



40º Simpósio Brasileiro de Bancos de Dados
40th Brazilian Symposium on Databases



SBBD
2 0 2 5



Anais SBBD 2025
Proceedings SBBD 2025

De 29 de setembro a 2 de outubro de 2025
September 29th to October 2nd, 2025





Organização do SBBD 2025

SBBD 2025 Organization

Program Chair

Carlos Eduardo Santos Pires (UFCG, Brasil)

Short Papers Chair

Elaine Parros M. de Sousa (USP, Brasil)

Demos and Applications Chair

Luiz Celso Gomes Jr (UTFPR, Brasil)

WTDBD Chair

Maria Camila Nardini Barioni (UFU, Brasil)

CTDBD Chair

Valeria Cesário Times (UFPE, Brasil)

Short Courses Chair

Dênio Duarte (UFFS, Brasil)

Tutorials Chair

Eduardo Cunha de Almeida (UFPR, Brasil)

Workshops Chair

Leonardo Andrade Ribeiro (UFG, Brasil)

WTAG Chair

Alessandreia Marta de Oliveira (UFJF, Brasil)

Proceedings Chair

Michele A. Brandão (UFMG, Brasil)

Marathon Chair

Arlino H. M. de Araújo (UFPI, Brasil)



Organização Geral *General Organization*

Chairs

Ticiania L. Coelho da Silva (UFC, Brasil)

Regis Pires Magalhães (UFC, Brasil)

Lívia Almada Cruz (UFC, Brasil)

José Antônio Fernandes de Macêdo (UFC, Brasil)



Os Anais do XL Simpósio Brasileiro de Bancos de Dados (SBBD 2025) incluem os artigos selecionados e apresentados na edição do evento realizado de 29 de setembro a 2 de outubro de 2025. Nesta edição, os anais incluem 54 artigos completos e 36 artigos curtos. Esses trabalhos foram selecionados através de um processo de revisão por pares do tipo duplo-anônimo, de um total de 135 artigos completos e 88 artigos curtos submetidos, o que implicou em taxas de aceitação de 40% (54/135) e 41% (36/88), respectivamente. Um artigo, originalmente submetido para a trilha de artigos completos, foi aceito como artigo curto, e um artigo originalmente submetido como artigo curto foi aceito na trilha de demos e aplicações.

Tanto os artigos completos quanto os artigos curtos incluem trabalhos das categorias pesquisa, indústria e visão, em diversas áreas de bancos de dados, ciência de dados e *big data*. O comitê de programa de Artigos Completos foi coordenado pelo Prof. Carlos Eduardo Santos Pires (UFCG, Brasil) e o referente aos Artigos Curtos foi coordenado pela Profa. Elaine Parros M. de Sousa (USP, Brasil), que organizaram este volume.

Os artigos completos que integram os anais foram submetidos até 05/05/2025, com aceitação em 11/06/2025 (sem *rebuttal*) e 25/06/2025 (com *rebuttal*). A versão final foi submetida em 15/07/2025. Os artigos curtos, por sua vez, foram submetidos até 20/06/2025, aceitos para publicação em 28/07/2025, e tiveram a versão final submetida em 02/08/2025.

A qualidade desta edição é devida essencialmente aos autores e revisores dos artigos publicados. Expressamos nossos fortes agradecimentos pelas contribuições e discussões durante o SBBD 2025. Para mais informações sobre o SBBD 2025, visite o site desta edição do evento em <https://sbbd.org.br/2025>.



The Proceedings of the 40th Brazilian Symposium on Databases (SBBD 2025) include the papers selected and presented at the event from September 29th to October 2nd, 2025. In this edition, the proceedings include 54 full papers and 36 short papers. These papers were selected through a double-anonymous peer review process, with 135 full papers and 88 short papers submitted, resulting in acceptance rates of 40% (54/135) and 41% (36/88), respectively. One paper originally submitted to the full paper track was accepted as a short paper, and one paper originally submitted as a short paper was accepted to the demos and applications track.

Both full and short papers include works from the categories of research, industry, and vision, in various areas of databases, data science, and big data. The Full Papers Program Committee was coordinated by Prof. Carlos Eduardo Santos Pires (UFCEG, Brazil), and the Short Papers Program Committee was coordinated by Prof. Elaine Parros M. de Sousa (USP, Brazil), who organized this volume.

The full papers that comprise these proceedings were submitted by 05/05/2025 and accepted on 06/11/2025 (without rebuttal) and 06/25/2025 (with rebuttal). The final version of all papers was submitted on 07/15/2025. The short papers, in turn, were submitted by 06/20/2025, accepted for publication on 07/28/2025, and the final version was submitted on 08/02/2025.

The quality of this edition is primarily due to the authors and reviewers of the published papers. We are deeply grateful for their contributions and discussions during SBBD 2025. For more information about SBBD 2025, please visit the event website at <https://sbbd.org.br/2025>.

