

ESTUDO COMPARATIVO ENTRE OS MODELOS DE BANCO DE DADOS NOSQL E SQL

COMPARATIVE STUDY BETWEEN NOSQL AND SQL DATABASE MODELS

Larissa V. Souza¹, Marcelo T. Boer²

¹Faculdade de Tecnologia Prof. José Camargo – Fatec Jales, larissa.souza39@fatec.sp.gov.br

²Faculdade de Tecnologia Prof. José Camargo – Fatec Jales, marcelo.boer@fatec.sp.gov.br

Trabalho de Graduação apresentado à Faculdade de Tecnologia Prof. José Camargo - Fatec Jales, como requisito parcial para obtenção do título de Tecnólogo em Sistemas para Internet

RESUMO

Devido à expansão acelerada da utilização da rede mundial de computadores e ao uso de aparelhos eletrônicos para acessar aplicações como *e-commerces*, streaming de multimídia e redes sociais, têm gerado um volume cada vez maior de dados, complexos e não estruturados. Dessa forma, a necessidade de armazenar grandes quantidades de dados está sempre aumentando e isso requer melhores soluções para gerenciar os dados. Com isso, novas abordagens de manipulação de dados são necessárias a fim de atender às crescentes demandas do mercado. Dessa forma, os bancos de dados NoSQL estão sendo cada vez mais reconhecidos como alternativas ao modelo relacional para manipulação de dados. O presente artigo consiste em uma pesquisa bibliográfica exploratória e descritiva que visa fazer um estudo comparativo entre bancos de dados relacionais e não relacionais com o objetivo de descrever sobre as principais características que constituem cada modelo. Dessa forma, este trabalho tem o intuito de contribuir para o conhecimento sobre cada abordagem e auxiliar na escolha do que melhor se encaixa em cada projeto. O resultado observado é que os bancos de dados não relacionais tem foco no desempenho, sendo mais atrativo para alto volumes de dados, enquanto os bancos de dados relacionais prezam por uma estrutura bem definida, com foco na consistência. Assim, concluiu que tanto os bancos de dados relacionais quanto os não relacionais têm particularidades específicas, e a escolha entre um ou outro dependerá das necessidades particulares de cada aplicação.

Palavras-chave: banco de dados; relacional; SQL; não relacional; NoSQL.

ABSTRACT

Due to the accelerated expansion of the use of the World Wide Web and the use of electronic devices to access applications such as e-commerce, multimedia streaming and social networks, they generate an increasing volume of data, complex and not secure. Therefore, the need to store large amounts of data is ever increasing and this requires better solutions to manage the data. With this, new approaches to data manipulation are necessary in order to meet the growing demands of the market, so NoSQL databases are being increasingly recognized as alternatives to the relational model for data manipulation. This article consists of an exploratory and descriptive bibliographical research that aims to make a comparative study between relational and non-relational databases with the objective of describing the main characteristics, advantages and disadvantages that constitute each model, with the intention of contributing to the knowledge about each approach and help in choosing what best fits each project. The observed result is that non-relational databases focus on performance, being more encouraging for high volumes of data, while relational databases value a well-defined structure, with a focus on consistency. Thus, he concluded that both relational and non-

relational databases have specific characteristics, and the choice between one or the other will depend on the particular needs of each application.

Keywords: database; relational; SQL; non-relational; NoSQL.

1 INTRODUÇÃO

Atualmente, a tecnologia se encontra em um contexto que abrange todos os processos, como na saúde, educação, vendas, com isso o crescimento da quantidade de informações na web é perceptível, devido a popularização da internet. Com o advento da pandemia de COVID-19, fez se necessário uma mudança nos hábitos, e a partir daí, o meio virtual apresentou uma transformação.

A internet teve maior demanda durante a pandemia devido a migração de atividades essenciais para o meio digital, e o resultado mostrou a resiliência da rede em um momento de crise sanitária (BARBOSA, 2021). O número de usuários de internet no Brasil é de 152 milhões de pessoas, em torno de 81% da população (CETIC, 2021).

O isolamento, ensino e o trabalho à distância ocorridos na pandemia deixaram marcas permanentes no relacionamento com a área de TI, e o esperado é que seja uma integração que potencializa as habilidades humanas e com as digitais (MEIRELLES, 2022).

Diante dessa nova conjuntura, com a ampliação do uso da internet, de softwares e aplicativos, os dados sofreram aumento significativo. Segundo a Corporação Internacional de Dados (INTERNATIONAL DATA CORPORATION – IDC, 2022), no ano, mais de cem mil exabytes (EB, unidade de medida de armazenamento de informações digitais para indicar o tamanho dos dados, que equivale a 1 bi de gigabytes, GB) de dados foram gerados e estima-se que esse número ultrapasse os 221.000 EB até 2026.

Com isso, o armazenamento dos dados e informações também tiveram uma mudança relevante, surgindo a necessidade de aprimorar a relação com os bancos de dados, tornando-os indispensáveis, principalmente com o seu armazenamento nas nuvens (*cloud computing*) e a forma como essas informações são armazenadas quer em imagens, vídeos e sons.

Hajjaji et al. (2021) considera que com o início do big data (conjunto de grandes quantidades de dados) em que os dados são oriundos de diversas fontes com diferentes estruturas, análises sobre alternativas de armazenamento de dados estão sendo imprescindíveis, em que é posto em comparativo o modelo tradicional de banco de dados relacional com modelos não relacional, onde o uso de tabelas e relacionamentos definidos vai dando espaço para chave-valor, coluna, documentos e grafos.

Portanto, essa pesquisa se faz necessária para o conhecimento dos principais conceitos sobre os bancos de dados Não Relacionais, chamados de NoSQL assim como suas diferenças em relação aos bancos de dados Relacionais, chamados de SQL.

Deste modo, este trabalho tem por objetivo fazer uma análise teórica a partir da pesquisa bibliográfica em obras existentes, sobre as características, usabilidade e funcionamento do modelo NoSQL em relação ao SQL.

Assim sendo, o estudo visa mostrar os fundamentos dos modelos relacionais e não relacionais, comparando-os, com o intuito de entender suas particularidades e contribuir para uma escolha apropriada de qual banco de dados será mais adequado ao perfil de cada aplicação.

2 REFERENCIAL TEÓRICO

Segundo Elmasri e Navathe (2018, p. 19), modelo de dados é um “conjunto de conceitos que podem ser usados para descrever a estrutura de um banco de dados”. É importante pois representa o ambiente observado e fornece processos de validação, é a documentação do projeto das bases de dados, onde são descritos as representações, os relacionamentos, tipos e restrições.

Banco de dados (BD) “é uma coleção de dados inter-relacionados, representando informações sobre um domínio específico” (KORTH; SILBERSCHATZ; SUNDARSHAN, 2019, p. 22). São conjuntos de dados relacionados, isto é, considerando dados como toda informação que pode ser armazenada com um propósito dentro de um contexto de interesse.

É uma coleção lógica e coerente de dados com algum significado, que atende alguma proposta específica, possuindo um grupo de usuários definidos e algumas aplicações pré-concebidas, conforme a necessidade desses usuários (ELMASRI; NAVATHE, 2018). Os objetivos para criação de um banco de dados são reduzir o espaço ocupado pela informação, facilitar a atualização da informação, aumentar a velocidade de pesquisa e evitar a repetição da informação.

