การศึกษาเปรียบเทียบระบบการจัดการฐานข้อมูลเชิงสัมพันธ์และ NoSQL : กรณีศึกษาฐานข้อมูลนักศึกษามหาวิทยาลัยราชภัฏกำแพงเพชร

Comparative Analysis of Database Management Approaches: Relational and NoSQL Databases - A Case Study of the Student Database at Kamphaeng Phet Rajabhat University

นรุตม์ บุตรพลอย[1] และ พรหมเมศ วีระพันธ์[2]
[1]คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏกำแพงเพชร, E-Mail: narut@kpru.ac.th
[2]คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏกำแพงเพชร E-Mail: phrommate_v@kpru.ac.th

วันที่รับบทความ 16 พฤษภาคม 2567
วันแก้ไขบทความ 19 มิถุนายน 2567
วันตอบรับบทความ 26 มิถุนายน 2567

บทคัดย่อ

ปัจจุบันมหาวิทยาลัยต้องเผชิญกับความท้าทายในการขยายตัวของข้อมูลภายในมหาวิทยาลัยแนวคิดการทบทวนการปรับฐานข้อมูลเชิงสัมพันธ์ไปเป็นฐานข้อมูล NoSQL ถูกนำเสนอขึ้นมาเพื่อรองรับแนวโน้มข้อมูลที่ขยายขึ้น ผู้วิจัยจึงศึกษาแนวทางการจัดการฐานข้อมูลของซอฟต์แวร์จัดการฐานข้อมูล mysql และ mongodb ผ่านข้อมูลงานที่เกี่ยวข้องกับนักศึกษา วัตถุประสงค์ของงานวิจัยได้แก่ (1) เพื่อทดลองเวลาการตอบสนองของการสืบค้นข้อมูลที่ซับซ้อนและความยืดหยุ่นในการพัฒนาสคีมาซอฟต์แวร์จัดการฐานข้อมูลเชิงสัมพันธ์ด้วย mysql และซอฟต์แวร์จัดการฐานข้อมูล NoSQL ด้วย mongodb (2) เพื่อวิเคราะห์เชิงเปรียบเทียบการจัดการฐานข้อมูลเชิงสัมพันธ์กับฐานข้อมูล NoSQL จากการทดลองพบว่า mongodb แสดงเวลาตอบสนองสืบค้นข้อมูลที่เร็วกว่าและการพัฒนาสคีมาที่ยืดหยุ่นมากกว่าเมื่อเปรียบเทียบกับ mysql แต่ถึงอย่างไรการศึกษาชี้ให้เห็นว่าฐานข้อมูล NoSQL ยังไม่ใช่ทางเลือกสำหรับแทนที่ในงานฐานข้อมูลของกรณีศึกษาเนื่องจากการศึกษารายละเอียดของระบบงานพบว่าข้อมูลเพิ่มขึ้นเป็นช่วงเวลาไม่เกิดขึ้นทุกวัน แต่ทั้งนี้ผู้วิจัยเล็งเห็นถึงประโยชน์ของการจัดการข้อมูล NoSQL ในการเก็บข้อมูลที่อาจก่อให้เกิดข้อมูลขนาดใหญ่ได้แก่ข้อมูลการใช้งานเครือข่ายอินเทอร์เน็ตของบุคลากรและนักศึกษา ผลการทดลองนี้จะถูกนำไปเสนอต่อผู้บริหารในเรื่องการพัฒนาฐานข้อมูลของมหาวิทยาลัยต่อไป

คำสำคัญ: ฐานข้อมูลเชิงสัมพันธ์, ฐานข้อมูล NoSQL, สคีมา, คุณสมบัติอะตอมมิก, MongoDB

ABSTRACT

Nowadays, universities are confronted with the challenge of managing expanding data within their campuses. To address this issue, there has been a proposal to reconsider the transition from relational databases to NoSQL databases, aiming to support the growing trend of data expansion. Consequently, this study examines the database management methodologies of MySQL and MongoDB software within the context

of student-related information. The research objectives are twofold: (1) to evaluate the response time of complex queries and the flexibility in schema development between MySQL, a relational database management software, and MongoDB, a NoSQL database management software; (2) to conduct a comparative analysis between relational database management and NoSQL databases. Experimental findings indicate that MongoDB demonstrates quicker query response times and greater schema development flexibility compared to MySQL. Nevertheless, investigations suggest that adopting NoSQL databases as a replacement for the current database system in the case study is not feasible due to the observed irregular data growth patterns. However, the researcher recognizes the potential benefits of employing NoSQL data management for capturing data that may lead to substantial data volumes, such as internet usage data of both staff and students. The outcomes of this experiment will be presented to administrators to guide further enhancements in the university's database infrastructure.

1. บทนำ

เทคโนโลยีการสื่อสารในปัจจุบันเจริญก้าวหน้าไปอย่างรวดเร็ว ความก้าวหน้านี้ส่งผลให้เกิดนวัตกรรมใหม่เกี่ยวกับการสร้างข้อมูล การแลกเปลี่ยนข้อมูล ก่อเกิดข้อมูลขนาดใหญ่ตามมา ข้อมูลขนาดใหญ่มีความหลากหลายในหลายรูปแบบทั้งที่เป็นข้อมูลภาพ เสียง ตัวเลข ตัวอักษร ข้อมูลต่าง ๆ มีการเปลี่ยนแปลงอย่างรวดเร็ว และมีคุณค่าต่อการนำไปใช้งาน การก่อเกิดข้อมูลที่มหาศาลนี้หลายหน่วยงานที่เป็นเจ้าของข้อมูลต้องทบทวนถึงรูปแบบการจัดการข้อมูลเนื่องจากนับวันขนาดข้อมูลจะใหญ่ขึ้น ฐานข้อมูลเชิงสัมพันธ์ (RDBMS) ถูกนำมาพัฒนาสำหรับการจัดเก็บและการจัดการข้อมูลมาอย่างช้านานการจัดการข้อมูลเชิงสัมพันธ์นี้มีการบังคับสคีมาที่ต้องกำหนดไว้ล่วงหน้าเพื่อให้มั่นใจถึงความสมบูรณ์ ความสัมพันธ์ระหว่างตาราง และการรักษาโครงสร้างที่ชัดเจนเป็นไปตามหลักการจัดการข้อมูลแบบฐานข้อมูลเชิงสัมพันธ์ แต่ทั้งนี้เมื่อถึงยุคข้อมูลขนาดใหญ่ทำให้เกิดการเปลี่ยนแปลงกระบวนทัศน์ซึ่งท้าทายการจัดการฐานข้อมูลเชิงสัมพันธ์ ซึ่งปัจจุบันมีการพัฒนาฐานข้อมูลแบบ NoSQL ที่ช่วยในการจัดการข้อมูลที่มีขนาดใหญ่เปลี่ยนแปลงรวดเร็วและมีความหลากหลาย ฐานข้อมูล NoSQL ใช้แนวทางแบบไม่ใช้สคีมาทำให้แตกต่างจากสคีมาที่เข้มงวดของฐานข้อมูลเชิงสัมพันธ์ทำให้การเข้าถึงข้อมูลทำได้อย่างรวดเร็วส่งผลให้แอปพลิเคชันที่เรียกใช้ฐานข้อมูลนี้มีการตอบสนองที่รวดเร็ว