Carlos Eduardo Santos Pires
(UFCEG, Brazil)

Elaine Parros M. Souza
(USP, Brazil)





Program Committee of Full Papers

- *Agma Traina (ICMC-USP)*
- *Alessandreia Oliveira (Universidade Federal de Juiz de Fora)*
- *Alex Sandro da Cunha Rêgo (Instituto Federal da Paraíba)*
- *Altigran Soares da Silva (Universidade Federal do Amazonas)*
- *Ana Carolina Almeida (Universidade do Estado do Rio de Janeiro)*
- *Ana Carolina Salgado (UFPE)*
- *André Santanchè (UNICAMP)*
- *Andreza Alencar (UFRPE)*
- *Angelo Brayner (Federal University of Ceara – UFC)*
- *Angelo Frozza IFC (Instituto Federal Catarinense – Campus Camboriú)*
- *Arlino Magalhães (Universidade Federal do Piauí)*
- *Bernadette Loscio (TNO)*
- *Caetano Traina Júnior (ICMC – USP)*
- *Carina F. Dorneles (Universidade Federal de Santa Catarina)*
- *Carmem Hara (Universidade Federal do Paraná)*
- *Celso Hirata (ITA)*
- *Claudio Campelo (Universidade Federal de Campina Grande)*
- *Clodoveu Davis (UFMG)*
- *Cristina Aguiar (Universidade de São Paulo)*
- *Damires Souza (Instituto Federal da Paraíba – IFPB)*
- *Daniel de Oliveira (Universidade Federal Fluminense)*
- *Daniel Kaster (UEL – Universidade Estadual de Londrina)*
- *Daniela Barreiro Claro (Federal University of Bahia)*
- *Débora Pina (COPPE/UFRJ)*
- *Demetrio Mestre (Universidade Federal de Campina Grande)*
- *Denio Duarte (UFFS)*
- *Diego Carvalho (CEFET-RJ)*
- *Diego Pessoa (IFPB)*
- *Dimas Cassimiro Nascimento (UFAPE)*
- *Edleno Moura (Federal University of Amazonas)*
- *Eduardo Bezerra (CEFET/RJ)*
- *Eduardo de Almeida (Universidade Federal do Paraná)*
- *Eduardo Ogasawara (CEFET/RJ)*
- *Eduardo Pena Federal (University of Technology – UTFPR)*
- *Elaine Sousa (University of Sao Paulo)*
- *Fabio Porto (LNCC)*
- *Felipe Domingos da Cunha (PUC-Minas)*



- *Felipe Louza (Universidade Federal de Uberlândia)*
- *Gabriella Costa (Federal Center for Technological Education of Minas Gerais)*
- *Genoveva Vargas-Solar (French Council on Scientific Research)*
- *Geomar Schreiner (UFFS)*
- *Humberto Razente (Universidade Federal de Uberlândia)*
- *Janderson Aguiar (Universidade Estadual da Paraíba – UEPB)*
- *Javam Machado Federal (University of Ceará)*
- *João Eduardo Ferreira (IME/USP)*
- *João Felipe Pimentel (Universidade Federal Fluminense)*
- *José Monteiro Universidade Federal do Ceará (UFC)*
- *José Palazzo Moreira de Oliveira (UFRGS)*
- *José Antonio Macêdo (UFC)*
- *Jose Rodrigues Jr (University of São Paulo)*
- *Julio dos Reis (Instituto de Computação da Universidade Estadual de Campinas)*
- *Julio Cesar Duarte (Instituto Militar de Engenharia)*
- *Karin Becker (UFRGS)*
- *Kelly Braghetto (IME/USP)*
- *Larissa Shimomura (Hasselt University)*
- *Leonardo Andrade Ribeiro (Federal University of Goiás)*
- *Liliane Kunstmann (COPPE/UFRJ)*
- *Luiz Celso Gomes Jr (UTFPR)*
- *Luiz Gustavo Dias (UFF)*
- *Marco Antonio Casanova (PUC-Rio)*
- *Marcos Bedo (Universidade Federal Fluminense)*
- *Marcos Gonçalves (Universidade Federal de Minas Gerais)*
- *Maria Camila Nardini Barioni (UFU)*
- *Maristela Holanda (Universidade de Brasília)*
- *Mariza Ferro (Universidade Federal Fluminense)*
- *Michele A. Brandão (UFMG)*
- *Mirela Cazzolato (University of São Paulo)*
- *Mirella M. Moro (Universidade Federal de Minas Gerais)*
- *Mirian Halfeld-Ferrari (Université d'Orleans – LIFO)*
- *Nadia Kozievitch (UTFPR)*
- *Rafaelli Coutinho (CEFET/RJ)*
- *Rebecca Salles (INRIA)*
- *Regina Braga (Universidade Federal de Juiz de Fora)*
- *Renata Galante (UFRGS)*
- *Renato Fileto (UFSC)*
- *Ricardo Oliveira (Universidade Federal de Campina Grande)*
- *Ricardo Torres (Wageningen University and Research)*