Um banco de dados precisa ser eficiente, possibilitando ao usuário o acesso de suas informações, para isso, segue os seguintes fatores: controle de redundância, para não ter o mesmo dado armazenado em duplicidade; consistência, onde o dado representa um atributo real do sistema modelado e restrição de integridade, que são as regras e condições estabelecidas para satisfazer as relações no banco de dados (AMADEU, 2015).

Na criação de um banco de dados é necessário a utilização de um sistema gerenciador de banco de dados (SGBD). Segundo Meira (2014, p. 4) “é uma coleção de programas que permitem ao usuário definir, construir e manipular bases de dados para as mais diversas finalidades”. É um software com recursos específicos para facilitar a manipulação das informações, é o responsável por tornar o banco de dados gerenciável.

Os SGBDs surgiram para atender à necessidade de armazenamento e de recuperação de grandes volumes de informações, propiciando um ambiente seguro e adequado. Seu principal objetivo é retirar da aplicação cliente a responsabilidade de gerenciar o acesso, manipulação e organização dos dados. Estes sistemas disponibilizam uma interface para que os seus usuários possam incluir, alterar ou consultar dados.

2.1 BANCO DE DADOS RELACIONAL

O modelo relacional foi introduzido por Codd (1970) no artigo “*A Relational Model of Data for Large Shared Data Banks*” teve sua base teórica na matemática a partir da teoria dos conjuntos, que trata das propriedades dos conjuntos que são coleções de objetos chamados elementos, e a lógica de predicado, que denota uma relação entre objetos de um determinado contexto de discurso.

Segundo Elmasri e Navathe (2018, p. 90), “o modelo relacional representa o banco de dados como uma coleção de relações,” onde cada relação é representada por uma tabela, sendo as linhas os registros e as colunas os atributos dos dados. Para Date (2015, p. 100), “um banco de dados relacional é percebido pelo usuário como uma coleção de relações normalizadas de vários graus que se modificam ao longo do tempo”.

São características dos bancos de dados relacionais a integridade referencial, que garante a precisão e a consistência dos dados dentro de um relacionamento entre tabelas (CHAPPLE, 2020). A normalização dos dados, um processo que busca otimizar e consistir os bancos de dados, evitando a duplicação de informações e as dividindo em tabelas de acordo com o tipo de elemento que está sendo armazenado (KOKAY, 2012), o uso das propriedades ACID para transações e a utilização de uma linguagem própria, SQL (*Structured Query Language*).

Uma transação representa um conjunto de operações de leitura ou escrita que são realizadas no banco de dados. A execução de transações tem o fim de garantir o correto funcionamento do sistema e a respectiva consistência dos dados, a partir das propriedades ACID (TRUICA et al., 2015).

O termo ACID significa: atomicidade, onde as transações são atômicas, executam completamente ou não executam; consistência, em que as transações criam estados válidos,

devem manter o estado consistente; isolamento, garante que uma transação em andamento não sofra interferência de outras transações; durabilidade, os dados estarão disponíveis na sua forma correta mesmo se o sistema falhar ou reiniciar, não serão perdidos (KORTH, SILBERSCHATZ, SUNDARSHAN, 2019).

O modelo relacional possui uma linguagem padronizada para o gerenciamento de dados, que é fundamental, a SQL (*Structured Query Language*). Foi criada originalmente pela IBM inspirada na álgebra relacional e tem seu grande destaque pela facilidade do seu uso (CALANCA, 2023).

Structured Query Language (SQL), traduzida como linguagem de consulta estruturada, é uma linguagem padrão para criação e manipulação de bancos de dados, foi criada para facilitar o gerenciamento de dados armazenados em bancos relacionais. É reconhecida pela ANSI (*American National Standards Institute*) e pela ISO (*International Organization for Standardization*) como linguagem universal de consultas.

A linguagem SQL oferece uma interface facilitada para interação com o banco (WADE; CHAMBERLIN, 2012). Segundo França (2022), essa linguagem apresenta uma organização estrutural definida que divide seus comandos em 5 subconjuntos diferentes, da seguinte maneira:

a. DML (*Data Manipulation Language*) é a linguagem de manipulação de dados que define os comandos usados para manipular os dados armazenados em um banco, inclui os operadores inserir, excluir e alterar os registros de uma tabela.

b. DDL (*Data Definition Language*) é a linguagem de definição de dados onde apresenta comandos usados para gerenciar as estruturas do banco de dados, é usada para especificar relações, domínios, regras de integridade, inclui comandos para definir, atualizar e remover tabelas e índices;

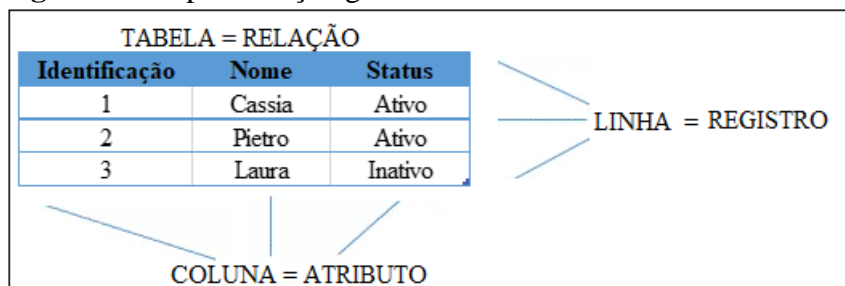
c. DQL (*Data Query Language*) é a linguagem de consulta de dados, define o comando utilizado para selecionar e filtrar os dados de uma tabela;

d. DCL (*Data Control Language*) é a linguagem de controle de dados, que estabelece restrições e permissões de acesso, são comandos para controlar a segurança.

e. DTL (*Data Transaction Language*) é a linguagem de transação de dados que faz o gerenciamento e controle das transações feitas no banco, são comando que permitem iniciar, confirmar e desfazer determinadas alterações.

A Figura 1 é utilizada para exemplificar como é o desenho do modelo relacional de dados baseado em uma tabela de valores que representa o relacionamento dos dados, onde a linha é o registro de cada dado relacionado e a coluna é o atributo que define a relação dos dados.

Figura 1 – Representação gráfica do modelo relacional de dados



Fonte: Elaborada pela autora baseado em PETERSON, 2023.

2.2 BANCO DE DADOS NÃO RELACIONAL (NoSQL)

O modelo de dados não relacional surgiu a partir das limitações do modelo relacional envolvendo a escalabilidade em relação ao crescimento de dados. A arquitetura relacional

não se encaixava mais na demanda de volume de dados, e uma única arquitetura não seria mais a solução para todos os casos (STONEBRAKER et al., 2018).

É uma nova geração de banco de dados, chamada NoSQL, que é um acrônimo para “*Not Only SQL*”, que em português se traduz “Não Apenas SQL”, foi introduzido inicialmente por Carlo Strozzi em 1998 a partir de uma solução de banco de dados relacional de código aberto que não oferecia uma interface SQL (CRUMMENAUER, 2020).

O termo ganhou a conotação usada hoje, quando Eric Evans e Johan Oskarsson em 2009, o utilizaram em um evento para se referir a bancos de dados distribuídos. A expressão em si concerne ao agrupamento de alguns modelos de dados que se distanciam da abordagem relacional, podendo ter em comum as características de serem de código-aberto, distribuídos, horizontalmente escaláveis (NOSQL DATABASE, 2013).

