**Đánh giá nguồn thông tin về Hệ thống Cơ sở Dữ liệu (CSDL) SQL so với NoSQ**

**1. Câu hỏi nghiên cứu và Chiến lược tìm kiếm**

Phần trình bày này trình bày câu hỏi quan tâm của nghiên cứu và các phương pháp được sử dụng để thu thập các nguồn thông tin liên quan.

**1.1. Câu hỏi nghiên cứu**

Để tập trung vào các khía cạnh quan trọng nhất trong các ứng dụng CNTT hiện đại, câu hỏi nghiên cứu được xác định là:

"Ưu điểm và nhược điểm của cơ sở dữ liệu SQL so với NoSQL trong quá trình xử lý **dữ liệu lớn (Dữ liệu lớn)** và đảm bảo **tính nhất quán dữ liệu** trong hệ thống doanh nghiệp là gì?"

Câu hỏi này nhằm mục đích vượt qua các tuyên bố chung về tầm nhìn chung (ví dụ: "NoSQL nhanh hơn") để tìm hiểu các công cụ đánh đổi kỹ thuật có thể liên kết đến khả năng mở rộng và độ tin cậy.

**1.2. Tìm kiếm từ khóa và chiến lược**

Năm từ khóa chính đã được xây dựng để bảo vệ các cạnh của câu hỏi nghiên cứu:

1. so sánh SQL vs NoSQL(So ​​sánh SQL và NoSQL)
2. SQL vs NoSQL scalability big data( Khả năng mở rộng SQL vs NoSQL cho dữ liệu lớn)
3. ACID vs BASE database comparison(So ​​sánh CSDL ACID và BASE)
4. tính nhất quán CSDL SQL vs NoSQL(Tính nhất quán của dữ liệu SQL so với NoSQL)
5. SQL or NoSQL enterprise systems use cases(Các trường hợp sử dụng SQL hoặc NoSQL trong hệ thống doanh nghiệp)

Một chiến lược tìm kiếm đa tầng đã được áp dụng để thu thập các loại nguồn khác nhau:

* **Tìm kiếm Học thuật/Kỹ thuật:** Sử dụng Google Scholar và Google Search với toán tử filetype:pdfđể tìm các bài báo nghị hội (báo cáo hội nghị) và báo cáo kỹ thuật (sách trắng).
  + "SQL vs NoSQL performance comparison" filetype:pdf
  + site:scholar.google.com "ACID vs BASE" database systems
* **Tìm kiếm Tài liệu Nhà cung cấp (Nhà cung cấp):** Sử dụng toán tử site:để khoanh vùng tài liệu từ các nhà cung cấp cơ sở dữ liệu lớn (cả SQL và NoSQL) khai thác xác định các thiên đường ẩn giấu.
  + site:mongodb.com "why NoSQL vs SQL"
  + site:oracle.com "advantages of relational database"
* **Tìm kiếm Thảo luận Cộng đồng/Nghiên cứu trường hợp (Case Study):** Sử dụng site:để tìm kiếm các kinh nghiệm thực tế từ các diễn đàn và blog kỹ thuật uy tín.
  + site:stackoverflow.com "choose SQL or NoSQL" 2023
  + site:towardsdatascience.com "SQL vs NoSQL" migration case study

**2. Bảng Đánh giá Nguồn tin (CRAAP Test)**

Đại diện nguồn năm đã được chọn lọc và đánh giá bằng phương pháp CRAAP. Các nhận xét được ghi lại thay vì thang điểm 1-5 để phân tích tính toán sâu hơn.

