บทที่ 3

แนวคิดและทฤษฎี

3.1 ฐานข้อมูล

ฐานข้อมูลเป็นการเก็บรวบรวมข้อมูลต่าง ๆ ในองค์กรมาไว้เป็นแหล่งข้อมูลเดียวกัน ข้อมูลสามารถถูกดึงขึ้นมาใช้งานร่วมกัน และเชื่อมโยงกันได้ง่าย ประเภทของฐานข้อมูลในปัจจุบันมี 3 ประเภท ได้แก่ ประเภทรีเลชันนัล (Relational) ลำดับชั้น (Hierarchical) และเนทเวอร์ค (Network) โดยจำแนกตามโครงสร้างฐานข้อมูล

3.1.1 ฐานข้อมูลแบบรีเลชันนัล

ข้อมูลถูกเก็บอยู่ในรูปแบบของตาราง

3.1.2 ฐานข้อมูลแบบลำดับชั้น

ข้อมูลถูกเก็บโดยมีโครงสร้างเป็นแบบต้นไม้ (Tree) สัมพันธ์กันเป็นลำดับชั้นแบบแม่ (Parent) และลูก (Child) ข้อมูลในระดับบนสุดจะถูกเรียกว่าราก (Root)

3.1.3 ฐานข้อมูลแบบเนทเวอร์ค

โครงสร้างข้อมูลแบบนี้จะมีความซับซ้อน การเก็บข้อมูลเป็นลักษณะลิสท์เชื่อมโยง (Link List) หรือ ตัวชี้ (Pointers) จากเรคคอร์ดแม่ (Parent Record) เชื่อมโยงไปหาเรคคอร์ดลูก (Children Record)

3.2 สถาปัตยกรรมระบบฐานข้อมูล (Database Architecture)

ได้มีการกำหนดมาตรฐานขึ้นโดย ANSI (American National Standards Institute) ในปี 1975 เรียกว่าสถาปัตยกรรมสามระดับ (Tree-Level Architecture) ดังนี้คือ

3.2.1 ระดับโครงร่างภายใน (Internal Schema)

เป็นระดับที่แสดงรายละเอียดถึงการเก็บข้อมูลภายใน ของระบบคอมพิวเตอร์ว่าถูกแทนที่อย่างไรในระดับกายภาพ (Physical) เช่นการเก็บข้อมูลในระดับเรคคอร์ดเป็นอย่างไร และมีดัชนี (Index) อะไรบ้างเป็นต้น

3.2.2 ระดับโครงร่างเชิงความคิด (Conceptual Schema)

แสดงในส่วนข้อมูลทั้งหมดรวมกันของกลุ่มผู้ใช้ในองค์กร โดยไม่เกี่ยวข้องกับรายละเอียดการเก็บภายในคอมพิวเตอร์ นอกจากนี้ยังบรรจุความต้องการการใช้ข้อมูล กฎเกณฑ์การดำเนินการ บูรณภาพของข้อมูล (Data Integrity) เป็นต้นอีกด้วย

3.2.3 โครงร่างภายนอก (External Schema)

โครงร่างระดับนี้จะแสดงโครงสร้างข้อมูลของแต่ละกลุ่มผู้ใช้ข้อมูลแยกส่วนกัน ซึ่งข้อมูลในมุมผู้ใช้กลุ่มหนึ่งไม่สนใจหรือคำนึงถึงข้อมูลของผู้ใช้ในกลุ่มอื่น ๆ

โครงร่างฐานข้อมูลทั้ง 3 ระดับ จะมีคุณสมบัติของความเป็นอิสระของข้อมูล (Data Independence) ด้วยกล่าวคือ เมื่อมีการเปลี่ยนแปลงในโครงร่างฐานข้อมูลระดับหนึ่งจะไม่มีผลกระทบกับโครงร่างระดับที่สูงกว่า ดังนี้

1. ความเป็นอิสระของข้อมูลเชิงตรรก (Logical Data Independence)

การเปลี่ยนแปลงข้อมูลในระดับโครงร่างเชิงความคิด ไม่มีผลทำให้เกิดการเปลี่ยนแปลงข้อมูลในโครงร่างภายนอกหรือโปรแกรมประยุกต์ใช้งาน ตัวอย่างเช่น กรณีเกิดการเปลี่ยนแปลงของขนาดฐานข้อมูล โดยการเพิ่มเรคคอร์ดใหม่ หรือ ส่วนข้อมูล (Data Item) เข้าไป หรือลดขนาดฐานข้อมูลลง ด้วยการลบเรคคอร์ดหรือตัดส่วนข้อมูลออกเป็นต้น

2. ความเป็นอิสระข้อมูลเชิงกายภาพ (Physical Data Independence)

หมายถึงการเปลี่ยนแปลงข้อมูลในระดับโครงร่างภายใน ไม่มีผลทำให้ต้องเปลี่ยนแปลงข้อมูลในโครงร่างเชิงความคิด หรือในโครงร่างภายนอก ตัวอย่างเช่น ต้องการเพิ่มเติมวิธีที่ใช้ในการดึงข้อมูล เพื่อปรับปรุงความเร็วในการเข้าถึงและปรับปรุงข้อมูล ซึ่งไม่จำเป็นต้องทำการเปลี่ยนแปลงโครงร่างเชิงความคิด หากข้อมูลมีอยู่แล้วในฐานข้อมูล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.3 โมเดลข้อมูล

การนำเทคนิคโมเดลข้อมูล (Data Model) มาใช้ในการพัฒนาระบบข้อมูลมีความสำคัญในปัจจุบัน โมเดลข้อมูลจะรวบรวมข้อมูลที่สำคัญในระบบทั้งหมด ได้แก่ การสร้างรูปภาพแสดงความสัมพันธ์แบบต่าง ๆระหว่างข้อมูล อธิบายชนิดของข้อมูล (Data Type) รวมทั้งข้อกำหนดที่ควรมีสำหรับโมเดลข้อมูล ซึ่งส่วนมากมีการกำหนดวิธีการในการดึงและปรับปรุงข้อมูลในฐานข้อมูล

3.3.1 ประเภทโมเดลข้อมูล

ประเภทของโมเดลข้อมูลที่ถูกจำแนก ในที่นี้ยึดหลักความคิดตามโครงสร้างฐานข้อมูล แบ่งเป็น 2 ประเภทที่สำคัญดังนี้

3.3.1.1 โมเดลข้อมูลระดับสูง (High-Level Data Model) หรือโมเดลข้อมูลเชิงตรรก (Logical Data Model)

