

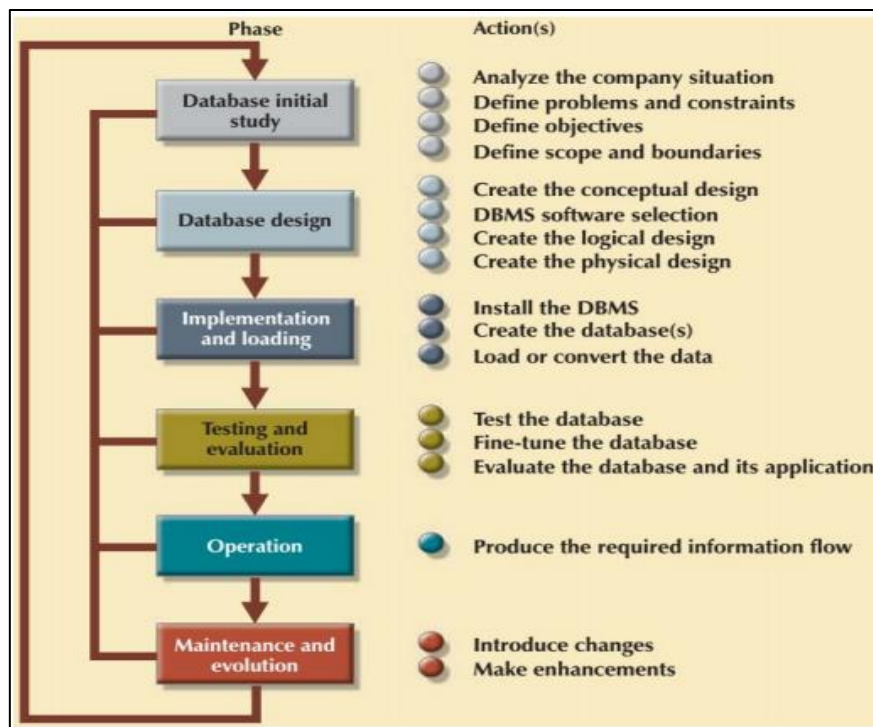
## บทที่ 3

### การออกแบบฐานข้อมูล

ปัจจุบันองค์กรต่างๆได้มีการประยุกต์ใช้ระบบสารสนเทศ (Information System) สำหรับการดำเนินการทางธุรกิจ โดยส่วนใหญ่ระบบสารสนเทศจะมีการประยุกต์ใช้ฐานข้อมูลเพื่อจัดเก็บข้อมูลต่างๆ ซึ่งการออกแบบฐานข้อมูลที่ดีจะส่งผลให้ระบบสารสนเทศสามารถจัดการกับสารสนเทศได้อย่างถูกต้องและมีประสิทธิภาพ ในบทนี้จะทำการศึกษาเกี่ยวกับวงจรการออกแบบฐานข้อมูล และวงจรการพัฒนา ระบบสารสนเทศ (Systems Development Life Cycle: SDLC) ที่มีกระบวนการคล้ายกันของกระบวนการสร้าง ตรวจสอบ ปรับปรุงพัฒนา และการปรับเปลี่ยนระบบสารสนเทศให้สามารถทำงานได้อย่างมีประสิทธิภาพ

#### 3.1 วงจรการพัฒนาฐานข้อมูล

ภายใต้การออกแบบระบบสารสนเทศ การออกแบบฐานข้อมูลเป็นขั้นตอนหนึ่งที่สำคัญ วงจรการพัฒนาฐานข้อมูล (Database Development Life Cycle: DBLC) ประกอบด้วย 6 ขั้นตอนหลัก คือ ศึกษาฐานข้อมูลขององค์กรเบื้องต้น (Database Initial Study) ออกแบบฐานข้อมูล (Database Design) สร้างฐานข้อมูลและการโหลดข้อมูลเข้าสู่ฐานข้อมูล (Implementation and Loading) ทดสอบและประเมินการทำงานของฐานข้อมูล (Testing and Evaluation) ประยุกต์ใช้ฐานข้อมูล (Operation) และดูแลรักษาและพัฒนาให้ดียิ่งขึ้น (Maintenance and Evolution) ดังแสดงในภาพที่ 3.1

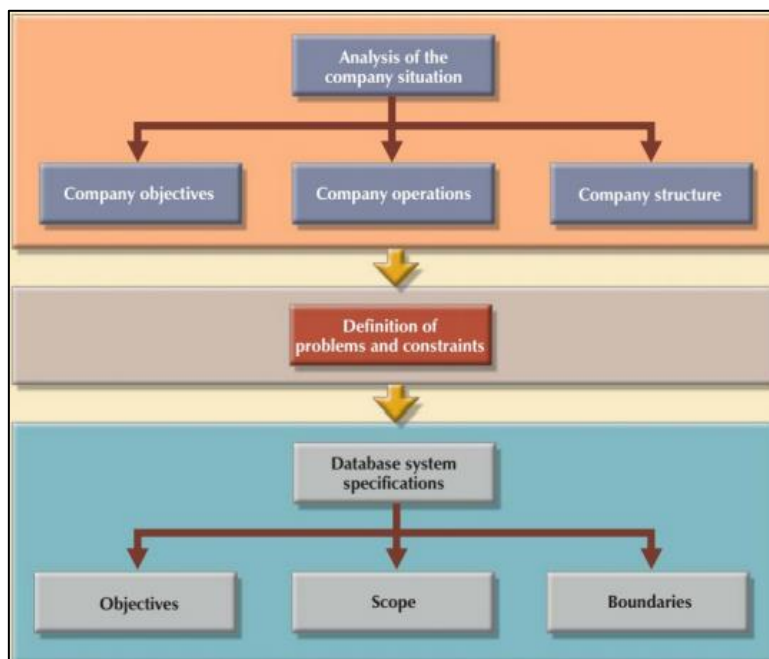


ภาพที่ 3.1 วงจรการพัฒนาฐานข้อมูล

#### ศึกษาฐานข้อมูลขององค์กรเบื้องต้น (Database Initial Study)

การศึกษาเกี่ยวกับฐานข้อมูลขององค์กรเพื่อทำการปรับปรุง หรือสร้างระบบสารสนเทศขึ้นใหม่ ผู้ออกแบบฐานข้อมูล/ผู้ออกแบบระบบต้องเข้าใจถึงสาเหตุที่ทำให้ระบบที่ใช้อยู่เกิดความล้มเหลวซึ่งสามารถทำความเข้าใจได้โดยการพูดคุย สัมภาษณ์ และ/หรือสอบถามจากผู้ใช้งานระบบ โดยการศึกษา มีขั้นตอนหลัก

3 ส่วนคือ การวิเคราะห์สถานการณ์ของบริษัท (Analysis of The Company Situation) การระบุถึงปัญหาและเงื่อนไข (Define Problems and Constraints) และการระบุวัตถุประสงค์และขอบเขต (Define Objectives, Scopes and Boundaries)



ภาพที่ 3.2 สรุปเกี่ยวกับขั้นตอนการศึกษาระบบข้อมูลขององค์กรเบื้องต้น

มีรายละเอียดดังนี้

#### การวิเคราะห์สถานการณ์ของบริษัท

สถานการณ์ขององค์กรจะบ่งบอกถึงการดำเนินการ โครงสร้าง และเป้าหมายขององค์กร จึงเป็นเหตุให้ผู้ออกแบบฐานข้อมูลต้องพิจารณาถึงการดำเนินการพื้นฐาน เป้าหมายที่ต้องดำเนินการ และยังรวมถึงการพิจารณาโครงสร้างหลักขององค์กรที่ทำให้ทราบถึงลำดับขั้นของการบริหารงาน ซึ่งจะมีประโยชน์ในการระบุถึงความต้องการของตำแหน่งงาน ความต้องการรายงาน และการสอบถาม (Query) สารสนเทศต่าง ๆ

#### การระบุถึงปัญหาและเงื่อนไข

การศึกษากิจการดำเนินงานขององค์กรทำให้รู้ถึงฟังก์ชันการทำงาน อินพุตที่ถูกป้อนเข้าไปในระบบ เอกสารหรือรายงานที่ระบบสร้าง ตลอดจนวิธีการและบุคคลที่ใช้เื่อท์พุด ส่วนการระบุถึงปัญหาการดำเนินการผู้ออกแบบฐานข้อมูลต้องทำการจัดเก็บคำอธิบายถึงปัญหาอย่างกว้างๆ หลังจากทราบถึงปัญหาและคำอธิบายถึงปัญหาในเบื้องต้นแล้วผู้ออกแบบฐานข้อมูลต้องไตร่ตรองอย่างละเอียดซึ่งจะได้ข้อมูล สำหรับระบุถึงปัญหา เมื่อทราบถึงปัญหาที่แน่ชัดแล้วต้องทำการค้นหาคำตอบหรือวิธีการแก้ปัญหาเหล่านั้นให้สอดคล้องกับแนวทาง หรือวิธีการในการดำเนินธุรกิจ

#### การระบุวัตถุประสงค์และขอบเขต

หลังจากการระบุถึงปัญหาและเงื่อนไข ผู้ออกแบบฐานข้อมูลต้องทำให้แน่ใจว่าวัตถุประสงค์ของการสร้างระบบฐานข้อมูล/ระบบสารสนเทศมีความสอดคล้องกับความต้องการหรือปัญหาของผู้ใช้งาน โดยต้องพิจารณาและหาคำตอบดังนี้

- วัตถุประสงค์เบื้องต้นของระบบที่จะสร้างคืออะไร?
- ระบบที่จะสร้างสามารถเชื่อมต่อกับระบบเดิมที่มีอยู่หรือสามารถเชื่อมต่อกับระบบที่จะสร้างขึ้นในอนาคตได้หรือไม่?

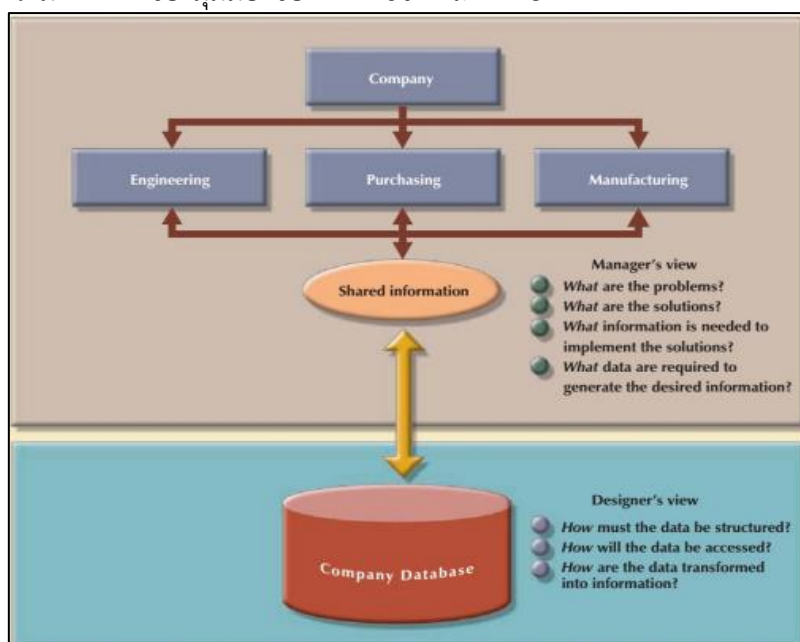
- ระบบที่จะสร้างมีความสามารถในการแบ่งปันข้อมูลให้กับระบบอื่นหรือผู้ใช้งานอื่นหรือไม่?

