



## Boas práticas de modelagem de dados em bancos relacionais: uma revisão rápida de literatura

Eugenio Lopes Fernandes Lima<sup>1</sup>; Pedro Henrique de Meneses Bittencourt Lopes<sup>2</sup>; Francisco Abud Nascimento<sup>3</sup>

### Como Citar:

LIMA, Eugenio Lopes Fernandes; LOPES, Pedro Henrique de Meneses Bittencourt; NASCIMENTO, Francisco Abud. Boas práticas de modelagem de dados em bancos relacionais: uma revisão rápida de literatura. Revista Sociedade Científica, vol. 9, n. 1, p. 1188-1202, 2026. <https://doi.org/10.61411/rsc2026136219>

DOI: 10.61411/rsc2026136219

### Área do conhecimento:

Ciências Exatas e da Terra

### Sub-área:

Ciência da Computação; Sistemas de Informação

### Palavras-chave:

Modelagem de Dados; Bancos de Dados Relacionais; Normalização. Boas Práticas; Governança de Dados.

Publicado: 7 de maio de 2026.

### Resumo

A modelagem de dados desempenha papel central no desenvolvimento de sistemas de informação robustos, eficientes e escaláveis. Esta revisão de literatura analisou estudos publicados entre 2015 e 2025 nas bases ACM Digital Library e IEEE Xplore, com o objetivo de identificar e sintetizar boas práticas de modelagem aplicadas a bancos de dados relacionais. Após triagem de 1.088 registros, 10 estudos foram incluídos e examinados em profundidade. Os resultados indicam que os fundamentos clássicos do modelo relacional, normalização, identificação de dependências funcionais, definição adequada de entidades e chaves e decomposições sem perda, permanecem essenciais para garantir integridade, consistência e clareza semântica. Ao mesmo tempo, a literatura evidencia a incorporação de novas diretrizes relacionadas a automação, inteligência artificial, governança, segurança, proveniência e integração com modelos não relacionais. Observou-se também que a modelagem contemporânea exige alinhamento às características de uso dos dados, contemplando desnormalização controlada, decisões de particionamento e estruturas orientadas a desempenho. Conclui-se que as boas práticas atuais combinam rigor formal, ferramentas inteligentes, modularidade arquitetural e mecanismos de segurança e rastreabilidade, reforçando que o modelo relacional permanece relevante e adaptável às demandas de ecossistemas de dados cada vez mais heterogêneos e distribuídos.

## Best practices in data modeling for relational databases: a rapid literature review

### Abstract

<sup>1</sup>Centro Universitário Mario Pontes Jucá (UMJ) - Maceió-AL, Brasil. Email: ✉

<sup>2</sup>Centro Universitário Mario Pontes Jucá (UMJ) - Maceió-AL, Brasil. Email: ✉

<sup>3</sup>Centro Universitário Mario Pontes Jucá (UMJ) - Maceió-AL, Brasil. Email: ✉



Data modeling is essential for developing robust, efficient, and scalable information systems. The objective of this literature review was to identify and synthesize best practices in data modeling applied to relational databases by analyzing studies published between 2015 and 2025 in the ACM Digital Library and IEEE Xplore. After screening 1,088 records, 10 studies were selected for in-depth examination. The findings suggest that classical relational model principles, such as normalization, identifying functional dependencies, properly defining entities and keys, and performing lossless decompositions, remain essential for ensuring integrity, consistency, and semantic clarity. At the same time, the literature emphasizes the adoption of new guidelines concerning automation, artificial intelligence, governance, security, provenance, and integration with non-relational models. The review also shows that contemporary modeling requires alignment with actual data usage patterns, including controlled denormalization, careful partitioning decisions, and performance-oriented structures. Overall, the integration of formal rigor, intelligent tools, architectural modularity, and security and traceability mechanisms reinforces the relevance and adaptability of the relational model to the demands of increasingly heterogeneous and distributed data ecosystems.

