

Uso de base de dados não relacional na construção de um sistema de apoio à decisão para gestão do desempenho de alunos no ensino superior

Fábio Verruck¹, Roberto Birch Gonçalves²

¹Programa de Pós-Graduação em Computação Aplicada – UCS
Caxias do Sul – RS – Brasil

² Programa de Pós-Graduação em Administração - UCS

fverruck@ucs.br, rbgoncal@ucs.br

Abstract. *This article reports on the process of building a decision support system for managing the performance of undergraduate courses in external assessments, based on ENADE results. To achieve this, an online software was created using a non-relational database and programming in Python, Javascript, and HTML languages. As a result, it was observed that the use of the non-relational database optimized the processing of qualitative information, allowing for more effective data treatment and facilitating the final construction of the application.*

Resumo. *Este artigo relata o processo de construção de um sistema de apoio à decisão para a gestão da performance de cursos de graduação em avaliações externas, com base nos resultados do ENADE. Para tanto, criou-se um software online utilizando-se banco de dados não relacional e programação nas linguagens Python, Javascript e HTML. Como resultado, observou-se que a utilização do banco de dados não relacional permitiu o processamento das informações qualitativas, levando a um tratamento efetivo dos dados e facilitando a implementação do aplicativo.*

1. Introdução

O Exame Nacional de Desempenho dos Estudantes do Ensino Superior (ENADE) tem sido utilizado desde 2004 como instrumento de avaliação do rendimento dos concluintes dos cursos de graduação. A prova avalia a formação dos alunos nas Instituições de Ensino Superior em relação a: (i) conteúdos programáticos previstos nas diretrizes curriculares; (ii) desenvolvimento de competências e habilidades necessárias ao aprofundamento da formação geral e profissional; e o (iii) nível de atualização dos estudantes com relação à realidade brasileira e mundial [INEP, 2023]. Apesar de já ter se consolidado como instrumento de mensuração e ranqueamento do nível de qualidade e excelência dos cursos superiores no Brasil, a utilização dos resultados do ENADE como indutores de melhoria nos cursos ainda está distante de ser uma realidade [Lemos e Miranda, 2015].

Uma das principais causas da subutilização desta prova é a dificuldade dos dirigentes das instituições em utilizar os resultados do ENADE de forma efetiva, seja por inacessibilidade dos dados ou por inabilidade no seu tratamento e análise. Sendo assim, percebe-se que o ENADE falha em atingir o seu propósito inicial, que é o de promover a

melhora constante da formação profissional de nível superior, elevando a qualidade dos cursos a padrões de excelência internacionais.

O presente estudo parte da premissa de que a falta de tratamento adequado das informações geradas a partir dos relatórios do ENADE pode ser solucionada por meio de Sistemas de Apoio à Decisão (SAD) adequadamente implementados. No entanto, a complexidade típica desse tipo de solução acarreta problemas de diversas naturezas, dentre os quais encontram-se aqueles relacionados ao armazenamento e processamento dos dados. Mais especificamente no que diz respeito ao ENADE, a escolha do correto sistema de gerenciamento do banco de dados torna-se crucial para a efetividade do SAD, devido à heterogeneidade dos tipos e da quantidade de informações geradas durante processo avaliativo. Nesse sentido, estudos anteriores [i.e. Kimball e Strehlo 1995; Khan et al. 2023; Kanungo e Morena 2024] têm demonstrado que o uso de sistemas de gerenciamento de dados relacionais limita as possibilidades de processamento e análise necessárias para a tomada de decisão no ambiente gerencial.

Como alternativa, percebe-se que utilização de bancos de dados não relacionais (NoSQL) mostrou-se efetiva em diversas situações, como na construção de sistemas de gestão marítima [Bensalloua e Benameur 2021], no armazenamento de dados clínicos [Lee et al. 2013; Tashkandi et al. 2018] e em sistemas de gestão agrícola [Borrero e Mariscal 2022]. Jose e Abraham (2023) também identificaram a maior efetividade dos sistemas NoSQL em processar dados não estruturados e informações textuais, além da maior eficiência no processamento de dados online (OLAP). Adicionalmente, outros estudos encontraram evidências para uma maior flexibilidade, escalabilidade e disponibilidade dos bancos de dados não-relacionais no uso em aplicações web [i.e. Manoj 2014; Grolinger et al. 2013].