O modelo de dados NoSQL é uma abordagem semi ou não estruturada de gerenciamento de dados que preza desempenho, escalabilidade e disponibilidade (MARTINS FILHO, 2015), e não exigem esquemas fixos como o modelo relacional (SOUZA et al., 2014). Ficaram populares entre as empresas geradoras de conteúdo, e são difundidos na comunidade *open source* e software livre (OLIVEIRA, 2015).

Os bancos de dados não relacionais são mais flexíveis quanto às propriedades ACID e são fundamentados nas propriedades BASE, (BINANI; GUTTI; UPADHYAY, 2016):

- a. Basicamente disponível - todos os dados são distribuídos, mesmo quando há uma falha, o sistema continua em funcionamento;
- b. Estado leve - não há garantia de consistência, isto é, não precisa ser consistente o tempo todo;
- c. Eventualmente consistente – o sistema garante que caso não ocorra novas mudanças nos dados, eventualmente haverá consistência dos dados.

Segundo Sadalage e Fowler (2019), Gonzalez et al. (2016), Gesseter e Ritter (2016) e Rockenbach et al. (2018), os bancos de dados não relacionais são classificados em bancos de dados de chave-valor, banco de dados de documentos, armazenamento em famílias de colunas e banco de dados de grafos.

a. Chave-valor - os dados são armazenados como pares chave-valor, são semelhantes a um dicionário de palavras, onde os dados são endereçados por uma chave (palavra), os valores são isolados e independentes, e o relacionamento é tratado pela lógica da aplicação. Como sua estrutura é simples, as consultas tendem a serem mais rápidas comparadas as bases de dados relacionais, além de suportar uma grande quantidade de dados e concorrência, é útil em diversas situações como armazenar imagens, documentos, dados de sessões de usuários.

b. Documentos - os documentos são conjuntos de atributos e valores, onde um atributo pode ser multivalorado, as chaves dentro dos documentos são únicas, cada documento contém um identificador, que é único dentro do conjunto. O banco de dados armazena e recupera documentos, os quais podem ser XML, JSON, BSON, entre outros. Esses documentos são estruturas de dados na forma de árvores hierárquicas e auto descritivas, constituídas de mapas, coleções e valores escalares. É um modelo ideal para armazenar dados de páginas da internet, catalogação de documentos gerenciamento de inventários entre outros.

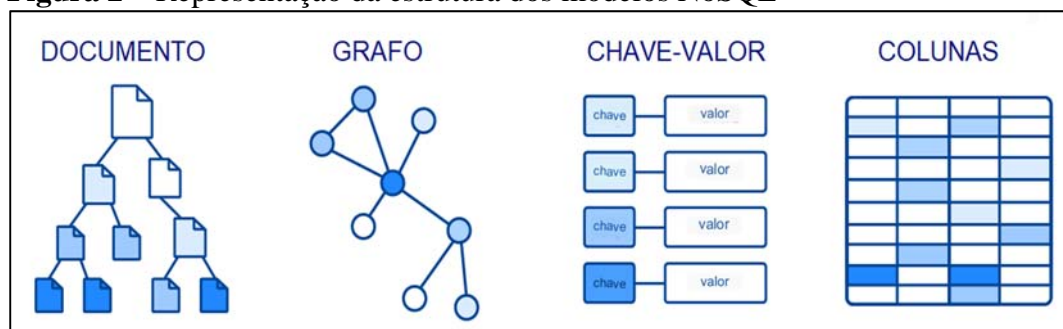
c. Colunas - pode ser considerado como um esquema que organiza os dados com base em uma distribuição de colunas, os grupos de dados são relacionados e, frequentemente, são acessados juntos. Usa uma tabela como modelo de dados, mas não permite associação entre as tabelas. Os dados são armazenados por colunas, cada coluna é um índice do banco de dados, processa consultas concorrentemente, todos os registros fazem parte da mesma tabela, mas cada um deles pode ter colunas diferentes; um dos modelos mais complexos, oferece flexibilidade e escalabilidade para o banco de dados.

d. Grafos - aqui os dados são dispostos no formato de arcos conectados por arestas (relacionamento), como um conjunto de linhas conectadas por vértices (nós). A organização do

grafo permite que os dados sejam armazenados uma vez e depois interpretados de formas diferentes baseadas em relacionamentos. O foco desse tipo de banco de dados é a criação de relacionamentos entre os dados, através desse tipo de banco é possível a construção de redes complexas de relacionamentos e com inferência para revelação de novos dados, é o mais especializado dos tipos de armazenamento NoSQL, tendo aplicações mais específicas como, as redes sociais.

Por meio da Figura 2 exemplifica-se os quatros modelos de banco de dados NoSQL em relação a sua estrutura.

Figura 2 – Representação da estrutura dos modelos NoSQL



Fonte: BANHUDO, 2020.

3 METODOLOGIA

A metodologia utilizada neste estudo foi definir, examinar, avaliar, analisar, interpretar e oferecer uma resposta ao objetivo estabelecido no trabalho: explicar sobre o NoSQL e SQL, a partir da pesquisa bibliográfica.

Foi utilizada a pesquisa bibliográfica, que deve se concentrar no aprofundamento detalhado dos trabalhos correlatos ao assunto da pesquisa. A proposta foi reunir informações relevantes sobre o assunto abordado e usá-las para sua fundamentação (IMÃNA-ENCINAS; SANTANA, 2019). Realizou a leitura e análise em torno de 50 pesquisas condizentes ao tema.

A investigação do tema foi feita a partir de materiais já publicados, como livros, revistas, jornais, panfletos, monografias, artigos científicos, dissertações, teses, publicações em periódicos, simpósios, conferências, internet para consulta em bases de dados online, repositórios, buscadores e ferramentas de pesquisa, como o Google Acadêmico, Scielo, RIC-CPS, IEEE Xplore, entre outros.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

As pesquisas realizadas sobre os modelos de bancos de dados relacional e não relacional evidenciou que para o desenvolvimento de aplicações existe lugar para utilização de ambos os modelos, devendo levar em consideração as características e requisitos do projeto.

A Figura 3 exemplifica em que cada modelo de banco de dados é usualmente mais bem utilizado, de acordo com as necessidades do projeto desenvolvido e seus objetivos, como por exemplo na área de jogos, internet das coisas e redes sociais, o NoSQL é mais procurado, já na área de dados financeiros, o SQL é o mais comum, sendo também possível usar os dois tipos de bancos de dados, como em casos de aplicativos móveis, desktop, web (ALMEIDA, 2020).

Figura 3 – Cenários de utilização dos bancos SQL e NoSQL



Fonte: ALMEIDA, 2020.

Com base nas pesquisas de Perez (2022), Knijnik (2022) percebe -se que em aplicações de jogos online e internet das coisas (IoT), onde se tem uma alta quantidade de dados em tempo real e múltiplos usuários simultaneamente, o banco de dados NoSQL possui a escalabilidade e velocidade necessárias para um melhor desempenho.

Para aplicativos de compartilhamento de conteúdo, as redes sociais, o uso do NoSQL também é mais vantajoso pois permite armazenar e recuperar facilmente dados não estruturados, tendo uma manipulação rápida, sendo possível também uma distribuição geográfica, como exemplo, o Facebook e Twitter que utilizam esse modelo de banco de dados (CRUMMENAUER, 2020).