จากคำถามดังกล่าวจะทำให้เข้าใจถึงวัตถุประสงค์และลักษณะ/คุณลักษณะของระบบที่จะสร้างซึ่งช่วยให้มีความแน่ใจขึ้นเกี่ยวกับวัตถุประสงค์ของการสร้างระบบ ขอบเขตของระบบจะเป็นตัวกำหนดขอบเขตของการออกแบบฐานข้อมูล ด้วยเหตุนี้จำเป็นต้องรู้ว่าการออกแบบนี้จะเป็นการออกแบบสำหรับทั้งองค์กรหรืออาจเป็นการออกแบบสำหรับแผนกใดแผนกหนึ่ง หรือเป็นการออกแบบสำหรับฟังก์ชันการทำงานภายใต้แผนกหนึ่งเท่านั้น การทราบถึงขอบเขตจะช่วยในการกำหนดโครงสร้างข้อมูลที่ต้องการใช้ ชนิด และจำนวนเอนทิตีที่ต้องการใช้

ในส่วนขอบเขตอีกขอบเขตหนึ่งจะเกี่ยวข้องกับขอบเขตทางด้านเวลาในการพัฒนาระบบ ขอบเขตทางด้านงบประมาณ และยังรวมถึงขอบเขตของฮาร์ดแวร์และซอฟต์แวร์ที่ใช้ในการสร้างระบบด้วยเช่นกัน ดังนั้นในการศึกษาเกี่ยวกับฐานข้อมูลขององค์กรเราจำเป็นต้องทราบถึงขอบเขตของระบบและเงื่อนไขต่างๆ ทั้งในด้านเวลาและงบประมาณที่จะช่วยให้เราตัดสินใจได้ว่าเราจะเลือกใช้ฮาร์ดแวร์ ซอฟต์แวร์ และส่วนประกอบอื่นๆอย่างไร

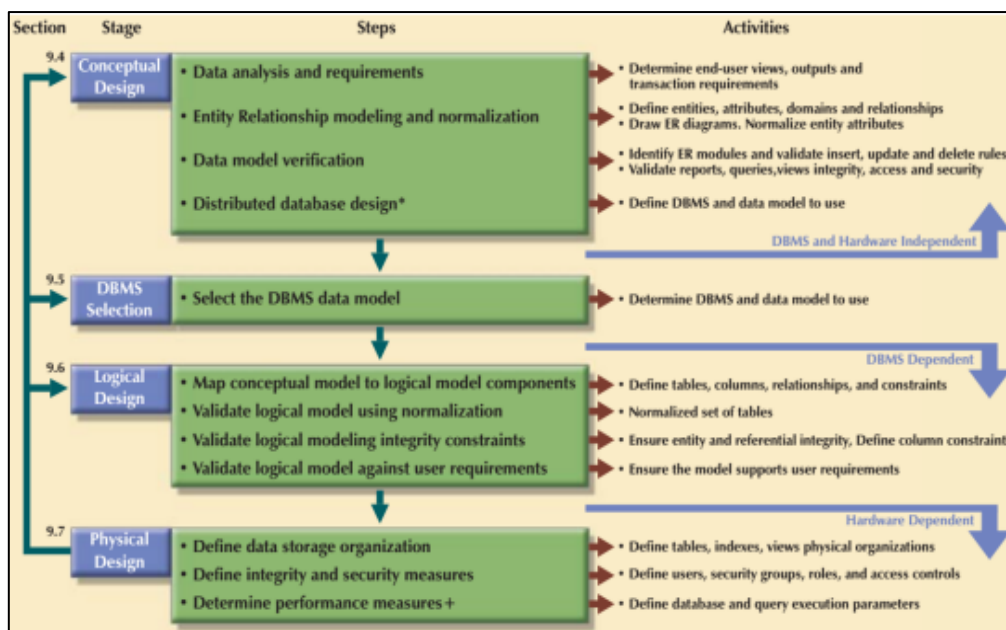
### การออกแบบฐานข้อมูล (Database Design)

ขั้นตอนการออกแบบฐานข้อมูลเป็นขั้นตอนสำคัญที่ใช้สำหรับออกแบบฐานข้อมูลเพื่อสนับสนุนการดำเนินการและวัตถุประสงค์ขององค์กร โดยผู้ออกแบบต้องให้ความสนใจกับลักษณะ/คุณลักษณะของข้อมูลที่ทำให้การจัดเก็บในฐานข้อมูล ซึ่งในการพิจารณาข้อมูลต่างๆจะมีสองมุมมองที่เกี่ยวข้องคือ มุมมองทางธุรกิจที่มองข้อมูลเป็นแหล่งกำเนิดของสารสนเทศ และมุมมองของการออกแบบที่จะพิจารณาโครงสร้างของข้อมูล การเข้าถึง/จัดการกับข้อมูล และกระบวนการต่างๆในการปรับเปลี่ยนข้อมูลให้กลายเป็นสารสนเทศ ดังภาพที่ 3.3 แสดงความแตกต่างของมุมมองของ “What” และ “How”



ภาพที่ 3.3 มุมมองของข้อมูล—มุมมองทางธุรกิจและมุมมองทางการออกแบบ

ในการออกแบบฐานข้อมูลมีขั้นตอนการออกแบบฐานข้อมูล 4 ขั้นตอนหลักคือ การออกแบบกรอบความคิด (Conceptual Design) การออกแบบเชิงตรรกะ (Logical Design) การออกแบบเชิงกายภาพ (Physical Design) และรวมถึงการเลือกใช้ระบบจัดการฐานข้อมูล (DBMS Selection) ดังแสดงในภาพที่ 3.4



ภาพที่ 3.4 ขั้นตอนการออกแบบฐานข้อมูล

การออกแบบเริ่มจากการออกแบบกรอบความคิดแล้วทำการปรับเปลี่ยนไปเป็นเชิงตรรกะ และการออกแบบทางกายภาพ โดยแต่ละขั้นตอนเป็นการเพิ่มรายละเอียดเกี่ยวกับคุณลักษณะของข้อมูล โครงสร้างการจัดเก็บข้อมูล ความสัมพันธ์ต่างๆ และยังรวมถึงการจัดทำเอกสารเพื่อแสดงถึงรายละเอียดต่างๆ การออกแบบกรอบความคิดจะเกี่ยวข้องกับข้อมูลทั้งหมดในมุมมองของผู้ใช้งาน (มุมมองทางธุรกิจ) ส่วนการออกแบบเชิงตรรกะเกี่ยวข้องกับข้อมูลภายใต้การประยุกต์ใช้ระบบจัดการฐานข้อมูล และการออกแบบเชิงกายภาพจะเกี่ยวข้องกับข้อมูลที่ถูกดำเนินการโดยระบบปฏิบัติการและอุปกรณ์ฮาร์ดแวร์ สำหรับการ จัดเก็บข้อมูล

### การสร้างฐานข้อมูลและการโหลดข้อมูลเข้าสู่ฐานข้อมูล (Implementation and Loading)

หลังจากขั้นตอนการออกแบบฐานข้อมูลเราจะทราบถึงตารางข้อมูล (เอนทิตี) แอททริบิวต์ ความสัมพันธ์ ขอบเขต มุมมอง ดัชนี และรายละเอียดที่เกี่ยวข้องกับการจัดเก็บข้อมูล จากนั้นเราจะดำเนินการ สร้างฐานข้อมูลตามทีออกแบบไว้ โดยขั้นตอนการสร้างฐานข้อมูลประกอบด้วยขั้นตอนย่อย ดังนี้

#### การติดตั้งระบบจัดการฐานข้อมูล

ในหลายครั้งของการสร้างระบบสารสนเทศ/ระบบฐานข้อมูล ต้องทำการติดตั้งระบบจัด ฐานข้อมูลในเซิร์ฟเวอร์ใหม่หรือเซิร์ฟเวอร์เดิมที่มีอยู่ แต่ปัจจุบันผู้สร้างมักประยุกต์ใช้เทคนิค “Virtualization” ที่ทำการสร้างเซิร์ฟเวอร์เสมือน พื้นที่ในการจัดเก็บข้อมูลเสมือน และเครือข่ายเสมือน เพื่อให้การแสดงผลในเชิงตรรกะไม่ขึ้นกับฮาร์ดแวร์ต่างๆ หลังจากการสร้างเซิร์ฟเวอร์เสมือนจะทำให้สามารถ ติดตั้งระบบจัดการฐานข้อมูลในเซิร์ฟเวอร์เสมือนได้ ฐานข้อมูลที่ถูกติดตั้งในเซิร์ฟเวอร์เสมือนว่า “Database virtualization”

#### การสร้างฐานข้อมูล

หลังจากทำการติดตั้งระบบจัดการฐานข้อมูลจะเป็นหน้าที่ของผู้ดูแลระบบ (System Administration SYSADM) และผู้บริหารฐานข้อมูล (Database Administration: DBA) ที่จะทำการสร้าง ฐานข้อมูล และตารางข้อมูลในขั้นตอนแรก จากนั้น DBA จะทำการกำหนดสิทธิ์ในการใช้งานฐานข้อมูลให้กับ ผู้ใช้งาน และทำการกำหนดสิทธิ์ในการเข้าถึง/การจัดเก็บตารางข้อมูล เช่น ถ้าต้องการกำหนดสิทธิ์ให้

Shannon Scott สามารถเรียกดูข้อมูลในตาราง PROFESSOR ได้จะใช้คำสั่ง GRANT ดังนี้ GRANT SELECT ON PROFESSOR TO USER SCOTT; เป็นต้น