- *Robson Fidalgo (Universidade Federal de Pernambuco)*
- *Rogério Luis Costa (Polytechnic University of Leiria)*
- *Ronaldo Goldschmidt (Instituto Militar de Engenharia)*
- *Ronaldo Mello (Universidade Federal de Santa Catarina)*
- *Sergio Lifschitz (PUC-Rio)*
- *Simone Dominico (UFPR)*
- *Thiago Moura (IFPB)*
- *Thiago Nóbrega (UFCG)*
- *Tiago Brasileiro Araújo (IFPB)*
- *Tiago de Melo (Universidade do Estado do Amazonas)*
- *Valéria C. Times UFPE (Universidade Federal de Pernambuco)*
- *Vanessa Lago Machado (Instituto Federal Sul-rio-grandense)*
- *Victor Aguiar Evangelista de Farias (Universidade Federal do Ceará)*
- *Wagner Meira Jr. (UFMG)*
- *Zoubida Kedad (Université de Versailles Paris Saclay)*

External reviewers:

- *Daniel Saad Nogueira Nunes (Instituto Federal de Brasília)*
- *Felipe Timbó Brito (Federal University of Ceará)*
- *Iago Chaves (Federal University of Ceará)*
- *Nicolas Hiot (Université d'Orléans)*



Program Committee of Short Papers

- Ana Carolina Almeida (UERJ)
- Anderson Chaves Carniel (Huawei Technologies R&D – UK)
- André Souza (WiKKi Soluções em Engenharia/ICMC-USP)
- Andreza Alencar (UFRPE)
- Carina F. Dorneles (UFSC)
- Carmem Hara (UFPR)
- Claudio Campelo (UFCG)
- Cristina Aguiar (ICMC-USP)
- Damires Souza (IFPB)
- Daniel de Oliveira (UFF)
- Daniel Kaster (UEL)
- Demetrio Mestre (UFCG)
- Denio Duarte (UFFS)
- Edimar Manica (IFRS)
- Eduardo Almeida (UFPR)
- Eduardo Ogasawara (CEFET-RJ)
- Eduardo Pena (UTFPR)
- Fabio Lobato (UFOPA)
- Fabio Porto (LNCC)
- Fabrício Pereira (UNIRIO)
- Felipe Brito (UFC)
- Humberto Razente (UFU)
- Jorge Soares (Cefet-RJ)
- José Monteiro (UFC)
- Karin Becker (UFRGS)
- Kelly Braghetto (IME-USP)
- Levy Souza (UFPE)
- Lucas Valem (ICMC-USP)
- Lucio Fernandes Dutra Santos (IFNMG)
- Luiz Celso Gomes Jr (UTFPR)
- Luiz Olmes Carvalho (UNIFEI)
- Marcela Ribeiro (UFSCar)
- Marcelo Reis (UNICAMP)
- Marcos Arrais (PUC-MG)
- Marcos Bedo (UFF)
- Maria Camila Nardini Barioni (UFU)
- Michele A. Brandão (UFMG)



- *Mirela Cazzolato (ICMC-USP)*
- *Pedro Bugatti (UTFPR)*
- *Rafael Pereira de Oliveira (Aditum)*
- *Rafaelli Coutinho (CEFET-RJ)*
- *Raquelina Penteado (UEM)*
- *Renan Souza (Oak Ridge National Laboratory)*
- *Renata Galante (UFRGS)*
- *Renato Marroquin (Oracle)*
- *Rocio Zorrilla (LNCC)*
- *Rogério Luis Costa (Polytechnic University of Leiria)*
- *Ronaldo Mello (UFSC)*
- *Sergio Lifschitz (PUC-RJ)*
- *Sergio Mergen (UFSM)*
- *Tarlis Tortelli Portela (UFSC)*
- *Thiago Nóbrega (UFCEG)*
- *Thiago Henrique Silva (UTFPR)*
- *Tiago Brasileiro Araújo (IFPB)*
- *Valéria C. Times (UFPE)*
- *Vinicius Dias (UFLA)*

External reviewer:

- *Afonso Matheus Sousa Lima (USP)*

FULL

Aprendizado Federado Incremental e Sensível ao Risco para Modelos de Ranqueamento em Cenários com Distribuições Heterogêneas de Dados	1	sbbd:01
<i>Gestefane Rabbi, Celso França, Daniel Xavier Sousa, Thierson Couto Rosa, Jussara M. Almeida, Marcos André Gonçalves</i>		
Explorando uma Nova Métrica para Calcular a Precisão Acurada de Resultados de Blocação em Tarefas de Resolução de Entidades	15	sbbd:02
<i>Dimas Cassimiro Nascimento, Vítor Alan Bezerra Silva</i>		
Enhancing ML Explainability with Multi-Agent LLMs: A Context-Aware XAI Approach	28	sbbd:03
<i>Renato Okabayashi Miyaji, Pedro Luiz Pizzigatti Corrêa</i>		
Uma Arquitetura para Big Data Geoespacial de Dispositivos IoT com Níveis Heterogêneos de Mobilidade	42	sbbd:04
<i>Dennis Sávio Martins Silva, Maristela Holanda</i>		
Dataflow Analysis of Serverless Scientific Applications using Provenance Data ..	56	sbbd:05
<i>Marcello Willians Messina Ribeiro, Uiratam De Paula, Liliane Kunstmann, Yuri Frota, Isabel Rosseti, Daniel de Oliveira</i>		
Uso de Shadow Pipelines para Garantir a Adequação de Pipelines ETL à LGPD ..	70	sbbd:06
<i>Raíza Albuquerque, Marcos Bedo, José Maria da Silva Monteiro Filho, Lenora Schwaitzer, Daniel de Oliveira</i>		
Detecting Anomalous Vessel Trajectories: A Collaborative Clustering-Based Approach	84	sbbd:07
<i>Cláudio Ribeiro, Marcos Bedo, Ronaldo dos Santos Mello, Aline Paes, Daniel de Oliveira</i>		
Exploiting Surrogate Submodular and Cost-Effective Lazy Forward Algorithms for Calibrated Recommendations	98	sbbd:08
<i>Diego Correa da Silva, Joel Machado Pires, Frederico Araújo Durão</i>		
Analysis of online and offline classification algorithms for human activity recognition using IMU sensors	112	sbbd:09
<i>Brena Rodrigues Machado, Regis Magalhães, Livia Almada Cruz, Criston P. de Souza, José Macedo, César Lincoln C. Mattos</i>		
Heuristic-Guided Text-to-SQL Translation with LLMs: Optimizing Natural Language Interfaces for Relational Databases	126	sbbd:10
<i>Laura Petrola, Angelo Brayner, Wellington Franco</i>		
Leis de Escala para Text-to-SQL: Um Estudo sobre a Relação entre Tamanho e Desempenho de Modelos de Linguagem	140	sbbd:11
<i>Letícia Oliveira Silva, Paulo Henrique Carneiro Silva, Fabrício Aguiar Silva</i>		
Desigualdades Educacionais no Brasil: Uma Análise por Clusterização de Indicadores Educacionais e de Desempenho Escolar	154	sbbd:12
<i>Matheus Leandro de Melo Silva, Livia Almada Cruz, Regis Magalhães, Tatieures Gomes Pires, José Macedo, Rossana M. C. Andrade</i>		
Mineração de Dados para Caracterizar Indivíduos Hipertensos com Doenças Cardiovasculares no Brasil	168	sbbd:13

<i>Gustavo Costa, Luis Enrique Zarate</i>	
Governança de Dados em Sistemas-de-Sistemas: Uma Abordagem Orientada à Da-	sbbd:14
dos de Proveniência	182
<i>Jéssica Monçôres de Almeida, Vanessa Braganholo, Daniel de Oliveira</i>	
Leveraging Large Language Models for Time Series Prediction on Low-Frequency	sbbd:15
Data	196
<i>Rodrigo Parracho, Fernando Alexandrino, Matheus de Souza Figueiredo, Lucas Pe-</i>	
<i>reira da Silva, Bruno Dutra de Macedo, Arthur Lamblet Vaz, Davi Louback, Victor</i>	
<i>Coculilo Desouzart, Rebecca Salles, Fabio Porto, Diego Carvalho, Eduardo Oga-</i>	
<i>sawara</i>	
Fine-Tuning Detection Criteria for Enhancing Anomaly Detection in Time Series	sbbd:16
.....	209
<i>Edson Pinto Sobrinho, Jessica da Silva Costa, Jânio Lima, Lucas Tavares, Eduardo</i>	
<i>Bezerra, Rafaelli de Carvalho Coutinho, Lais Baroni, Esther Pacitti, Fabio Porto,</i>	
<i>Kele Belloze, Eduardo Ogasawara</i>	
Redução da Subjetividade na Identificação da Variação do Nível de Valores de Ati-	sbbd:17
vos Financeiros	222
<i>José Jeovane R. Cordeiro, Arlino Magalhaes, Victor Gabriel C. Rodrigues, Gui-</i>	
<i>lherme A. Avelino</i>	
Efficient Reuse of Metric Indexes for Multi-resolution Queries	sbbd:18
	236
<i>Agma Juci Machado Traina, Caetano Traina Jr., Rodrigo César Arboleda</i>	
Aprimorando Geração Aumentada por Recuperação via Ajuste Fino Sequencial de	sbbd:19
Modelos de Linguagem Pequenos	250
<i>Ronaldinho Vega Centeno Olivera, Frances A. Santos, Julio Cesar Dos Reis, Allan</i>	
<i>M. de Souza</i>	
Muitas Classes Desbalanceadas? Não Classifique - Ranqueie! Uma Abordagem Ba-	
seada em Retrieval-Augmented Generation (RAG)-labels para Classificação Textual	sbbd:20
Multi-classe	264
<i>Celso França, Ian Nunes, Thiago Salles, Washington Cunha, Gabriel Jallais, Leo-</i>	
<i>nardo Rocha, Marcos André Gonçalves</i>	
Agente Autônomo Guiado por LLM para Extração de Notícias	sbbd:21
	278
<i>João Victor Cosme Neres de Sousa, Lucas Malta Mingardo, Carlos Eduardo Torres</i>	
<i>Freire, Agma Juci Machado Traina, Caetano Traina Jr.</i>	
Nova Base de Dados Brasileira para Sistemas de Recomendação de Artigos Cientí-	sbbd:22
ficos	289
<i>João Vitor Felipe dos Santos, Ricardo Marçal de Andrade Nascimento, Adriano Cé-</i>	
<i>sar de Melo Camargo, Sergio Daniel Carvalho Canuto, Gustavo de Assis Costa,</i>	
<i>Daniel Xavier Sousa</i>	
Repositório de Dados para Ciência Aberta na Região Amazônica	sbbd:23
	303
<i>André Nascimento Maia, Caio Nascimento Maia, Pedro Pizzigatti Corrêa</i>	
KV-RAPTOR: Scalable Tree-Structured Retrieval with KV Cache Compression for	sbbd:24
Question-Answering Systems	316
<i>João Gabriel Junqueira da Silva, Sávio Salvarino Teles de Oliveira, Lucas Alexan-</i>	
<i>dria Alves, Nicolás Eiris, Arlindo Rodrigues Galvão Filho</i>	
OndePublicar.com: A Collaborative Web Platform about Academic Events and	sbbd:25
Journals Accepting Paper Submissions	330
<i>William Jones beckhauser, Renato Fileto</i>	