ในขณะที่มหาวิทยาลัยต้องมีระบบจัดเก็บและจัดการข้อมูลนักศึกษาทั้งในด้านลงทะเบียนเรียน การยืมคืนหนังสือในห้องสมุด งานด้านกิจกรรมนักศึกษา การเข้าใช้ระบบเครือข่ายที่นักศึกษาใช้ช่องทางการสื่อสารผ่านระบบเครือข่ายของมหาวิทยาลัย ข้อมูลเหล่านี้กลายเป็นเรื่องสำคัญในการพิจารณาการจัดการฐานข้อมูล ความสามารถในการกระจายข้อมูลไปยังหลายโหนด การปรับขนาดแนวนอนของเอกสารทำให้การจัดการฐานข้อมูลแบบ NoSQL เป็นเรื่องที่น่าสนใจสำหรับแอปพลิเคชันที่เกี่ยวข้องกับชุดข้อมูลจำนวนมหาศาลและปริมาณงานแบบไดนามิก ในการทดสอบหาซอฟต์แวร์จัดการข้อมูลที่มีประสิทธิภาพให้แก่มหาวิทยาลัย ตัวเลือกซอฟต์แวร์มีบทบาทสำคัญต่อการตัดสินในการลงทุนการพัฒนาระบบ

ดังนั้นเพื่อเป็นการทดสอบสอบซอฟต์แวร์สำหรับบริหารจัดการข้อมูลสำหรับเป็นข้อมูลสำหรับตัดสินใจเลือกระบบฐานข้อมูลผู้วิจัยจึงทดลองการจัดการข้อมูลของซอฟต์แวร์บริหารจัดการข้อมูลเชิงสัมพันธ์ด้วย mysql และซอฟต์แวร์บริหารจัดการข้อมูลแบบ NoSQL ด้วย mongodb เพื่อทดลองและวิเคราะห์เชิงเปรียบเทียบซอฟต์แวร์จัดการฐานข้อมูลทั้งสอง ผ่านข้อมูลระบบนักศึกษา ได้แก่ ระบบลงทะเบียน ระบบการยืมคืนหนังสือ และระบบกิจกรรมนักศึกษา โดยทดลองเชิงเปรียบเทียบในสองมุมมองแก่ เวลาการตอบสนองของการสืบค้นข้อมูลที่ซับซ้อนและความยืดหยุ่นในการพัฒนาสคีมา ผ่านการจำลองสถานการณ์ การทดสอบ ผลการทดลอง การวิเคราะห์เปรียบเทียบและการสรุปผลการทดลองเพื่อนำมาเป็นข้อมูลสำหรับการพัฒนาระบบจัดการข้อมูลให้กับมหาวิทยาลัยต่อไป

2. วัตถุประสงค์ของการวิจัย

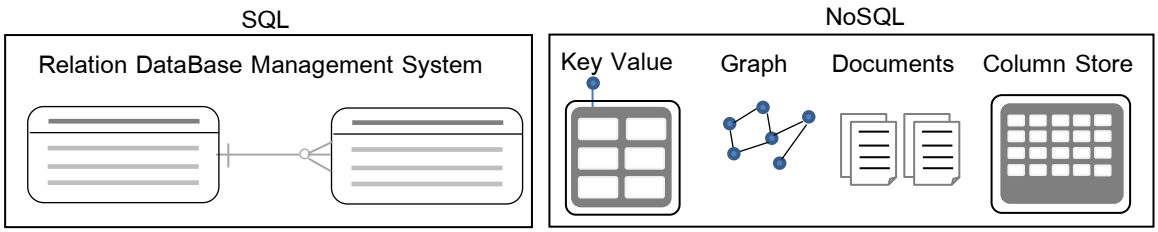
2.1 เพื่อทดลองเวลาการตอบสนองของการสืบค้นข้อมูลที่ซับซ้อนและความยืดหยุ่นในการพัฒนาสคีมาซอฟต์แวร์จัดการฐานข้อมูลเชิงสัมพันธ์ด้วย mysql และซอฟต์แวร์จัดการฐานข้อมูล NoSQL ด้วย mongodb

2.2 เพื่อวิเคราะห์เชิงเปรียบเทียบจัดการฐานข้อมูลเชิงสัมพันธ์กับฐานข้อมูล NoSQL

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ฐานข้อมูลเชิงสัมพันธ์ : ฐานข้อมูลเชิงสัมพันธ์เป็นส่วนสำคัญในการจัดการข้อมูลโดยมีแนวทางจัดการโครงสร้างการจัดระเบียบข้อมูล ฐานข้อมูลผ่านตารางที่มีสคีมาที่กำหนดไว้ล่วงหน้า (Netinant,2023) ฐานข้อมูลเชิงสัมพันธ์จัดระเบียบข้อมูลในตารางที่มีโครงสร้างการบังคับใช้ความสัมพันธ์ระหว่างเอนทิตี ความสอดคล้องของธุรกรรมจะถูกรักษาไว้ด้วยคุณสมบัติอะตอมมิก (Lotfy, 2015) ซึ่งเรียกว่าคุณสมบัติ ACID (Atomicity) ความสม่ำเสมอ (Consistency) การแยกตัว (Isolation) และความทนทาน (Durability) ภาษาที่ใช้สำหรับจัดการฐานข้อมูลเชิงสัมพันธ์ได้แก่ภาษา SQL ซอฟต์แวร์จัดการฐานข้อมูล mysql เป็นระบบจัดการฐานข้อมูลเชิงสัมพันธ์แบบเปิดเผยรหัส มีบทบาทสำคัญสำหรับเป็นตัวจัดการฐานข้อมูลให้แก่หน่วยงานทั้งขนาดเล็กและขนาดใหญ่ มีชื่อเสียงในด้านความน่าเชื่อถือ ความสามารถในการปรับขนาด (Gopi,2023) มีคุณสมบัติ ACID เป็นตัวเลือกอันดับต้น ๆ สำหรับแอปพลิเคชันที่ต้องการการจัดการข้อมูลที่มีโครงสร้างใช้ภาษา SQL สำหรับการสืบค้นและรองรับการสืบค้นที่ซับซ้อนในสถานการณ์ที่ต้องการความสอดคล้อง ทำให้เป็นตัวเลือกที่แพร่หลายในอุตสาหกรรมซอฟต์แวร์ ตั้งแต่ระบบอีคอมเมิร์ซไปจนถึงระบบการศึกษา