**Keywords:** Data Modeling; Relational Databases; Normalization; Best Practices; Data Governance.

## 1. Introdução

A modelagem de dados ocupa um papel central no desenvolvimento de sistemas de informação, sendo responsável por definir a estrutura lógica que sustenta o armazenamento, a organização e a integridade dos dados em bancos de dados. No contexto dos bancos de dados relacionais, essa atividade é fundamental para assegurar consistência, eficiência e manutenibilidade ao longo do ciclo de vida das aplicações.



Apesar da crescente adoção de abordagens NoSQL (Not Only SQL – Não Apenas SQL), especialmente em cenários que demandam alta escalabilidade e flexibilidade estrutural, o modelo relacional permanece amplamente utilizado em aplicações corporativas, científicas e governamentais. Tal permanência decorre de sua sólida fundamentação teórica, previsibilidade comportamental, suporte robusto a transações e aderência a padrões consolidados da indústria.

Nesse contexto, o SQL (Structured Query Language – Linguagem de Consulta Estruturada) desempenha papel essencial como linguagem padrão para manipulação e consulta de dados em bancos relacionais, os quais se caracterizam por estruturas organizadas em tabelas com *schema* (esquema) bem definido. Em contraste, os bancos NoSQL englobam modelos não relacionais — como os orientados a documentos, chave-valor e grafos — que priorizam flexibilidade de *schema* e escalabilidade horizontal, sendo amplamente adotados em sistemas distribuídos e aplicações de grande volume de dados. Entretanto, a crescente complexidade dos sistemas, aliada à heterogeneidade das fontes de dados e à necessidade de evolução contínua dos esquemas, torna cada vez mais desafiadora a definição de boas práticas que assegurem qualidade, consistência e robustez no projeto de bancos relacionais.

Estudos recentes indicam que problemas recorrentes em projetos de banco de dados estão frequentemente associados a decisões inadequadas nas fases conceitual e lógica da modelagem. Trabalhos como os de DiScala e Abadi [4] evidenciam que práticas como normalização (normalization), identificação rigorosa de dependências funcionais (functional dependencies – FDs) e definição adequada de entidades permanecem fundamentais para reduzir redundância, melhorar a integridade e preservar a semântica dos dados, inclusive em cenários que envolvem dados semiestruturados ou aninhados. De forma complementar, Kumar e Azad [6] reforçam o papel da normalização como método sistemático para evitar anomalias de inserção, atualização e exclusão, destacando a importância da identificação de chaves candidatas, preservação de dependências e realização de decomposições sem perda de informação.



Outro desafio relevante refere-se à evolução dos esquemas ao longo do ciclo de vida das aplicações. A necessidade de manter sistemas em operação contínua impulsiona abordagens como a proposta por De Jong *et al.* [3], que introduzem mecanismos para evolução de *schema* com zero downtime (tempo de indisponibilidade zero). Esse cenário evidencia que boas práticas de modelagem não se restringem ao desenho inicial do banco de dados, mas também envolvem estratégias para modificações estruturais seguras, compatíveis com múltiplas versões da aplicação e que preservem a integridade referencial.

Além disso, aspectos semânticos e expressivos da modelagem conceitual têm ganhado destaque na literatura recente. Li, Currim e Ram [7] apontam que modelos tradicionais frequentemente assumem dados completos e consistentes, o que não reflete a realidade de muitos sistemas atuais. Nesse sentido, os autores defendem a ampliação das capacidades do modelo entidade-relacionamento para lidar com entidades complexas, relações hierárquicas e diferentes níveis de completude dos dados, fatores que impactam diretamente a qualidade do modelo lógico e do esquema relacional resultante.

Diante desse contexto, torna-se necessária uma revisão sistemática das boas práticas de modelagem de dados, contemplando tanto princípios clássicos — como normalização, definição de chaves, minimização de redundância e preservação da integridade — quanto diretrizes emergentes relacionadas à evolução de esquemas, integração de múltiplas fontes de dados e uso de métodos automatizados para apoio ao processo de modelagem.

Dessa forma, este estudo tem como objetivo reunir e sintetizar as principais boas práticas de modelagem de dados descritas na literatura contemporânea, oferecendo um panorama estruturado que contribua para o desenvolvimento de esquemas relacionais mais consistentes, escaláveis e alinhados às demandas atuais.