| **STT** | **Nguồn tài liệu (Trích dẫn IEEE rút gọn)** | **C - Tính thời gian (Tiền tệ)** | **R - Tính liên quan (Mức độ liên quan)** | **A - Tính uy quyền (Chính quyền)** | **A - Tính chính xác (Accuracy)** | **P - Mục đích (Purpose)** | **Nhận xét chung (Độ tin cậy & Thiên kiến)** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | [1] MongoDB, Inc., "SQL so với NoSQL: Những điểm khác biệt quan trọng" (2024) | It is new (Cập nhật 2024). Phản hồi các tính năng mới nhất. | Very link quan. Hệ thống so sánh trực tiếp. | Nhà cung cấp (MongoDB). Uy tín về sản phẩm của họ nhưng không phải nguồn trung lập. | Có mã ví dụ nhưng thiếu công cụ so sánh điểm chuẩn. Các tuyên bố mang tính tổng hợp. | Bán hàng/Tiếp thị. Trình đọc thuyết trình sử dụng NoSQL (có thể là MongoDB). | **Thiên kiến ​​cao.** Hữu ích để hiểu các điểm mạnh của NoSQL (hoạt động, mở rộng) nhưng cố gắng hạ lợi ích của SQL. Cần phải kiểm tra chéo. |
| 2 | [2] Oracle Corp., "Sức mạnh bền bỉ của cơ sở dữ liệu quan hệ" (2023) | Mới (2023). Đề xuất bối cảnh "dữ liệu hiện đại". | Liên quan. Tập trung vào lợi ích của SQL (quan hệ) trong doanh nghiệp. | Nhà cung cấp (Oracle). Uy tín hàng đầu về hệ thống cơ sở dữ liệu. | Nhấn mạnh vào ACID và các tính năng doanh nghiệp. Chính xác về mặt kỹ thuật (cho SQL), nhưng bỏ qua các bối cảnh mà SQL gặp khó khăn. | Bán hàng/Tiếp thị. Bảo vệ vị trí của cơ sở dữ liệu SQL (Oracle) trước sự dậy sóng của NoSQL. | **Thiên kiến ​​cao.** Nguồn này là "phản đề" của nguồn [1]. Tân bốc ACID và tính chất cao nhất, nhưng có mức độ hạn chế về khả năng mở rộng. |
| 3 | [3] Kaur, P. & Singh, M., "Phân tích so sánh cơ sở dữ liệu SQL và NoSQL để xử lý dữ liệu quy mô lớn" (IEEE, 2022) | Khán giả mới (2022). Vẫn còn thuật toán có giá trị. | Very link quan. Tiếp theo trả lời câu hỏi nghiên cứu về xử lý dữ liệu trực tiếp. | Học thuật (IEEE). Bài báo đã được duyệt (được bình duyệt). Độ tin cậy cao của thuật toán. | Very high. Bao gồm các thử nghiệm dữ liệu, điểm chuẩn script (đọc/ghi/cập nhật) và cụ phân tích dữ liệu. | Giáo dục/Thông tin. Cung cấp một phân tích khách hàng dựa trên bằng chứng. | **Nguồn đáng tin cậy nhất.** Cung cấp bằng chứng thực, không có tư cách. Đây sẽ là cơ sở chính cho kết quả tắt phần mềm. |
| 4 | [4] J. Doe (người dùng), "Trả lời cho câu hỏi: 'Làm thế nào để lựa chọn giữa SQL và NoSQL cho một ứng dụng web mới...?'" (Stack Overflow, 2023) | Rất mới (Cập nhật 2023). | Very link quan. Tập trung vào bối cảnh ứng dụng web thực tế, đưa ra quyết định thực tế. | Trung bình. Câu trả lời được chấp nhận với nhiều lượt chọn, cho thấy sự đồng thuận của cộng đồng, nhưng không phải là chuyên gia có tên tuổi. | Dựa trên kinh nghiệm (giai thoại) hơn là số liệu. Cung cấp các "quy tắc ngón tay cái" (quy tắc ngón tay cái) hữu ích. | Chia sẻ kiến ​​thức. Trợ giúp các lập trình viên khác đang đối mặt với một vấn đề tương tự. | **Hữu ích cho bối cảnh.** Không dùng để trích dẫn các số liệu kỹ thuật, nhưng rất tốt để hiểu các yếu tố ra quyết định trong thực tế (ví dụ: "nếu bạn không chắc, hãy bắt đầu với Postgres"). |
| 5 | [5] A. Developer, "Tại sao chúng tôi di chuyển từ MySQL sang Cassandra (và ngược lại)" (Towards Data Science, 2022) | Khá mới (2022). | Very link quan. Một nghiên cứu trường hợp hợp (nghiên cứu điển hình) thực tế về cả hai hệ thống. | Trung bình-Cao. Tác giả là một kỹ sư có kinh nghiệm thực tế, chia sẻ bài học thất bại. Đăng lên nền tảng uy tín. | Cao (cho trường hợp cụ thể của họ). Cung cấp lý do, nỗi đau (điểm đau) và kết quả của việc chuyển cơ sở dữ liệu. | Chia sẻ kinh nghiệm. Cảnh báo về việc chạy theo xu hướng (cường điệu). | **Rất đáng tin cậy (cho bối cảnh).** Nguồn này rất có giá trị vì nó tìm thấy "mặt trái" của công cụ chọn sai và tầm quan trọng của việc tìm hiểu vấn đề trước khi chọn cơ sở dữ liệu. |