No entanto, quando a necessidade é de segurança e recursos avançados de consultas estruturadas aos dados, como em casos que envolvem a análise de dados financeiros, exemplo *business intelligence* e *data warehouse*, o modelo de banco de dados SQL é mais recomendado, pois garante a integridade e consistência necessária (MARCHIORI, 2022).

A Tabela 1 é o ranking dos SGBS mais conhecidos de acordo com o DB-Engines (2023a), onde a classificação é feita através da popularidade, tendo como métrica o número de pesquisas e menções, relevância nas redes sociais, oferta de empregos e perfis profissionais que envolvam o sistema.

Tabela 1 – Classificação dos sistemas gerenciadores de banco de dados em junho/2023

Classificação	DBMS	Modelo	Pontuação	Licença
1	Oracle	Relacional	1231,48	Comercial
2	MySQL	Relacional	1163,94	Código aberto
3	Microsoft SQL Server	Relacional	930,06	Comercial
4	PostgreSQL	Relacional	612,82	Código aberto
5	MongoDB	Não Relacional (Documento)	425,36	Código aberto
6	Redis	Não Relacional (Chave Valor)	167,35	Código aberto
7	IBM Db2	Relacional	144,89	Comercial
8	Elasticsearch	Não Relacional (Motor de pesquisa)	143,75	Código aberto
9	Microsoft Access	Relacional	134,45	Comercial
10	SQLite	Relacional	131,21	Código aberto
11	Snowflake	Relacional	114,13	Comercial
12	Cassandra	Não Relacional (Coluna)	108,55	Código aberto
13	MariaDB	Relacional	97,31	Código aberto
14	Splunk	Não Relacional (Motor de pesquisa)	89,45	Comercial
15	Amazon DynamoDB	Não Relacional (Documento/Chave Valor)	79,9	Comercial

Fonte: Adaptado de DB-ENGINES, 2023b.

Na Figura 4 é exemplificado alguns dos sistemas gerenciadores de banco de dados existentes no mercado de acordo com seu modelo, relacional e não relacional, sendo o Oracle, MySQL, PostgreSQL e SQLServer os mais conhecidos na área do SQL e MongoDB, Cassandra, Redis e Neo4j os mais utilizados dentro de suas categorias, documento, coluna, chave-valor e grafos, respectivamente, no âmbito do NoSQL.

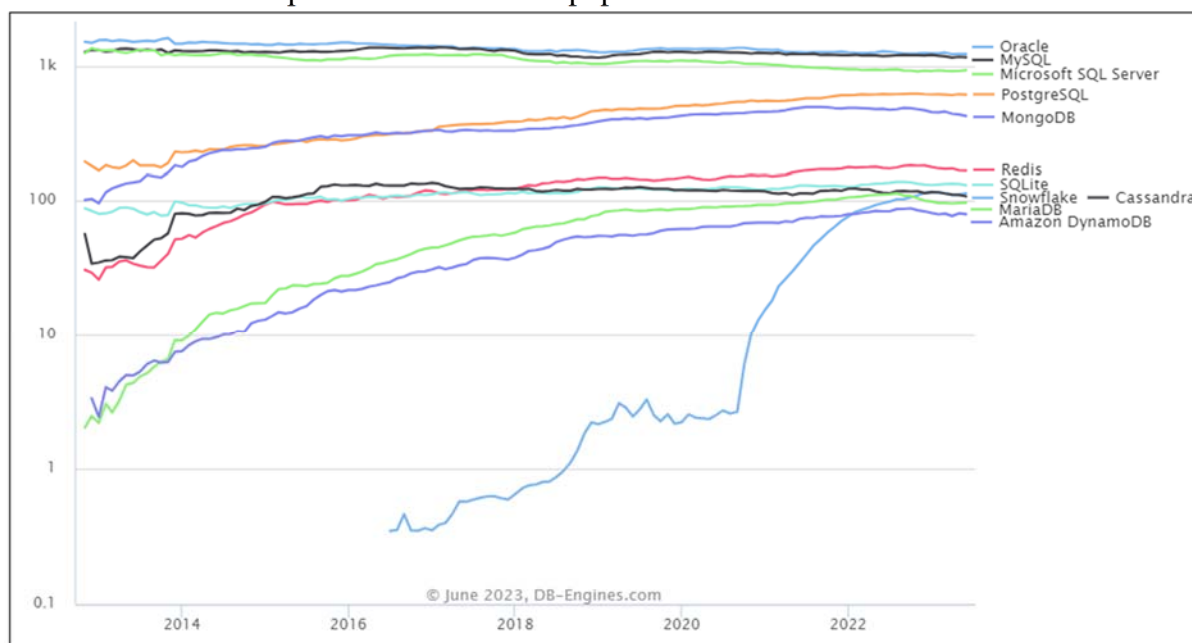
Figura 4 – Sistemas gerenciadores de banco de dados relacional e não relacional



Fonte: Elaborada pelos autores.

O Gráfico 1 demonstra a evolução da popularidade nos últimos 10 anos dos principais banco de dados. Percebe se que os principais SGBDs de modelo relacional, Oracle, MySQL, SQLServer e PostgreSQL se mantêm no topo, porém os modelos não relacionais vem em ascensão ao longo do tempo, como o MongoDB, Redis, Cassandra e Amazon DynamoDB, corroborando as informações vistas na Tabela 1.

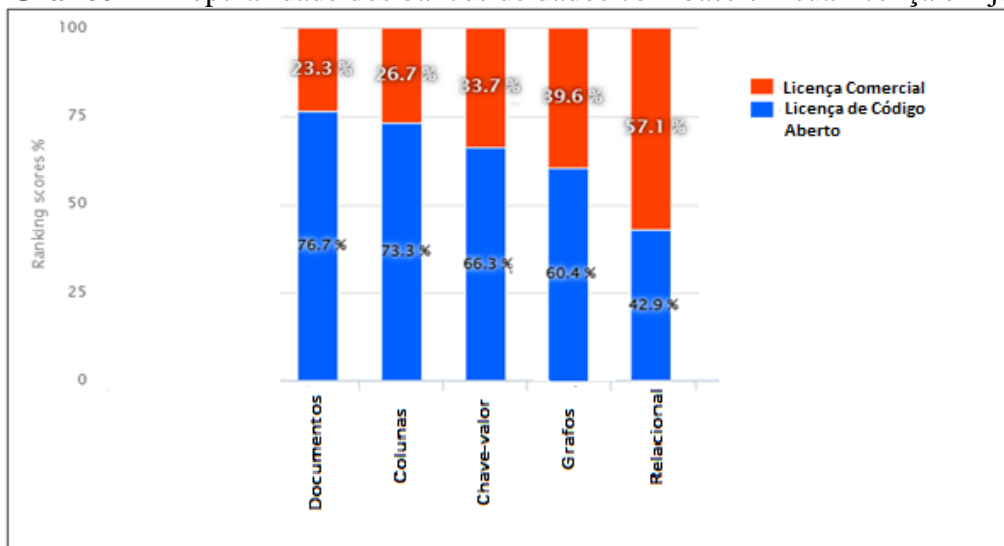
Gráfico 1 – Série temporal da tendência de popularidade dos SGBDs



Fonte: DB-ENGINES, 2023b.