### **การโหลดข้อมูลเข้าสู่ฐานข้อมูล**

หลังจากทำการสร้างฐานข้อมูลและตารางข้อมูลแล้วจะทำการโหลด (จัดเก็บ) ข้อมูลเข้าสู่ฐานข้อมูล โดยในการโหลดข้อมูลอาจทำการเคลื่อนย้ายข้อมูลจากระบบเดิมไปยังฐานข้อมูลที่สร้างขึ้นใหม่ ซึ่งจะอยู่ในรูปแบบฐานข้อมูลเชิงสัมพันธ์เหมือนกัน แต่อาจมีหลายกรณีที่ต้องทำการโหลดข้อมูลจากระบบอื่นหลายระบบที่ข้อมูลไม่ได้อยู่ในรูปแบบของฐานข้อมูลเชิงสัมพันธ์ เช่น ข้อมูลที่อยู่ในรูปแบบของแฟ้มข้อมูลต่างๆ สเปรดชีต หรือแม้แต่ข้อมูลที่ถูกจัดเก็บอยู่ในกระดาษ ด้วยเหตุนี้จึงต้องทำการสร้างโปรแกรมเพื่อทำการปรับเปลี่ยนข้อมูลที่อยู่ในรูปแบบอื่นให้กลายเป็นข้อมูลที่อยู่ในรูปแบบของฐานข้อมูลเชิงสัมพันธ์ เพื่อให้สามารถโหลดข้อมูลเหล่านั้นเข้าสู่ฐานข้อมูลที่สร้างขึ้นได้ แต่อาจมีบางกรณีที่ต้องทำการกรอกข้อมูลเอง (Manual) ซึ่งต้องตรวจสอบอย่างละเอียดเพื่อหลีกเลี่ยงข้อผิดพลาดในการโหลดข้อมูล

### **การทดสอบและประเมินการทำงานของฐานข้อมูล (Testing and Evaluation)**

ในขั้นตอนการออกแบบฐานข้อมูลต้องทำการตรวจสอบเกี่ยวกับความคงสภาพของฐานข้อมูล (Integrity) ความปลอดภัย ประสิทธิภาพ และความสามารถในการกู้คืนข้อมูลของฐานข้อมูลที่ถูกออกแบบ โดยขั้นตอนการตรวจสอบจะประกอบไปด้วยขั้นตอนดังต่อไปนี้

#### **การตรวจสอบฐานข้อมูล**

การตรวจสอบฐานข้อมูลจะทำให้แน่ใจได้ว่า 1) ตารางข้อมูลมีการกำหนดคีย์หลัก (Primary key) คีย์นอก (Foreign key) อย่างเหมาะสม 2) ตารางข้อมูลมีการกำหนดขอบเขตของแอททริบิวส์ ขอบเขตของ เงื่อนไข และ ทรiggerฐานข้อมูล (Database Trigger) ได้อย่างเหมาะสม และ 3) มีการกำหนดมาตรการความปลอดภัยที่เกี่ยวข้องกับการกำหนดสิทธิ์ในการเข้าถึง/ใช้งานข้อมูลในตารางข้อมูลตามลำดับ ในการทดสอบความปลอดภัยของฐานข้อมูล ต้องทดสอบในแง่มุมต่อไปนี้

- การรักษาความปลอดภัยทางกายภาพ (Physical Security) คือ การยินยอมให้ผู้ใช้ที่มีการยืนยันตัวตนสามารถเข้าใช้งานข้อมูล
- การรักษาความปลอดภัยด้วยรหัสผ่าน (Password Security) คือ การบังคับให้ผู้ใช้งานทำการล็อกอิน (Log in) เพื่อใช้งานข้อมูลตามที่ถูกกำหนดสิทธิ์
- สิทธิ์การเข้าถึง (Access rights) คือ การกำหนดสิทธิ์ในการใช้งานข้อมูลในตารางข้อมูลเช่น CREATE UPDATE DELETE INSERT SELECT เป็นต้น
- การติดตามการตรวจสอบ (Audit Trails) คือ การตรวจสอบการล่วงละเมิดสิทธิ์ในการดำเนินการกับข้อมูลโดยระบบจัดการฐานข้อมูล

#### **การประเมินการทำงานของฐานข้อมูล**

การประเมินประสิทธิภาพการทำงานของฐานข้อมูลจะทำได้ค่อนข้างยากเนื่องจากไม่มีมาตรฐาน ในการประเมิน แต่อย่างไรก็ตามการประเมินประสิทธิภาพของฐานข้อมูลเป็นสิ่งจำเป็นอย่างยิ่งที่จะทำให้รู้ว่าฐานข้อมูลมีการทำงานอย่างไร ระบบฐานข้อมูลที่แตกต่างกันอาจต้องการประสิทธิภาพของการทำงานในแง่มุมที่แตกต่างกัน เช่น บางระบบอาจต้องการให้ฐานข้อมูลสามารถเพิ่ม ปรับปรุง และลบข้อมูลปริมาณมากได้อย่างรวดเร็ว หรือ บางระบบอาจต้องการให้ระบบที่สร้างขึ้นสามารถทำการค้นคืนข้อมูลที่มีความซับซ้อนได้อย่างรวดเร็ว เป็นต้น

จากความต้องการข้างต้นจำเป็นต้องทำการประเมินการทำงานของฐานข้อมูลและโปรแกรมที่เกี่ยวข้องกับฐานข้อมูลซึ่งต้องทำการทดสอบประเมินส่วนประกอบย่อยของระบบเพื่อให้แน่ใจว่าส่วนต่างๆ

ของระบบสามารถทำงานได้อย่างเหมาะสมและสามารถตอบสนองต่อความต้องการของผู้ใช้งานได้นอกจากนั้นเรายังต้องทำการตรวจสอบเกี่ยวกับกระบวนการสำรองและกู้คืนข้อมูล ซึ่งทำได้หลายระดับดังนี้

- Full backup เป็นการสำรองข้อมูลทั้งหมดในฐานข้อมูล
- Differential backup การคัดลอกข้อมูลที่มีการเปลี่ยนแปลงล่าสุดจากการสำรองข้อมูลครั้งก่อนหน้า
- Transaction log backup เป็นการสำรองข้อมูลที่มีการดำเนินการที่มีการจดบันทึกใน log file เท่านั้น

จากการสำรองข้อมูลทั้ง 3 ระดับจะช่วยให้สามารถป้องกันอันตรายที่อาจเกิดจากไฟไหม้ ขโมย น้ำท่วม และอุบัติเหตุหรืออุบัติเหตุต่างๆได้ การสำรองข้อมูลจะช่วยให้องค์กรมั่นใจเกี่ยวกับการกู้คืนข้อมูลเมื่อ ประสบปัญหาทางด้านซอฟต์แวร์และฮาร์ดแวร์

### การประยุกต์ใช้ฐานข้อมูล (Operation)

หลังจากการประเมินและตรวจสอบการทำงานของฐานข้อมูลอย่างแน่ชัดแล้วจะเป็นการประยุกต์ใช้ฐานข้อมูลเพื่อการดำเนินการต่างๆขององค์กร โดยจะเริ่มจากการกำหนดให้ผู้ใช้งานใช้งานฐานข้อมูล เมื่อผู้ใช้งานเริ่มใช้งานจะทำให้ค้นพบปัญหาที่ไม่เคยพบมาก่อนในขั้นตอนการทดสอบและตรวจสอบการทำงานของฐานข้อมูล และเมื่อพบปัญหาจะต้องทำการระบุถึงวิธีการแก้ไขปัญหาในหลายแง่มุมซึ่งจะช่วยให้สามารถแก้ไขปัญหาได้อย่างสมบูรณ์

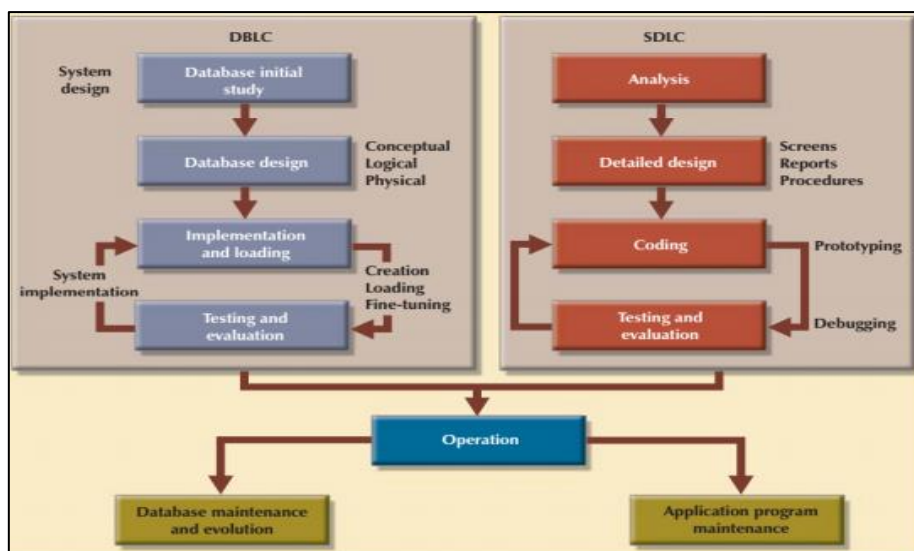
### การดูแลรักษาและพัฒนาให้ดียิ่งขึ้น (Maintenance and Evolution)

หลังจากทำการประยุกต์ใช้ฐานข้อมูลแล้ว ผู้ดูแลฐานข้อมูลควรต้องทำกิจกรรมเพื่อให้ฐานข้อมูลสามารถทำงานได้อย่างราบรื่นดังนี้

- การสำรองข้อมูล
- การกู้คืนข้อมูล
- การเพิ่มประสิทธิภาพการทำงานของฐานข้อมูล
- การเพิ่มเอนทิตี แอทริบิวส์ และอื่นๆ
- การกำหนดสิทธิ์ในการใช้งานข้อมูลให้กับผู้ใช้
- การจัดเก็บสถิติเกี่ยวกับการเข้าถึงข้อมูลเพื่อพิจารณาถึงการเพิ่มประสิทธิภาพการทำงานของระบบและ เพื่อติดตามเฝ้าสังเกตการทำงานของระบบ
- การตรวจสอบมาตรการความปลอดภัยจากสถิติที่ได้รับ นอกเหนือจากกิจกรรมข้างต้น อาจมีความต้องการจากผู้ใช้งานที่จะต้องข้อมูลสารสนเทศที่เพิ่มขึ้นจากเดิม หรือผู้ใช้อาจร้องขอรายงานเพิ่มเติมซึ่งต้องทำการกำหนดการสอบถามเพิ่มเติม หากการออกแบบฐานข้อมูลมีความยืดหยุ่นก็จะทำให้สามารถดำเนินการปรับเปลี่ยนระบบเพื่อสามารถตอบโต้ความต้องการของผู้ใช้งานได้โดยง่าย