จุดประสงค์ของโมเดลข้อมูลระดับนี้ แสดงให้เห็นถึงข้อมูลสำหรับส่วนโครงร่างมุมมองของผู้ใช้กลุ่มต่าง ๆ (Skeletal User View) โดยไม่คำนึงถึงรายละเอียดในการติดตั้ง (Implementation) สื่อความหมายโดยอาศัยแผนภาพประกอบ

ขั้นตอนการสร้างโมเดลข้อมูลเชิงตรรกที่ดีควรมีบรรทัดฐาน ดังนี้

- มีความถูกต้องตรงกันของความต้องการและการจัดองค์กรของสารสนเทศ
- ง่ายต่อการทำความเข้าใจสำหรับบุคคลทั่วไป
- ขจัดความซ้ำซ้อนของสารสนเทศ
- ข้อมูลสามารถถูกใช้ร่วมกันได้ระหว่างผู้ใช้กลุ่มต่าง ๆโดยคำนึงถึงผลประโยชน์รวมขององค์กรเป็นหลัก
- สามารถตอบสนองความต้องการในการเปลี่ยนแปลง เพิ่มหรือขยายขอบเขตที่อาจเกิดขึ้นในอนาคตได้
- มีบรูณภาพในแง่การนำไปใช้งานและการบริหารสารสนเทศ

3.3.1.2 โมเดลข้อมูลระดับต่ำ (Low-Level Data Model) หรือโมเดลข้อมูลเชิงกายภาพ (Physical Data Model)

โมเดลข้อมูลระดับนี้เป็นรายละเอียดในช่วงการเก็บลงเครื่องคอมพิวเตอร์ เช่น ฟอร์แมทของเรคคอร์ด การเรียงลำดับเรคคอร์ด และวิถีทางในการเข้าถึงข้อมูล (Access Path) อย่างมีประสิทธิภาพ การกู้ข้อมูล (Recovery) เป็นต้น ซึ่งเหล่านี้ไม่ใช่ข้อมูลสำหรับผู้ใช้ปลายทาง (End User)

3.4 รีเลชันนัลโมเดล

แพร่หลายเป็นที่รู้จักโดย Dr. E.F. Codd ในปี ค.ศ.1970 มาจากทฤษฎีคณิตศาสตร์ แต่โมเดลนี้มีโครงสร้างที่ง่าย ดังนั้นถึงแม้บุคคลที่ไม่คุ้นเคยกับทฤษฎีเหล่านั้นก็สามารถเข้าใจได้

รีเลชันนัลโมเดล มีส่วนประกอบที่สำคัญ 3 ส่วนคือ

3.4.1 โครงสร้างข้อมูล (Data Structure) ข้อมูลถูกรวบรวมเป็นความสัมพันธ์ กล่าวคือ เป็นตาราง (Table) 2 มิติ คือประกอบด้วยคอลัมน์ (Columns) และแถว (Rows)

และมีคุณสมบัติพิเศษดังนี้

- รายการในแต่ละคอลัมน์เป็นค่าเดี่ยว (Single-Valued)
- รายการในแต่ละคอลัมน์เป็นชนิดเดียวกัน หรืออีกนัยหนึ่งรายการในแต่ละคอลัมน์มาจากโดเมน (Domain) เดียวกัน
- รายการแต่ละแถวมีความเป็นเอกลักษณ์ (Unique) ไม่ซ้ำกับแถวอื่น ๆ
- ลำดับการเรียงคอลัมน์ (จากซ้ายไปขวา) ไม่มีความสำคัญ
- ลำดับการเรียงแถว (จากบนลงล่าง) ไม่มีความสำคัญ
- แต่ละคอลัมน์มีชื่อเป็นเอกลักษณ์ สำหรับการอ้างถึง

3.4.2 <u>การดำเนินการโดยผู้ใช้บนตาราง</u> (Data Manipulation) เช่นการดำเนินการเลือก (Select) การดำเนินการรวมหรือตัดตอนข้อมูลจากแถวหนึ่งของตารางหนึ่งกับแถวอื่น ๆในตารางเดียวกันหรือตารางอื่น ๆ (Join) เป็นต้น

3.4.3 <u>บรูณภาพของข้อมูล</u> (Data Integrity)

ในโมเดลข้อมูลแบบรีเลชันนัลนี้กฎบรูณภาพ (Intergrity Rules) เป็นสิ่งที่มีความสำคัญมาก ข้อกำหนดที่กำหนดช่วยให้ข้อมูลในฐานข้อมูลมีความถูกต้องน่าเชื่อถือ และสมบรูณ์โดยเฉพาะในระบบผู้ใช้หลายคน (Multi Users) ซึ่งใช้ข้อมูลร่วมกัน หากปราศจากการควบคุมที่ดีจะมีผลทำให้ข้อมูลผิดพลาด และขยายความผิดพลาดต่อเนื่องกันไป

กฎบรูณภาพข้อมูลที่สำคัญมี 3 ประเภทด้วยกันดังนี้

3.4.3.1 กฎบรูณภาพของเอนติตี (Entity Integrity Rule)

กำหนดให้แอตตริบิวต์ที่เป็นคีย์หลัก หรือส่วนประกอบของคีย์หลักต้องมีค่าเสมอ

3.4.3.2 กฎบรูณภาพในการอ้างอิง (Referential Intergrity Rule)

กำหนดให้ถ้าเอนติตีที่มีฟอร์เรนจ์คีย์ แล้วค่าของข้อมูลในแอตตริบิวต์หรือกลุ่มแอตตริบิวต์ที่เป็นฟอร์เรนจ์คีย์นั้นอาจมีค่าเป็นนัล (Null Value) ได้ หรือหากไม่มีค่าเป็นนัลก็ต้องสามารถอ้างอิงได้ตรงกันกับค่าเดียวกันนี้ในอีกเอนติตีหนึ่ง ซึ่งแอตตริบิวต์นี้หรือกลุ่มแอตตริบิวต์นี้เป็นคีย์หลักอยู่

3.4.3.3 กฎบรูณภาพของโดเมน (Domain Intergrity Rule)

กำหนดให้ควรมีการกำหนดโดเมนสำหรับทุก ๆคอลัมน์ไม่ว่าจะเป็นคีย์หลัก ฟอร์เรนจ์คีย์ หรือแอตตริบิวต์ที่ไม่ใช่คีย์ในตารางรีเลชันนัล

018297

โดเมนในที่นี้รวมความหมายถึงสิ่งต่าง ๆ ดังต่อไปนี้

- ชนิดของข้อมูล
- ความยาว
- ช่วงการตรวจสอบค่าที่เป็นไปได้
- ค่าโดยปริยาย (Default Values)
- ความเป็นเอกลักษณ์ (Uniqueness)
- ความไม่มีค่า (Nullability)