ฐานข้อมูล NoSQL : ฐานข้อมูล NoSQL นำเสนอการวิธีการจัดการฐานข้อมูลที่แตกต่างจากฐานข้อมูลเชิงสัมพันธ์ที่จะต้องกำหนดสคีมาที่กำหนดไว้ล่วงหน้าทำให้โครงสร้างข้อมูลมีความเข้มงวด แต่ฐานข้อมูลแบบ NoSQL นำเสนอแนวทางที่ยืดหยุ่นมากกว่า โดยเฉพาะอย่างยิ่งเหมาะสำหรับสถานการณ์ที่เกี่ยวข้องกับข้อมูลที่ไม่มีโครงสร้างอนุญาตให้มีโครงสร้างข้อมูลแบบไดนามิก และไม่มีสคีมาล่วงหน้า มีความสามารถในการปรับขนาดในแนวนอน (Ali,2023) เพื่อรองรับชุดข้อมูลที่เพิ่มมากขึ้น โมเดลข้อมูลใช้โมเดลข้อมูลที่เน้นเอกสาร โดยจัดเก็บข้อมูลไว้ในเอกสารลักษณะ JSON ที่มียืดหยุ่นต่อการนำมาใช้งาน ซอฟต์แวร์จัดการฐานข้อมูล mongodb เป็นซอฟต์แวร์สำหรับจัดการฐานข้อมูล NoSQL โมเดลข้อมูลเชิงเอกสารที่ให้ความยืดหยุ่นและความสามารถในการปรับขนาดได้ (Gyorödi,2023) ทำให้มีความสามารถในจัดการโครงสร้างสคีมาแบบไดนามิกหรือไม่มีโครงสร้าง ความสามารถในการปรับขนาดแนวนอนอย่างมีประสิทธิภาพ ทำให้ mongodb เป็นโซลูชันที่ตอบโจทย์สำหรับแอปพลิเคชันสมัยใหม่ รวมถึงระบบการจัดการเนื้อหา การวิเคราะห์แบบเรียลไทม์



(ก) (ข)

รูปที่ 1 รูปแบบฐานข้อมูลเชิงสัมพันธ์กับฐานข้อมูล NoSQL

รูปที่ 1 แสดงรูปแบบฐานข้อมูลเชิงสัมพันธ์กับฐานข้อมูล NoSQL โดยรูป 1(ก) แสดงรูปแบบฐานข้อมูลเชิงสัมพันธ์ซึ่งแสดงถึงตารางที่มีความสัมพันธ์ระหว่างกันผ่านตัวเชื่อมค่า index ที่มีมุมมองเป็นเส้นเชื่อมระหว่างกัน รูปที่ 1(ข) แสดงรูปแบบฐานข้อมูล NoSQL มีรูปแบบการจัดเก็บเอกสารมีการขยายตัวแบบคอลัมน์ภายในเอกสารมีคู่คีย์-ค่าการสืบค้นผ่านความสัมพันธ์กราฟ ซึ่งเปรียบเทียบแล้วฐานข้อมูลทั้งสองไม่เหมือนกันทั้งในการออกแบบสคีมาและการสืบค้น

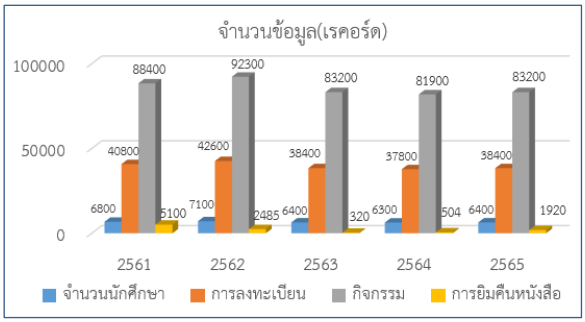
งานวิจัยที่เกี่ยวข้อง

งานวิจัยของ Carvalho (2023) ประเมินประสิทธิภาพและความสามารถในการปรับขนาดของฐานข้อมูลเอกสาร NoSQL สามอันดับแรก ได้แก่ couchbase, couchdb และ mongodb โดยใช้เกณฑ์มาตรฐานการให้บริการคลาวด์ (YCSB) การประเมินเกี่ยวข้องกับการประเมินประสิทธิภาพรันไทม์ของฐานข้อมูลภายใต้ปริมาณงานที่แตกต่างกัน ผลการทดลอง พบว่า couchdb ให้เวลาในการตอบสนองดีกว่าอีกสองระบบ แต่ mongodb มีประสิทธิภาพเหนือกว่าในการทดสอบงานอื่น ๆ เช่น การเปลี่ยนแปลงจำนวนบันทึกและเธรด ประสิทธิภาพรันไทม์ งานวิจัยของ Khan(2017) เปรียบเทียบระหว่างฐานข้อมูลเชิงสัมพันธ์กับฐานข้อมูล NoSQL วิธีการวิจัยเกี่ยวข้องกับการเปรียบเทียบฐานข้อมูลเชิงสัมพันธ์ของ Oracle กับฐานข้อมูลกราฟ NoSQL โดยใช้ชุดข้อมูล MedCare ซึ่งเป็นสคีมาที่พัฒนาขึ้นสำหรับระบบการดูแลสุขภาพของโรงพยาบาล ผลการทดลองระบุว่าฐานข้อมูล NoSQL มีประสิทธิภาพเหนือกว่า Oracle ทั้งนี้ขึ้นอยู่กับการออกแบบสคีมาที่ไม่มีโครงสร้าง งานวิจัยของ Ali (2019) สำรวจความแตกต่างระหว่างฐานข้อมูล SQL และ NoSQL โดยเน้นจุดแข็งและจุดอ่อนสำหรับการวิเคราะห์ข้อมูลขนาดใหญ่ (big data) ผลการทดลองพบว่าข้อมูลขนาดใหญ่มีลักษณะที่ไม่มีสคีมาซึ่งเหมาะสมกับการจัดการแบบ NoSQL ตรงกันข้ามกับฐานข้อมูลเชิงสัมพันธ์ซึ่งไม่รองรับงานประเภทข้อมูลขนาดใหญ่ นอกจากนี้ในแง่ของพื้นที่จัดเก็บข้อมูลที่เปลี่ยนแปลงไปของแอปพลิเคชันสมัยใหม่ ความสามารถในการปรับตัวของฐานข้อมูล NoSQL สามารถปรับขนาดแบบไดนามิกส่งผลต่อประสิทธิภาพในการวิเคราะห์ข้อมูลขนาดใหญ่ได้อย่างดี งานวิจัยของ Chen (2019) สำรวจฐานข้อมูล NoSQL สิบห้าหมวดหมู่เพื่อช่วยเป็นตัวเลือกที่เหมาะสมต่อการพิจารณาของงานอุตสาหกรรม ข้อสรุปงานวิจัยพบว่าฐานข้อมูล NoSQL วิเคราะห์ความเหมาะสมสำหรับการประมวลผลข้อมูลประเภทต่าง ๆ แสดงให้เห็นว่าอุตสาหกรรมต่าง ๆ สามารถเลือกฐานข้อมูล NoSQL ที่เหมาะสมเพื่อเพิ่มขีดความสามารถในการแข่งขันและการบริการลูกค้าได้อย่างไรโดยให้ข้อมูลเชิงลึกต่อการตัดสินใจพร้อมเสนอแนะข้อจำกัด ข้อดีและข้อเสียในการตัดสินใจย้ายฐานข้อมูลต่อไป งานวิจัยของ Khan (2022) การวิเคราะห์ประสิทธิภาพระหว่างเอกสาร SQL และ NoSQL ผลการทดลองเป็นการวิเคราะห์พบว่าฐานข้อมูล NoSQL ไม่ใช่ทางเลือกแทนฐานข้อมูล SQL แต่ละฐานข้อมูลมีข้อดีในสถานการณ์เฉพาะทั้งนี้เป็นไปตามโมเดลข้อมูลและสถาปัตยกรรมซอฟต์แวร์ที่พัฒนาขึ้น