จากวงจรการพัฒนาฐานข้อมูล (DBLC) ทั้งหมด พบว่ามีขั้นตอนคล้ายคลึงกับขั้นตอนในวงจรการพัฒนาาระบบ (SDLC) สามารถแสดงการทำงานร่วมกันระหว่าง DBLC และ SDLC ได้ดังภาพที่ 3.5





ภาพที่ 3.5 การทำงานร่วมกันระหว่างขั้นตอน DBLC และ SDLC

จากวงจรการพัฒนาฐานข้อมูลในขั้นตอนการออกแบบฐานข้อมูลได้แบ่งออกเป็น 3 ระดับคือ การออกแบบฐานข้อมูลในระดับแนวคิด การออกแบบฐานข้อมูลในระดับตรรก และการออกแบบฐานข้อมูลในระดับกายภาพ โดยมีรายละเอียดของแต่ละระดับดังนี้

### 3.2 การออกแบบฐานข้อมูลระดับแนวคิด

การออกแบบฐานข้อมูลระดับแนวคิด (Conceptual Design) เป็นขั้นตอนแรกของการออกแบบฐานข้อมูลที่มีเป้าหมายในการออกแบบฐานข้อมูลที่มีความอิสระจากซอฟต์แวร์ฐานข้อมูล และอุปกรณ์ที่ใช้สำหรับจัดเก็บข้อมูล โดยผลลัพธ์ที่ได้จากการออกแบบในระดับแนวคิดเป็นแบบจำลองกรอบความคิด (Conceptual Data Model) ที่บอกถึงเอนทิตี แอททริบิวส์ ความสัมพันธ์ และเงื่อนไขต่างๆในระบบ การสร้างแบบจำลองกรอบความคิดถูกใช้ในการกำหนดโครงสร้างของการจัดเก็บข้อมูลที่ทำให้ความเข้าใจในธุรกิจและฟังก์ชันการดำเนินการของการดำเนินธุรกิจ โดยระบุถึงข้อมูลทั้งหมดที่ใช้ในการดำเนินธุรกิจ แต่ไม่ระบุถึงชนิดของฮาร์ดแวร์และซอฟต์แวร์ที่ใช้ในการจัดเก็บและจัดการกับข้อมูล

การออกแบบฐานข้อมูลระดับแนวคิดมีการทำงาน คือ การวิเคราะห์ข้อมูลและความต้องการของผู้ใช้งาน การสร้างแบบจำลองข้อมูลเชิงสัมพันธ์เอนทิตี และการทวนสอบแบบจำลองข้อมูล รายละเอียดดังนี้

#### การวิเคราะห์ข้อมูลและความต้องการของผู้ใช้งาน

ขั้นตอนแรกของการออกแบบในระดับแนวคิดคือ การระบุถึงลักษณะของข้อมูล โดยข้อมูลเหล่านี้จะถูกปรับเปลี่ยนไปเป็นสารสนเทศเพื่อใช้ในการดำเนินการทางธุรกิจ ดังนั้นผู้ออกแบบควรพิจารณาสิ่งที่เกี่ยวข้องดังนี้

1. ความต้องการสารสนเทศ (Information Needs) โดยผู้ออกแบบต้องพิจารณาว่าข้อมูลสารสนเทศใดเป็นที่ต้องการของผู้ใช้และเอาต์พุตของระบบคืออะไร
2. ผู้ใช้สารสนเทศ (Information Users) ผู้ออกแบบต้องพิจารณาว่าใครเป็นผู้ใช้ข้อมูล/สารสนเทศ และถูกใช้อย่างไร ตลอดจนมุมมองข้อมูลของผู้ใช้มีอะไรบ้าง
3. แหล่งของสารสนเทศ (Information Sources) ผู้ออกแบบต้องพิจารณาว่าข้อมูล/สารสนเทศค้นหาได้จากที่ใด รวมถึงสามารถสกัดข้อมูลสารสนเทศได้อย่างไร

4. ส่วนประกอบของสารสนเทศ (Information Constitution) ผู้ออกแบบต้องพิจารณาว่าข้อมูลส่วนใดที่ถูกประยุกต์ใช้เพื่อสร้างเป็นสารสนเทศ แอทริบิวส์ของข้อมูลมีอะไรบ้าง ความสัมพันธ์ระหว่างข้อมูลมีอะไรบ้าง ปริมาณข้อมูลที่ถูกจัดเก็บเป็นเช่นไร และข้อมูลถูกเข้าถึง/ประยุกต์ใช้บ่อยแค่ไหน

สิ่งที่เกี่ยวข้องดังกล่าวนี้สามารถค้นหาคำตอบได้จากการการพัฒนาและรวบรวมมุมมองข้อมูลของผู้ใช้ซึ่งจะช่วยให้ระบุได้ถึงข้อมูลที่ถูกจัดเก็บในฐานข้อมูล การสังเกตการทำงานของระบบปัจจุบัน เอาท์พุตที่มีอยู่และเอาท์พุตที่ต้องการทำให้ทราบถึงรายงานต่างๆที่ผู้ใช้ต้องการและยังรู้ถึงข้อมูลที่สนับสนุนการสร้างรายงานเหล่านั้น รวมถึงการพูดคุยกับทีมผู้พัฒนาระบบ ในการพัฒนาแบบจำลองข้อมูลได้อย่างถูกต้อง ผู้ออกแบบควรมีความเข้าใจเกี่ยวกับชนิดของข้อมูล และการใช้ข้อมูลขององค์กร เพื่อให้สามารถระบุได้ถึงกฎเกณฑ์ทางธุรกิจที่ระบุได้ถึงเอนทิตี แอทริบิวส์ ความสัมพันธ์ จำนวนความสัมพันธ์ (Cardinalities) และเงื่อนไขต่างๆ เป็นต้น นอกจากนี้กฎเกณฑ์ทางธุรกิจจะช่วยให้ผู้ออกแบบฐานข้อมูลมีความเข้าใจเกี่ยวกับวิธีการดำเนินธุรกิจและบทบาทของข้อมูลที่มีส่วนเกี่ยวข้องกับการดำเนินธุรกิจ ดังนั้นสามารถสรุปประโยชน์ของกฎเกณฑ์ทางธุรกิจได้ดังนี้

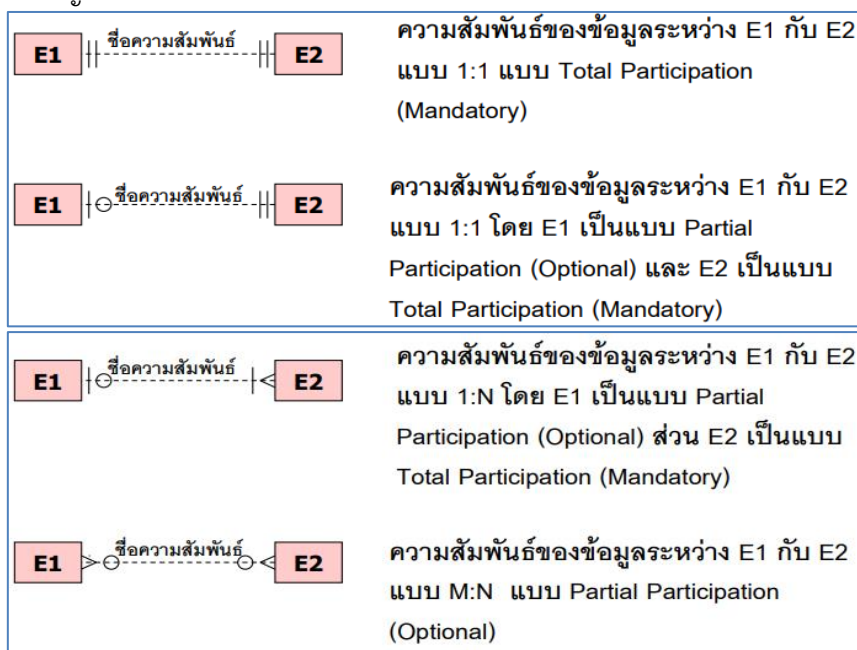
1. ช่วยในการเป็นตัวกลางในการสื่อสารระหว่างผู้ใช้และผู้ออกแบบ
2. ช่วยให้ผู้ออกแบบมีความเข้าใจในธรรมชาติ บทบาท และขอบเขตของข้อมูล
3. ช่วยให้ผู้ออกแบบมีความเข้าใจในขั้นตอนการดำเนินธุรกิจ
4. ช่วยให้ผู้ออกแบบสามารถพัฒนา/หรือระบุถึงความสัมพันธ์การมีส่วนร่วมกับความสัมพันธ์

และ คีย์นอก (Foreign key)

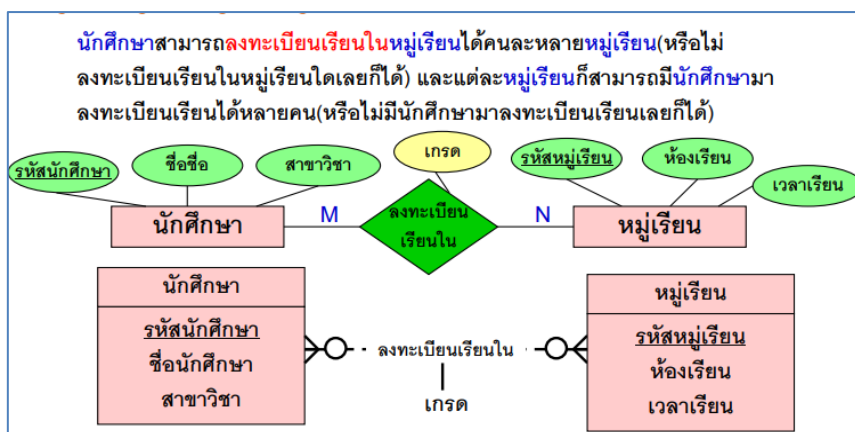
วิธีการแสดงการออกแบบได้หลายวิธี เช่น 1) ERD ที่คิดค้นจาก Chen และ 2) ERD ที่คิดค้นจาก Crow's Foot