3.5 โมเดลข้อมูลเชิงตรรก (Flemming and Von Halle, 1989)

โมเดลเชิงตรรกที่จะกล่าวถึงนี้จัดเป็นโมเดลประเภทรีเลชันนัลประเภทหนึ่ง โมเดลข้อมูลชนิดนี้เน้นเรื่องความหมายของข้อมูลในขอบเขตที่สนใจ ประกอบด้วย เอนติตี แอตตริบิวต์ และรีเลชันชิป จุดประสงค์โดยสำคัญเพื่อให้เข้าใจข้อมูลของผู้ใช้ได้ดี และง่ายยิ่งขึ้น อันได้แก่ โครงสร้างของฐานข้อมูล และการดำเนินงานต่าง ๆ ที่ต้องมี เช่น การดึง และการปรับปรุงข้อมูล

3.5.1 เอนติตี หมายถึง สิ่งที่อาจมีอยู่ในระบบ เช่น พนักงาน นิสิต ผู้สอน หรือเป็นสิ่งที่เป็นนามธรรม เช่น รายวิชา หรือเป็นเหตุการณ์เช่น การลงทะเบียนเรียนของนิสิต เป็นต้น โดยที่เอนติตีอาจเป็นหน่วยพื้นฐานที่มีอยู่แล้วเป็นหลัก หรือเป็นหน่วยที่พึ่งพิงกับหน่วยพื้นฐานหลัก เช่น เอนติตี "บุตร" พึ่งพิงอยู่กับเอนติตี "พนักงาน"เป็นต้น โดยปกติแล้วเอนติตีถูกแทนที่ด้วยรูปสี่เหลี่ยมในรูปภาพ

3.5.2 แอตตริบิวต์ หมายถึง ส่วนที่เป็นคุณสมบัติต่าง ๆ ที่ประกอบ หรืออธิบายเอนติตีว่าเป็นอย่างไร และเป็นหน่วยที่เล็กที่สุดในเอนติตี ตัวอย่างเช่น ในเอนติตี "ผู้สอน" ประกอบด้วย แอตตริบิวต์ "ชื่อ" แอตตริบิวต์ "สาขาที่ถนัด" แอตตริบิวต์ "ประสบการณ์การสอน" เป็นต้น

3.5.2.1 ประเภทของแอตตริบิวต์

3.5.2.1.1 คีย์หลัก (Primary Key Attribute) หมายถึง แอตตริบิวต์หรือกลุ่มของแอตตริบิวต์ที่มีจำนวนแอตตริบิวต์น้อยที่สุดของเอนติตีที่มีคุณสมบัติเป็นเอกลักษณ์ ทำให้สามารถใช้ชี้ระบุถึงสมาชิกหนึ่ง ๆในเอนติตีนั้นได้ ในเอนติตีหนึ่งจะมีแอตตริบิวต์ที่เป็นคีย์หลักได้เพียงชุดเดียวเท่านั้น

3.5.2.1.2 คีย์รอง (Alternative Key Attribute) ความหมายเหมือนกับคีย์หลัก แต่มีความสำคัญน้อย

กว่าเนื่องจากแอตตริบิวต์ หรือกลุ่มแอตตริบิวต์นั้นอาจไม่มีค่าก็ได้ จึงถูกเลือกให้เป็นคีย์รอง ในเอนติตีหนึ่งอาจมีแอตตริบิวต์ประเภทนี้ได้หลายชุด สัญญลักษณ์ที่ใช้แทนในรูปภาพคือ

AKn โดย n = 1,2,3.....

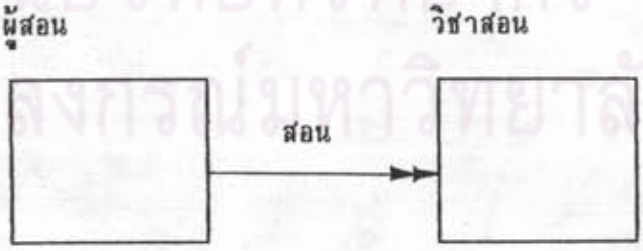
3.5.2.1.3 <u>ฟอร์เรนจ์คีย์</u> (Foreign Key Attribute) หมายถึง แอตตริบิวต์หรือกลุ่มแอตตริบิวต์ที่ปรากฏในเอนติตีลูก และแอตตริบิวต์หรือกลุ่มแอตตริบิวต์เดียวกันนี้เป็นคีย์หลักของเอนติตีอื่น ซึ่งเป็นเอนติตีแม่ สัญญลักษณ์ที่ใช้แทนในรูปภาพคือ

FKn โดย n = 1,2,3.....

3.5.2.1.4 <u>แอตตริบิวต์ที่ไม่ใช่คีย์</u> (Non-Key Attribute) ความหมายเหมือนกับคำนิยามทั่วไป ดังข้างบนที่ได้กล่าวไปแล้ว

3.5.2.1.5 <u>ดีไรฟ์แอตตริบิวต์</u> (Derived Attribute) หมายถึงแอตตริบิวต์ที่เกี่ยวพันกับแอตตริบิวต์อื่น การบันทึกค่าของแอตตริบิวต์นั้นอาจไม่จำเป็นเพราะสามารถหาได้จากแอตตริบิวต์อื่น ๆ ในเอนติตีเดียวกันหรือต่างเอนติตีกันได้ เช่นจากการคำนวณ เป็นต้น

3.5.3 <u>รีเลชันชิป</u> เป็นความสัมพันธ์ระหว่างเอนติตี ซึ่งแทนด้วยเส้นที่เชื่อมกันระหว่างเอนติตี แต่ละรีเลชันชิปมีทิศทางและอัตราส่วนคาร์ดินัลลิตี (Cardinality) ทิศทางจะแสดงให้ทราบถึงเอนติตีแม่ (Parent Entity) และเอนติตีลูก (Child Entity) โดยเอนติตีในทิศทางหัวลูกศรเป็นเอนติตีแม่ สำหรับอัตราส่วนคาร์ดินัลลิตี หมายถึง ค่าอัตราส่วนโดยเฉลี่ยหรือกะประมาณของสมาชิกที่เกิด (Occurence) ระหว่างเอนติตี ตัวอย่างของรีเลชันชิป เช่น เอนติตี "วิชาสอน" มีความสัมพันธ์กับเอนติตี "ผู้สอน" ด้วยรีเลชันชิป "สอน" เป็นต้น