4. วิธีดำเนินการวิจัย

แหล่งข้อมูลสำหรับทดลอง

ผู้วิจัยสร้างแบบจำลองฐานข้อมูลและจำลองข้อมูลเพื่อใช้เป็นกรณีศึกษาโดยอ้างอิงการเก็บข้อมูลมหาวิทยาลัยราชภัฏกำแพงเพชรทั้งข้อมูลนักศึกษา การลงทะเบียนเรียน ข้อมูลการยืมคืนหนังสือ ข้อมูลการร่วมกิจกรรมของนักศึกษา ย้อนหลัง 5 ปีการศึกษาเนื่องจากข้อมูลเหล่านี้ยังคงเป็นข้อมูลใช้ในระบบ จากนั้นทำการตัดส่วนที่เป็นข้อมูลสูญหายออกแล้วนำไปสร้างข้อมูลในฐานข้อมูลทดสอบ โดยมีข้อมูลที่นำมาทดลองดังรูปที่ 2



รูปที่ 2 จำนวนข้อมูลในแต่ละตาราง

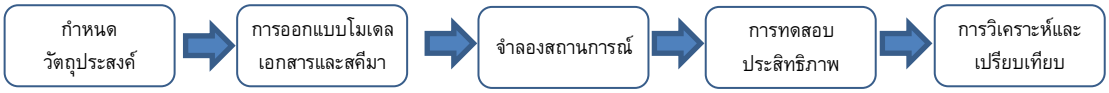
รูปที่ 2 แสดงข้อมูลที่ทำดรัมมี่ข้อมูลออกมาสู่ฐานข้อมูลทดลอง 4 ฐานข้อมูลซึ่งสามารถอธิบายได้ เช่น ข้อมูลนักศึกษาปี 2561 มีนักศึกษาทั้ง 4 ชั้นปีรวม 6,800 คน นักศึกษาแต่ละคนลงทะเบียนเรียนจำนวนวิชามากน้อยไม่เท่ากันแต่โดยเฉลี่ยอยู่ที่คนละ 6 วิชาจึงเกิดธุรกรรม 40,800 เรคอร์ด ในส่วนกิจกรรมนักศึกษาเช่นเดียวกันนักศึกษาลงทะเบียนเพื่อเช็คชื่อเข้าร่วมกิจกรรมเฉลี่ยคนละ 13 กิจกรรมต่อปี การยืมคืนหนังสือมีธุรกรรมน้อยสุด ในส่วนของการลงทะเบียนกิจกรรมปี 2563-2564 ถึงแม้จะมีสถานการณ์ระบาดของโรค covid-19 แต่มหาวิทยาลัยกำหนดให้นักศึกษาลงทะเบียนเข้าร่วมกิจกรรมในรูปแบบออนไลน์ได้ดังนั้นจึงยังมีธุรกรรมเกิดขึ้น

เครื่องมือ

การทดลองนี้ผู้วิจัยใช้เขียนโปรแกรมและสร้างฐานข้อมูลด้วย mongodb และ mysql และใช้ apache jmeter สำหรับทดสอบประสิทธิภาพเวลาการตอบสนองการคิวรี่ ใช้คอมพิวเตอร์ CPU core i7 หน่วยความจำหลักขนาด 16 GB ในการพัฒนาระบบ

ขั้นตอนการดำเนินการทดลอง

งานวิจัยนี้ออกแบบการทดลองโดยนำข้อมูลนักศึกษาจากหลายแหล่งข้อมูลทั้งจากสำนักทะเบียนและประมวลผลเพื่อใช้ข้อมูลการลงทะเบียน ข้อมูลหลักสูตร ข้อมูลการศึกษา แหล่งข้อมูลจากสำนักวิทยบริการเพื่อใช้ข้อมูลการยืมคืนหนังสือ แหล่งข้อมูลจากงานกิจการนักศึกษาเพื่อใช้ข้อมูลการร่วมกิจกรรมของนักศึกษา และออกแบบการดำเนินการทดลองดังรูปที่ 3