#### E-R Model แบบ Crow's Foot

มีสัญลักษณ์ที่ใช้ความหมาย และตัวอย่าง ดังภาพที่ 3.6







ภาพที่ 3.6 ERD แบบ Crow's Foot

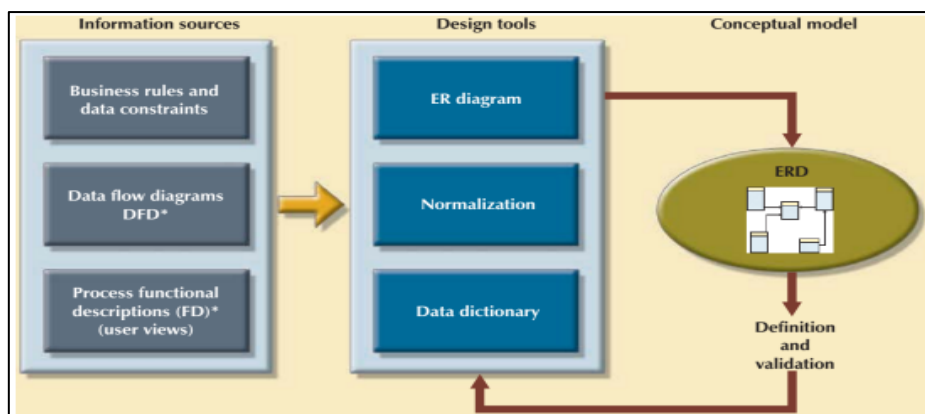
**ตัวอย่าง** การสร้างแบบจำลองข้อมูลสัมพันธ์ บริษัทให้เช่าภาพยนตร์จูลี่กูด (JollyGood Movie Rental Corporation) ซึ่งต้องการจัดเก็บข้อมูลการเช่าภาพยนตร์ของลูกค้า เมื่อวิเคราะห์ข้อมูลและความต้องการของผู้ใช้ พบกฎเกณฑ์ทางธุรกิจ คือ ลูกค้าสามารถทำการเช่าภาพยนตร์ได้หลายเรื่อง และ ภาพยนตร์เรื่องหนึ่งถูกเช่าได้โดยลูกค้าหลายคน จากกฎเกณฑ์ทางธุรกิจดังกล่าวทำให้ต้องจัดเก็บข้อมูลลูกค้าในเอนทิตีลูกค้า (CUSTOMER) เก็บข้อมูลภาพยนตร์ในเอนทิตีวิดีโอ (VIDEO) และจัดเก็บข้อมูลการเช่าภาพยนตร์ของลูกค้าในเอนทิตีการเช่า (RENTAL) ที่มีลักษณะเป็นคอมโพสิทเอนทิตี (Composite Entity) สามารถออกแบบแผนภาพอีอาร์ได้ดังภาพที่ 3.7



ภาพที่ 3.7 แผนภาพอีอาร์บริษัทให้เช่าภาพยนตร์

จากแผนภาพอีอาร์ภาพที่ 3.7 ควรพิจารณาถึงรายละเอียดของการจัดเก็บข้อมูลในส่วนของโครงสร้างการจัดเก็บข้อมูลที่ได้ออกแบบนั้นสามารถจัดเก็บข้อมูลทั้งหมดที่ต้องการได้หรือไม่ หากโครงสร้างที่กำหนดยังไม่สามารถจัดเก็บข้อมูลได้อย่างครบถ้วน ควรเพิ่มแอททริบิวต์เพื่อการจัดเก็บข้อมูลที่ สำคัญและทำการตรวจสอบการขึ้นต่อกัน (Dependency) ของแอททริบิวต์ จากนั้นควรต้องทำการตรวจสอบจนแน่ใจว่าแผนภาพที่ได้สามารถจัดเก็บข้อมูลที่ใช้ต้องการและมีความสอดคล้องกับการดำเนินธุรกิจ โดยแผนภาพอีอาร์จะเกี่ยวข้องกับการกำหนดเอนทิตี แอททริบิวต์ ความสัมพันธ์และมุมมองข้อมูลที่มีการจัดเก็บข้อมูลในพจนานุกรมข้อมูล สรุปขั้นตอนการออกแบบฐานข้อมูลระดับแนวคิด ดังนี้

1. กำหนดถึงเอนทิตี แอททริบิวต์ คีย์หลัก และคีย์นอก
2. พิจารณาถึงคุณสมบัติของคีย์ต่าง ๆ ที่เกี่ยวข้องกับโครงสร้างการจัดเก็บข้อมูลให้สอดคล้องกับข้อมูลที่จัดเก็บ ทั้งในส่วนของคีย์ประกอบ (Composite Attribute) และ แอททริบิวต์แบบหลายค่า (Multivalued Attribute) และ ดีไรฟ์แอททริบิวต์ (Derived Attribute) ที่จะสนับสนุนการประมวลผล
3. สร้างแผนภาพอีอาร์



ภาพที่ 3.8 เครื่องมือและแหล่งข้อมูลในการออกแบบแบบจำลองในระดับแนวคิด

นอกจากนี้การกำหนดมาตรฐานในการตั้งชื่อแอททริบิวต์ เอนทิตี ยังถือว่าเป็นสิ่งที่สำคัญในการพิจารณาเพื่อช่วยให้มีความเข้าใจเกี่ยวกับข้อมูลได้อย่างถูกต้อง

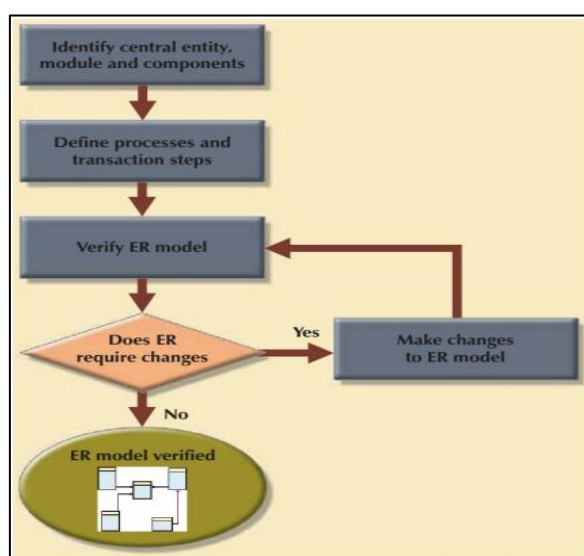
#### การทวนสอบแบบจำลองข้อมูล

การทวนสอบแบบจำลองข้อมูลเป็นขั้นตอนสุดท้ายที่สำคัญในการออกแบบแบบจำลองข้อมูล ซึ่งจะต้องตรวจสอบในหลายแง่มุม ดังนี้

1. การตรวจสอบมุมมองข้อมูลของผู้ใช้
2. การตรวจสอบการดำเนินการต่างๆกับฐานข้อมูล
3. การตรวจสอบสิทธิ์ในการใช้งานและมาตรการความปลอดภัย
4. การตรวจสอบความสอดคล้องการดำเนินการธุรกิจและเงื่อนไขในการออกแบบฐานข้อมูล

พิจารณาถึงมอดูล (Module) ที่เป็นส่วนประกอบย่อยที่เกี่ยวข้องกับการดำเนินการธุรกิจ เช่น คลังสินค้า การสั่งซื้อสินค้า การชำระเงิน เป็นต้น

5. ตรวจสอบความเปลี่ยนแปลงกฎการดำเนินการธุรกิจและทำการปรับปรุงแผนภาพอีอาร์ ดังแสดงในภาพที่ 3.9



ภาพที่ 3.9 การตรวจสอบแผนภาพอีอาร์

### การเลือกซอฟต์แวร์DBMS

การเลือกซอฟต์แวร์ระบบจัดการฐานข้อมูลเป็นกระบวนการหนึ่งที่สำคัญที่ส่งผลกระทบต่อประสิทธิภาพการทำงานของฐานข้อมูล ดังนั้นในการเลือกซอฟต์แวร์ระบบจัดการฐานข้อมูล ควรทำการพิจารณาถึงข้อดี-เสียของซอฟต์แวร์อย่างละเอียด ดังนี้

- ค่าใช้จ่าย รวมถึงค่าใช้จ่ายต่างๆ เช่น ค่าลิขสิทธิ์ซอฟต์แวร์ ค่าดำเนินการติดตั้งซอฟต์แวร์ ค่าใช้จ่ายสำหรับอบรมพนักงาน และค่าใช้จ่ายในการดำเนินการอื่น
- คุณลักษณะและเครื่องมือของระบบจัดการฐานข้อมูล ซอฟต์แวร์ระบบจัดการฐานข้อมูล ควรมีฟังก์ชันและเครื่องมือ เช่น query by example (QBE) การจัดการหน้าจอ (screen management) การสร้างรายงาน การสร้างแอปพลิเคชัน การจัดการพจนานุกรมข้อมูล เป็นต้น รวมถึงควรพิจารณาถึงการดูแลบำรุงรักษา การประมวลผลคิวรี ความยาก-ง่ายในการใช้งาน ประสิทธิภาพในการทำ งาน มาตรการความปลอดภัย เป็นต้น
- แบบจำลองข้อมูลที่สุดคล้องกับซอฟต์แวร์ระบบจัดการฐานข้อมูล ซอฟต์แวร์สำหรับแบบจำลองข้อมูลที่ไม่เหมือนกัน เช่น อาจสนับสนุนการทำงานแบบจำลองข้อมูลแบบลำดับชั้น แบบจำลองข้อมูลแบบเครือข่าย แบบจำลองข้อมูลเชิงสัมพันธ์ หรือแบบจำลองข้อมูลเชิงวัตถุ
- ความสามารถในการเคลื่อนย้าย ต้องพิจารณาว่าซอฟต์แวร์ระบบจัดการฐานข้อมูลที่ใช้สามารถเคลื่อนย้ายข้ามแพลตฟอร์ม
- ความต้องการฮาร์ดแวร์เพื่อสนับสนุนการทำงานของซอฟต์แวร์ระบบจัดการฐานข้อมูล รวมถึงโปรเซสเซอร์ หน่วยความจำ พื้นที่ในดิสก์ และอื่นๆ

### 3.3 การออกแบบเชิงตรรกะ

การออกแบบเชิงตรรกะจะนำข้อมูลในระดับแนวคิดมากำหนดโครงสร้างการจัดเก็บข้อมูลด้วยการใช้แบบจำลองข้อมูลหนึ่งๆ เช่น ถ้าต้องการออกแบบเชิงตรรกะด้วยการประยุกต์ใช้แบบจำลองข้อมูลเชิงสัมพันธ์ ต้องทำการระบุถึงตารางข้อมูล แอทริบิวต์ ความสัมพันธ์ และเงื่อนไขต่างๆ เป็นต้น การออกแบบเชิงตรรกะจะประกอบไปด้วย 4 ขั้นตอนหลักดังนี้

- ปรับเปลี่ยนแบบจำลองกรอบความคิดให้เป็นแบบจำลองเชิงตรรกะ
- ตรวจสอบแบบจำลองเชิงตรรกะด้วยการประยุกต์ใช้นอร์มัลไลเซชัน
- ตรวจสอบแบบจำลองเชิงตรรกะด้วยกฎความคงสภาพ (Integrity Constraints)
- ตรวจสอบแบบจำลองเชิงตรรกะกับความต้องการของผู้ใช้ โดยดำเนินการซ้ำหลายรอบเพื่อให้แน่ใจได้ว่าจะสามารถออกแบบแบบจำลองเชิงตรรกะที่มีความสมบูรณ์ และถูกต้องตรงความต้องการของผู้ใช้ รายละเอียดแต่ละขั้นตอนดังนี้