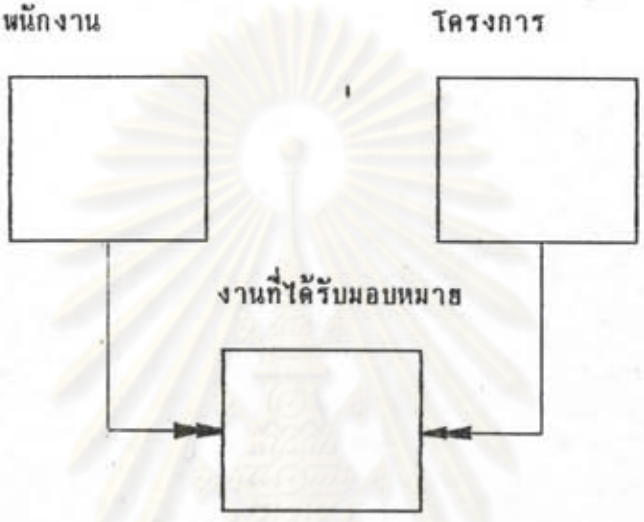
รูปที่ 3.1 แสดงความสัมพันธ์ของ 2 เอนติตี

สำหรับความสัมพันธ์ระหว่าง 2 เอนติตีมีด้วยกัน 3 รูปแบบ คือ

1. แบบหนึ่งต่อหนึ่ง (One to One, 1:1 Relationship)
2. แบบหนึ่งต่อหลาย (One to Many, 1:M Relationship)

3. แบบหลายต่อหลาย (Many to Many, M:N Relationship)

สำหรับความสัมพันธ์รูปแบบนี้ต้องทำการแปลงให้อยู่ในรูปแบบที่ง่ายขึ้น เนื่องจากระบบจัดการฐานข้อมูลในปัจจุบันจำนวนมากไม่สนับสนุนความสัมพันธ์ลักษณะนี้โดยตรง ตัวอย่างเช่นเอนติตี "พนักงาน" มีความสัมพันธ์กับ เอนติตี "โครงงาน" แบบ M:N ทำการแปลงโดยสร้างเอนติตีใหม่ขึ้นมาเอนติตีหนึ่ง ในที่นี้คือ เอนติตี "ความรับผิดชอบงาน" ดังนั้นจึงเกิดความสัมพันธ์แบบ 1:M ขึ้น 2 ความสัมพันธ์ ดังรูป



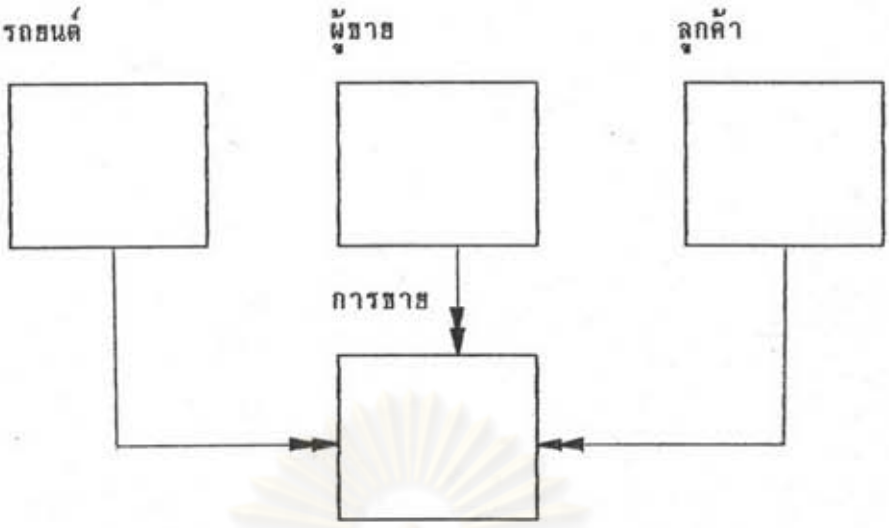
รูปที่ 3.2 แสดงความสัมพันธ์ของ 2 เอนติตีแบบหลายต่อหลาย

3.5.3.1 ประเภทของรีเลชันชิป

3.5.3.1.1 ความสัมพันธ์ระหว่าง 2 เอนติตี

3.5.3.1.2 ความสัมพันธ์ระหว่าง 3 เอนติตีหรือมากกว่า (Complex Relationship)

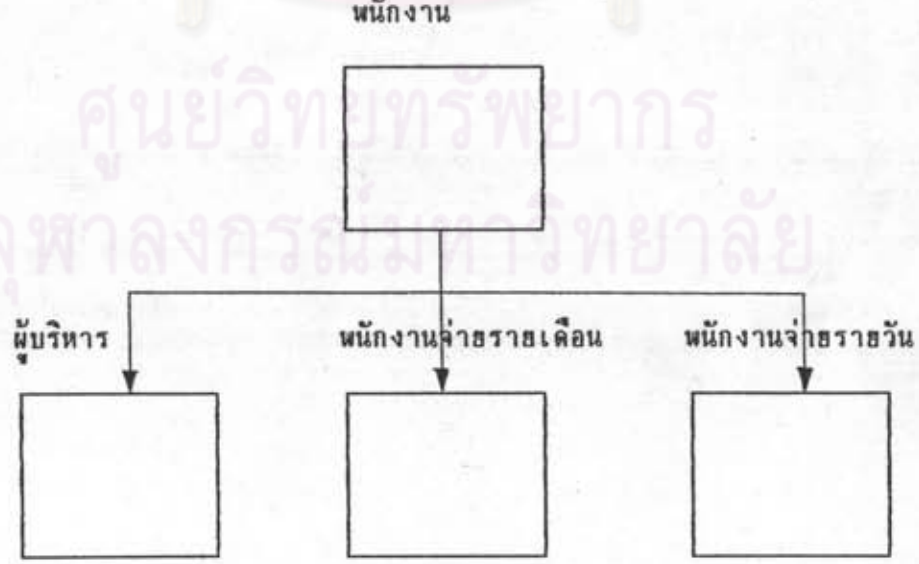
ตัวอย่างเช่น "การขายรถ" เกิดจากความสัมพันธ์ระหว่าง เอนติตี "รถยนต์" เอนติตี "ผู้ขาย" และ เอนติตี "ลูกค้า" ทั้งหมด 3 เอนติตี โมเดลถูกแปลงให้ง่ายขึ้นโดยสร้างเอนติตีใหม่ขึ้นมา ในที่นี้คือ เอนติตี "การขาย" ดังรูปที่ 3.3