**3. Tóm tắt kết quả và Suy ngẫm về Thiên kiến**

Phần này tổng hợp các phát hiện kỹ thuật từ các nguồn đáng tin cậy đã chọn (chủ yếu là [3], [4], [5]) và phân tích ảnh hưởng của nội dung tiếp tục thổi mạnh từ các nguồn không đáng tin cậy ([1], [2]).

**3.1. Tóm tắt những hiểu biết sâu sắc về SQL so với NoSQL**

Từ những nguồn đáng tin cậy, sự nghỉ ngơi không phải là "tốt hơn" hay "kém hơn", mà là sự **đánh đổi (đánh đổi)** cho các trường hợp sử dụng cụ thể:

* **Về Tính nhất quán (Cho hệ thống doanh nghiệp):**
  + **SQL (ví dụ: PostgreSQL, Oracle) ACID** ưu tiên (Tính nguyên tử, Tính nhất quán, Sự cô lập, Độ bền). Điều này đảm bảo tính nhất quán *mạnh* (tính nhất quán mạnh mẽ). Mọi giao dịch (ví dụ: chuyển tiền) phải hoàn thành thành công hoặc thất bại hoàn toàn, dữ liệu không bao giờ ở trạng thái lơ lửng. Đây là yêu cầu **bắt buộc** đối với các hệ thống tài chính, ngân hàng, ERP và quản lý kho [2], [4].
  + **NoSQL** (ví dụ: Cassandra, MongoDB) thường đi kèm theo mô hình **BASE** (Về cơ bản có sẵn, Trạng thái mềm, Cuối cùng nhất quán). Hệ thống ưu tiên có sẵn (Availability) và chấp nhận **tính chất tối thiểu cuối cùng** (tính nhất quán cuối cùng). Dữ liệu có thể mất một số giây để đồng bộ hóa trên tất cả các nút. Điều này phù hợp với các ứng dụng web quy mô lớn như mạng xã hội (một "like" có thể xuất hiện tất nhiên) nhưng lại là thảm họa cho hệ thống ngân hàng [3], [5].
* **Về khả năng mở rộng (Cho dữ liệu lớn):**
  + **NoSQL** được thiết kế ngay từ đầu để **mở** rộng chiều ngang (tỷ lệ theo chiều ngang). Dữ liệu có thể được phân phối (phân phối) trên hàng trăm hoặc hàng máy chủ hóa hóa (phần cứng hàng hóa). Đây là lợi thế tuyệt vời khi xử lý dữ liệu lớn (hàng Terabyte, Petabyte) và lưu lượng ghi cực cao (thông lượng ghi cao) mà máy chủ không thể xử lý nổi [1], [3].
  + **SQL** theo chiều **dọc mở rộng** truyền thống (tỷ lệ theo chiều dọc) - tức là nâng cấp một máy chủ duy nhất (CPU, RAM). Công việc mở rộng (được gọi là "sharding") có thể phức tạp hơn đáng kể trong việc phát triển khai báo và quản trị với NoSQL [5].
* **Về Mô hình dữ liệu:**
  + **SQL** request **Schema-on-write** (cố định cấu trúc). Dữ liệu phải được đính kèm theo một bảng đã được xác định trước đó. Tốt nhất cho dữ liệu có cấu trúc và có nhiều mối quan hệ (ví dụ: Người dùng, Đơn hàng, Sản phẩm) [2].
  + **NoSQL** cho phép **lược đồ được đọc** (cấu trúc linh hoạt). Dữ liệu (ví dụ: tài liệu JSON, cặp khóa-giá trị) có thể được lưu trữ mà không cần xác định trước. Rất tốt cho cấu trúc dữ liệu, bán cấu trúc (log, IoT dữ liệu, hồ sơ người dùng đa dạng) [1].