#### การปรับเปลี่ยนแบบจำลอง

ขั้นตอนแรกของการออกแบบจำลองข้อมูลเชิงตรรกะคือ การปรับเปลี่ยนแบบจำลองกรอบความคิดให้กลายเป็นแบบจำลองข้อมูลที่ถูกออกแบบเลือกไว้ ซึ่งเนื้อหานี้จะเน้นเกี่ยวกับแบบจำลองฐานข้อมูลเชิงสัมพันธ์ ดังนั้นจึงปรับเปลี่ยนแบบจำลองให้กลายเป็นตารางข้อมูล แอทริบิวต์ ความสัมพันธ์ และเงื่อนไขต่างๆ ประกอบไปด้วยขั้นตอนดังนี้

- พิจารณา strong entity
- พิจารณาความสัมพันธ์ระหว่างเอนทิตี supertype และ subtype
- พิจารณา weak entity
- พิจารณา binary relationship

- พิจารณา higher degree relationship การตรวจสอบแบบจำลองเชิงตรรกะด้วยการใช้ نرمัลไลเซชัน หลังจากการเปลี่ยนแบบจำลองกรอบความคิดให้อยู่ในรูปแบบของแผนภาพอาร์ จากนั้นควรที่จะพิจารณาถึงการขึ้นแก่กันหรือการพึ่งพาอาศัยกันของแอทริบิวในแต่ละเอนทิตีด้วยการใช้ نرمัลไลเซชัน เพื่อการปรับเปลี่ยนแบบจำลองข้อมูลให้อยู่ในรูปแบบ 3NF นอกจากนั้นยังจำเป็นต้องพิจารณาเกี่ยวกับกฎความคงสภาพของข้อมูลด้วย เพื่อลดความผิดปกติของข้อมูลและต้องทำการตรวจสอบสิทธิ์ในการเข้าถึง/ใช้งานข้อมูลของผู้ใช้งานที่จะช่วยให้ข้อมูลถูกใช้งานอย่างเหมาะสมและมีความปลอดภัย

### 3.4 การออกแบบเชิงกายภาพ

การออกแบบเชิงกายภาพเป็นขั้นตอนการออกแบบองค์ประกอบการจัดเก็บข้อมูล และลักษณะการเข้าถึง/การใช้งานข้อมูลเพื่อให้แน่ใจว่าฐานข้อมูลที่สร้างขึ้นมีความสมบูรณ์ ความปลอดภัย และมีประสิทธิภาพในการทำงาน การออกแบบเชิงกายภาพจะเป็นขั้นตอนที่เกี่ยวข้องกับฮาร์ดแวร์ และประสิทธิภาพการทำงานของระบบ ประกอบด้วย 3 ขั้นตอนหลักคือ การระบุถึงองค์ประกอบการจัดเก็บข้อมูล การประเมินความสมบูรณ์และความปลอดภัย การตรวจสอบประสิทธิภาพการทำงาน มีรายละเอียดดังนี้

#### การระบุถึงองค์ประกอบการจัดเก็บข้อมูล

ก่อนการระบุถึงองค์ประกอบของการจัดเก็บข้อมูล ควรพิจารณาถึงปริมาณข้อมูลที่ต้องจัดการและรูปแบบการใช้ข้อมูลจากฐานข้อมูล ซึ่งจะมีประโยชน์ดังนี้

1. ช่วยให้สามารถคำนวณการจองพื้นที่ในฐานข้อมูลในการจัดเก็บข้อมูล
2. การรู้ถึงความถี่ในการเพิ่ม อัปเดต ค้นหาข้อมูลจะช่วยให้รู้ถึงรูปแบบการใช้งานข้อมูล

เมื่อรู้ถึงปริมาณข้อมูลและรูปแบบข้อมูลจะทำให้สามารถดำเนินการดังต่อไปนี้ได้

1. ทำการตรวจสอบพื้นที่สำหรับจัดเก็บข้อมูลในแต่ละตารางข้อมูลและการจัดองค์ประกอบของตารางข้อมูล
2. ระบุถึงดัชนีที่จะใช้ในแต่ละตารางข้อมูล
3. ระบุถึงมุมมองและประเภทของมุมมองที่จะใช้ในแต่ละตารางข้อมูล

#### การประเมินความสมบูรณ์และความปลอดภัย

หลังจากทำการระบุองค์ประกอบของการจัดเก็บข้อมูลให้กับตารางข้อมูล การสร้างดัชนี และการสร้างมุมมองสำหรับตารางข้อมูลแล้ว ฐานข้อมูลจะมีความพร้อมให้ผู้ใช้สามารถใช้งานได้ แต่ก่อนที่จะใช้งานผู้ใช้ควรที่จะต้องระบุ/ยืนยันตัวตนเพื่อเข้าใช้งาน ด้วยเหตุนี้จึงเป็นเหตุให้ต้องทำการพิจารณาถึง 1) การกำหนดให้ผู้ใช้เข้าใช้งานและบทบาทของผู้ใช้งาน และ 2) กำหนดความสามารถในการดำเนินการให้กับผู้ใช้

#### การตรวจสอบประสิทธิภาพการทำงาน

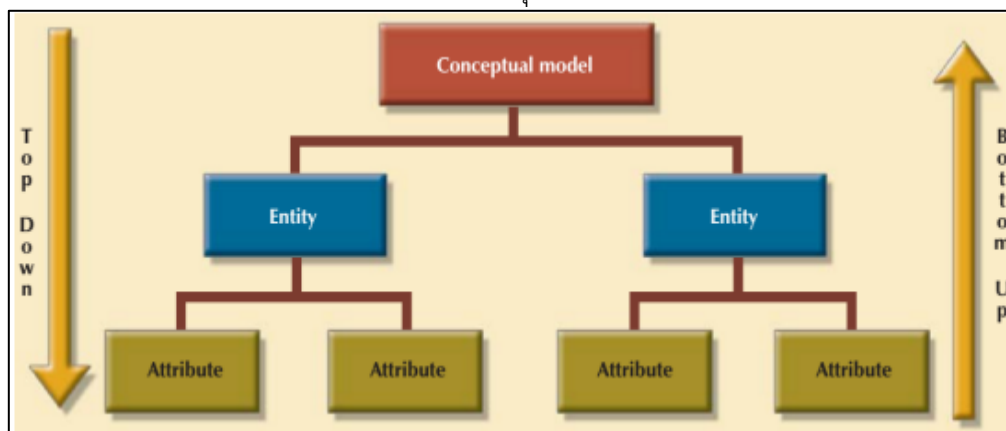
ก่อนเริ่มใช้งานควรต้องทำการตรวจสอบประสิทธิภาพการดำเนินการต่างๆของฐานข้อมูล ที่ต้องยุ่งเกี่ยวกับกระบวนการในการปรับฐานข้อมูลและการสอบถามเพื่อเพิ่มประสิทธิภาพการทำงาน

### 3.5 กลยุทธ์ในการออกแบบฐานข้อมูล

การออกแบบฐานข้อมูลจะมีวิธีการที่มีรายละเอียดดังนี้

1. การออกแบบจากบนลงล่าง (Top-down Design) เริ่มจากการระบุถึงข้อมูลอย่างคร่าวๆ แล้วทำการระบุถึงข้อมูลที่มีความละเอียดมากขึ้น เช่น เริ่มจากการระบุถึงเอนทิตีจากนั้นจึงระบุถึงแอทริบิวท์ในเอนทิตีเหล่านั้น

2. การออกแบบจากล่างขึ้นบน (Bottom-up Design) เริ่มจากการระบุถึงข้อมูลที่มีความละเอียดก่อนแล้วจึงระบุถึงกลุ่ม/ประเภทของข้อมูล เช่น การระบุถึงแอทริบิวต์ที่จำเป็นต้องทำการจัดเก็บในฐานข้อมูล จากนั้นทำการรวบรวมแอทริบิวต์ที่เกี่ยวข้องกันเพื่อระบุถึงเอนทิตีอื่น



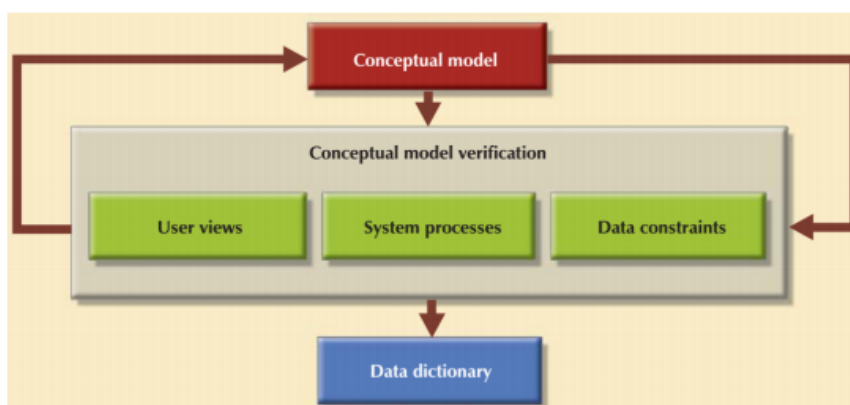
ภาพที่ 3.10 วิธีการออกแบบฐานข้อมูลแบบ top-down และ bottom-up

การเลือกวิธีการออกแบบฐานข้อมูลจาก 2 วิธีข้างต้น สามารถเลือกได้จากขอบเขตของปัญหาที่จะทำการออกแบบ และความถนัดของผู้ออกแบบ โดยการออกแบบในลักษณะ bottom-up จะมีประสิทธิภาพในการออกแบบฐานข้อมูลขนาดเล็กที่มีเอนทิตี แอทริบิวต์ ความสัมพันธ์ และเงื่อนไขไม่มากนัก แต่การออกแบบในลักษณะ top-down เหมาะกับการออกแบบฐานข้อมูลขนาดใหญ่ที่มีความหลากหลาย ซับซ้อน ซึ่งวิธีการนี้จะช่วยให้สามารถบริหารจัดการได้โดยง่าย

## 2.6 การออกแบบตามลักษณะการบริหารงานองค์กร

การเลือกกลยุทธ์การออกแบบฐานข้อมูลแบบใดนั้นขึ้นกับขอบเขตและขนาดของระบบฐานข้อมูลที่จะทำการออกแบบ แต่อย่างไรก็ตามควรพิจารณาเกี่ยวกับโครงสร้างขององค์กรที่มีลักษณะเป็นแบบรวมศูนย์ (Centralized) หรือ แบบกระจาย (Decentralized) ที่จะเกี่ยวเนื่องกับหลักการการออกแบบฐานข้อมูลมีรายละเอียดดังนี้