รูปที่ 3.3 แสดงความสัมพันธ์แบบ 3 เอนติตี

3.5.3.1.3 ความสัมพันธ์แบบซูเปอร์ไทป์-ซับไทป์ (Supertype - Subtype Relationship)

เอนติตีซับไทป์ (Subtype) เป็นเซตย่อยของเอนติตีซูเปอร์ไทป์ (Supertype) ซูเปอร์ไทป์และซับไทป์จะหมายความถึงสิ่งเดียวกัน แต่ซับไทป์จะบรรจุด้วยแอตตริบิวต์ที่เพิ่มเติมนอกเหนือจากในซูเปอร์ไทป์ตามคุณสมบัติพิเศษเพิ่มเติมที่จำแนกประเภทไปตามแต่ละซับไทป์นั้น ๆ ความสัมพันธ์ที่เกิดระหว่างเอนติตีซูเปอร์ไทป์ และเอนติตีซับไทป์เป็นแบบ 1:1 ตัวอย่างเช่น เอนติตีซูเปอร์ไทป์ "พนักงาน" มีเอนติตีซับไทป์คือ เอนติตีซับไทป์ "ผู้บริหาร" เอนติตีซับไทป์ "พนักงานจ่ายรายเดือน" และเอนติตีซับไทป์ "พนักงานจ่ายรายวัน" เป็นต้น



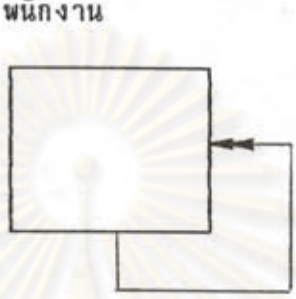
รูปที่ 3.4 ความสัมพันธ์แบบซูเปอร์ไทป์-ซับไทป์

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.5.3.1.4 ความสัมพันธ์แบบรีเคอซีฟ

(Recursive Relationship)

ความสัมพันธ์แบบนี้มีลักษณะคือเกิดความสัมพันธ์ในทิศทางทั้ง 2 ซึ่งตรงกันข้ามกันภายในเอนติตีเดียวกัน ดังตัวอย่างในรูป 3.5 พนักงานหนึ่งคนบริหาร พนักงานอายหลายคน และพนักงานคนหนึ่งถูกบริหารโดยพนักงานคนหนึ่ง



รูปที่ 3.5 ความสัมพันธ์แบบรีเคอซีฟ

3.6 ขั้นตอนในการออกแบบฐานข้อมูล

3.6.1 ขั้นตอนที่ 1

ทำการรวบรวมและวิเคราะห์ความต้องการใช้ข้อมูลของผู้ใช้ โดยการไปสัมภาษณ์ผู้ใช้ให้ทราบถึงรายละเอียดอินพุต เอาต์พุต เอกสารต่าง ๆ สภาพแวดล้อมที่จำเป็นการปฏิบัติงาน และกลุ่มผู้ใช้ข้อมูล เป็นต้น การวิเคราะห์ข้อมูลไม่ควรพิจารณาแต่เพียงความต้องการในปัจจุบันเท่านั้นแต่ควรคำนึงความต้องการในอนาคตด้วย

3.6.2 ขั้นตอนที่ 2

เป็นขั้นตอนการออกแบบฐานข้อมูลเชิงความคิด (Conceptual Database Design) ซึ่งจะเป็นอิสระจากระบบจัดการฐานข้อมูล สิ่งที่ได้จากขั้นตอนนี้คือโครงร่างเชิงความคิด (Conceptual Schema) ซึ่งก็คือการที่สามารถกำหนดเกี่ยวกับ เอนติตี รีเลชันชิป แอตตริบิวต์ และรายละเอียดอื่น ๆในระบบได้เช่น ชนิดของข้อมูล (Data Type) ข้อกำหนดต่าง ๆ วิธีการเข้าถึงข้อมูล ความถี่ในการเข้าถึง ปริมาณข้อมูล เป็นต้น

ในขั้นตอนนี้ต้องสามารถแน่ใจแล้วว่าความต้องการข้อมูลของผู้ใช้ที่เก็บรวบรวมได้ไม่มีความขัดแย้งกัน

3.6.3 ขั้นตอนที่ 3

ขั้นตอนนี้เป็นการแปลงโมเดลข้อมูลระดับสูง ที่ได้จากขั้นตอนที่แล้วไปสู่โมเดลข้อมูลที่อิงระบบจัดการฐานข้อมูลที่เลือกแล้ว

3.6.4 ขั้นตอนที่4

เป็นขั้นตอนสุดท้ายคือการออกแบบข้อมูลเชิงกายภาพ (Physical Database Design) ในขั้นตอนนี้ทำการกำหนดโครงสร้างการเก็บข้อมูลภายใน และการจัดองค์กรของไฟล์ (File Organization) ของฐานข้อมูลนั้น

3.7 ขั้นตอนการออกแบบโมเดลข้อมูล

3.7.1 ขั้นตอนในการออกแบบโมเดลข้อมูลเชิงตรรก

ขั้นตอนการสร้างโมเดลข้อมูลเชิงตรรกที่ดีจะเป็นผลให้ได้ฐานข้อมูลที่ดีได้เช่นเดียวกัน นอกจากการกำหนดส่วนประกอบต่าง ๆของโมเดลข้อมูลแล้ว การพิจารณาคุณสมบัติบูรณภาพของข้อมูล คุณสมบัตินอร์มัลไลเซชันของข้อมูล ได้ถูกนำมารวมเป็นขั้นตอนในการสร้างโมเดลข้อมูลดังต่อไปนี้

LDM1 กำหนดเอนติตีหลัก ๆ ในระบบ

LDM2 พิจารณากำหนดรีเลชันชิประหว่างเอนติตี รูปแบบ และประเภทของรีเลชันชิป รวมทั้งในกรณีรีเลชันชิปที่เกิดจากความสัมพันธ์ระหว่างเอนติตี ตั้งแต่ 2 เอนติตีขึ้นไปก็ต้องพิจารณากำหนดเป็นเอนติตีใหม่ขึ้นมา นอกจากนั้นควรทำการพิจารณาตั้งรีเลชันชิป แบบซูเปอร์ไทป์-ซับไทป์ ในกรณีความสัมพันธ์ที่เกิดขึ้นอยู่ในรูปแบบ 1:1 และ จำแนกประเภทได้

LDM3 พิจารณากำหนดคีย์หลัก และคีย์รอง

LDM4 พิจารณากำหนดฟอร์เรนจ์คีย์ในแต่ละรีเลชันชิป

LDM5 พิจารณากฎธุรกิจของคีย์ (Key Business Rules) เพื่อให้ข้อมูลมีคุณสมบัติความมีบูรณภาพ อันได้แก่กฎการเพิ่ม (Insert Rule) และกฎการลบ (Delete Rule)