## 2. Metodologia

Este estudo consiste em uma revisão de literatura focada em identificar, sintetizar e analisar boas práticas de modelagem de dados aplicadas ao contexto de bancos de dados relacionais. A metodologia seguiu um procedimento estruturado composto por definição de estratégias de busca, seleção de estudos e análise qualitativa do conteúdo.

### 2.1. Fontes de informação

A busca sistemática foi realizada em duas bases de dados amplamente reconhecidas na área de Ciência da Computação: ACM Digital Library e IEEE Xplore. Ambas são referências consolidadas para pesquisa em Bancos de Dados, Engenharia de Software e Sistemas de Informação, e incluem artigos de periódicos, conferências e *workshops* revisados por pares.

As consultas foram conduzidas em inglês, com recorte temporal entre 1º de janeiro de 2015 e 31 de dezembro de 2025, a fim de abranger a produção científica dos últimos dez anos. As *strings* (cadeias de busca) de busca utilizadas foram:

- ACM Digital Library:

*[[All: "data modeling"] OR [All: "database design"] OR [All: "relational schema"]] AND [[All: "best practices"] OR [All: guidelines] OR [All: standards] OR [All: patterns]] AND [[All: "relational database"] OR [All: sql] OR [All: rdbms]] AND [E-Publication Date: (01/01/2015 TO 12/31/2025)]]*

- IEEE Xplore:

*("data modeling" OR "database design" OR "relational schema") AND ("best practices" OR guidelines OR standards OR patterns) AND ("relational database" OR SQL OR RDBMS)*



As buscas foram executadas sem filtros adicionais além do período definido, abrangendo artigos completos, estudos empíricos, propostas metodológicas e investigações teóricas relacionadas ao tema.

## 2.2. Critérios de elegibilidade

Foram considerados elegíveis para inclusão: 1) *research papers* (artigos de periódicos indexados); 2) *conference proceedings* (trabalhos apresentados em conferências ACM e IEEE).

A inclusão de publicações oriundas de conferências se justifica pelo fato de que a área de Bancos de Dados historicamente concentra parte significativa de suas contribuições científicas nesses eventos, que adotam processos rigorosos de revisão por pares.

Foram excluídos: artigos fora do recorte temporal, estudos que não abordavam práticas de modelagem de dados, trabalhos focados exclusivamente em tecnologias NoSQL ou em aspectos não relacionados ao modelo relacional (SQL).

## 2.3. Processo de seleção dos estudos

As buscas resultaram em 1.089 registros (89 da IEEE Xplore e 1.000 da ACM Digital Library). Após a remoção de uma duplicata, permaneceram 1.088 registros únicos.

A triagem ocorreu em duas etapas: 1) leitura de títulos e resumos, para verificar alinhamento com o tema; 2) leitura do texto completo, aplicada aos estudos potencialmente relevantes.

Ao final do processo, 10 estudos atenderam a todos os critérios de inclusão e foram selecionados para a revisão.

## 2.4. Procedimentos de análise



Os estudos incluídos foram examinados qualitativamente, com foco na identificação de 1) boas práticas descritas explicitamente pelos autores; 2) recomendações metodológicas implícitas nos resultados; 3) padrões de modelagem, diretrizes e princípios aplicáveis ao projeto de esquemas relacionais.

A síntese dos achados foi organizada em tabelas, de modo a permitir comparação entre os estudos e facilitar a identificação de convergências e particularidades no uso de boas práticas.

### 3. **Desenvolvimento e discussão**

A busca sistemática nas bases ACM Digital Library e IEEE Xplore resultou em 1.088 registros após a remoção de duplicatas. Após triagem por título, resumo e leitura completa, 10 estudos atenderam aos critérios de inclusão e compuseram o corpus final desta revisão. A Tabela sintetizada reúne as boas práticas identificadas em cada artigo, abrangendo contribuições relacionadas à normalização, dependências funcionais, modelagem conceitual, automação, segurança, proveniência, arquitetura de dados e ensino de modelagem.