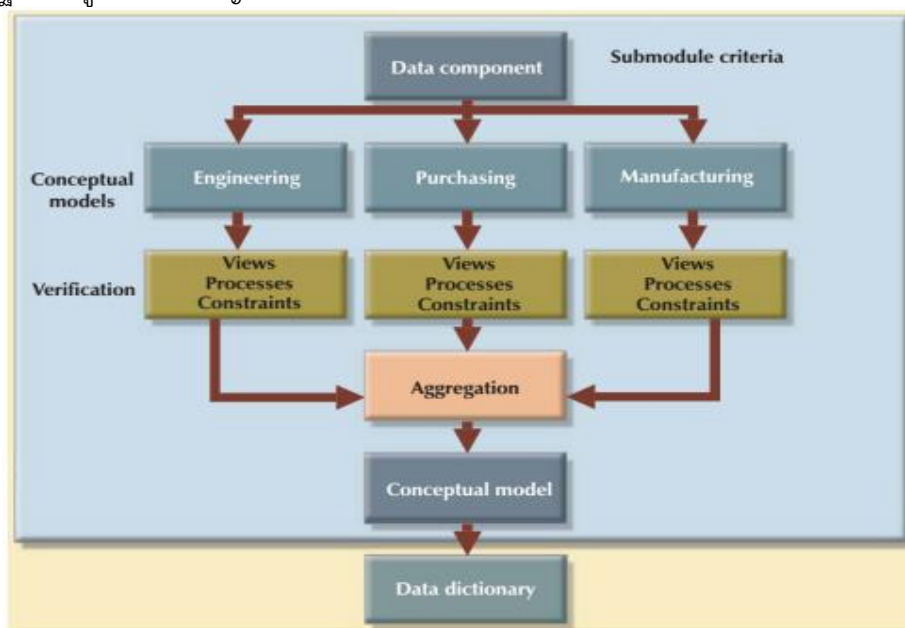
1. การออกแบบแบบรวมศูนย์ (Centralized Design) เป็นหลักการออกแบบฐานข้อมูลขนาดเล็กที่มีการดำเนินการกับฐานข้อมูลไม่มากนัก ดังนั้น การออกแบบฐานข้อมูลอาจดำเนินการด้วยผู้ออกแบบเพียงคนเดียวหรืออาจดำเนินการด้วยทีมผู้ออกแบบกลุ่มเล็กๆ การออกแบบแบบรวมศูนย์ จะประกอบไปด้วยขั้นตอนดังภาพที่ 3.11



ภาพที่ 3.11 ขั้นตอนการออกแบบฐานข้อมูลแบบรวมศูนย์

จากภาพที่ 3.11 ประกอบด้วยการระบุถึงปัญหา/การดำเนินการต่างๆ การออกแบบกรอบความคิด การตรวจสอบความคิดที่ออกแบบไว้ การกำหนดฟังก์ชันการทำงาน/การดำเนินการต่างๆของระบบ การกำหนดเงื่อนไขของข้อมูล และการตรวจสอบกรอบความคิดว่ามีความสอดคล้อง/สามารถตอบสนองต่อการดำเนินการทางธุรกิจและสิ่งที่ผู้ใช้ต้องการ ตามลำดับ เมื่อทำการออกแบบตามขั้นตอนข้างต้นทำให้ได้กรอบความคิดและพจนานุกรมข้อมูลที่บอกถึงโครงสร้างการจัดเก็บข้อมูล

2. การออกแบบแบบกระจาย (Decentralized Design) ถูกนำมาใช้ในกรณีที่ระบบฐานข้อมูลมีเอนทิตีที่ต้องพิจารณาจำนวนมาก ความสัมพันธ์ระหว่างเอนทิตีที่มีความซับซ้อน และการดำเนินการกับฐานข้อมูลที่มีความซับซ้อน และสำหรับองค์กรที่มีการดำเนินการในหลายพื้นที่ โดยแต่ละพื้นที่จะมีการจัดเก็บข้อมูล ระบบฐานข้อมูลมีขนาดใหญ่ และมีความซับซ้อน แสดงขั้นตอนดังภาพที่ 3.12



ภาพที่ 3.12 ขั้นตอนการออกแบบฐานข้อมูลแบบกระจาย

จากภาพการออกแบบนี้ต้องมีการทำงานเป็นทีมซึ่งจะแบ่งการทำงานออกเป็นมอดูลย่อยแล้วกำหนดให้แต่ละคนในทีมรับผิดชอบมอดูลนั้นๆที่ออกแบบบางส่วน of ระบบฐานข้อมูลเท่านั้น จากนั้นต้องทำการออกแบบกรอบความคิด และทำการตรวจสอบกรอบความคิดที่ได้ออกแบบไว้ เมื่อทุกกรอบความคิดย่อยได้ถูกตรวจสอบจนครบถ้วนแล้วจะถูกรวมเข้าด้วยกันเพื่อสร้างเป็นกรอบความคิดของทั้งระบบ และท้ายสุดเป็นการกำหนดพจนานุกรมข้อมูล



## ตัวอย่างการออกแบบฐานข้อมูล

### ความต้องการสารสนเทศรายงานผลการศึกษา

รายงานการศึกษา				
ภาคการศึกษาที่ 1/25XX			วันที่ dd/mm/yy	
รหัสนักศึกษา : 58700012345			สาขาวิชา : เทคโนโลยีสารสนเทศ	
ชื่อนักศึกษา : สุดสวย มงคลชัย				
รหัสวิชา	ชื่อวิชา	อาจารย์	ห้องเรียน	เกรด
204201	IS II (ระบบสารสนเทศ 2)	ผศ.สมถวิล	CI 101	A
204308	ISM (การจัดการระบบสารสนเทศ)	Prof. Peter	CI 201	B
204216	Media (การจัดการสื่อ)	hang ดร.สุขใจ	CI 202	C

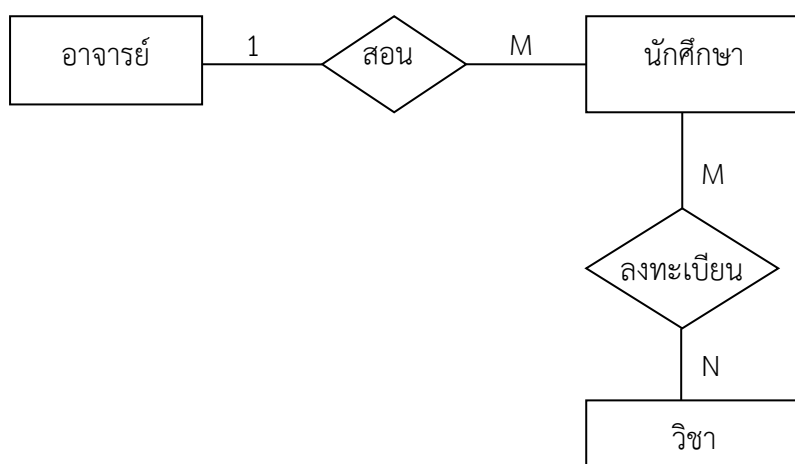
จากรายงานที่ต้องการ ข้อมูลที่ต้องใช้สำหรับแสดงผลรายงาน จะประกอบด้วย 8 แอททริบิวต์ ได้แก่ รหัสนักศึกษา, ชื่อนักศึกษา, วิชาเอก, รหัสวิชา, ชื่อวิชา, ชื่ออาจารย์, รหัสห้อง, เกรด

### แสดงชุดข้อมูล

รหัสนักศึกษา	ชื่อนักศึกษา	วิชาเอก	รหัสวิชา	ชื่อวิชา	ชื่ออาจารย์	รหัสห้อง	เกรด
B5870001	สุดหล่อ	IC	204201	IS II	ผศ.สมถวิล	CI 101	B
			204216	Media	ดร.สุขใจ	CI 202	C
B5871239	น้ำนิ่ง	MIS	204201	IS II	ผศ.สมถวิล	CI 101	A
			204308	ISM	โชค	B1128	B
			204216	Media	ดร.สุขใจ	CI 202	C
B5873333	มาลี	IS	204201	IS II	ผศ.สมถวิล	CI 101	C
			204314	Lib. Auto.	โชค	B1114	A

### การออกแบบฐานข้อมูลระดับแนวคิด

- พิจารณาเค้าร่างข้อมูล ประกอบด้วย เอนทิตีอาจารย์ เอนทิตีนักศึกษา เอนทิตีวิชา
- ความสัมพันธ์ คือ อาจารย์สอน นักศึกษาลงทะเบียน สร้าง ERD แสดงความสัมพันธ์ระหว่าง เอนทิตี



- ระบบโครงสร้างข้อมูลพิจารณาได้ ดังนี้ คือ ตารางหรือเอนทิตี เป็นหน่วยที่ใช้จัดเก็บชุดข้อมูลในระบบ อาจมีได้หลายตารางหรือขึ้นอยู่กับความสัมพันธ์ของรายการข้อมูล
- แอทริบิวต์หรือรายการข้อมูลที่อยู่ในรีเลชันอาจมีลักษณะเหมือนกันหรือซ้ำกัน ซึ่งจะต้องแก้ไขโดยการสร้างรีเลชันใหม่ เช่น นักศึกษา 1 คน สามารถเรียนได้หลายวิชา
- การสร้างคีย์ (key) เพื่อระบุความสัมพันธ์ของรีเลชัน

### การออกแบบฐานข้อมูลระดับตรรกะ

**การนอร์มัลไลเซชัน** เป็นกระบวนการนำรีเลชันมาแตกให้อยู่ในรูปแบบบรรทัดฐานที่เหมาะสม (normal form) โดยมีวัตถุประสงค์เพื่อ

1. ลดความซ้ำซ้อนของข้อมูล เมื่อลดความซ้ำซ้อนก็ทำให้ลดเนื้อหาที่ใช้ในการจัดเก็บข้อมูล
2. ลดปัญหาความไม่ถูกต้องของข้อมูล เมื่อข้อมูลไม่เกิดความซ้ำซ้อนทำให้การปรับปรุงข้อมูลสามารถทำได้จากแหล่งข้อมูลเพียงแหล่งเดียว ทำให้ช่วยลดความผิดพลาดที่อาจเกิดจากการปรับปรุงข้อมูล (Update Anomalies) ซึ่งประกอบด้วย ข้อผิดพลาดจากการเพิ่มข้อมูล (Insertion Anomalies) การลบข้อมูล (Deletion Anomalies) และการแก้ไขข้อมูล (Modification Anomalies)