eTRUMiner: Mining Multivariate Temporal Rules from Heterogeneous and Incomplete Time Series	sbbd:26 344
<i>Eliane Gniech Karasawa, Elaine Parros Sousa</i>	
Unveiling Power on Combining Prompt Engineering Techniques: An Experimental Evaluation on Code Generation	sbbd:27 357
<i>Cristofer Girardi, Damires Souza, Alex Rego</i>	
Tudo em Todo Lugar ao Mesmo Tempo: Rastreabilidade de Dados em Cidades Inteligentes por meio de Proveniência	sbbd:28 371
<i>Maria Luiza Falci, Débora Pina, Liliane Kunstmann, Vanessa Braganholo, Daniel de Oliveira</i>	
Intervening in problematic data regions to improve machine learning models ..	sbbd:29 385
<i>Gregully Willian da Silva Lima, Fabio Porto, Eduardo Henrique Monteiro Pena</i>	
Effectiveness of Small and Large Language Models for PL/SQL Bad Smell Detection	sbbd:30 399
<i>Vinicius Ferreira de Sousa, Cláudio de Souza Baptista, André Luiz Firmino Alves, Hugo Feitosa de Figueirêdo</i>	
Mitigando Impactos de Distribuições Não-IID em Aprendizagem Federada para Sistemas de Recomendação	sbbd:31 413
<i>Arthur Negrão de Faria Martins da Costa, Guilherme Carolino Rodrigues e Rocha, Lucas Gomes Dos Santos, Pedro Igor de Souza Malaquias, Rodrigo César Pedrosa Silva, Reinaldo Silva Fortes, Pedro Henrique Lopes Silva</i>	
TWIX: Balancing Fairness and Utility in Item Exposure for Recommendation Systems	sbbd:32 427
<i>Maria de Lourdes M. Silva, Iago de Castro Chaves, Eduardo R.D. Neto, André Luís da Costa Mendonça, Javam de Castro Machado</i>	
Abordagem de migração de bases relacionais para bases orientadas a documentos apoiada por LLM	sbbd:33 441
<i>Luan Felipe Marmentini, Evandro Miguel Kuszera</i>	
Advanced Chunking Techniques: a Novel Approach for Semantic Splitters	sbbd:34 455
<i>Bernardo Ramos Toresan, Viviane Pereira Moreira, Felipe Soares Fagundes Paula, Luciana Regina Bencke</i>	
Equitable Diabetes Diagnosis: Tackling Ethnic and Gender Disparities	sbbd:35 466
<i>Lívia Ruback, Luisa Felix, Ariel Soares Teles</i>	
Language Models are not a Panacea: Combining them with Domain Knowledge and Efficient Indexes for Entity Linking	sbbd:36 479
<i>Daniel Lucas Albuquerque, Vitória Soares dos Santos, Pedro Nack, Renato Fileto, Carina F. Dorneles</i>	
Domínio Delimitado, Ódio Exposto: O Uso de Prompts para Identificação de Discurso de Ódio Online com LLMs	sbbd:37 493
<i>Laryssa Paiva, Gabriel Assis, Annie Amorim, Luiz Gustavo Dias, Aline Paes, Daniel de Oliveira</i>	
Estudo Comparativo de Banco de Dados NoSQL para Gerenciamento de Séries Espaço-Temporais	sbbd:38 507
<i>Luís Eduardo Damasceno, Melise Maria Veiga de Paula, Vanessa Cristina Oliveira de Souza</i>	
Data Dependent Itemset Mining Under Local Differential Privacy	sbbd:39 521
<i>Renan Ribeiro Oliveira, José Costa Filho, José Maria da Silva Monteiro Filho, Ja-</i>	

vam de Castro Machado

Uma Avaliação Comparativa entre o DepreBERTBR e Modelos de Linguagem para Classificação de Textos Depressivos [sbbd:40](#)
535