กฎการเพิ่ม สำหรับควบคุมเงื่อนไขการเพิ่มรายการของแต่ละรีเลชันชิปสามารถจำแนกได้ 6 แบบดังนี้คือ

1. แบบขึ้นต่อกัน (Dependent) คือการเพิ่มสมาชิกในเอนติตีลูกจะต้องสอดคล้องกับในเอนติตีแม่ด้วย

2. แบบอัตโนมัติ (Automatic) คือหากเมื่อต้องการเพิ่มสมาชิกในเอนติตีลูกแล้ว ปรากฏว่าไม่มีค่าที่สอดคล้องกันในเอนติตีแม่ที่สัมพันธ์กันอยู่แล้ว กำหนดให้มีการสร้างสมาชิกใหม่ในเอนติตีแม่โดยอัตโนมัติ

3. แบบกำหนดให้มีค่าเป็นนัล (Nullify) คืออนุญาตให้เพิ่มสมาชิกในเอนติตีลูกได้ ถึงแม้ไม่มีสมาชิกใดในเอนติตีแม่ที่สอดคล้องกัน และจะกำหนดให้ฟอร์เรนจ์คีย์ในเอนติตีลูกให้มีค่าเป็นนัล (ไม่มีค่า)

4. แบบกำหนดค่าโดยปริยาย (Default) ในทำนองเดียวกันกับแบบที่ 3 แต่ค่าที่กำหนดให้ในฟอร์เรนจ์คีย์จะมีค่าโดยปริยาย

5. แบบกำหนดให้มีค่าตามต้องการ (Customized) คือจะอนุญาตให้เพิ่มสมาชิกในเอนติตีลูกได้เมื่อเป็นไปตามเงื่อนไขที่กำหนดขึ้นเท่านั้น

6. แบบไม่มีผลกระทบ (No effect) คือจะอนุญาตให้เพิ่มสมาชิกในเอนติตีลูกโดยอิสระ โดยไม่จำเป็นต้องสอดคล้องกับค่าในเอนติตีแม่ หรือมีการตรวจสอบเงื่อนไขใด ๆ

กฎการลบข้อมูล สำหรับควบคุมเงื่อนไขการลบรายการในแต่ละรีเลชันชิปจำแนกได้ 6 แบบคือ

1. แบบมีข้อจำกัด (Restrict) คือจะอนุญาตให้ลบสมาชิกในเอนติตีแม่ได้ก็ต่อเมื่อในเอนติตีลูกที่สัมพันธ์กัน ไม่มีค่าที่สอดคล้องกันอยู่

2. แบบต่อเนื่อง (Cascade) คือการลบสมาชิกในเอนติตีแม่จะมีผลทำให้สมาชิกในเอนติตีลูกที่สัมพันธ์กันถูกลบไปด้วย

3. แบบกำหนดให้มีค่าเป็นนัล (Nullify) คือการลบสมาชิกในเอนติตีแม่จะมีผลทำให้ฟอร์เรนจ์คีย์ของสมาชิกในเอนติตีลูกที่สอดคล้องกันถูกกำหนดให้มีค่าเป็นนัล

4. แบบกำหนดค่าโดยปริยาย (Default) คือการลบสมาชิกในเอนติตีแม่จะมีผลทำให้ฟอร์เรนจ์คีย์ของสมาชิกในเอนติตีลูกที่สอดคล้องกัน ถูกกำหนดให้มีค่าเป็นค่าโดยปริยาย

5. แบบกำหนดให้มีค่าตามต้องการ (Customized) คือการลบสมาชิกในเอนติตีแม่จะทำได้ก็ต่อเมื่อข้อมูลเป็นไปตามเงื่อนไขที่ถูกกำหนดขึ้น

6. แบบไม่มีผลกระทบ (No effect) คือการลบสมาชิกในเอนติตีแม่จะทำได้ทุกกรณี

<u>หมายเหตุ</u> ทั้งนี้มีข้อควรหลีกเลี่ยง และกระทำดังนี้

- ควรใช้กฎเกณฑ์แบบกำหนดให้มีค่าโดยปริยายสำหรับการเพิ่มหรือลบข้อมูลมากกว่ากฎเกณฑ์แบบกำหนดให้มีค่าเป็นนัล
- หลีกเลี่ยงการใช้วิธีแบบกำหนดให้มีค่าเป็นนัลอย่างเด็ดขาดในกรณีที่ฟอร์เรนจ์คีย์เป็นส่วนหนึ่งของคีย์หลักในเอนติตีลูก

- ควรกำหนดกฎเกณฑ์การเพิ่มสำหรับความสัมพันธ์แบบซูเปอร์ไทป์-ซับไทป์ให้เป็นแบบอัตโนมัติ หรือแบบขึ้นต่อกัน และกำหนดกฎเกณฑ์การลบในแบบความสัมพันธ์นี้เป็นแบบต่อเนื่อง

LDM6 เพิ่มแอตตริบิวต์ต่าง ๆที่สอดคล้องกันมากที่สุดลงไปในเอนติตีให้สมบูรณ์โดยที่แอตตริบิวต์เหล่านั้นมีคุณสมบัติขึ้นกับคีย์หลักทั้งหมดในเอนติตีที่บรรจุ ในกรณีที่เป็นดีไรฟ์แอตตริบิวต์ควรบันทึกสูตรที่เป็นที่มาของแอตตริบิวต์นั้นด้วย และนอกจากนี้ทำการพิจารณารวมเอนติตีที่มีคีย์หลักเหมือนกันเข้าด้วยกัน

LDM7 พิจารณาปรับให้เอนติตีที่มีคุณสมบัติตามกฎนอร์มัลไลเซชันในระดับที่ 1, 2, 3 ถึง 5 ตามลำดับ แต่ไม่ทำการย่อยเอนติตีที่มีคุณสมบัตินอร์มัลไลซ์อย่างสมบูรณ์แล้วออกเป็นเอนติตีย่อย ๆอีก รวมทั้งทำการวิเคราะห์โมเดลข้อมูลที่มีคุณสมบัตินอร์มัลไลซ์แล้วอีกครั้งในแง่กฎการเพิ่ม การลบข้อมูล และการเก็บข้อมูลเป็นประวัติ