No Gráfico 2 é exposto a popularidade das ferramentas de software de acordo com sua licença de uso. Percebe-se que os quatros tipos de banco de dados presentes no NoSQL têm sua maioria de SGBDs como *open source*, isto é, o código fonte é disponível gratuitamente. No caso dos SGBDs que pertencem ao relacional, ocorre o inverso, a maioria é de uso comercial, isto é, com licença paga. Evidenciando que a popularidade das ferramentas de software está ligada ao modelo de banco de dados pertencente.

Gráfico 2 – Popularidade dos bancos de dados com base em sua licença em junho/2023



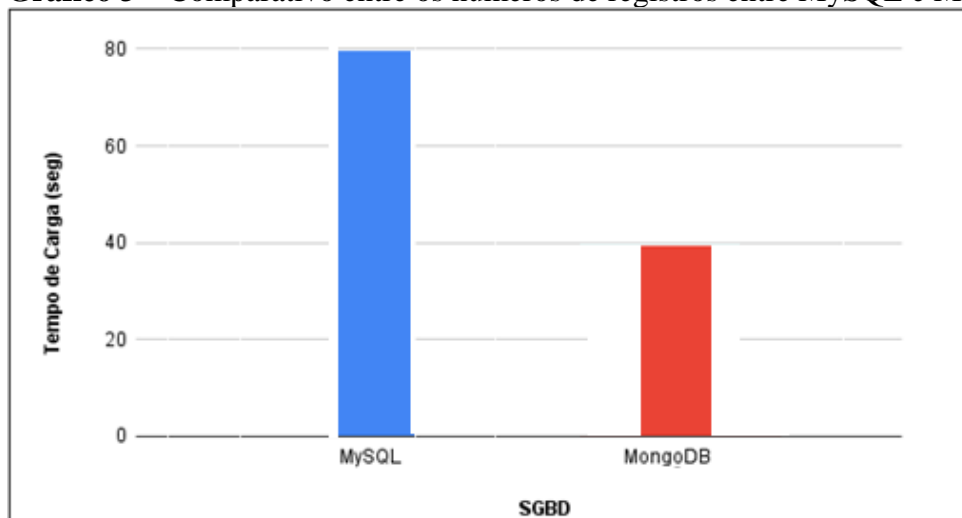
Fonte: DB-ENGINES, 2023c.

Foram analisados trabalhos correlatos na área que explanavam sobre o uso prático dos sistemas gerenciadores de banco de dados SQL e NoSQL e comparavam o desempenho dos dois por meio de testes usando os mais variados fatores.

Lau (2021) fez uma comparação entre o MySQL e MongoDB considerando para a análise a modelagem e projeto, criação do banco de dados, inserção de dados, consultas e espaço de armazenamento, utilizou como base dados o BDQueimadas do Inpe.

O Gráfico 3 apresenta o total de números de registros em cada um dos bancos de dados, onde fica perceptível que o número de registros total foi menor para o MongoDB do que para o MySQL.

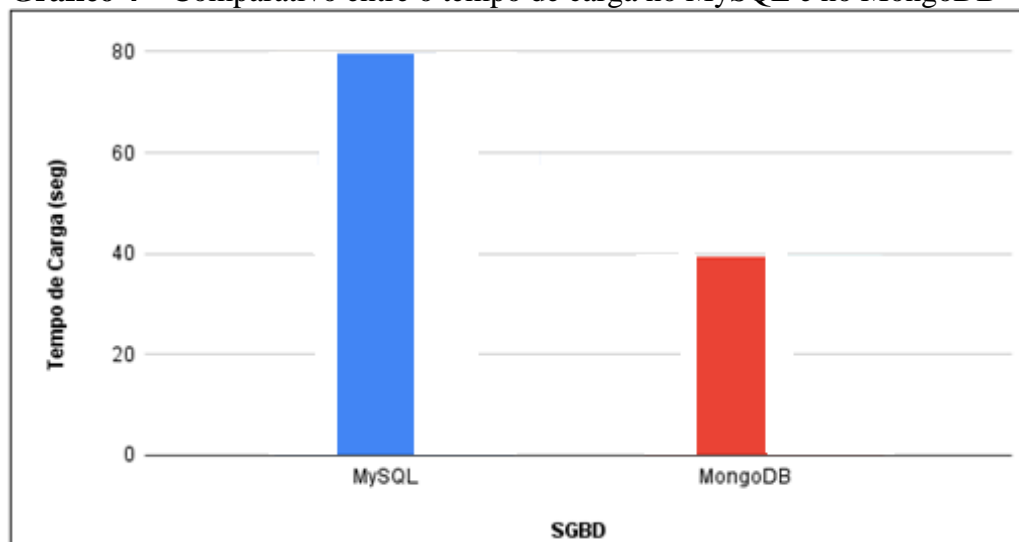
Gráfico 3 – Comparativo entre os números de registros entre MySQL e MongoDB



Fonte: LAU, 2021.

O Gráfico 4 apresenta o tempo de carga na inserção dos dados nos dois bancos de dados, MySQL e MongoDB, em que o MongoDB inseriu os dados de forma mais rápida se comparado ao tempo obtido pelo MySQL.

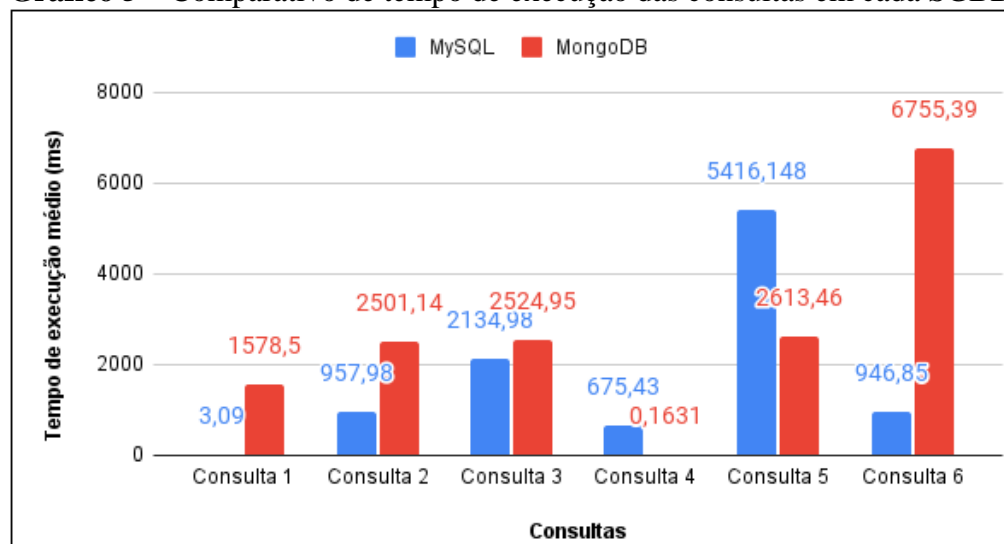
Gráfico 4 – Comparativo entre o tempo de carga no MySQL e no MongoDB



Fonte: LAU, 2021.

O Gráfico 5 apresenta o tempo de execução para todas as consultas criadas entre o MySQL e o MongoDB, onde o MySQL obteve de forma geral uma performance mais eficiente e rápida que o MongoDB. O MySQL possui uma tecnologia mais madura que torna mais otimizado as consultas, porém, com uma carga de trabalho maior, devido aos padrões do SQL, já que o MongoDB possui uma estrutura e linguagem mais simples para essa operação.