De modo geral, os estudos reforçam que boas práticas de modelagem de dados envolvem tanto princípios tradicionais do modelo relacional (normalização, identificação correta de chaves e dependências, eliminação de redundâncias, definição clara de entidades e relacionamentos) quanto diretrizes contemporâneas ligadas à realidade atual dos sistemas: uso de IA, automação, governança, modelagem híbrida SQL/NoSQL, segurança e proveniência de dados.

Entre os resultados observados, destacam-se:

- I. Normalização e dependências funcionais continuam centrais [9,6], reforçando sua relevância como práticas fundamentais para reduzir redundância, aprimorar integridade e orientar decomposições corretas.



II. Automação e uso de ferramentas inteligentes emergem como tendência [11], incluindo NLP (Natural Language Processing – Processamento de Linguagem Natural) para apoio ao modelador, geração automática de rótulos, identificação de dependências, sugestão de estruturas e atualização contínua.

III. Modelagem orientada a múltiplas tecnologias aparece de forma robusta [12,1,2], indicando que o esquema relacional serve como referência para práticas em ambientes NoSQL, *analytics*, MapReduce e Spark.

IV. Arquiteturas de dados modernas, como data mesh (malha de dados) e pipelines (fluxos de processamento) reconfiguráveis, enfatizam interoperabilidade, modularidade e governança [5].

V. Segurança na modelagem e no acesso aos dados também surge como dimensão essencial [8], destacando *prepared statements* (consultas parametrizadas), validação de entrada, princípios de privilégio mínimo e auditoria.

VI. Proveniência de dados e rastreabilidade aparecem como elemento estruturante para qualidade e auditabilidade [10], especialmente em ambientes onde a modelagem precisa incorporar histórico e preservação de mudanças.

VII. Modelagem baseada nos padrões dos dados, e não na estrutura física do armazenamento, é enfatizada para ambientes com dados aninhados ou semiestruturados [4].

As boas práticas identificadas variam em escopo, indo do nível pedagógico (ensino de normalização) ao nível arquitetural (*data mesh*), passando por modelagem lógica e física, mas convergem para a necessidade de rigor, clareza semântica, interoperabilidade e governança.

**Tabela 1: Boas práticas citadas nos estudos incluídos**

ID do estudo	Boas práticas citadas
Storey, Lukyanenko	Uso de IA (Inteligência Artificial) e NLP para apoiar boas práticas de modelagem; modelos mais acessíveis (narrativas, multimídia, instâncias); uso de instâncias para