รูปที่ 3 ขั้นตอนการดำเนินการทดลอง

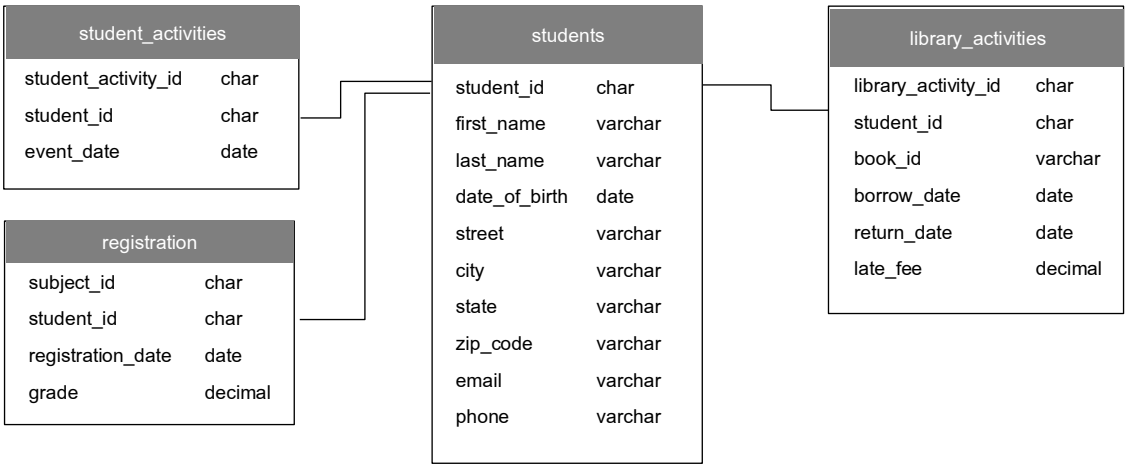
รูปที่ 3 แสดงขั้นตอนการดำเนินการทดลองเริ่มจากการ 1. กำหนดวัตถุประสงค์ด้วยการกำหนดเป้าหมายเฉพาะของการทดสอบโดยระบุตัวชี้วัดประสิทธิภาพ ได้แก่ เวลาการตอบสนอง ความสามารถในการขยายขนาด การแทรกและอัปเดทปริมาณงานและความยืดหยุ่นในการพัฒนาและสคีมา 2. ออกแบบโมเดลเอกสารของ NoSQL และออกแบบสคีมาของฐานข้อมูลเชิงสัมพันธ์เมื่อได้โมเดลเอกสารและสคีมา 3. จำลองสถานการณ์ด้วยการเพิ่มข้อมูลเพื่อรองรับการทดสอบ
4. ทดสอบประสิทธิภาพตามตัวชี้วัดประสิทธิภาพ 5. นำผลที่ได้มาวิเคราะห์และเปรียบเทียบเพื่อเขียนรายงานการวิจัย

5. ผลการวิจัยและอภิปรายผลการวิจัย

การทดลองการผู้วิจัยจะทำการออกแบบ สร้างเอกสารและสคีมา จากนั้นจำลองการทดลองด้วยซอฟต์แวร์ mongodb และ mysql ซึ่งเอกสารและสคีมาสามารถแสดงได้ดังรูปที่ 4 และ 5

```
{ “_id”: ObjectId(“uniqueIdentifier”), “student_id”: “61xxxxxxx”, “personal_info”: { “first_name”: “Nattawut”, “last_name”: “Meemak”, “date_of_birth”: “2005-01-01”, “address”: { “street”: “123 Main St”, “city”: “Cityville”, “state”: “State”, “zip_code”: “12345”}, “contact”: { “email”: “nattawut@example.com”, “phone”: “+1234567890” } }, “registration”: { “registration_date”: “2022-01-01”, “subject_id”: [“565xxxx”, “412xxxx”, “335xxxx”] ,“library_activities”: { “borrowed_books”: [{“book_id”: “BK001”, “borrow_date”: “2022-02-15” }, { “book_id”:“BK002”, “borrow_date”: “2022-03-01” }], “returned_books”: [{“book_id”: “B001”,“return_date”: “2022-03-10”, “late_fee”: 0}]}, “student_activities”: {“ student_activity_id”: [“A001”, “A007”],“ event_date”: [“2022-08-20” , “2022-08-25”] } }
```

รูปที่ 4 การออกแบบโมเดลเอกสารและสคีมาฐานข้อมูล NoSQL



รูปที่ 5 การออกแบบสคีมาฐานข้อมูลเชิงสัมพันธ์

รูป 4 แสดงการออกแบบโมเดลเอกสารและสคีมา NoSQL ด้วย mongodb ซึ่งจะเห็นว่าการสร้างเอกสารไม่จำเป็นต้องกำหนดโครงสร้างตารางก่อน แต่สามารถกำหนดคู่ คีย์(key) ค่า(value) ได้ตั้งแต่การสร้างเอกสาร รูป 5 แสดงการออกแบบสคีมาฐานข้อมูลเชิงสัมพันธ์ของระบบนักศึกษาซึ่งประกอบด้วย ตารางนักศึกษา ตารางกิจกรรมนักศึกษา ตารางการลงทะเบียนเรียน และตารางการยืมคืนหนังสือ ทดลองด้วยการเติมชุดข้อมูลโดยเติมข้อมูลในฐานข้อมูล mongodb จำนวน 12,000 เอกสาร ในส่วนของฐานข้อมูล mysql เติมข้อมูลที่ได้จากรูปที่ 2 จากนั้นทดสอบประสิทธิภาพด้านต่าง ๆ ตามประเด็นที่กำหนด เหตุผลที่เอกสารในเรื่องจำนวนคน mongodb จะมีมากกว่า mysql เนื่องจากแต่ละเอกสารแทนนักศึกษาหนึ่งคน แต่ละคนมีกิจกรรมที่เป็นการลงทะเบียน การยืมคืนหนังสือ และกิจกรรมนักศึกษาซึ่งยังมีข้อมูลของนักศึกษารุ่นเก่า แต่ทั้งนี้ในแต่ละเอกสารของนักศึกษาในแต่ละรุ่นจำนวนคอลัมน์จะไม่เท่ากัน นักศึกษารุ่นเก่า ๆ จะไม่มีการขยายตัวของคอลัมน์เนื่องจากนักศึกษาบางคนจบการศึกษาแล้ว

ผลการทดสอบตามประเด็นที่กำหนด

ผู้วิจัยออกแบบประเด็นการเปรียบเทียบประสิทธิภาพเป็น 2 ประเด็นได้แก่

1. เวลาการตอบสนองของการสืบค้นข้อมูลที่ซับซ้อน เป็นการเปรียบเทียบเวลาที่ฐานข้อมูลทั้งสองใช้ในการดำเนินการสืบค้นข้อมูลทดสอบด้วย apache jmeter โดยเปรียบเทียบเวลาตอบสนองระหว่างฐานข้อมูลเชิงสัมพันธ์ด้วย mysql กับฐานข้อมูล NoSQL ด้วย mongodb ในส่วนของฐานข้อมูลเชิงสัมพันธ์สืบหาความสัมพันธ์ระหว่างตารางนักศึกษากับตารางการลงทะเบียน การยืมคืนหนังสือ และการกิจกรรมของแต่ละคน ด้วยวิธีการ JOIN ดังนี้

```
SELECT s.student_id, s.first_name, s.last_name, r.grade, COUNT(la.library_activity_id) AS num_library_activities, COUNT(sa.student_activity_id) AS num_student_activities FROM students s LEFT JOIN registration r ON s.student_id = r.student_id LEFT JOIN libraryactivities la ON s.student_id = la.student_id LEFT JOIN studentactivities sa ON s.student_id = sa.student_id WHERE s.date_of_birth BETWEEN '2003-01-01' AND '2005-12-31' GROUP BY s.student_id, s.first_name, s.last_name, r.grade;
```