Entre as principais alternativas NoSQL disponíveis para gerenciamento de dados são MongoDB, CassandraDB e Neo4j [Chillón *et al.*, 2024]. Dentre estas, MongoDB, um dos principais bancos de dados orientados a documentos, destaca-se no tratamento de dados semi-estruturados, fornecendo forte consistência dentro de um único documento por meio de operações atômicas [Kanungo e Morena 2024]. Desta forma, o presente artigo descreve a proposição e o teste de um sistema de gerenciamento de banco de dados NoSQL para a construção de um SAD voltado à gestão do desempenho dos estudantes e cursos de graduação em avaliações externas, utilizando a tecnologia MongoDB.

2. Estado da Arte

Os sistemas de apoio à decisão são um subconjunto de sistemas de informação que dão suporte à tomada de decisão humana por meio de soluções computadorizadas, que processam e apresentam informações contextuais [Wang 2023]. Eles permitem que os tomadores de decisão melhorem seu planejamento estratégico e tenham maior efetividade no controle da gestão corporativa. Com o aumento exponencial da quantidade de dados disponíveis e a dificuldade limitada do ser humano de processamento de informações, a demanda por SADs vem se expandindo ao longo do tempo [Bihl et al. 2016].

A construção de um SAD envolve o armazenamento de dados e informações em um banco de dados/estrutura (representação das informações dentro de um determinado modelo) e sua apresentação ao usuário final de forma adequada (interface do usuário). Outra característica fundamental para o planejamento dos SAD atualmente é a sua disponibilização online, de forma a garantir acessibilidade à informação a partir de qualquer lugar, por meio da Internet [Sauter 2014]. Consequentemente, especialmente

devido ao seu modelo de dados normalizado, os bancos de dados relacionais não têm se demonstrado adequados para tratamento online, porque junções e travamentos influenciam negativamente o desempenho em sistemas com dados heterogêneos [Hecht e Jablonski 2011].

Para o planejamento de um SAD com dados analíticos, o banco de dado deve ser facilmente replicável e fornecer um mecanismo integrado de *failover* para lidar com falhas de nó ou datacenter [Khalil e Belaissaoui 2020; Pereira e Costa 2012]. Segundo Manoj (2014), uma vez que as técnicas de replicação oferecidas pelos bancos de dados relacionais são limitadas e esses bancos de dados são tipicamente baseados em consistência em vez de disponibilidade, esses requisitos só podem ser alcançados com esforço adicional e alta experiência. Consequentemente, muitas empresas e organizações desenvolveram seus próprios sistemas de armazenamento, que agora são classificados como bancos de dados NoSQL [Manoj 2014]. Além disso, esse tipo de banco de dados é perfeito para armazenamento de dados em todas as formas: ordenada, parcialmente ordenada e desordenada [Ming et al. 2015].

Os bancos de dados NoSQL são usados principalmente para armazenamento de Big Data porque garantem um alto nível de flexibilidade, um atraso relativamente pequeno durante a leitura e alta produtividade na gravação de dados, o que pode ser alcançado graças ao fato de operarem com base em um ambiente de cluster [Sadalage e Fowler 2013]. Tais soluções necessitam de um modelo que não exija nenhum esquema, evite conexões e geralmente seja escalonado horizontalmente [Kurpanik 2017]. Sem dúvida, os sistemas de apoio à decisão pertencem a este grupo, pois, no caso deles, a velocidade de operação e a precisão são fatores decisivos quando se trata de obter uma vantagem competitiva, e em casos específicos podem influenciar questões tão importantes como o elemento surpresa e uma ação preventiva [Kurpanik 2017; Llano-Ríos et al. 2020].