ID do estudo	Boas práticas citadas
e Castellanos [11]	capturar semântica; assistentes automáticos para melhorar rótulos e estruturas; automação da geração de modelos; melhoria do processo de criação/atualização; uso de ferramentas CASE.
Pavkovic e Brdjanin [9]	Uso de diagramas de dependências funcionais para tornar o processo de normalização mais sistemático e visível; representação gráfica explícita de dependências parciais e transitivas; introdução de novos símbolos Dados Persistentes Avançados (APD) e seta dupla para representar dependências relacionadas a BCNF (Boyce-Codd Normal Form – Forma Normal de Boyce-Codd) e 4NF (Fourth Normal Form – Quarta Forma Normal); ensino da normalização passo a passo até a 4NF; uso dos diagramas aprimorados como ferramenta didática.
Wu [12]	Uso de SQL como “linguagem pivô” para ensinar diferentes modelos (relacional, NoSQL, MapReduce, Spark); tradução sistemática de SQL para álgebra relacional e planos físicos; boas práticas de otimização (empurrar seleções antes de junções, algoritmos adequados de <i>join</i> ); modelagem cuidadosa em NoSQL a partir do esquema relacional; atenção às consequências da desnormalização; uso de pipelines que espelham operações da álgebra relacional.
Anderson [1]	Ênfase na modelagem desde o início em sistemas intensivos em dados; escolha criteriosa de tecnologias de persistência; projeto cuidadoso de chaves de partição/ <i>row keys</i> ; desnormalização controlada quando necessária; uso combinado de processamento em lote e fluxo; ciclos de desenvolvimento iterativos; composição de equipes multidisciplinares.
Carey et al. [2]	Abordagem <i>data-first</i> com modelagem conceitual ER (Entity-Relationship – Entidade-Relacionamento) e a UML (Unified Modeling Language – Linguagem de Modelagem Unificada) antes do desenho físico do banco de documentos; mapeamento de entidades e relacionamentos para coleções; boas práticas de <i>embedding vs. Linking</i> (incorporação vs referência); uso de arrays e objetos uniformes; incorporação de <i>weak entities</i> como subdocumentos; projeto orientado a consultas e indexação; evitar <i>anti-patterns</i> estruturais; uso da flexibilidade do JSON para evolução controlada do esquema.
Dončević et al. [5]	Arquitetura padronizada <i>mask-mediator-wrapper</i> ; separação clara de responsabilidades; princípios de <i>consume-transform-serve</i> ; uso de modelos de dados generalizados para interoperabilidade; evitar dependência tecnológica; componentes modulares (“arquitetural quanta”); governança federada; uso de metadados, versionamento e esquemas claros; pipelines reusáveis e reconfiguráveis; observabilidade ( <i>logging</i> (registro de eventos), métricas, monitoramento); prototipação rápida para adoção incremental.
Markulin et al. [8]	Uso de <i>prepared statements/parameterized queries</i> ; validação de entrada; princípio do menor privilégio; tratamento adequado de dados ausentes; uso de WAF (Web Application Firewall – Firewall de Aplicação Web); utilização de ORMs (Object-Relational Mappers – Mapeadores Objeto-Relacional) para reduzir vulnerabilidades; sanitização e tipagem correta de parâmetros; <i>logging</i> e exceções significativas; auditorias, revisões de código e comunicação segura.
Rani, Goyal e Gadia [10]	Uso de proveniência para garantir qualidade, auditabilidade e rastreabilidade; manutenção de histórico via ZILD (Zero Information Loss Data – Dados com Perda Zero de Informação); evitar redundância e perda de informação com <i>nested tables</i> ; <i>query inversion</i> para rastrear derivação; uso de grafos para armazenar proveniência; captura de proveniência de consultas históricas; registro sistemático por <i>query table</i> ; suporte a <i>forward/backward tracing</i> ; separação entre banco principal e mecanismo de



ID do estudo	Boas práticas citadas
DiScala e Abadi [4]	proveniência; armazenamento incremental para evitar explosão de dados. Redução de redundância por normalização; detecção de dependências funcionais (incluindo <i>soft FDs</i> ); agrupamento de atributos por dependência; identificação de entidades equivalentes; modelagem orientada aos padrões dos dados e não ao formato físico; criação de tabelas para entidades estáveis; evitar normalização excessiva; uso de <i>surrogate keys</i> ; deduplicação pós-carga; avaliação de cardinalidade para decisões de decomposição; separação entre atributos estáveis e instáveis; unificação de entidades com domínios sobrepostos; avaliação de <i>skew</i> ; manutenção de clareza semântica; inclusão de

Fonte: Elaborado pelos autores (2026).

A análise dos 10 estudos revela que a evolução das boas práticas de modelagem de dados em bancos relacionais ocorre em duas frentes simultâneas: preservação dos princípios fundamentais e expansão para cenários modernos de dados heterogêneos, distribuídos e de larga escala.

### 3.1. A permanência do núcleo clássico da modelagem relacional

Os estudos de Pavkovic & Brdjanin [9], Kumar & Azad [6] e DiScala & Abadi [4] demonstram que os fundamentos teóricos do modelo relacional permanecem essenciais. Normalização, dependências funcionais e decomposição *lossless* (sem perda de informação) continuam sendo práticas insubstituíveis para assegurar integridade, minimizar anomalias e manter clareza estrutural.

Mesmo em contextos contemporâneos, como integração de dados, pipelines analíticos e sistemas distribuídos, os autores reforçam que esquemas mal normalizados amplificam custos, erros e dificuldades de evolução.