รูปที่ 6 คำสั่งการคิวรี่ข้อมูลฐานข้อมูลเชิงสัมพันธ์ mysql ด้วยเงื่อนไข COUNT และ GROUP BY

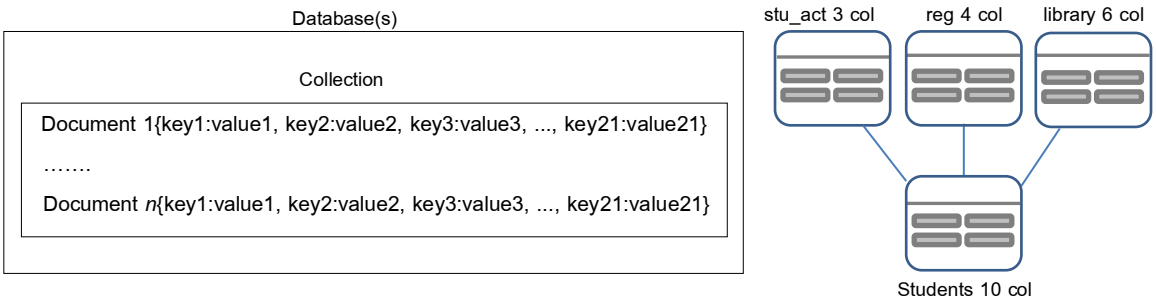
รูปที่ 6 แสดงการคิวรี่ข้อมูลเพื่อหาความสัมพันธ์ของฐานข้อมูล mysql ระหว่างตารางนักศึกษาว่านักศึกษาแต่ละคนลงทะเบียนวิชาใด ได้เกรดเท่าไหร่ ยืมคืนหนังสือเล่มใด และทำกิจกรรมใดระหว่างเรียนด้วยการสืบค้นข้อมูลด้วยการJOIN ตารางเพื่อสืบค้นข้อมูลซึ่งเพิ่มเงื่อนไขเพื่อให้เกิดความซับซ้อนได้แก่ count และ group by

```
db.student_activity.aggregate([{ $match: {"personal_info.date_of_birth": {$gte: "2003-01-01",$lte: "2005-12-31" } }},  {$group: {_id: { student_id: "$student_id", first_name: "$personal_info.first_name", last_name: "$personal_info.last_name", grade: "$registration.grade" }, num_library_activities: { $sum: { $cond: [{ $gt: ["$library_activities", 0] }, 1, 0] } }, num_student_activities: { $sum: { $cond: [{ $gt: ["$student_activites", 0] }, 1, 0] } } } } ]);
```

รูปที่ 7 การคิวรี่ด้วย mongodb ด้วยเงื่อนไข $group

รูปที่ 7 แสดงการสร้างไปป์ไลน์คิวรี่ด้วยการทำ aggregate ใส่เงื่อนไข {$gte: "2003-01-01",$lte: "2005-12-31" } และจัดกลุ่มด้วย $group

2. ความยืดหยุ่นในการพัฒนาสคีมา ผู้วิจัยทดลองสร้างโมเดลเอกสารและสคีมาฐานข้อมูลจากรูปที่ 4 พบว่าความสามารถในการขยายแนวนอนของ mongodb ช่วยให้สามารถจัดการชุดข้อมูลขนาดใหญ่และมีปริมาณงานสูงโดยการกระจายภาระงานไปยังหลายโหนดมีประโยชน์สำหรับแอปพลิเคชันที่ข้อมูลเติบโตอย่างรวดเร็ว ในทางตรงกันข้าม mysql มีความสามารถในการปรับขนาดตามแนวตั้งเป็นหลักทำให้การการจัดการบางอย่างมีความยุ่งยาก เช่น การแทรกข้อมูลต้องแทรกข้อมูลที่ตารางข้อมูลหลัก(master) ก่อนแล้วค่อยไปทำตารางธุรกรรม(transaction) ได้ หรือบางฐานข้อมูลถ้าออกแบบฐานข้อมูลไม่ดีไม่ทำ normalization ก่อน แล้วมีการนำข้อมูลบางอย่างไปอ้างอิง เช่น การนำชื่อไปอ้างอิงซึ่งถ้าเป็นแบบนั้นต้องตามไปแก้ทุกตารางที่มีชื่อนั้นปรากฏอยู่

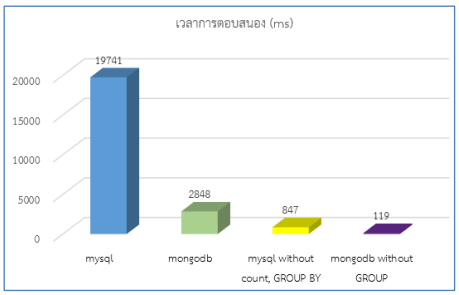


รูปที่ 8 ภาพรวมจำนวนคอลัมน์ของฐานข้อมูล mongodb และ mysql

รูปที่ 8 แสดงจำนวนคอลัมน์ของฐานข้อมูล mongodb และ mysql โดยรูปที่ 8(ก) แสดงผลการสร้างฐานข้อมูล mongodb ซึ่งได้ผลลัพธ์ขนาด 23 คู่คีย์-ค่า (23 คอลัมน์) รูปที่ 8(ข) แสดงผลการสร้างฐานข้อมูล mysql ซึ่งประกอบด้วยตาราง student_activities ขนาด 3 คอลัมน์, registration 4 คอลัมน์, students 10 คอลัมน์ และ library_activities 6 คอลัมน์ รวมทั้ง 4 ตาราง 23 คอลัมน์ ตัด student_id ที่ซ้ำกันออกเหลือ 21 คอลัมน์ ซึ่งการสร้างฐานข้อมูล mongodb จะนำแต่ละคอลัมน์ในแต่ละตารางมาเรียงต่อกันให้เป็น 1 เอกสาร (document) ดังนั้นจึงไม่มีตารางข้อมูลหลักและตารางธุรกรรมการปรับปรุงข้อมูลทำได้ในแต่ละเอกสารแยกอิสระ และเมื่อต้องการคิวรี่ข้อมูลจึงไม่จำเป็นต้องสร้างการJOIN ระหว่างตารางเหมือนกับ mysql ทำให้การประมวลผลทำได้อย่างรวดเร็วนอกจากนี้ในแต่ละเอกสารไม่ขึ้นต่อกันจึงทำให้การออกแบบสคีมาเป็นแบบไดนามิกยืดหยุ่นได้