การนอร์มัลไลเซชันมีได้ถึง 5 ระดับ ในระดับที่ 3 ก็จัดว่าเพียงพอ

1. 1NF - กำจัดกลุ่มของข้อมูลที่มีความซ้ำซ้อน (Repeating Group)
  2. 2NF - กำจัดการขึ้นต่อกันบางส่วน (Partial Dependency)
  3. 3NF - กำจัดการขึ้นต่อกันของนอลคีย์ (Transitive Dependency)
- สำหรับขั้นที่ 4 และ 5 เป็นนอร์มัลไลเซชันแบบ Boyce-Codd Normal Form (BCNF)  
จากหน่วยที่ใช้จัดเก็บชุดข้อมูลในระบบ คือ รีเลชันนักศึกษา ที่มีโครงสร้าง ดังนี้

รีเลชันนักศึกษา (รหัสนักศึกษา , ชื่อนักศึกษา , วิชาเอก , รหัสวิชา , ชื่อวิชา , ชื่ออาจารย์ , รหัสห้อง , เกรด)

### การทำนอร์มัลไลเซชัน ขั้นที่ 1 (1NF- First Normal Form)

- นำรายการข้อมูลที่ซ้ำซ้อนไปสร้างรีเลชันใหม่
- นำคีย์หลักของรีเลชันเดิมไปด้วย
- กำหนดคีย์หลักของรีเลชันใหม่ตามความสัมพันธ์ของระเบียบจากรีเลชันนักศึกษา
- รหัสนักศึกษา 1 คน มีชื่อนักศึกษาได้ 1 และวิชาเอกได้ 1 และใช้เป็นคีย์หลัก
- รหัสนักศึกษา 1 คน สามารถเรียนได้ M รหัสวิชา ดังนั้น รหัสวิชาจึงเป็นกลุ่มข้อมูลที่ซ้ำซ้อนต้องนำไปสร้างรีเลชันใหม่และกำหนดคีย์ใหม่ เมื่อเสร็จนอร์มัลไลเซชันขั้นที่ 1 จะได้รีเลชันดังนี้

**รีเลชันนักศึกษา (รหัสนักศึกษา , ชื่อนักศึกษา , วิชาเอก)**

**รีเลชันนักศึกษา\_วิชาเรียน (รหัสนักศึกษา , รหัสวิชา , ชื่อวิชา , ชื่ออาจารย์ , รหัสห้อง , เกรด)**

ใช้ รีเลชันนักศึกษา\_วิชาเรียน ไปทำนอร์มัลไลเซชัน ขั้นที่ 2

### การทำนอร์มัลไลเซชัน ขั้นที่ 2 (2NF- Second Normal Form)

- นำรายการข้อมูลที่เกี่ยวข้องกับคีย์หลักในรีเลชันไปสร้างรีเลชันใหม่
- นำคีย์หลักของรีเลชันเดิมไปด้วย
- กำหนดคีย์หลักของรีเลชันใหม่

จากรีเลชันนักศึกษา - วิชาเรียน ที่ต้องนำไปสร้างรีเลชันใหม่ให้ชื่อรีเลชันการลงทะเบียน

- รหัสนักศึกษา 1 คน เมื่อลงทะเบียนเรียน 1 วิชา จะมีได้ 1 เกรด
- ข้อมูลที่เหลือนำไปสร้างรีเลชันใหม่ ซึ่งได้แก่ ชื่อวิชา ,ชื่ออาจารย์ ,รหัสห้อง พร้อมกำหนดคีย์หลัก จะได้รีเลชันดังนี้

รีเลชันการลงทะเบียน (รหัสนักศึกษา , รหัสวิชา , เกรด)

รีเลชันวิชา\_ผู้สอน\_ห้องบรรยาย (รหัสวิชา , ชื่อวิชา , ชื่ออาจารย์ , รหัสห้อง)

ใช้ รีเลชันวิชา\_ผู้สอน\_ห้องบรรยาย ไปทำนอร์มัลไลเซชัน ขั้นที่ 3

การทำนอร์มัลไลเซชัน ขั้นที่ 3 (3NF - Third Normal Form)

- นำรายการข้อมูลที่เกี่ยวข้องกับคีย์หลักและรายการข้อมูลอื่น ๆ ไปสร้างรีเลชันใหม่
- กำหนดคีย์หลักในรีเลชันใหม่
- กำหนดคีย์หลักรีเลชันใหม่ ให้เป็นคีย์นอกของรีเลชันเดิม เพื่อให้เชื่อมโยงระหว่าง 2 รีเลชันได้

จากรีเลชันวิชา\_ผู้สอน\_ห้องบรรยาย ที่ต้องนำไปสร้างรีเลชันใหม่

- รหัสวิชา 1 จะมีชื่อวิชาได้ 1 และสัมพันธ์กับชื่ออาจารย์
- ข้อมูลที่เหลือ รหัสห้อง
- ใช้ชื่ออาจารย์เป็นคีย์เพื่อเชื่อมโยงระหว่างรีเลชัน จะได้รีเลชันดังนี้

รีเลชันรายวิชา (รหัสวิชา , ชื่อวิชา , ชื่ออาจารย์)

รีเลชันอาจารย์ (ชื่ออาจารย์ , รหัสห้อง)

เมื่อจบนอร์มัลไลเซชันขั้นที่ 3 แล้วจะได้รายการข้อมูลหรือแอททริบิวต์แต่ละรายการเป็นอิสระต่อกัน เมื่อมีการแทรกข้อมูลเพิ่ม ลบ หรือทำการปรับเปลี่ยนค่าในรายการใดรายการหนึ่งจะไม่ทำให้ข้อมูลผิดปกติ กล่าวคือ มีข้อมูลชนิดเดียวกันในหลายแห่งจะต้องตรงกัน ถ้าหากว่าข้อมูลยังผิดปกติอีกก็ต้องจัดทำนอร์มัลไลเซชันในขั้นที่ 4 ขั้นที่ 5 ต่อไป

เมื่อสิ้นสุดนอร์มัลไลเซชันขั้นที่ 3 แล้วจะได้รีเลชันหรือตารางพร้อมแอททริบิวต์รวม 4 รีเลชัน แสดง เอนทิตีและแอททริบิวต์ด้านการลงทะเบียนเรียน

การลงทะเบียน	รหัสนักศึกษา รหัสวิชา เกรด	นักศึกษา	รหัสนักศึกษา ชื่อนักศึกษา วิชาเอก
รายวิชา	รหัสวิชา ชื่อวิชา ชื่ออาจารย์	อาจารย์	ชื่ออาจารย์ รหัสห้อง

ระดับที่ 1 พิจารณาความสัมพันธ์แบบเดียว นักศึกษา 1 คน มีได้ 1 วิชาเอก จะได้รีเลชันนักศึกษา

รีเลชันนักศึกษา

รหัส นักศึกษา	ชื่อนักศึกษา	วิชาเอก

(1)

รีเลชันนักศึกษา - วิชาเรียน

ยังมีความสัมพันธ์แบบกลุ่มของแอททริบิวต์บางตัวอยู่

รหัสนักศึกษา	รหัสวิชา	ชื่อวิชา	ชื่ออาจารย์	รหัสห้อง	เกรด

(2)

**ระดับที่ 2** จากรีเลชัน (2) จัดความสัมพันธ์แบบเดี่ยวได้ นักศึกษา 1 คน เรียนวิชา มีได้ 1 เกรด ทำให้ได้ ตาราง (3)

รีเลชันการลงทะเบียน

รหัสนักศึกษา	รหัสวิชา	เกรด

(3)

รีเลชันวิชา\_ผู้สอน\_ห้องบรรยาย ยังมีความสัมพันธ์แบบกลุ่มของแอททริบิวต์บางตัวอยู่

รหัสวิชา	ชื่อวิชา	ชื่ออาจารย์	รหัสห้อง

(4)

**ระดับที่ 3** จากรีเลชัน (4) วิชา 1 วิชา มีได้ 1 ชื่อ จะได้ตาราง (5) และ (6)

รีเลชันรายวิชา

รหัสวิชา	ชื่อวิชา	ชื่ออาจารย์

(5)

รีเลชันอาจารย์

ชื่ออาจารย์	รหัสห้อง

(6)

เมื่อจบการนอร์มัลไลซ์ระดับที่ 3 จะได้รายการข้อมูลแต่ละรายการเป็นอิสระต่อกัน คือแอททริบิวต์ จะมีค่า 1 ค่า ระบบฐานข้อมูลข้างต้นจะประกอบด้วย 4 เอนทิตี เขียนเป็นรูปของทัวเพิล (tuple) ได้แก่

รีเลชัน (1) รีเลชันนักศึกษา (รหัสนักศึกษา, ชื่อนักศึกษา, วิชาเอก)

รีเลชัน (3) รีเลชันการลงทะเบียน(รหัสนักศึกษา, รหัสวิชา, เกรด)

รีเลชัน (5) รีเลชันรายวิชา(รหัสวิชา, ชื่อวิชา, ชื่ออาจารย์ )

รีเลชัน (6) รีเลชันอาจารย์(ชื่ออาจารย์, รหัสห้อง )

ความสัมพันธ์ที่ใช้เชื่อมโยง คือ คีย์หลัก (primary key) ใช้ขีดเส้นใต้ และคีย์นอก (foreign key) ใช้จุดประ ซึ่ง ใช้เชื่อมโยงและจะเป็นคีย์หลักของอีกรีเลชันหนึ่ง

รีเลชัน (1) คีย์หลัก คือ รหัสนักศึกษา

รีเลชัน (3) คีย์หลัก คือ รหัสนักศึกษา คีย์นอก คือ รหัสวิชา

รีเลชัน (5) คีย์หลัก คือ รหัสวิชา คีย์นอก คือ ชื่ออาจารย์

รีเลชัน (6) คีย์หลัก คือ ชื่ออาจารย์

### คำถามทบทวน

1. SDLC คืออะไร?
2. DBLC คืออะไร?
3. กฎเกณฑ์ทางธุรกิจคืออะไร? และมีความสำคัญอย่างไร?
4. ขั้นตอนการออกแบบ ER diagram ประกอบไปด้วยขั้นตอนอะไรบ้าง?
5. การเลือกซอฟต์แวร์ระบบจัดการฐานข้อมูลมีความสำคัญอย่างไร?
6. ขั้นตอนการออกแบบ logical design ประกอบไปด้วยขั้นตอนอะไรบ้าง?
7. ขั้นตอนการออกแบบ physical design ประกอบไปด้วยขั้นตอนอะไรบ้าง?
8. ความแตกต่างระหว่างการออกแบบฐานข้อมูลแบบ bottom-up และ top-down เป็นอย่างไร?