### 3.2. Modelagem orientada ao uso e ao comportamento dos dados

Estudos como Carey *et al.* [2] e Anderson [1] mostram que, em ambientes híbridos (SQL/NoSQL), a modelagem de dados precisa ser guiada pelos padrões de acesso e pelas operações dominantes. O esquema relacional deixa de ser apenas um artefato estrutural e passa a servir como referência para decisões sobre desnormalização



controlada, definição de chaves de partição, e escolha entre *embedding* e *linking* em modelos de documentos.

Isso reforça a importância de uma modelagem orientada ao uso, não apenas à teoria.

### 3.3. A ascensão da automação e da inteligência artificial

Storey *et al.* [11] indicam uma tendência significativa: a modelagem está se tornando cada vez mais assistida, seja por ferramentas CASE avançadas, por NLP, ou por assistentes que sugerem entidades, nomes, relacionamentos e regras.

Essa automação não substitui o modelador, mas reduz erros, melhora consistência e acelera iterações, especialmente relevante em ambientes de desenvolvimento ágeis.

### 3.4. Governança, interoperabilidade e arquiteturas modernas

O estudo de Dončević *et al.* [5] amplia o escopo tradicional da modelagem ao situá-la dentro de arquiteturas de dados complexas, modularizadas e federadas. A organização *Mask–Mediator–Wrapper* e os princípios de *consume–transform–serve* refletem uma visão na qual a modelagem não é apenas estrutural, mas parte de um ecossistema maior de interoperabilidade e governança.

Esse panorama evidencia um movimento em direção a esquemas projetados para coexistir, evoluir e se recompor, característica essencial em *data mesh* e sistemas distribuídos.

### 3.5. Segurança e proveniência como elementos de modelagem

Markulin *et al.* [8] e Rani *et al.* [10] destacam que segurança e proveniência não são apenas responsabilidades da camada de aplicação, mas devem fazer parte das decisões de modelagem. Isso inclui:



- definição de estruturas para rastreamento (*query tables*, grafos de proveniência);
- escolhas de tipos fortes e parâmetros bem definidos;
- restrições de acesso e minimização de superfícies de ataque;
- mecanismos que permitam auditoria completa e análises históricas.

Essas práticas tornam o modelo relacional não mais seguro, e também mais transparente e confiável.

### 3.6. Convergência entre teoria e prática

Apesar da diversidade dos estudos, há uma convergência clara: fundamentos relacionais continuam indispensáveis; práticas modernas expandem, mas não substituem a teoria; modelagem hoje exige compreensão semântica, arquitetural e operacional.

Assim, os resultados sugerem que o conjunto de boas práticas contemporâneas deve combinar 1) rigor teórico; 2) automação e ferramentas inteligentes; 3) governança e rastreabilidade; 4) segurança; 5) integração entre modelos relacionais e modelos não relacionais; 6) arquiteturas modulares e evolutivas.

## 4. Considerações finais

Esta revisão de literatura evidencia que as boas práticas de modelagem de dados em bancos relacionais permanecem fundamentais para a qualidade, integridade e longevidade dos sistemas de informação, mesmo diante da crescente complexidade dos ecossistemas de dados contemporâneos. Os estudos analisados mostram que princípios clássicos, como normalização, identificação de dependências funcionais, eliminação de redundâncias e definição rigorosa de entidades e relacionamentos, continuam a servir como alicerces estruturais, orientando decisões técnicas independentemente do contexto tecnológico.

Ao mesmo tempo, a literatura demonstra que essas práticas se expandiram para além do escopo tradicional, incorporando diretrizes relacionadas à automação, uso de



inteligência artificial, governança de dados, segurança, proveniência e integração entre modelos relacionais e não relacionais. Em particular, observa-se que a modelagem relacional desempenha um papel central na transposição e no desenho de esquemas para ambientes NoSQL, *pipelines* analíticos e arquiteturas distribuídas, reforçando sua relevância como referência conceitual.