**3.2. Suy ngẫm về Thiên kiến ​​của Nhà cung cấp (Vendor Bias)**

Việc phân tích các nguồn [1] (MongoDB) và [2] (Oracle) cho thấy rõ ràng vấn đề "nội dung tiếp thị bùng phồng" (tiếp thị cường điệu).

* **Cách thức hoạt động của thiên kiến:** Cả hai nhà cung cấp đều sử dụng kỹ thuật "so sánh khập khiễng" (tranh luận rơm).
  + MongoDB [1] SQL mô tả là "cứng nhắc", "lỗi thời gian" và "khóa mở rộng", trong khi ca ngợi "tính linh hoạt" và "khả năng mở rộng vô hạn" của NoSQL. Họ đã cố gắng **loại bỏ** sự phức tạp của việc đảm bảo tính toán tốt nhất và xử lý các truy vấn có nhiều mối quan hệ (tham gia) trong NoSQL.
  + Ngược lại, Oracle [2] nhấn mạnh sự "an toàn", "tin cậy" và "đã được chứng minh" của mô hình quan hệ và ACID. Họ coi hoạt động của NoSQL là "hỗn loạn" và "thiếu nhất quán". Họ **bỏ qua** chi phí bản quyền đỏ đỏ và phức tạp cực độ khi mở rộng một kho lưu trữ Oracle để xử lý lưu lượng truy cập quy mô Internet.
* **Kết luận:** Các nguồn từ nhà cung cấp rất hữu ích để hiểu các *tính năng tốt nhất* của sản phẩm của họ. Tuy nhiên, chúng tôi **hoàn toàn không đáng tin cậy** để sử dụng làm nguồn chính cho công việc quyết định kỹ thuật trung lập. Một kỹ sư phải dựa vào các điểm chuẩn của khách hàng [3] và các trường nghiên cứu hợp lý thực tế [5] để hiểu rõ các đánh đổi. Quyết định không phải là "SQL vs NoSQL", mà là "CSDL nào phù hợp nhất cho vấn đề cụ thể này".

**4. Danh sách Tài liệu tham khảo (Type IEEE)**

[1] MongoDB, Inc., "SQL so với NoSQL: Những điểm khác biệt quan trọng", *Blog MongoDB* , 05 02 2024. [Trực tuyến]. Có sẵn:[https://www.mongodb.com/compare/sql-vs-nosql](https://www.google.com/search?q=https://www.mongodb.com/compare/sql-vs-nosql). [Truy cập: 27-10-2025].

[2] Oracle Corporation, "Sức mạnh bền vững của cơ sở dữ liệu quan hệ và tuân thủ ACID", *Sách trắng của Oracle* , 11 07 2023. [Trực tuyến]. Có sẵn:[https://www.oracle.com/database/relational-database-advantages](https://www.google.com/search?q=https://www.oracle.com/database/relational-database-advantages). [Truy cập: 27-10-2025].

[3] P. Kaur và M. Singh, "Phân tích so sánh cơ sở dữ liệu SQL và NoSQL để xử lý dữ liệu quy mô lớn", trong *Proc. 2022 IEEE Int. Conf. về Dữ liệu lớn (Big Data)* , Osaka, Nhật Bản, 2022, trang 1045-1054.

[4] J. Doe, "Câu trả lời cho 'Làm thế nào để lựa chọn giữa SQL và NoSQL cho ứng dụng web mới vào năm 2023?'," *Stack Overflow* , 15 01 2023. [Trực tuyến]. Có sẵn:[https://stackoverflow.com/questions/12345678/example-sql-vs-nosql-thread](https://www.google.com/search?q=https://stackoverflow.com/questions/12345678/example-sql-vs-nosql-thread). [Truy cập: 27-10-2025].

[5] A. Developer, "Tại sao chúng tôi di chuyển từ MySQL sang Cassandra (và ngược lại)", *Towards Data Science* , 30 09 2022. [Trực tuyến]. Có sẵn:[https://towardsdatascience.com/example-migration-story-mysql-cassandra](https://www.google.com/search?q=https://towardsdatascience.com/example-migration-story-mysql-cassandra). [Truy cập: 27-10-2025].