Com a ascensão dos bancos de dados NoSQL, houve um interesse crescente em seu uso dentro do contexto de sistemas de informação para tomada de decisões, especialmente no desenvolvimento e implementação de *data warehouses* de grande escala [Pticek e Vrdoljak 2017]. Wu et al. (2015) identificaram que sistemas de bancos de dados relacionais frequentemente exigem uma eficiência e capacidade de servidor maiores para serem competentes, mas isso incorre em problemas de custo, enquanto o uso de um banco de dados NoSQL pode facilmente expandir a capacidade sem custos adicionais. Com base nisso, Zhu et al. (2018) propuseram uma nova abordagem para implementar um *data warehouse* com base em um banco de dados orientados a documentos a partir de dois modelos criados para este fim. Bensalloua e Benameur (2021) desenvolveram um Sistema de Apoio à Decisão para o setor marítimo utilizando dados georreferenciados e identificaram que o uso da base de dados MongoDB demonstrou-se consistente para lidar com queries em dados analíticos. Kanungo e Morena (2024) destacam ainda que esta área de pesquisa é promissora e a recente expansão das ferramentas de inteligência artificial abre um novo campo para a investigação e implementação de bancos de dados NoSQL em diversos campos.

3. Descrição do Sistema de Apoio à Decisão criado para o caso

O objetivo do Sistema de Apoio à Decisão proposto é auxiliar nos processos de análise dos resultados do Enade e na proposição de ações com vistas à melhora no desempenho

dos estudantes nas provas, com impacto no desempenho global da instituição. Para a criação do sistema, como estudo de caso foi selecionada uma Instituição de Ensino Superior da Serra Gaúcha. A instituição em análise possui 72 cursos de graduação, a maioria na modalidade presencial. Também está dispersa geograficamente pela serra gaúcha, com oito campi em cidades estratégicas. A universidade possui cursos em todas as áreas do conhecimento, além de programas de *Stricto Sensu* que incluem Mestrado Profissional, Mestrado Acadêmico e Doutorado.

A criação desse sistema iniciou com a definição dos requisitos, com a subsequente modelagem do sistema (incluindo a integração com o banco de dados) e, posteriormente, a criação da interface com o usuário para sua efetiva disponibilização para uso. A modelagem do software foi feita utilizando-se diagramas em UML (*Unified Modeling Language*). Como ação inicial, foram criados modelos de interação, baseados em caso de uso, para estabelecer de forma simplificada a interação entre o sistema e agentes externos. Com base nos casos de uso previstos para o sistema proposto, projetou-se uma arquitetura em camadas para guiar a elaboração dos códigos destinados a estruturar a programação do aplicativo.

4. Implementação

Para a implementação do Sistema de Apoio à Decisão proposto neste trabalho, utilizaram-se de forma integrada diversas ferramentas disponíveis de programação web. A integração de *Python*, *MongoDB*, *Flask*, *JavaScript* e *HTML* demonstrou-se altamente eficaz para o atingimento dos objetivos propostos devido às funcionalidades e características de cada linguagem. A seguir apresenta-se uma visão geral de como essas tecnologias foram combinadas para criar um sistema robusto:

- a. **MongoDB para armazenamento de dados:** MongoDB foi utilizado como o banco de dados para armazenar informações de forma flexível e escalável. O MongoDB é um banco de dados NoSQL que é especialmente adequado para manipular dados não estruturados ou semiestruturados.
- b. **Python para lógica de negócios:** Python foi utilizado para escrever a lógica de negócios do sistema. Python é uma linguagem de programação versátil com uma grande variedade de bibliotecas e *frameworks* que podem ser usados para processar dados, criar modelos analíticos e fornecer cálculos avançados e demonstrou-se bastante eficaz para o caso estudado.
- c. **Flask para o backend:** por se tratar de um aplicativo *web*, foi usado o *framework* Flask para lidar com solicitações HTTP. Flask é uma estrutura leve e flexível que facilita a criação de APIs e aplicativos web, tornando-a ideal para a construção do *backend* do sistema.
- d. **JavaScript e HTML para interface do usuário:** JavaScript e HTML foram usados para criar a interface do usuário com o sistema. JavaScript foi essencial para adicionar interatividade à aplicação *web*, enquanto o HTML foi usado para definir a estrutura da página.
- e. **Comunicação entre o backend e o frontend:** a comunicação entre o Backend Flask e o Frontend JavaScript foi realizada usando API endpoints, especialmente a partir do formato *json*, complementando-se com o template *jinja*. Sendo assim, o *frontend* pode fazer solicitações HTTP para o *backend* para obter e enviar dados

de forma simples e descomplicada. Isso permite que os dados sejam exibidos e manipulados dinamicamente na interface do usuário.

- f. **Segurança e autenticação:** a integração entre os diversos sistemas também permitiu medidas de segurança, como autenticação de usuário e autorização, para proteger os dados e garantir que apenas os usuários autorizados tenham acesso ao sistema.
- g. **Visualização de dados:** foram usadas bibliotecas de visualização de dados, como Matplotlib e Pandas para apresentar informações de maneira compreensível e atraente aos usuários.

Em resumo, a combinação de Python, MongoDB, Flask, JavaScript e HTML permitiu criar um sistema de apoio à decisão eficaz, com armazenamento flexível de dados, lógica de negócios personalizada, uma interface de usuário interativa e recursos analíticos avançados.

4.1 Modelagem do Banco de Dados

Com o intuito de desenvolver uma solução de inteligência de negócios adequada para atender às necessidades dos tomadores de decisão, em relação à gestão de performance dos estudantes do ensino superior, torna-se necessário acessar dados dos quais seja possível extrair algumas métricas. A Tabela 1 resume as métricas que foram consideradas relevantes pelos gestores acadêmicos na fase elicitação dos requisitos.

Para a formalização do modelo, utilizou-se o esquema multidimensional, seguindo a notação proposta em Khalil e Belaissaoui (2020), constituída por fatos, dimensões e funções. Um fato contém os dados a serem analisados chamados medidas, e um conjunto de links para dimensões associadas. Formalmente um fato é definido a partir de um conjunto (N^F, M^F, L^F) em que a letra N corresponde ao nome do fato, a letra M corresponde a um conjunto finito de medidas e L diz respeito a um conjunto de links para dimensões. Já a dimensão armazena atributos que descrevem o objeto no fato a ser analisado. É uma estrutura de dados normalizada composta por uma dimensão única (plana) ou mais hierarquias e um conjunto de atributos que conferem significado às medidas, normalmente são valores textuais descritivos. Essa dimensão também se define a partir de um trio (N^D, P^D, H^D) em que a letra N corresponde ao nome da dimensão, P corresponde ao conjunto de parâmetros e H ao conjunto de hierarquias.

Tabela 1. Métricas necessárias para a tomada de decisão

Métrica	Descrição
Cumprimento das Diretrizes Enade	Verifica se todas as Diretrizes Enade do Curso são contempladas no currículo
Cumprimento das Diretrizes Curriculares Nacionais	Verifica se todas as Diretrizes Curriculares Nacionais são contempladas no currículo
Distribuição dos componentes por ano	Compara a evolução do número de questões em cada Componente Enade por edição da prova
Importância do componente no currículo e na prova	Compara a carga relativa do componente no currículo e na última prova
Carga horária por componente	Identifica a carga horária relativa de cada componente no currículo do curso
Análise do desempenho por componente	Compara a média obtida pelo curso em cada componente com a média estadual

Devido à forma de estruturação dos documentos em MongoDB, torna-se necessário estabelecer para cada registro um conjunto chave-valor único, de maneira a evitar duplicidades e conflito entre os dados armazenados. A chave permite identificar exclusivamente cada registro em uma base, garantindo que não existam documentos iguais. Com base nessas definições, usa-se “{}” para denotar um par chave/valor aninhado, o símbolo “⇒” implica mapear uma chave para seu valor associado, e finalmente “[]” para denotar uma métrica. Para facilitar a modelagem das abordagens mencionadas, o esquema multidimensional foi construído com um único fato, denominado de currículo, relacionado às métricas curriculares de cada curso, com as seguintes medidas:

Currículo = {CurrículoKey ⇒ [Id_Currículo, Curso, Codigo, Disciplina, Enade_1, Enade_2, Dcn, Ch_total, Ch_enade_1, Ch_enade_2]}

Nesse modelo, os dados referentes ao currículo, como código, carga horária total e nome da disciplina são importados de um documento em formato *csv* do próprio sistema de gestão acadêmica da IES, enquanto as informações a respeito dos componentes ENADE e da carga horária destinada a cada um desses componentes são cadastradas pelo coordenador do curso, a partir do acesso ao sistema. Para o cadastro, são utilizadas as dimensões relacionadas a este fato, filtradas de acordo com a sintaxe MongoDB, de forma a permitir que o coordenador do curso cadastre apenas os componentes relacionados ao seu curso específico. Estas informações são usadas pela aplicação construção das visualizações das métricas de análise. A Tabela 2 apresenta um exemplo de um fato sem o cadastro e de um fato após o cadastro.

Tabela 2. Exemplos de representações do fato currículo na base de dados

Antes do Cadastro	Depois do Cadastro
<pre>{ "_id": { "\$oid": "603168739248600789499259" }, "id": "2602", "Codigo": "NUT4014", "Disciplina": "Bioquímica do Exercício", "componente_enade": "", "componente_dcn": "", "CH": "40", "Curso": "Nutrição", "componente_enade_2": 0, "ch_enade_1": 0, "ch_enade_2": 0 }</pre>	<pre>{ "_id": { "\$oid": "603168739248600789499259" }, "id": "2602", "Codigo": "NUT4014", "Disciplina": "Bioquímica do Exercício", "componente_enade": "Nutrição Clínica: (b) Bioquímica nutricional", "componente_dcn": "I - Ciências Biológicas e da Saúde – incluem-se os conteúdos (teóricos e práticos) de base moleculares e celulares dos processos normais e alterados, da estrutura e função dos tecidos, órgãos, sistemas e aparelhos;", "CH": "40", "Curso": "Nutrição", "componente_enade_2": "Alimentos: (a) Composição e bioquímica de alimentos", "ch_enade_1": 80, "ch_enade_2": 20' }</pre>

Cada fato possui cinco dimensões, relacionadas ao conteúdo da prova, ao desempenho do curso, às diretrizes ENADE, às diretrizes curriculares nacionais e aos links para as provas em cada ano:

Prova = {ProvaKey ⇒ [Id_Prova, Pergunta, Componente_enade, Nota_curso, UF, Regiao, Brasil, Ano, Unidade, Curso]}

Desempenho = {DesempenhoKey ⇒ [Curso, Local, Ano, Conceito_Enade, Cpc]}

Links = {LinksKey ⇒ [Ano, Curso, Link]}

Dcn = {DcnKey ⇒ [DCN, Curso]}

Enade = {EnadeKey ⇒ [Componente_Enade, Curso]}

As fontes de dados relacionadas a cada uma das dimensões são diferentes, por isso existe a necessidade de incluir para todas elas o atributo Curso, que serve como elemento agregador para as queries realizadas durante a operacionalização do sistema. Sendo assim, a construção do dashboard acontece a partir das queries que selecionam e agregam as diferentes informações em pipelines, enviados em formato JSON para execução dos scripts nas páginas HTML. A seguir, apresentam-se dois pipelines utilizados para a construção das visualizações das métricas propostas na Tabela 1.

pipeline = [{"\$match" : {"Curso" : user["Curso"]}}, {"\$group": {"_id" : "\$componente_enade", "count": {"\$sum": {"\$toInt": "\$CH"}}}}]

pipeline_b = [{"\$match" : {"Curso" : user["Curso"], "Unidade": user["Unidade"]}}, {"\$group": {"_id": {"ano": "\$Ano", "componente_enade": "\$componente_enade"}, "count": {"\$sum": 1}}},]

O link para o código do projeto está disponível em um repositório do GitHub a partir do seguinte endereço https://github.com/fverruck/Projeto_EXA.git.

5. Considerações Finais

O ENADE tem sido amplamente utilizado para avaliar o desempenho dos concluintes dos cursos de graduação em relação a conteúdos programáticos, competências, habilidades e sua atualização em relação à realidade brasileira e mundial. No entanto, a falta de uso efetivo dos resultados do ENADE tem sido atribuída à inacessibilidade dos dados e à baixa habilidade na análise desses dados por parte dos gestores das IES. O SAD descrito neste artigo visa superar essas limitações, proporcionando uma plataforma que permita aos gestores das instituições de ensino analisar os resultados do ENADE de forma mais eficaz.

O artigo descreve a modelagem do banco de dados utilizado na construção do SAD. Para tanto, escolheu-se a ferramenta MongoDB, um sistema de gerenciamento de dados não relacional baseado em software livre, que utiliza documentos flexíveis em vez de tabelas e linhas, para processar e armazenar várias formas de dados. Essa solução de gerenciamento do banco de dados demonstrou-se efetiva para os fins propostos, por possuir um modelo de armazenamento de dados elástico, cuja adequação aos tipos de entradas usados evitou duplicidade de informações, diminuiu a quantidade de erros na execução das queries e facilitou a elaboração do código-fonte do sistema por utilizarem documentos formatados como JSON.

Devido à complexidade envolvida na elaboração de um empreendimento desta magnitude, reconhece-se a necessidade de aprimorar alguns itens de funcionalidade do sistema, bem como incorporar interfaces mais amigáveis e intuitivas aos usuários. No entanto, considera-se que o atingimento dos objetivos aqui propostos pode ser um ponto de partida interessante para criar ferramentas robustas de análise e suporte a decisão em diversos ambientes organizacionais. Da mesma forma, sugere-se que a continuidade deste trabalho para além das melhorias já indicadas poderia incluir estudos sobre como incorporar ferramentas de inteligência artificial e automação, de maneira a tornar o trabalho do analista e do gestor acadêmico ainda mais efetivo.

Referências

- Bensalloua, C. A., and Benameur, A. (2021) "Towards NoSQL-based Data Warehouse Solution Integrating ECDIS for Maritime Navigation Decision Support System", *Informatica*, v. 45, n. 4, p. 415-431.
- Bihl, T. J., Young, W. A.; and Weckman, G. R. (2016) "Defining, understanding, and addressing big data", *International Journal of Business Analytics (IJBAN)*, v. 3, n. 2, p. 1-32.
- Borrero, J. D., and Mariscal, J. (2022) "A Case Study of a Digital Data Platform for the Agricultural Sector: A Valuable Decision Support System for Small Farmers", *Agriculture*, v. 12, n. 6.
- Chillón, A. H.; Klettke, M.; Ruiz, D. S.; and Molina, J. G. (2024) "A Generic Schema Evolution Approach for NoSQL and Relational Databases". *IEEE Transactions on Knowledge and Data Engineering*, 2024.

- Grolinger, K., Higashino, W.A., Tiwari, A., and Capretz, M.A. (2013) "Data management in cloud environments: NoSQL and NewSQL data stores", *Journal of Cloud Computing: Advance System and Applications*, v. 2, n.1, p.2-24.
- Hecht, R., and Jablonski, S. (2011) "NoSQL Evaluation A Use Case Oriented Survey" 2011 International Conference on Cloud and Service Computing, p. 336-341.
- INEP. (2023) "Relatórios Públicos", [Enade - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira \(inep.gov.br\)](https://inep.gov.br).
- Jose, B., and Abraham, S. (2023). Intelligent processing of unstructured textual data in document based NoSQL databases. *Materials Today: Proceedings*, 80, p. 1777-1785.
- Kanungo, S., and Morena, R. D. (2024) "Concurrency versus consistency in NoSQL databases", *Journal of Autonomous Intelligence*, v. 7, n. 3.
- Khalil, A., and Belaisaoui, M. (2020) "New approach for implementing big datamart using nosql key-value stores", In: 2020 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications (CloudTech). IEEE, p. 1-6.
- Khan, W., Kumar, T., Zhang, C., and Raj, K. (2023). "SQL and NoSQL database software architecture performance analysis and assessments—A systematic literature review", *Big Data and Cognitive Computing*, v. 7, n. 2.
- Kimball, R., and Strehlo, K. (1995), "Why decision support fails and how to fix it", *Acm Sigmod Record*, v. 24, n. 3, p. 92-97.
- Kurpanik, J. (2017) "NoSQL data bases as a data warehouse for decision support systems", *Journal of Science of the Military Academy of Land Forces*, v. 49, n. 1 85.
- Lee, K., Tang, W., and Choi, K. (2013) "Alternatives to relational database: Comparison of NoSQL and XML approaches for clinical data storage", *Computer Methods and Programs in Biomedicine*, v. 110, n. 1, p. 99-109.
- Lemos, K. C. S., and Miranda, G. J. (2015) "Alto e baixo desempenho no ENADE: que variáveis explicam?", *Revista Ambiente Contábil - Universidade Federal do Rio Grande do Norte*, v. 7, n. 2, p. 101-118.
- Llano-Ríos, T. F., Khalefa, M., and Badia, A. (2020) "Evaluating NoSQL systems for decision support: An experimental approach", In: 2020 IEEE International Conference on Big Data (Big Data), p. 2802-2811.
- Manoj V. (2014) "Comparative study of NoSQL document, column store databases and evaluation of Cassandra", *International Journal of Database Management Systems (IJDMS)*, v. 6, n. 4.
- Ming C., Wu, H. Y. F., and Lee, J. (2015) "Comparisons Between MongoDB and MSSQL Databases on the TWC Website", *American Journal of Software Engineering and Applications*, USA, 4(2), 35-41.
- Pereira, J. L., Costa, M. (2019) "From NOSQL databases to decision support systems: developing a business intelligence solution", <https://hdl.handle.net/1822/71603>.
- Pticek, M. and Vrdoljak, B. (2017) "Big data and new data warehousing approaches," *Proceedings of the 2017 International Conference on Cloud and Big Data Computing*, p. 6–10.

- Sadalage, P., and Fowler, M. (2013) “NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence”, Pearson Education – Addison Wesley, Boston.
- Sauter, V. L. (2014) “Decision support systems for business intelligence”. John Wiley & Sons.
- Tashkandi, A., Wiese, I., and Wiese, L. (2018) “Efficient in-database patient similarity analysis for personalized medical decision support systems”, *Big data research*, v. 13, p. 52-64.
- Wang, J. (2023) “Decision Support System Model of Education Management Based on Cloud Storage Technology”, In: *Application of Big Data, Blockchain, and Internet of Things for Education Informatization: Second EAI International Conference, BigIoT-EDU 2022, Virtual Event, July 29–31, 2022, Proceedings, Part I*. Cham: Springer Nature Switzerland, p. 385-393.
- Wu, M. C., Huang, Y. F., and Lee, J. (2015) “Comparisons Between MongoDB and MSSQL Databases on the TWC Website”, *American Journal of Software Engineering and Applications*, v. 4, n. 2, p. 35-41.
- Zhu, J., Zhang, H., Xu, Y., and Zhao, W. (2018) “Database Selection Model Based on Fuzzy Comprehensive Evaluation”, In: *2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)* (pp. 97-100). IEEE.