Gráfico 5 – Comparativo de tempo de execução das consultas em cada SGBD



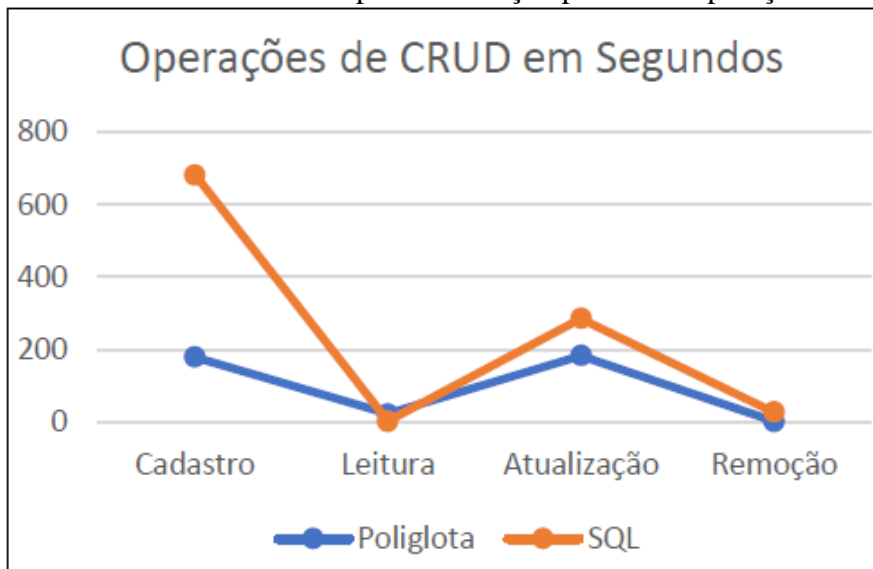
Fonte: LAU, 2021.

Com os gráficos é possível verificar que o resultado obtido foi que apesar do SQL ter uma melhor performance nas consultas, o NoSQL apresentou melhor resultado nos outros aspectos avaliados. Concluiu que o “MongoDB tende a ser um SGBD que aceita uma maior flexibilidade dos dados a serem armazenados, estes podendo variar em seu formato e tipo de informação, e tudo isso é armazenado de uma forma simples e fácil” (LAU, 2021, p. 55).

No artigo escrito por Santos e Alencar (2020), desenvolveram uma aplicação utilizando a persistência poliglota (uso de várias bases de dados e linguagens para lidar com diferentes necessidades de armazenamento), com MongoDB, Redis e MySQL como bases de dados, adotaram como metodologia um estudo de caso para verificarem o consumo de memória, CPU e tempo de execução das operações, em um processo de CRUD (cadastro, leitura, atualização e remoção).

O Gráfico 6 apresenta o tempo de execução das operações em CRUD, o SQL levou o maior tempo em todas as operações realizadas, com exceção da leitura, no cadastro, o tempo médio para inserção dos dados do SQL chegou até 3 vezes mais tempo.

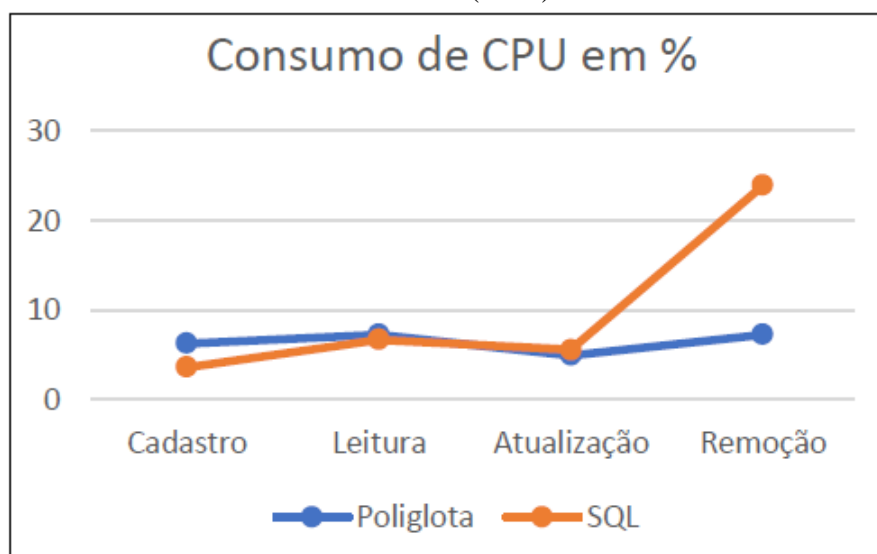
Gráfico 6 – Médias do tempo de execução para cada operação de CRUD



Fonte: SANTOS; ALENCAR, 2020.

O Gráfico 7 apresenta o consumo de CPU, que foi moderado e baixo durante os testes, destacando o maior consumo de CPU na operação de remoção de dados e o menor consumo de CPU na operação de cadastro de dados.

Gráfico 7 – Recursos de Hardware (CPU)



Fonte: SANTOS; ALENCAR, 2020.

Tendo como resultado que o MongoDB e Redis tiveram um nível de desempenho na execução dos testes maior que o MySQL, somente tendo consumido mais recursos de hardware comparado ao MySQL, que na maioria das operações de CRUD, apresentou um tempo maior na execução (SANTOS; ALENCAR, 2020).

Na pesquisa anual com desenvolvedores do Stack Overflow (2023) na categoria de banco de dados, o MongoDB é usado por 25,66% de desenvolvedores profissionais e é o segundo banco de dados mais popular para aqueles que estão aprendendo a codificar atrás apenas do MySQL. Na parte de admirados e desejados, o Redis é um dos bancos de dados mais admirados, com uma porcentagem de 69,92% (STACK OVERFLOW, 2023).

Rudecindo (2022) analisou os bancos de dados relacionais e não relacionais fazendo uma comparação de suas características, fundamentos e desempenho através da gestão da informação de grupos de resgate internacional em desastres, usando os sistemas MySQL e CASSANDRA para realização dos testes práticos com os indicadores de uso de RAM, uso de CPU e resposta.

Na Tabela 2 é exposto o resultado dos testes de desempenhos realizados com o MySQL e Cassandra, observa-se que o Cassandra obteve o menor índice em todos eles, tendo um desempenho consideravelmente menor no uso da CPU ao executar as instruções no banco de dados e na RAM alocada em cada operação do sistema comparado ao MySQL.

Tabela 2 – Desempenho dos bancos de dados MySQL e Cassandra

SGBD	Tempo de consulta	Uso médio da RAM	Utilização média do CPU
MySQL	0,04%	0,12%	1,98%
Cassandra	0,01%	0,05%	0,33%

Fonte: Adaptado de RUDECINDO, 2022.

No trabalho de Rudecindo (2022), o banco de dados Cassandra apresentou melhor desempenho no tratamento das informações, sendo esse sistema recomendado para lidar com grandes volumes de informação.

Fiss, Ferreira, Perez (2020) em seu trabalho apresentou uma análise do desempenho de bancos de dados relacionais, MySQL e orientados a grafos, Neo4j, com base nas operações de inserção e consulta de dados de produção agropecuária com características espaço-temporais.