Ayrton Herculano, Laerty Santos da Silva, Damires Souza, Alex Rego [sbbd:41](#)
Particionamento Dinâmico Espaço-Temporal para Séries de Dados Matriciais . 549

Geovani Pereira dos Santos, Daniel S. Kaster
Dinâmicas de Grupos de WhatsApp da Extrema Direita no Brasil: Uma Análise Comparativa Pré e Pós-Eleição de 2022 [sbbd:42](#)
563

Fernanda Ferreira do Nascimento, José Maria da Silva Monteiro Filho, Javam de Castro Machado

Twinscie-Prov: Gerenciando a Proveniência sobre o Ciclo-de-vida de ML no Sistema Twinscie [sbbd:43](#)
576

Júlia Neumann Bastos, Fabio Porto, Fábio Levi Siqueira, Edson Gomi, Ismael Santos, Rodrigo Barreira, Isabela Siqueira, Eduardo Ogasawara

QualiBus: implementando métricas de qualidade para dados de transporte coletivo [sbbd:44](#)
589

Rafael Luciano Silva, Diêgo Correia, Letícia Mendes, Ruan Tenório de Melo, Fabio J. Coutinho

Scalable privacy-preserving record linkage: Evaluating MultiBit tree indexing in Atyimo [sbbd:45](#)
602

Victor Dantas da Rocha Orrico, Fernanda Eustáquio, Bethânia Almeida, Mirlei Silva, Robespierre Pita

Algoritmo Paralelo Eficiente para Ordenação Chave-Valor [sbbd:46](#)
616

Michel Brasil Cordeiro, Rodrigo Morante Blanco, Wagner M. Nunan Zola [sbbd:47](#)
Uma Abordagem para a Gestão da Linhagem de Dados Heterogêneos 630

Hudson Afonso Batista Silva, José Eduardo Medeiros Jochem, João Vitor dos Santos, Eduardo F. R. Sousa, Ronaldo dos Santos Mello, Carina F. Dorneles, Renato Fileto

Training-Free Hybrid Evidence Retrieval for Question Answering: Dynamic Fusion of Knowledge-Graph Triples and Dense Text Embeddings [sbbd:48](#)
644

Otávio Calaçá Xavier, Anderson da Silva Soares

Database Modeling Automation from Natural Language Requirements [sbbd:49](#)
658

Júlia Menezes, Claudio Calazans Campelo Elizio
Análise de Sentimentos em Avaliações de Livros Utilizando a API Gemini para Recomendação Personalizada [sbbd:50](#)
671

Jose Alves da Silva Fagner, Sebastião Alves Filho Emidio, Raul Paradedá Benites, Lenardo e Silva Chaves

Criando Portfólios de Alto Desempenho: Otimização de Portfólios de Ativos de Alta Volatilidade Através da Previsão de Retorno Baseada em CNN+BiLSTM [sbbd:51](#)
685

Israel Rodrigues Soares, Sergio Daniel Carvalho Canuto

Data-level sampling for dealing with imbalanced datasets: better protection against membership inference attacks [sbbd:52](#)
699

KARLA Felicia Carvalho da Silva, ANTONIO de Abreu Batista Jr, Jesus P. Mena-Chalco, LUCIANO Reis Coutinho

Subnational Foreign Trade Data in Latin America: Diagnostic, Challenges, and Technical Pathways [sbbd:53](#)
711

Felipe Ramon de Britto Redondo, Orlando Junior

Preliminary Ranking-Based Selection for Optimized Retriever Configuration in RAG Systems [sbbd:54](#)
725
Salvador Ludovico Paranhos, Jonatas Novais Tomazini, Celso Junior Camilo, Sávio Salvarino Teles de Oliveira