<u>นอร์มัลไลซ์เซชัน</u> เป็นทฤษฎีที่ทำการวิเคราะห์และแยกส่วนโครงสร้างข้อมูลออกเป็นชุด ๆของความสัมพันธ์ ในทางปฏิบัติเป็นเทคนิคซึ่งประกอบด้วยหลายขั้นตอน อันก่อให้เกิดประโยชน์ต่าง ๆ ดังนี้

1. ลดขนาดที่ว่างที่ใช้เก็บข้อมูล
2. ลดความขัดแย้งกันภายในฐานข้อมูล
3. ลดปัญหาที่อาจเกิดขึ้นเมื่อทำการลบ หรือปรับปรุงข้อมูล
4. ทำให้โครงสร้างข้อมูลมีเสถียรภาพสูงสุด

ในที่นี้จะกล่าวถึงนอร์มัลฟอร์มระดับ 1 ถึงระดับที่ 3 โดยสังเขปดังนี้

<u>นอร์มัลฟอร์มระดับที่1</u> (First Normal Form หรือ 1NF)

ทุกสมาชิกในเอนติตีจะมีค่าในแอตตริบิวต์หนึ่ง ๆได้เพียงค่าเดียวหรือกล่าวคือ จะมีกลุ่มของค่าในแอตตริบิวต์นั้นในสมาชิกหนึ่งไม่ได้

<u>นอร์มัลฟอร์มระดับที่2</u> (Second Normal Form หรือ 2NF)

ความสัมพันธ์จะมีคุณสมบัตินอร์มัลฟอร์มระดับ 2 ได้ก็ต่อเมื่อมีคุณสมบัติของนอร์มัลฟอร์มระดับ 1 แล้ว และแอตตริบิวต์ที่ไม่ใช่คีย์ทุก ๆแอตตริบิวต์ต้องขึ้นตรงกับคีย์หลักทั้งชุด

<u>นอร์มัลฟอร์มระดับที่3</u> (Third Normal Form หรือ 3NF)

ความสัมพันธ์จะมีคุณสมบัตินอร์มัลฟอร์มระดับ 3 ก็ต่อเมื่อมีคุณสมบัติของนอร์มัลฟอร์มระดับ 2 แล้ว และแอตตริบิวต์ที่ไม่ใช่คีย์ใด ๆ จะขึ้นกับแอตตริบิวต์อื่น ๆ ที่ไม่ใช่คีย์หลักหรือกลุ่มของคีย์หลักไม่ได้

LDM8 พิจารณากำหนดโดเมนของแต่ละแอตตริบิวต์รวมทั้งโดเมนของคีย์หลัก คีย์รอง และฟอร์เรนจ์คีย์ ตัวอย่างเช่น กำหนดชนิดของข้อมูล ฟอร์แมท ความยาว ความเป็นเอกลักษณ์ของแอตตริบิวต์ หรือกลุ่มของแอตตริบิวต์ เป็นต้น

LDM9 พิจารณากำหนดทริกเกอร์ดำเนินการ (Trigger Operations) เพื่อรักษาให้ค่าแอตตริบิวต์ในโมเดลข้อมูลมีความเป็นบูรณภาพ และไม่ขัดแย้งกันของค่าแอตตริบิวต์

ทริกเกอร์ดำเนินการ หมายถึง กฎในการดำเนินการเพิ่ม ลบ ปรับปรุง หรือแม้แต่การสืบค้นข้อมูล ซึ่งจะมีผลกระทบต่อแอตตริบิวต์ในเอนติตีอื่น หรือแอตตริบิวต์อื่น ๆ ในเอนติตีเดียวกัน ดังนี้ได้แก่

- กำหนดทริกเกอร์ดำเนินการสำหรับทุกแอตตริบิวต์ซึ่งเป็นต้นกำเนิดของดีไรฟ์วแอตตริบิวต์
- กำหนดทริกเกอร์ดำเนินการของเอนติตีซับไทป์ ตัวอย่างเช่นสมาชิกในซับไทป์ถูกลบไป จะมีผลทำให้สมาชิกของซูเปอร์ไทป์เอนติตีที่สอดคล้องกันถูกลบไปด้วย
- กำหนดทริกเกอร์ดำเนินการที่เกี่ยวข้องกับเวลา ตัวอย่างเช่น กำหนดให้ทำการลบสมาชิกในเอนติตีหนึ่ง เมื่อมีค่าในแอตตริบิวต์หนึ่งซึ่งเป็นวันที่คำนวณถึงวันปัจจุบันมีค่ามากกว่า 1 ปี เป็นต้น

LDM10 พิจารณารวมเอนติตี รีเลชันชิป แอตตริบิวต์ที่สามารถรวมกันได้เข้าด้วยกันกล่าวคือ

1. ทำการรวมเอนติตีซึ่งมีคีย์หลักเหมือนกันและมีโดเมนของคีย์หลักที่เหมือนกันเป็นเอนติตีเดียวกัน
2. ทำการรวมรีเลชันชิป ซึ่งเกิดจากเอนติตีที่มีการรวมกันนั้นเข้าด้วยกันเช่นกัน โดยมีเงื่อนไขว่ารีเลชันชิปเหล่านั้นจะต้องมีความหมายเดียวกัน
3. ทำการรวมแอตตริบิวต์ซึ่งมีความหมายเหมือนกันเข้าไว้ด้วยกัน

LDM11 พิจารณาโมเดลข้อมูลเชิงตรรกชุดใหม่ที่สร้างให้เข้ากับฐานข้อมูลในระดับโครงร่างเชิงความคิดที่มีอยู่เดิมแล้ว

LDM12 วิเคราะห์ความมีเสถียรภาพ และการเปลี่ยนแปลงของโมเดลข้อมูลที่อาจเกิดขึ้นในอนาคต

3.7.1 ขั้นตอนในการออกแบบโมเดลข้อมูลเชิงกายภาพ

เป็นขั้นตอนการแปลงโมเดลข้อมูลเชิงตรรกที่รวบรวมมุมมองของผู้ใช้ต่าง ๆ แล้วไปสู่ฐานข้อมูลที่มีเสถียรภาพ โดยใช้สิ่งที่เอื้ออำนวยให้ใช้ในระบบจัดการฐานข้อมูลแบบรีเลชันนัลที่ใช้นั้น สิ่งที่จะกล่าวถึงเป็นเพียงขั้นตอนเริ่มแรกในการออกแบบฐานข้อมูลเบื้องต้นเท่านั้นได้แก่ กระบวนการแปลง (Translation Process) ซึ่งมีลักษณะคือ ยังไม่คำนึงถึงประสิทธิภาพการใช้งาน (Performance) นัก