O Neo4j apresentou melhores resultados, concluindo que “[...] para dados que devem ser consultados tanto por sua chave quanto por valores de seus atributos, como é o caso em sistemas produtivos agropecuários, o desempenho de bancos de dados orientados a grafos, quando as operações são nativas, é significativamente melhor” (FISS; FERREIRA; PEREZ, 2020, p. 216).

O Quadro 1 resume as principais características de cada modelo de banco de dados em que se diferem sobre os fundamentos de como os dados devem ser guardados e recuperados, sendo a comparação entre eles que auxilia na avaliação e escolha de qual estrutura implementar nos projetos adequadamente.

Quadro 1 – Características dos modelos de banco de dados SQL e NoSQL

SQL	NOSQL
Armazenamento de dados estruturados por tabelas	Armazenamento de dados estruturados e não estruturados por colunas, grafos, documentos e chave valor
Esquema fixo	Esquema flexível
Necessidade de predefinição de um esquema de tabela antes da adição de qualquer dado	Fácil adição de colunas e campos de dados não estruturados
Escalabilidade vertical	Escalabilidade horizontal

Linguagem padronizada SQL	Linguagem diversificada e própria de acordo com cada SGBD
O desempenho não é alto em todas as consultas.	Alto desempenho em consultas
Controle de redundância – uso de chaves estrangeiras	Não há preocupação com a redundância
Propriedade ACID com foco na consistência	Propriedade BASE com foco na velocidade
Confidencialidade, integridade e autenticação voltados a segurança	Não possui mecanismos específicos a segurança
Maturidade de suporte maior	Suporte por comunidade independente

Fonte: Adaptado de VICTÓRIA, 2019.

5 CONSIDERAÇÕES FINAIS

Tendo em vista as mudanças no processamento, geração, armazenamento, utilização e natureza dos dados com o passar do tempo, houve a necessidade de alterar o modo de gerenciamento para satisfazer as novas demandas do mercado. Essa foi a grande motivação para o movimento NoSQL, resolver o problema de escalabilidade dos bancos relacionais, com isso, ficando populares na indústria.

O intuito desta pesquisa foi proporcionar conhecimento sobre uma alternativa de gerenciamento de dados através da descrição e fundamentação dos modelos SQL e NoSQL, devido que cada tecnologia de armazenamento de dados atende necessidades específicas, sendo importante conhecer os tipos de banco de dados, suas características e como escolher a melhor opção dada a sua realidade.

Os bancos de dados NoSQL têm características próprias no que diz respeito ao gerenciamento das informações, possuem estruturas dinâmicas baseadas em chave-valor, orientado a colunas, documentos e grafos, enquanto os dados SQL são estruturados de forma estática em tabelas. As diferenças marcantes em relação ao modelo não relacional são velocidade, flexibilidade e escalabilidade no armazenamento e acesso a alto volume de dados não estruturados, que proporcionam um alto desempenho. Já no modelo relacional o que sobressai são os padrões amplamente aceitos e reconhecidos e a linguagem estruturada própria SQL, além da consistência e integridade que possibilitam maior segurança.

Portanto, cada tecnologia atende necessidades específicas de forma a suprir os objetivos de cada aplicação. Dessa forma, a manipulação adequada de dados utilizando tecnologias convenientes para cada situação tem o potencial de tornar os sistemas mais eficientes, seguros e acessíveis para suportar o volume crescente de dados.

Para trabalhos futuros sugere-se a pesquisa quantitativa em relação aos bancos de dados, com uma análise prática sobre como é o funcionamento do SQL e NoSQL em uma aplicação real e um estudo comparativo entre os modelos do NoSQL para melhor entendimento das suas diferenças.

REFERÊNCIAS

ALMEIDA, W. **NoSQL -Já ouviu falar?** 2020. Disponível em: <https://blog.grancursosonline.com.br/nosql-ja-ouvir-falar/>. Acesso em: 15 abr. 2023.

AMADEU, C. V. **Banco de dados**. São Paulo: Pearson, 2015.

BANHUDO, G. **Exploring the NoSQL family**: a (long) primer on a growing requirement for data scientists. 2020. Disponível em: <https://pub.towardsai.net/exploring-the-nosql-family-49e9f23313ad>. Acesso em: 15 abr. 2023.

BARBOSA, A. **TIC Domicílios**. 2021. Disponível em: <https://cetic.br/en/noticia/cresce-o-uso-de-internet-durante-a-pandemia-e-numero-de-usuarios-no-brasil-chega-a-152-milhoes-e-o-que-aponta-pesquisa-do-cetic-br/>. Acesso em: 24 ago. 2022.

BINANI, S.; GUTTI, A.; UPADHYAY, S. SQL vs NoSQL vs NewSQL: a comparative study. **Communications on Applied Electronics**, New York, USA, v. 6, n. 1, p. 43-46, 2016. Disponível em: <https://www.caeaccess.org/archives/volume6/number1/669-669-2016652418>. Acesso em: 20 maio 2023.

CALANCA, P. **SQL e NoSQL**: trabalhando com bancos relacionais e não relacionais. 2023. Disponível em: <https://www.alura.com.br/artigos/sql-nosql-bancos-relacionais-nao-relacionais#o-que-e-um-banco-de-dados-relacional?|Alura>. Acesso em: 27 maio 2023.

CENTRO REGIONAL DE ESTUDOS PARA O DESENVOLVIMENTO DA SOCIEDADE DA INFORMAÇÃO – CETIC. **Cresce o uso de Internet durante a pandemia e número de usuários no Brasil chega a 152 milhões, é o que aponta pesquisa do Cetic.br**. 2021. Disponível em: <https://cetic.br/pt/pesquisa/domicilios/>. Acesso em: 24 ago. 2022.

CHAPPLE, M. **How referential integrity ensures database consistency**. 2020. Disponível em: <https://www.lifewire.com/referential-integrity--definition-1019181>. Acesso em: 27 nov. 2022.

CODD, E. F. A relational model of data for large shared data banks. **Communications of the ACM**, New York, v. 13, n. 6, p. 377-387, jun. 1970.

CRUMMENAUER, C. A. **Motivações para o uso de sistemas NoSQL**. 2020. Trabalho de Conclusão de Curso (Tecnólogo em Sistemas para Internet) – Universidade Federal de Santa Maria, Santa Maria, RS, 2020.

DATE, C. J. **Projeto de banco de dados e teoria relacional**: formas normais e tudo o mais. São Paulo: Novatec, 2015.

DB-ENGINES. **DB-Engines Ranking**. Disponível em: <https://db-engines.com/en/ranking>. Acesso em: 7 jun. 2023a.

DB-ENGINES. **DB-Engines ranking**: trend popularity. Disponível em: https://db-engines.com/en/ranking_trend. Acesso em: 7 jun. 2023b.

DB-ENGINES. **Popularity of open-source DBMS versus commercial DBMS**. Disponível em: https://db-engines.com/en/ranking_osvsc. Acesso em: 7 jun. 2023c.

ELMASRI, R.; NAVATHE, S. B. **Sistema de banco de dados**. 7. ed. São Paulo: Pearson, 2018.

FISS, R. E.; FERREIRA, A. P. L.; PEREZ, N. B. Análise de consultas SQL e CYPHER em dados de produção agrícola. In: MIRANDA, P. (ed.). **Conferências IADIS ibero-americanas WWW/internet e computação aplicada 2020**. [S.l.: s.n.], 2020. p. 212-2016. Disponível em: <https://iadisportal.org/digital-library/an%C3%A1lise-de-consultas-sql-e-cypher-em-dados-de-produ%C3%A7%C3%A3o-agr%C3%ADcola>. Acesso em: 15 maio 2023.