SHORT

Enhancing Text-to-SQL with In-Context Learning: A Multi-Agent Approach Based on CHESS	sbbd:55 739
<i>Renato Okabayashi Miyaji, Rafael Fernandes, Krysthian Martins, Jorge Melegati, Pedro Luiz Pizzigatti Corrêa</i>	
Aplicação do Modelo ARIMA no Vertica para Previsão da Velocidade do Vento	sbbd:56 746
<i>Gabriel Ciriaco Fornitano, Flávio Belizário da Silva Mota, Vanessa Cristina Oliveira de Souza, Arcilan Assireu, Melise Maria Veiga de Paula</i>	
RPA's e Data Lakes para a Indústria 4.0: Um Estudo de Caso de Ecossistema de Dados Integrados	sbbd:57 753
<i>Arthur Lucas dos Santos Bezerra, Iranildo Santos Batalha, Luís Ricardo Arantes Filho, Clarice Martins Almeida, Matheus Inácio Nascimento Dantas, Nelson Alexandre Gouvêa</i>	
How Culture Shapes Customers: A Cross-Continent Analysis of Apps Reviews Using NLP Techniques	sbbd:58 760
<i>Maria Fernanda Azolin Kotsifas, Thiago Henrique Silva, Ricardo Lüders</i>	
Análise Teórica do Impacto de Dados Faltantes em Atributos Sensíveis sobre a Métrica de Fairness p%-rule	sbbd:59 767
<i>Dalton da Silva, Dimas Cassimiro Nascimento, Luis Filipe Alves Pereira</i>	
Arquitetura Agnóstica para Transcrição, Armazenamento e Análise de Áudios Operacionais no Setor Elétrico Utilizando Serviços de Nuvem e Elasticsearch	sbbd:60 774
<i>Patrícia de Castro Wang, Maurício Silva de Magalhães, Kleyton Pontes Cotta, Hugo Diniz Rebelo, Leonardo Tavares Vianna, Hugo Reiser Vieira Portuita, Tássio Simioni, Rodrigo Mazo Rocha</i>	
Assessing the impact of missing value mechanisms on anomaly detection in healthcare wearable data	sbbd:61 781
<i>Afonso Matheus Sousa Lima</i>	
Quando os Erros Informam: Apoio ao Diagnóstico de Diabetes em Cenários de Alta Incerteza	sbbd:62 788
<i>Samuel Norberto Alves, Celso França, Regina Bernal, Crizian Gomes, Oluwatoyin Omole, Deborah Malta, Marcos Gonçalves, Jussara Almeida</i>	
Drift Detection for Identifying Training Patterns Prior to Performance Improvement in Runners	sbbd:63 795
<i>Nathália Tito, Lucas Tavares, Glauco Fiorott Amorim</i>	
Collaborative Classification for Object Labeling on Expansible Datasets	sbbd:64 802
<i>Bruno Padilha, João Eduardo Ferreira</i>	
Avaliando as Limitações e Potenciais do Algoritmo k-Vizinhos Mais Próximos (kNN) na Imputação de Dados Clínicos	sbbd:65 809
<i>izadora monken ganem, Guilherme Dal Bianco, José Carlos Serufo Filho, Luciano Romero Soares de Lima, Leonardo Chaves Dutra da Rocha, Marcos André Gonçalves</i>	
Representação Probabilística de Trajetórias Veiculares como Entrada para Redes Neurais Artificiais	sbbd:66 816

Bianca Lahm Gomes, Kame Haung Zhu

Metadata Management in Data Mesh: Toward Federated Discovery and Governance
[sbbd:67](#)
823

Rafael Hideki Suguimoto, Paulo Meirelles, Kelly Rosa Braghetto

Data Augmentation for Medical Image Segmentation: A Comparative Analysis of
Traditional Techniques and Synthetic Data Generation [sbbd:68](#)
830

*Mariana Aya Suzuki Uchida, Erikson Júlio de Aguiar, Caetano Traina Jr., Agma
Juci Machado Traina*

Estudo do Impacto de Dados Sintéticos e Paráfrases na Mitigação do Desbalancea-
mento em Tarefas de Classificação de Textos em Português com Baixa Amostragem
[sbbd:69](#)
837

*Claudio Moisés Valiense Andrade, Gestefane Rabbi, Raiane Asevedo, Julia Paes,
Isaias José Ramos Oliveira, Adriana Pagano, Zilma Reis, Marcos André Gonçalves*

ARANI: Uma Abordagem Baseada em Linha de Experimento para Preservação de
Privacidade em Data Lakes [sbbd:70](#)
844

Thiago da Costa Jordão, Daniel de Oliveira, Marcos Bedo

Quantifying the RAG Advantage: A Multi-Metric Benchmark for LLM-based Code
Generation [sbbd:71](#)
851

*Gabriel Machado Lunardi, Guilherme Machado Medeiros, José Palazzo Moreira de
Oliveira, Gabriel Souza Baggio*

Towards Enabling the Analysis of Visual Exploration Processes through Interaction
Provenance [sbbd:72](#)
858

*Lycoln Sousa de Oliveira, Gustavo Moreira, Fábio Miranda, Marcos Lage, Daniel
de Oliveira*

A Data Augmentation and Validation Pipeline for Improving Emotion Classification
in Mobile App Reviews [sbbd:73](#)
865

*Kalidsa de Oliveira Buzzatti, Gabriel Machado Lunardi, Williamsom Silva Alison,
Thiago Trugillo da Silveira Lopes, Adriano de Oliveira Quilião*

DiveScan: um módulo para Recuperação de dados por Similaridade com Diversidade
em Postgres [sbbd:74](#)
872

Anna Júlia Costa Lauton, Agma Juci Machado Traina, Caetano Traina Jr.