ฐานข้อมูลประกอบด้วย ส่วนสำคัญ 2 ส่วนคือ

1. โครงสร้างข้อมูล (Data Structure) เป็นตาราง และคอลัมน์
   คอลัมน์ หมายถึงหน่วยหนึ่งของข้อมูล อาจเรียกอีกอย่างว่า เขตข้อมูล
   ตาราง หมายถึงการรวมกลุ่มของคอลัมน์จำนวนหนึ่ง
2. บูรณภาพของข้อมูล ได้แก่ การออกแบบกฎธุรกิจ (Business Rules) สำหรับเอนติตี รีเลชันชิป และ แอตตริบิวต์

กระบวนการแปลงโมเดลข้อมูลเชิงตรรกไปสู่ฐานข้อมูลเบื้องต้น มีดังนี้

RDD1 กำหนดตารางขึ้นจากเอนติตี

RDD2 กำหนดคอลัมน์ขึ้นภายในแต่ละตาราง จากแอตตริบิวต์ในเอนติตี

RDD1 และ RDD2 ยังไม่ใช่ขั้นตอนการออกแบบฐานข้อมูลรีเลชันนัล เนื่องจากยังคงเป็นอิสระจากระบบจัดการฐานข้อมูลที่จะใช้ แต่ในส่วน RDD3 ถึง RDD6 จะมีวิธีที่การออกแบบควบคุมแตกต่างกันไปตามระบบจัดการฐานข้อมูลที่ใช้ ใน RDD4 ถึง RDD6 จะเกี่ยวข้องกับการรักษาความมีบูรณภาพของข้อมูล หากระบบจัดการฐานข้อมูลนั้นไม่มีภาษาจำกัดความ (DDL : Data Definition Language) ให้ใช้ได้ อาจทำได้โดยสร้างรูทีนการบำรุงรักษา (Maintenance Routines) ขึ้นควบคุมเอง

RDD3 กำหนดการจัดเรียงคอลัมน์ในตาราง พิจารณาเนื้อที่ที่ใช้จัดเก็บในฐานข้อมูล และพิจารณาพารามิเตอร์ในการล็อค (Locking Parameter) สำหรับควบคุมกลไกการล็อคข้อมูลในฐานข้อมูล

RDD4 ออกแบบกฎธุรกิจควบคุมเกี่ยวกับเอนติตี ได้แก่ การกำหนดควบคุมคุณสมบัติของคีย์หลัก และคีย์รอง เช่น ความเป็นเอกลักษณ์ของคีย์หลัก มีค่าเป็นนัลไม่ได้หรือ คอลัมน์ใดที่เป็นส่วนหนึ่งของชุดคีย์หลักจะต้องมีคุณสมบัติไม่เป็นเอกลักษณ์ เป็นต้น

RDD5 ออกแบบกฎธุรกิจควบคุมเกี่ยวกับรีเลชันชิป ได้แก่ การเพิ่ม การลบ และการปรับปรุง ซึ่งมีผลกระทบถึงคีย์ที่เชื่อมกัน หรือที่หมายถึงกฎบูรณภาพในการอ้างถึง

RDD6 ออกแบบกฎธุรกิจควบคุมเกี่ยวกับแอตตริบิวต์เพิ่มเติม ได้แก่ เรื่อง โดเมน ทริกเกอร์ดำเนินการ

3.8 การรักษาความปลอดภัยของข้อมูล

วัตถุประสงค์การรักษาความปลอดภัยของฐานข้อมูลคือป้องกันบุคคลใด ๆ ทำการเปลี่ยนแปลงแก้ไข หรือรับทราบข้อมูลในฐานข้อมูลนอกเหนือขอบเขตหน้าที่ความรับผิดชอบ หรือโดยบุคคลที่ไม่เกี่ยวข้องทั้งที่โดยจงใจหรือไม่ก็ตามเพื่อให้ข้อมูลถูกต้องเชื่อถือได้ โดยเฉพาะอย่างยิ่งในระบบผู้ใช้หลายคน

ระบบรักษาความปลอดภัยของฐานข้อมูลต้องสามารถแยกแยะ หรือตัดสินได้ว่าขณะนั้น ๆ บุคคลใดเป็นผู้เข้าสู่ และใช้ข้อมูลในฐานข้อมูลอยู่ และประเภทการใช้ข้อมูลเป็นอะไร

วิธีการในการรักษาความปลอดภัยข้อมูลมีทั้งที่เกี่ยวข้องโดยระบบการวางระเบียบปฏิบัติในการใช้คอมพิวเตอร์ในองค์กร ระบบปฏิบัติการ (Operating System) และระบบจัดการฐานข้อมูล

3.8.1 โดยการสร้างชื่อผู้ใช้ และรหัสผ่าน (User Name / Password)

อาจสร้างสำหรับผู้ใช้คนหนึ่ง ๆ หรือเป็นกลุ่มของผู้ใช้ ซึ่งมีสิทธิในการใช้ข้อมูลเท่ากันเมื่อผู้ใช้คีย์ชื่อ และรหัสเพื่อเข้าสู่ระบบฐานข้อมูล ระบบจัดการฐานข้อมูลจะทำการตรวจสอบชื่อและรหัสว่าถูกต้องมีสิทธิเข้าสู่ระบบหรือไม่ ระบบควรจะสามารถเก็บรายละเอียดการเข้าสู่ระบบฐานข้อมูลว่าโดยบุคคลใด เข้าสู่ระบบเมื่อเวลาใด และออกเวลาใด และนอกจากนั้นสามารถเก็บรายละเอียดเกี่ยวกับการเปลี่ยนแปลงข้อมูลในฐานข้อมูล เกิดขึ้นอย่างไร โดยใครเป็นต้น

3.8.2 โดยการสร้างวิว (View)

ทำโดยการสร้างวิวแล้วใช้ภาษาในการควบคุม (DCL : Data Control Language) ได้แก่คำสั่ง "GRANT" และ "REVOKE" เพื่อสร้างความปลอดภัยให้แก่ข้อมูลในโครงสร้างแบบรีเลชันนัล

ลักษณะวิธีการสร้างความปลอดภัยข้อมูลโดยการสร้างวิว มีดังนี้

- ควบคุมความสามารถในการใช้ข้อมูล ซึ่งได้แก่ การสืบค้น การเพิ่ม การลบ และ การปรับปรุง ของข้อมูลทั้งตารางตามหน้าที่ความรับผิดชอบของผู้ใช้

- ควบคุมความสามารถในการใช้ข้อมูลบางส่วนของตารางหนึ่ง ในระดับคอลัมน์ หรือแถว ตามหน้าที่ความรับผิดชอบของผู้ใช้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย