**Báo cáo Bài 4: Đánh giá nguồn thông tin về Hệ thống Cơ sở Dữ liệu (CSDL) SQL so với NoSQL**

1. **Câu hỏi nghiên cứu**

**Câu hỏi nghiên cứu:** "Ưu điểm và nhược điểm của CSDL SQL so với NoSQL trong việc xử lý dữ liệu lớn trong IT là gì?"

**Từ khóa tìm kiếm:**

* "SQL vs NoSQL comparison"
* "advantages disadvantages SQL NoSQL big data"
* "scalability databases IT"
* "relational vs non-relational databases"
* "SQL NoSQL big data handling"

**Chiến lược tìm kiếm:**

* Sử dụng công cụ tìm kiếm web (tương tự Google) với các truy vấn cơ bản như "SQL vs NoSQL comparison advantages disadvantages big data IT" để tìm các bài viết phổ thông và blog.
* Tìm kiếm học thuật với toán tử site:scholar.google.com "SQL vs NoSQL" để lấy bài báo học thuật.
* Tìm kiếm trên diễn đàn và blog ngành: site:stackoverflow.com "SQL vs NoSQL" và site:towardsdatascience.com "SQL vs NoSQL" .
* Tìm tài liệu từ nhà cung cấp: "IBM SQL vs NoSQL comparison" và "Microsoft SQL vs NoSQL".
* Sử dụng browse\_page để đọc chi tiết nội dung các nguồn chọn lọc, tập trung vào tóm tắt so sánh, ưu nhược điểm, và thông tin meta (ngày, tác giả, bias).
* Tìm kiếm thêm filetype:pdf để lấy báo cáo hoặc bài nghiên cứu dạng PDF, ví dụ từ các trường đại học.
* Tổng cộng tìm được khoảng 20-30 kết quả, từ đó chọn 6 nguồn đa dạng: 2 bài viết phổ thông/blog (GeeksforGeeks, Coursera), 1 bài học thuật (Rochester PDF), 1 bài thực hành (Towards Data Science), và 2 từ nhà cung cấp (IBM, Microsoft).

**3. Bảng đánh giá CRAAP**

Tôi áp dụng bài kiểm tra CRAAP cho 6 nguồn được chọn. Đánh giá dựa trên thang điểm 1-5 (1: kém, 5: xuất sắc) cho từng tiêu chí, kèm nhận xét. Bảng có cột: STT, Nguồn tài liệu, Currency, Relevance, Authority, Accuracy, Purpose, và Nhận xét chung.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| STT | Nguồn tài liệu | Currency (Thời sự) | Relevance (Liên quan) | Authority (Uy quyền) | Accuracy (Chính xác) | Purpose (Mục đích) | Nhận xét chung |
| 1 | GeeksforGeeks: "Advantages and Disadvantages of using SQL vs. NoSQL Databases" (2025) | 5 (Cập nhật tháng 7/2025, bao gồm các ví dụ hiện đại như cloud-native apps) | 5 (Trực tiếp so sánh ưu nhược điểm cho big data, scalability trong IT) | 4 (Nền tảng giáo dục công nghệ, nhưng tác giả không nêu rõ) | 4 (Bảng so sánh chi tiết, tham chiếu CAP theorem, nhưng thiếu dữ liệu thực nghiệm) | 4 (Giáo dục, cân bằng, nhưng nhấn mạnh NoSQL cho big data) | Nguồn tốt cho tổng quan, đáng tin cậy nhưng thiếu dẫn chứng cụ thể. |
| 2 | Coursera: "SQL vs. NoSQL: The Differences Explained + When to Use Each" (2025) | 5 (Cập nhật tháng 10/2025, đề cập cloud và big data hiện đại) | 5 (So sánh scalability, use cases cho big data như IoT) | 4 (Đội ngũ biên tập Coursera, nền tảng giáo dục uy tín) | 4 (Mô tả chính xác với ví dụ công ty lớn, nhưng thiếu số liệu đo lường) | 3 (Giáo dục nhưng quảng cáo khóa học Coursera, có thể thiên vị) | Nguồn hữu ích nhưng cần chú ý quảng cáo. |
| 3 | Rochester PDF: "SQL vs NoSQL: A Performance Comparison" (không ngày) | 3 (Không ngày, nhưng nội dung học thuật từ 2017, có thể lỗi thời với công nghệ mới) | 5 (Tập trung performance cho big data, với benchmark) | 5 (Tác giả sinh viên Đại học Rochester, term paper học thuật) | 5 (Benchmark cụ thể với số liệu tốc độ, tham chiếu sách và hội nghị) | 5 (Nghiên cứu trung lập, không quảng cáo) | Nguồn học thuật chất lượng cao, đáng tin cậy cho dữ liệu thực nghiệm. |
| 4 | Towards Data Science: "SQL vs NoSQL in 8 Examples" (không ngày) | 3 (Không ngày, nội dung từ 2021, ví dụ cơ bản chưa cập nhật mới) | 4 (Ví dụ thực hành cho big data, nhưng tập trung query hơn scalability) | 4 (Nền tảng Towards Data Science, cộng đồng data scientist) | 4 (Code ví dụ chính xác, nhưng thiếu phân tích sâu big data) | 4 (Giáo dục thực hành, cân bằng) | Tốt cho ví dụ code, nhưng không sâu về big data. |
| 5 | IBM: "SQL vs. NoSQL Databases: What's the Difference?" (không ngày) | 3 (Không ngày, nhưng đề cập công nghệ hiện đại như IBM DB2) | 5 (So sánh scalability, use cases cho big data unstructured) | 5 (Tác giả kỹ sư IBM Cloud Databases, công ty công nghệ lớn) | 5 (Tham chiếu ACID/CAP, ví dụ công ty thực tế như Netflix) | 3 (Giáo dục nhưng quảng cáo sản phẩm IBM như DB2, bias bán hàng) | Đáng tin cậy kỹ thuật, nhưng thiên vị sản phẩm IBM. |
| 6 | Microsoft: "Relational vs. NoSQL data" (2022) | 4 (Cập nhật 2022, đề cập Azure và NewSQL hiện đại) | 5 (Tập trung cloud-native, scalability cho big data cao volume) | 5 (Microsoft Learn, tài liệu chính thức cho .NET/Azure) | 5 (Giải thích CAP/PACELC, ví dụ Azure Cosmos DB với số liệu latency) | 3 (Hướng dẫn kỹ thuật nhưng quảng cáo Azure, bias vendor) | Xuất sắc cho IT cloud, nhưng thiên vị Microsoft. |

**4. Tổng hợp kết quả**Từ bảng CRAAP, tôi chọn 3 nguồn đáng tin cậy nhất: Rochester PDF (học thuật, chính xác cao), IBM (kỹ thuật sâu, dù có bias), và Microsoft (cập nhật cloud, bias vendor nhưng chi tiết). Các nguồn không đáng tin cậy hơn như Datascientest (không ngày, quảng cáo đào tạo) và Towards Data Science (cũ, thiếu sâu) bị loại bỏ vì thiếu thời sự và chính xác đo lường.

Tóm tắt những hiểu biết sâu sắc về SQL so với NoSQL trong IT (dựa trên nguồn chọn):

* SQL (Relational Databases): Ưu điểm: Cấu trúc bảng cố định đảm bảo tính nhất quán dữ liệu (ACID), phù hợp cho giao dịch phức tạp và dữ liệu có quan hệ (ví dụ: ngân hàng, e-commerce). Nhược điểm: Mở rộng dọc (vertical scaling) giới hạn xử lý dữ liệu lớn, khó khăn với dữ liệu không cấu trúc, và overhead cao cho join/query phức tạp trong big data (theo Rochester: MySQL chậm hơn NoSQL ở read/write lớn).
* NoSQL (Non-Relational Databases): Ưu điểm: Mở rộng ngang (horizontal scaling) dễ dàng, linh hoạt với dữ liệu không cấu trúc (document, key-value), tốc độ cao cho big data phân tán (ví dụ: IoT, social media; IBM nhấn mạnh CAP theorem ưu tiên availability). Nhược điểm: Tính nhất quán cuối cùng (eventual consistency) có thể gây lỗi dữ liệu, thiếu ACID đầy đủ, và khó query phức tạp (Microsoft: cần de-normalize để tránh join).
* Khác biệt chính trong big data IT: SQL tốt cho tính nhất quán và báo cáo, NoSQL ưu việt cho tốc độ và quy mô lớn (Rochester benchmark: NoSQL nhanh hơn 2-3 lần ở dataset triệu records).

**Suy ngẫm về thiên kiến của nhà cung cấp CSDL trong các nguồn thông tin:** Nhà cung cấp như IBM và Microsoft thường thổi phồng lợi ích của sản phẩm họ (ví dụ: IBM nhấn mạnh DB2 hỗ trợ ACID trong NoSQL, Microsoft quảng bá Azure Cosmos DB cho scalability), dẫn đến bias bán hàng thay vì phân tích trung lập. Điều này có thể làm người dùng bỏ qua nhược điểm như chi phí cao hoặc phức tạp triển khai. Ngược lại, nguồn học thuật như Rochester trung lập hơn, dựa trên benchmark thực nghiệm, giúp tránh nội dung tiếp thị. Tổng thể, cần kết hợp nguồn đa dạng để giảm bias và đảm bảo thông tin cân bằng.

**4. Danh sách tài liệu tham khảo**

|  |  |
| --- | --- |
| [1] | GeeksforGeeks, “Advantages and Disadvantages of using SQL vs. NoSQL Databases,,” Jul 2025. [Trực tuyến]. Available: https://www.geeksforgeeks.org/data-engineering/what-are-the-advantages-and-disadvantages-of-using-sql-vs-nosql-databases. |
| [2] | C. E. Team, “SQL vs. NoSQL: The Differences Explained + When to Use Each,,” Coursera Oct 2025. [Trực tuyến]. Available: https://www.coursera.org/articles/nosql-vs-sql. |
| [3] | R. W. a. Z. Yang, “SQL vs NoSQL: A Performance Comparison,,” Univ. Rochester, [n.d.]., [Trực tuyến]. Available: https://www.cs.rochester.edu/courses/261/fall2017/termpaper/submissions/06/Paper.pdf. |
| [4] | “SQL vs NoSQL in 8 Examples,,” Towards Data Science, [n.d.]., [Trực tuyến]. Available: https://towardsdatascience.com/sql-vs-nosql-in-8-examples-25aebcf49922. |
| [5] | B. Anderson and B. Nicholson, “SQL vs. NoSQL Databases: What's the Difference?,,” IBM Blog, [n.d.]., [Trực tuyến]. Available: https://www.ibm.com/think/topics/sql-vs-nosql. |
| [6] | “Relational vs. NoSQL data,,” Microsoft Learn, Apr 2022. [Trực tuyến]. Available: https://learn.microsoft.com/en-us/dotnet/architecture/cloud-native/relational-vs-nosql-data. |