Subdomain Identification Strategies for Efficient Machine Learning Models [sbbd:75](#)
879

*Samuel R. Torres, Rocio Zorrilla, Raphael de Freitas Saldanha, Victor Riberiro,
Eduardo Henrique Monteiro Pena, Fabio Porto*

UniChat: Arquitetura e Avaliação de um Agente RAG de Baixo Custo para o Con-
texto Universitário [sbbd:76](#)
886

*Ana Clara Boniatti Bordin, Gabriel Machado Lunardi, Eduardo Kesler Piveta, Le-
onardo Ramos Emmendorfer*

Large Language Models para detecção de conluios em licitações [sbbd:77](#)
893

Jorge Nelson de Souza Pavão, Diego Nunes Brandão, Kele Belloze

Adaptando Indicadores para Avaliação da Saúde de Projetos de Infraestrutura Ci-
entífica: Uma Aplicação de Business Intelligence nas Linhas de Luz do Sirius . [sbbd:78](#)
900

Dante Paes Brito, Mauri Pereira dos Santos Júnior, Marcelo Moreira Xavier

PANDORA: Sistema Estatístico para Previsão de Eventos Climáticos Extremos [sbbd:79](#)
907

Flávia Maristela S. Nascimento, Luís Fernando Cezar Santos

Utilização do Apache Superset para Visualização Escalável de Dados Educacionais
Públicos: Um Estudo de Caso com o Censo Escola [sbbd:80](#)
914

<i>João Pedro Vicente Ramalho, João Armênio Silveira, Thales Gabriel Carvalho de Lima, Mateus dos Santos Herbele, Josiney de Souza, Guilherme Alex Derenievicz, Letícia Mara Peres, Simone Dominico</i>	
Um Estudo Comparativo de Estratégias de Seleção de Exemplos para In-Context Learning aplicado à Classificação Automática de Texto com Grandes Modelos de Linguagem	sbbd:81 921
<i>Gabriel Silva Prenassi, Davi dos Reis de Jesus, Washington Cunha, Marcos André Gonçalves, Leonardo Chaves Dutra da Rocha, Guilherme Fonseca do Nascimento</i>	
Pondere e Expanda: Impacto e Limitações de Representações Contextual-Esparsas na Modelagem de Tópicos	sbbd:82 928
<i>Ana Cláudia Machado, Celso França, Ian Nunes, Marcos André Gonçalves, Leonardo Rocha</i>	
Extração Automática de Atributos de Sinais de Emissão Acústica com Redes Neurais Autocodificadoras para Predição de Integridade em Tubulações	sbbd:83 935
<i>Bernardo Lemos, Antônio Caetano Neves, Marcos Carvalho, Sergio Daniel Carvalho Canuto, Jorge Ribeiro, Rodrigo Pires, Jussara M. Almeida, Marcos André Gonçalves, Douglas Santos, André Fonseca</i>	
A Comparative Analysis of Denoising Methods for Deep Learning-Based Audio Event Detection in Noisy Agricultural Environments	sbbd:84 942
<i>André Moreira Souza, Guilherme Augusto Moreira, Lucas Eduardo Gulka Pulcinelli</i>	
Estendendo Consultas Contínuas por Abrangência aos Dados Métricos	sbbd:85 949
<i>Enzo Seraphim, Thatyana de Faria Piola Seraphim, Lucio F.D. Santos, Edmilson Marmo Moreira, Luiz Olmes Carvalho</i>	
Entity Matching com Large Language Models: estudo comparativo com abordagem de Entity Blocking	sbbd:86 956
<i>Rodolfo Bolconte Donato, Tiago Brasileiro Araújo</i>	
Exploring Temporal Text-to-SQL Challenges in Brazilian Portuguese: Lessons from Educational Data	sbbd:87 963
<i>Karina de Carvalho Fróes, Kelly Rosa Braghetto</i>	
Pipeline Distribuído para Análise Espacial em Larga Escala: Avaliação da Regra 3 do Índice 3-30-300 em Fortaleza com Apache Spark e Sedona	sbbd:88 970
<i>Lucas Lopes Silva, Lucas Feitosa de Albuquerque Lima Babadopulos, Jorge Barbosa Soares, Lara Sucupira Furtado, Marta Gonzalez</i>	
Addressing Database-Related Issues in Digital Social Network Data Analysis ..	sbbd:89 977
<i>Alexandre Augusto Poggio Heine, Mariana Barreto, Rodrigo Motta Ferreira, Edward Hermann Haeusler, Sérgio Lifschitz</i>	
Deep Learning for Urban Flood Prediction: An LSTM Model Integrating Satellite Reanalysis and Historical Weather Data in Curitiba	sbbd:90 984
<i>Lucas Iuri dos Santos, Luiz Celso Gomes Jr</i>	