Outro aspecto importante revelado pelos estudos é a crescente necessidade de alinhar a modelagem ao comportamento real dos dados e aos padrões de acesso das aplicações. Práticas como desnormalização controlada, definição criteriosa de chaves de partição, uso consciente de *embedding* e *linking* e adoção de esquemas evolutivos mostram que o modelo relacional continua sendo adaptado para atender às demandas de escalabilidade, desempenho e interoperabilidade.

Por fim, a convergência observada entre teoria e prática sugere que um conjunto robusto de boas práticas deve combinar rigor formal, clareza semântica, automação, modularidade arquitetural e mecanismos de segurança e rastreabilidade. Assim, esta revisão indica que, embora o cenário tecnológico esteja em constante transformação, a modelagem relacional, quando fundamentada em princípios sólidos e atualizada com as necessidades contemporâneas, permanece essencial para o desenvolvimento de sistemas de dados confiáveis, eficientes e duráveis.

## 5. Declaração de direitos

Os autores declaram ser detentores dos direitos autorais da presente obra, que o artigo não foi publicado anteriormente e que não está sendo considerado por outra(o) Revista/Journal. Declaram que as imagens e textos publicados são de responsabilidade dos autores, e não possuem direitos autorais reservados a terceiros. Textos e/ou imagens de terceiros são devidamente citados ou devidamente autorizados com concessão de direitos para publicação quando necessário. Declaram respeitar os direitos de terceiros e de Instituições públicas e privadas. Declaram não cometer plágio ou autoplágio e não ter considerado/gerado conteúdos falsos e que a obra é original e de responsabilidade dos autores.



## 6. Referências

1. ANDERSON, Kenneth M. Embrace the challenges: software engineering in a big data world. *In: BIGDSE '15*. IEEE Press, 2015.
2. CAREY, Michael *et al.* Towards Principled, Practical Document Database Design. *Proc. VLDB Endow.*, v. 18, n. 12, p. 4804-4816, ago. 2025.
3. DE JONG, Michael; VAN DEURSEN, Arie; CLEVE, Anthony. Zero-downtime SQL database schema evolution for continuous deployment. *In: ICSE-SEIP '17*. IEEE Press, 2017. DOI: <https://doi.org/10.1109/ICSE-SEIP.2017.5>.
4. DISCALA, Michael; ABADI, Daniel J. Automatic Generation of Normalized Relational Schemas from Nested Key-Value Data. *In: SIGMOD '16*. New York, NY, USA: Association for Computing Machinery, 2016. DOI: <https://doi.org/10.1145/2882903.2882924>.
5. DONČEVIĆ, J. *et al.* Mask–Mediator–Wrapper Architecture as a Data Mesh Driver. *IEEE Transactions on Software Engineering*, v. 50, n. 4, p. 900-910, 2024.
6. KUMAR, K.; AZAD, S. K. Database normalization design pattern. *In: 2017*.
7. LI, Yuanxia; CURRIM, Faiz; RAM, Sudha. Data Completeness and Complex Semantics in Conceptual Modeling: The Need for a Disaggregation Construct. *J. Data and Information Quality*, v. 14, n. 4, nov. 2022.
8. MARKULIN, H. *et al.* Using ChatGPT in Designing Web Applications Resistant to SQL Injection Attacks. *In: 2025*.
9. PAVKOVIC, Z. S.; BRDJANIN, D. A UML-based approach to forward engineering of SQLite database. *In: 2016*.
10. RANI, Asma; GOYAL, Navneet; GADIA, Shashi K. Data Provenance for Historical Queries in Relational Database. *In: Compute '15*. New York, NY, USA: Association for Computing Machinery, 2015. DOI: <https://doi.org/10.1145/2835043.2835047>.



11. STOREY, Veda C.; LUKYANENKO, Roman; CASTELLANOS, Arturo. Conceptual Modeling: Topics, Themes, and Technology Trends. *ACM Comput. Surv.*, v. 55, n. 14s, jul. 2023.
12. WU, Wensheng. SQL2X: Learning SQL, NoSQL, and MapReduce via Translation. *In: SIGCSE '21*. New York, NY, USA: Association for Computing Machinery, 2021. DOI: <https://doi.org/10.1145/3408877.3432541>.