *SIMPÓSIO BRASILEIRO DE BANCO DE DADOS (SBBD)*, 40. , 2025, Fortaleza/CE. Anais [...]. Porto Alegre: Sociedade

Brasileira de Computação, 2025 . p. 589-601. ISSN 2763-8979. DOI: https://doi.org/10.5753/sbbd.2025.247287

- Almeida, A., Brás, S., Sargento, S. et al. Time series big data: a survey on data stream frameworks, analysis and algorithms. J Big Data 10, 83 (2023). https://doi.org/10.1186/s40537-023-00760-1
- Inoubli, W., Aridhi, S., Mezni, H., Maddouri, M., & Mephu Nguifo, E. (2018). An experimental survey on big data frameworks. *Future Generation Computer Systems*, *86*, 546-564. https://doi.org/10.1016/j.future.2018.04.032
- R. Vidgen, S. Shaw and D. B. Grant, Management challenges in creating value from business analytics, European Journal of Operational Research 261 (2017) 626–639.
- S. Bahri, N. Zoghlami, M. Abed and J. M. R. S. Tavares, Big Data for healthcare: A survey, IEEE Access 7 (2019) 7397–7408.