การวิเคราะห์เปรียบเทียบและอภิปรายผลการวิจัย

ความสามารถในการปรับขนาดในแนวนอนของ mongodb มีความยืดหยุ่นและความสามารถในการปรับขนาดที่มากกว่าเมื่อเปรียบเทียบกับวิธีการปรับขนาดในแนวตั้งของ mysql อย่างไรก็ตามการใช้งานและการจัดการคลัสเตอร์ mongodb ที่ปรับขนาดได้ในแนวนอนอาจต้องใช้ความเชี่ยวชาญและโครงสร้างพื้นฐานเพิ่มเติม เมื่อเทียบกับ mysql ซึ่งอาจจัดการได้ง่ายกว่า นอกจากนี้ถ้าการสืบค้นข้อมูลไม่ซับซ้อน เช่น ตัดเงื่อนไข count , group by ออกเวลาการตอบสนองรวดเร็วขึ้นอย่างมากดังแสดงรูปที่ 9



รูปที่ 9 เปรียบเวลาการตอบสนอง

รูปที่ 9 แสดงเวลาการเปรียบเทียบเวลาการตอบสนองการสืบค้นข้อมูลของ mysql กับ mongodb ด้วยโปรแกรม apache jmeter ซึ่งเวลาการประมวลผล mysql เมื่อการ JOIN ร่วมกับคำสั่งที่เพิ่มรอบการประมวลผลเช่น count , group by การประมวลผลใช้เวลาการตอบสนองที่ 19,741 มิลลิวินาที (19.741 วินาที) ในส่วน mongodb ใช้เวลา 2,848 มิลลิวินาที (2.848 วินาที) แต่เมื่อแก้ไขคำสั่งโดยไม่ใช้คำสั่งที่เพิ่มรอบการประมวลผล mysql ใช้เวลาการตอบสนองใช้เวลาเพียง 847 มิลลิวินาที (0.847 วินาที) และ mongodb ใช้เวลา 119 มิลลิวินาที (0.119 วินาที) ทั้งนี้การสืบค้นข้อมูลที่ใช้คำสั่งการเพิ่มรอบ (count, group by) ซึ่งถือว่าเป็นการประมวลผลซับซ้อนที่เป็นความสามารถที่โดดเด่นของ mysql ในการหาความสัมพันธ์ระหว่างตาราง ในขณะที่ mongodb ถึงแม้จะมีความรวดเร็วในการสืบค้นข้อมูลแต่คำตอบที่ได้มาไม่อยู่ในรูปความสัมพันธ์แต่จะเป็นการแสดงข้อมูลทั้งเอกสาร เมื่อวิเคราะห์ในเชิงการใช้งาน ทั้งสองระบบโปรแกรมเมอร์ยังคงต้องเขียนโปรแกรมเพื่อกรองข้อมูลเฉพาะที่จะใช้งานสำหรับมาทำรายงาน นอกจากนี้จากการศึกษาการเติบโตของข้อมูลซึ่งดูได้จากรูปที่ 2 จะเห็นว่าในแต่ละปีข้อมูลเหล่านี้มีธุรกรรมเพิ่มขึ้นปริมาณไม่สูง ดังนั้นผู้วิจัยจึงเห็นว่าฐานข้อมูลเชิงสัมพันธ์ mysql ยังเป็นตัวเลือกที่เหมาะสำหรับการพัฒนาระบบงานนักศึกษา เนื่องจากข้อมูลนักศึกษาแต่ละคนไม่ได้เพิ่มขึ้นในทุกวันทำให้อัตราการขยายตัวยังคงเป็นช่วง ๆ เช่น ช่วงการลงทะเบียน เป็นต้น แต่ทั้งนี้ผู้วิจัยมองเห็นประโยชน์ของ mongodb และมีแผนที่จะนำฐานข้อมูล NoSQL มาพัฒนาสำหรับจัดเก็บข้อมูลงานด้านเครือข่าย เช่น การเข้าใช้บริการในเว็บไซต์ต่าง ๆ เนื่องจากมีข้อมูลการใช้งานแต่ละคนมีปริมาณการใช้ที่มหาศาลในแต่ละวันมีอัตราการขยายตัวในแนวนอนจึงเหมาะแก่การพัฒนาด้วยฐานข้อมูล NoSQL ต่อไป

6. สรุปผลการวิจัย

การวิจัยนี้มีจุดมุ่งหมายเพื่อทดลองและวิเคราะห์ซอฟต์แวร์การจัดการฐานข้อมูลเชิงเปรียบเทียบ ได้แก่ ซอฟต์แวร์การจัดการฐานข้อมูลเชิงสัมพันธ์ด้วย mysql และซอฟต์แวร์การจัดการฐานข้อมูล NoSQL ด้วย mongodb จุดเน้นอยู่ที่การทดสอบนี้ในบริบทของการจำลองระบบข้อมูลนักศึกษา ได้แก่ การลงทะเบียน การยืมคืนหนังสือ และกิจกรรมของนักศึกษา วัตถุประสงค์งานวิจัยเพื่อทดสอบประสิทธิภาพการจัดการฐานข้อมูลทั้งสองระบบโดยเปรียบเทียบจากเวลาการตอบสนองของการสืบค้นข้อมูลที่ซับซ้อนและความยืดหยุ่นในการพัฒนาสคีมา