FRANÇA, R. **SQL: O que é e como usar os principais comandos básicos SQL**. 2022. Disponível em: <https://blog.betrybe.com/sql/>. Acesso em: 13 mar. 2023.

GESSETER, F.; RITTER, N. Scalable data management: NoSQL data stores in research and practice. *In: INTERNATIONAL CONFERENCE ON DATA ENGINEERING – ICDE, 32nd.*, 2016, Helsing, Finland. **Proceeding** [...], Helsing, Finland, 2016. p. 1420-1423. Disponível em: <https://doi.ieeecomputersociety.org/10.1109/ICDE.2016.7498360>. Acesso em: 15 nov. 2022.

GONZALEZ, M. T. *et al.* A new model for testing crud operations. *In: INTERNATIONAL CONFERENCE ON ADVANCED INFORMATION NETWORKING AND APPLICATIONS – AINA, 30th.*, 2016, Crans-Montana, Switzerland. **Proceeding** [...], Crans-Montana, Switzerland, 2016. p. 79-86. Disponível em: <https://ieeexplore.ieee.org/document/7474073>. Acesso em: 15 nov. 2022.

HAJAJI, Y. A. *et al.* Big Data and IoT-Based Applications in smart environments: a systematic review. **Computer Science Review**, n. 39, 2021.

IMÃNA-ENCINAS, J.; SANTANA, O. A. **O trabalho científico na metodologia científica**. Brasília: Universidade de Brasília, 2019.

INTERNATIONAL DATA CORPORATION – IDC. **Dados: vantagem operacional ou imperativo estratégico?** 2022. Disponível em: <https://blogs.idc.com/2022/10/17/data-operational-advantage-or-strategic-imperative/>. Acesso em: 15 nov. 2022.

KNIJNIK, D. M. **Comparação de Sgbd's MongoDB e PostgreSQL para jogos digitais**. 2022. Monografia (Graduação em Ciência da Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2022.

KOKAY, M. C. Banco de dados NoSQL: um novo paradigma. **Revista SQL Magazine 102**, Rio de Janeiro, 2012. Disponível em: <http://www.devmedia.com.br/banco-de-dados--nosql-um-novoparadigma-revista-sql-magazine-102/25918>. Acesso em: 15 ago. 2022.

KORTH, H. F.; SILBERSCHATZ, A.; SUDARSHAN, S. **Conceitos de sistemas de banco de dados**. 7. ed. São Paulo: McGraw-Hill, 2019.

LAU, N. A. **Análise comparativa de performance e modelagem entre um banco de dados relacional e um de documentos para a base de dados do BDQueimadas (INPE)**. 2021. Monografia (Graduação em Engenharia da Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2021.

MARCHIORI, L. **Data warehouse: o que é, como funciona e por que usar?** 2022. Disponível em: <https://blog.betrybe.com/dados/data-warehouse/>. Acesso em: 01 jun. 2023.

MARTINS FILHO, M. A. P. **SQL X NoSQL: análise de desempenho do uso do MongoDB em relação ao uso do PostgreSQL**. 2015. Monografia (Graduação em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 2015.

MEIRA, R. **Apostila de banco de dados**. 2014. Disponível em: <http://www.regilan.com.br/>. Acesso em: 3 maio 2023.

- MEIRELLES, F. S. **Panorama do uso de TI no Brasil 2022**. 2022. Disponível em: <https://portal.fgv.br/artigos/panorama-uso-ti-brasil-2022>. Acesso em: 10 set. 2022.
- NOSQL DATABASES. **Your ultimate guide to the non-relational universe!** 2013. Disponível em: <http://nosql-database.org/>. Acesso em: 26 mar. 2023.
- OLIVEIRA, D. **Analytics: comece pequeno e depois amplie, aconselha IBM**. 2015. Disponível em: <https://itforum.com.br/noticias/analytics-comece-pequeno-e-depois-amplie-aconselha-ibm/>. Acesso em: 10 set. 2022.
- PEREZ, L. A. S. **Estudio comparativo entre bases de datos relacional y no relacional**. 2022. Trabalho de Conclusão de Curso (Graduação em Engenharia de Sistemas) – Universidad Tecnica de Babahoyo, Equador, 2022.
- PETERSON, R. **Relational Data Model in DBMS: database concepts & example**. 2023. Disponível em: <https://www.guru99.com/relational-data-model-dbms.html>. Acesso em: 10 maio 2023.
- ROCKENBACH, D. *et al.* **Estudo comparativo de bancos de dados NoSQL**, 2018. Disponível em: <https://revistas.setrem.com.br/index.php/reabtic/article/view/286/131>. Acesso em: 15 out. 2022.
- RUDECINDO, V. C. C. **Análisis comparativo de sistemas gestores de base de datos relacional y no relacional en el contexto del manejo de información de grupos de rescate internacional en desastres**. 2022. Tese (Doutorado em Engenharia de Sistemas) – Universidad Señor de Sipán, Pimentel, Peru, 2022.
- SADALAGE, P.; FOWLER, M. **NoSQL essencial: um guia conciso para o mundo emergente da persistência**. São Paulo: Novatec, 2019.
- SANTOS, K. L.; ALENCAR, A. L. Aplicabilidade da persistência poliglota em sistemas potencialmente escaláveis. **Revista de Engenharia e Pesquisa Aplicada**, v. 5, n. esp., p. 78-87, 2020.
- SOUZA, A. M. *et al.* Critérios para seleção de sgbd nosql: o ponto de vista de especialistas com na base na literatura. *In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO*, 10., Londrina, 2014. **Anais [...]**. Londrina: UEL, 2014.
- STACK OVERFLOW. **Pesquisa anual com desenvolvedores do Stack Overflow**. 2023. Disponível em: <https://insights.stackoverflow.com/survey>. Acesso em: 27 maio 2023.
- STONEBRAKER, M. *et al.* The end of an architectural era: It's time for a complete rewrite. *In: KLAS, W.; NEUHOLD, E. J. (ed.). VLDB '07: proceedings of the 33rd international conference on Very large data bases. [S.l.]: VLDB, 2007. p. 463–489.*
- TRUICA, C. O. *et al.* Performance Evaluation for CRUD Operations in Asynchronously Replicated Document Oriented Database. *In: INTERNATIONAL CONFERENCE ON CONTROL SYSTEMS AND COMPUTER SCIENCE*, 20th., 2015, Bucharest, Romania. **Anais eletrônicos [...]**. Bucharest, Romania, 2015. p. 191-196. Disponível em: <https://ieeexplore.ieee.org/document/7168428>. Acesso em: 15 nov. 2022.

VICTÓRIA, P. **Banco de dados NoSQL**: um manual prático e didático. 2019. Disponível em: <https://blog.geekhunter.com.br/banco-de-dados-nosql-um-manual-pratico-e-didatico/>. Acesso em: 15 abr. 2023.

WADE, B.; CHAMBERLIN, D. **IBM relational database systems**: the early years. **IEEE Annals of the History of Computing**, v. 34, n. 4, p. 38-48, 2012. Disponível em: <https://ieeexplore.ieee.org/document/6297962>. Acesso em: 10 set. 2022.