ผลการทดลองพบว่าเวลาการตอบสนองของการสืบค้นข้อมูลฐานข้อมูลด้วย mongodb ใช้เวลาตอบสนองได้รวดเร็วกว่าฐานข้อมูลเชิงสัมพันธ์ mysql การพัฒนาสคีมาฐานข้อมูล mongodb มีความยืดหยุดมากกว่าเนื่องจากรูปแบบการสร้างสคีมาที่เน้นการนำข้อมูลของบุคคลมารวมเป็นเอกสารเดียวในลักษณะข้อมูลแนวนอน แต่ทั้งนี้เมื่อวิเคราะห์เปรียบเทียบในแง่มุมต่าง ๆ สามารถสรุปได้ว่า ฐานข้อมูลเชิงสัมพันธ์มีข้อดี ได้แก่ โมเดลข้อมูลที่มีโครงสร้างช่วยให้มั่นใจในความสมบูรณ์ของข้อมูลผ่านสคีมาที่กำหนดไว้ล่วงหน้าอย่างเข้มงวดซึ่งต้องมีความสอดคล้องกันในฐานข้อมูล นอกจากนี้ยังปฏิบัติตามคุณสมบัติ ACID เพื่อให้มั่นใจว่าความสมบูรณ์ของข้อมูลในฐานข้อมูลเชิงสัมพันธ์เหมาะสำหรับการจัดการสืบค้นที่มีความสัมพันธ์ระหว่างข้อมูล อย่างไรก็ตามฐานข้อมูลเชิงสัมพันธ์มีข้อจำกัดในการปรับขนาดฐานข้อมูล โดยเฉพาะชุดข้อมูลขนาดใหญ่และปริมาณการรับส่งข้อมูลที่สูงอาจเป็นเรื่องท้าทาย การปรับเปลี่ยนสคีมาอาจต้องใช้เวลาและต้องสร้างดรัมมีข้อมูลซึ่งเกิดความยุ่งยากและเสียค่าใช้จ่ายเพิ่ม ในทางกลับกันฐานข้อมูล NoSQL ให้ข้อดีด้านความยืดหยุ่นและความสามารถในการปรับขนาดได้ สคีมาที่ยืดหยุ่นช่วยให้สามารถปรับโครงสร้างข้อมูลที่กำลังพัฒนาได้ง่าย ซึ่งรองรับสภาพแวดล้อมข้อมูลแบบไดนามิกได้ดี ฐานข้อมูล NoSQL สามารถปรับขนาดในแนวนอนเพิ่มความสะดวกให้กับฐานข้อมูลแบบกระจายในหลายโหนดและช่วยให้สามารถจัดการรูปแบบข้อมูลที่ไม่มีโครงสร้างหรือกึ่งโครงสร้าง เช่น JSON, XML หรือคู่คีย์-ค่าได้อย่างมีประสิทธิภาพ แต่เนื่องจากฐานข้อมูล NoSQL ไม่มีธุรกรรม ACID ซึ่งอาจนำไปสู่ปัญหาความสอดคล้องข้อมูล ข้อจำจัดการสืบค้นและความสัมพันธ์ที่ซับซ้อนเมื่อเปรียบเทียบกับฐานข้อมูลเชิงสัมพันธ์ซึ่งอาจส่งผลต่อความสอดคล้องและความสมบูรณ์ของข้อมูลในบางสถานการณ์ นอกจากนี้ในมุมมองนักพัฒนาที่ไม่คุ้นเคยกับโมเดลลักษณะนี้ยังมีข้อจำกัดในหลายประการทั้งคำสั่งการได้มาของข้อมูลที่เป็นเอกสารเฉพาะบุคคลที่ต้องหาความสัมพันธ์ระหว่างข้อมูลต่อเป็นการทำงานที่ซ้ำซ้อน ดังนั้นการเลือกฐานข้อมูล NoSQL อาจยังไม่ใช่ทางเลือกแทนสำหรับฐานข้อมูลเชิงสัมพันธ์ของระบบงานของกรณีตัวอย่างแต่จะเหมาะสำหรับฐานข้อมูลที่มีข้อมูลไหลเข้าออกตลอดเวลาซึ่งในอนาคตจะกลายเป็นข้อมูลขนาดใหญ่ สอดคล้องกับงานวิจัยของ Ali et al. (2019), Chen & Lee (2019) และ Khan et al. (2022) ซึ่งในอนาคตผู้วิจัยจะทดลองจัดเก็บข้อมูลด้วย NoSQL ของระบบเครือข่ายมหาวิทยาลัยเนื่องจากมีข้อมูลธุรกรรมไหลตลอดเวลาในอนาคตจะกลายเป็นฐานข้อมูลขนาดใหญ่ต่อไปได้

7. กิตติกรรมประกาศ

ขอขอบคุณผู้บริหารมหาวิทยาลัยราชภัฏกำแพงเพชรที่ให้ความอนุเคราะห์ให้หน่วยงานที่เกี่ยวข้องในการให้ข้อมูล

8. เอกสารอ้างอิง

Ali, A., Naeem, S., Anam, S., & Ahmed, M. M. (2023). A State of Art Survey for Big Data Processing and NoSQL Database Architecture. **International Journal of Computing and Digital Systems, 14**(1). http://dx.doi.org/10.12785/ijcds/140124.

Ali, W., Shafique, M. U., Majeed, M. A., & Raza, A. (2019). Comparison between SQL and NoSQL Databases and Their Relationship with Big Data Analytics. **Asian Journal of Research in Computer Science, 4**(2), 1-10.

Carvalho, I., Sá, F., & Bernardino, J. (2023). Performance Evaluation of NoSQL Document Databases: Couchbase, CouchDB, and MongoDB. **Algorithms, 16**(2), 78. https://doi.org/10.3390/a16020078.

Chen, J.-K., & Lee, W.-Z. (2019). An Introduction of NoSQL Databases Based on Their Categories and Application Industries. **Algorithms, 12**(5), 106; https://doi.org/10.3390/a12050106.

Gopi, K., Mazumder, D., Crawford, J., Gadd, P., Tadros, C. V., Atanacio, A., Saintilan, N., & Sammut, J. (2023). Developing a MySQL Database for the Provenance of Black Tiger Prawns (Penaeus monodon). **Foods, 12**(14), 2677. https://doi.org/10.3390/foods12142677.

Gyorödi, C. A., DumŞe-Burescu, D. V., Zmaranda, D. R., & Gyorödi, R. Ş. (2022). A Comparative Study of MongoDB and Document-Based MySQL for Big Data Application Data Management. **Big Data and Cognitive Computing, 6**(2), 49. https://doi.org/10.3390/bdcc6020049

Khan, W., Ahmed, E., & Shahzad, W. (2017). Predictive Performance Comparison Analysis of Relational & NoSQL Graph Databases. **International Journal of Advanced Computer Science and Applications, 8**(5), 523-530.

Khan, W., Kumar, T., Cheng, Z., Raj, K., Roy, A. M., & Luo, B. (2022). **SQL and NoSQL Databases Software architectures performance analysis and assessments -- A Systematic Literature review**. [Online]. Retrieved from: https://arxiv.org/abs/2209.06977 [2024, February 15].

Lotfy, E., Saleh, A., El-Ghareeb, H. & Ali, H. (2015). A Middle layer solution to support ACID properties for NoSQL databases. **Journal of King Saud University - Computer and Information Sciences, 28**(1), 133-145.

Netinant, P., Saengsuwan, N., Rukhiran, M., & Pukdesree, S. (2023). Enhancing Data Management Strategies with a Hybrid Layering Framework in Assessing Data Validation and High Availability. **Sustainability, 15**(20), 15034. https://doi.org/10.3390/su15