บทที่ 3

แนวคิดและทฤษฎี

วัตถุประสงค์หนึ่งของการออกแบบ และพัฒนางานประมาณราคาระบบส่งไฟฟ้านั้น ก็เพื่อจัดเก็บข้อมูลที่เกี่ยวกับราคางานระบบส่งไฟฟ้าของทุกหน่วยงานที่เกี่ยวข้องให้มีมาตรฐานเดียวกัน เพื่อให้สามารถควบคุมความถูกต้องและลดความซ้ำซ้อนของข้อมูลได้ เพื่อบรรลุถึงเป้าหมายดังกล่าว จึงใช้หลักการออกแบบฐานข้อมูล ซึ่งเน้นการจัดการกับข้อมูล (data oriented) มาใช้แทนการออกแบบแฟ้มข้อมูล ซึ่งเน้นวิธีประมวลผลข้อมูล (process oriented) ดังจะได้กล่าวถึงขั้นตอนการสร้างฐานข้อมูลต่อไป

การจัดเก็บข้อมูลอย่างมีระเบียบแบบแผน และวิธีการนำข้อมูลส่วนที่ต้องการออกมาใช้ให้ทันต่อเหตุการณ์ จะยังประโยชน์ให้แก่ผู้นำมาใช้ได้อย่างสูง ไม่ว่าจะเป็นการนำมาใช้ในการวางแผนหรือตัดสินใจด้านงบประมาณ การลงทุน และอื่น ๆ ความยากลำบากในการจัดเก็บข้อมูลอย่างมีระเบียบแบบแผน และวิธีการนำข้อมูลส่วนที่ต้องการออกมาใช้ให้ทันต่อเหตุการณ์ เมื่อนำระบบคอมพิวเตอร์มาใช้ในการประมวลผลข้อมูล ทำให้ระบบการจัดเก็บข้อมูลเป็นไปอย่างสะดวกขึ้น แต่การควบคุมความถูกต้องปลอดภัยของข้อมูล ตลอดจนการเรียกใช้ข้อมูล ยังอยู่ภายใต้ความสามารถของโปรแกรมเมอร์ สิ่งที่ตามมาก็คือต้องใช้เจ้าหน้าที่และเวลาในการสร้างโปรแกรม เพื่อประยุกต์ใช้งานตามความต้องการเป็นจำนวนมาก และนี่คือจุดกำเนิดของฐานข้อมูลที่เกิดขึ้นมา เพื่อแก้ปัญหาดังกล่าว โดยได้มีการออกแบบและพัฒนาระบบการจัดการฐานข้อมูล เพื่อทำหน้าที่ควบคุมและจัดการความถูกต้องและประสิทธิภาพของการบริหารข้อมูล

การเก็บข้อมูลจำนวนมากไว้ แล้วสามารถนำออกมาใช้ได้อย่างมีประสิทธิภาพ ต้องอาศัยฐานข้อมูลที่ดี แต่ทว่าขั้นตอนในการสร้างฐานข้อมูลที่ดีนั้นต้องอาศัยเทคนิคต่าง ๆ รวมถึงประสบการณ์ด้วย การออกแบบฐานข้อมูลเป็นเรื่องยากและซับซ้อน เพราะต้องออกแบบระบบข้อมูลที่ครอบคลุมทั้งหน่วยงาน แต่ก็ต้อง

สามารถป้องกันให้มีแต่ผู้ที่มีหน้าที่รับผิดชอบ ให้เรียกใช้เฉพาะข้อมูลส่วนของตนได้เท่านั้น จะอ่านและเรียกใช้ข้อมูลของผู้อื่นไม่ได้ ทั้ง ๆ ที่ข้อมูลเหล่านี้อยู่ในฐานข้อมูลเดียวกัน งานส่วนที่ยุ่งยากที่สุดของการออกแบบคือ การทำความเข้าใจข้อมูลที่ผู้ใช้ต้องการใช้ในงานต่าง ๆ และการแทนข้อมูลเหล่านี้ในฐานข้อมูล

ความหมายของระบบฐานข้อมูล

ระบบฐานข้อมูล หมายถึงการรวบรวมข้อมูลที่มีความสัมพันธ์กันและอาจอยู่ต่างที่กัน ให้เสมือนอยู่ร่วมกัน เพื่อให้สามารถรับใช้งานที่มีวัตถุประสงค์แตกต่างกันของหน่วยงานต่าง ๆ โดยที่ผู้ใช้ฐานข้อมูลไม่ได้รับรู้ข้อมูลทั้งหมดในฐานข้อมูล แต่รับรู้เฉพาะส่วนที่เกี่ยวข้องกับงานของตนเท่านั้น ฉะนั้นผู้ใช้แต่ละคนจะรับรู้ข้อมูลที่แตกต่างกันจากฐานข้อมูลเดียวกัน เช่น ผู้ใช้ที่อยู่ฝ่ายเงินเดือน จะรู้ชื่อ และเงินเดือนของพนักงาน ในขณะที่ผู้บริหารรับรู้ข้อมูลเกี่ยวกับผลงานของพนักงาน ทำนองเดียวกับคนตาบอดคลำช้าง คนที่คลำถูกหางช้าง คิดว่าช้างมีลักษณะเหมือนเชือก ในขณะที่คนที่คลำถูกลำตัว คิดว่าช้างเหมือนฝาผนัง เป็นต้น ฉะนั้นผู้ออกแบบช้าง (ฐานข้อมูล) จะต้องแน่ใจว่า สิ่งที่ออกแบบนั้นสามารถตอบสนองทัศนะที่แตกต่างกันของผู้ใช้ได้

เหตุผลในการพัฒนาระบบฐานข้อมูล

ข้อมูลจำนวนมากมายที่ต้องการรวบรวมไว้ที่เดียวกันนั้นสมควรที่จะพัฒนาเป็นระบบฐานข้อมูลเป็นอย่างยิ่ง เนื่องจากมีข้อดีที่เห็นได้ชัดหลายประการดังนี้

1. สามารถลดความซ้ำซ้อนของข้อมูล เนื่องจากการเก็บข้อมูลในรูปของฐานข้อมูลจะรวมเอาข้อมูลมาเก็บไว้ที่เดียวกัน ทำให้สามารถลดปัญหาความซ้ำซ้อนของข้อมูลได้ แต่ในบางกรณีก็จำเป็นจะต้องมีความซ้ำซ้อนของข้อมูลในฐานข้อมูลบ้าง อาจด้วยเหตุผลที่เกี่ยวกับประสิทธิภาพ (performance)หรืออะไรก็ตาม ในกรณีเช่นนี้ผู้บริหารฐานข้อมูลจะต้องบอกให้ระบบจัดการฐานข้อมูลรับรู้เพื่อที่จะตรวจสอบและแก้ไขข้อมูลที่ซ้ำซ้อนกันนี้ให้สอดคล้องกันเสมอ

2. สามารถแก้ปัญหาการที่ข้อมูลไม่สอดคล้องกัน คือเมื่อปัญหาความซ้ำซ้อนลดลง หรือหากมีแต่ระบบจัดการฐานข้อมูลได้รับรู้ ปัญหาที่ว่าข้อมูลไม่สอด

คล้องกันก็จะไม่เกิดขึ้น

3. สามารถควบคุมความถูกต้องของข้อมูล (integrity) นอกจากความสอดคล้องกันของข้อมูลในกรณีที่มีความซ้ำซ้อนแล้ว ความถูกต้องของข้อมูลที่จะต้องคำนึงถึงมีมากมาย หากในระบบฐานข้อมูลนั้นใช้ระบบจัดการฐานข้อมูลที่มีความสามารถ ผู้บริหารฐานข้อมูลก็จะสามารถกำหนดกฎเกณฑ์ให้ระบบจัดการฐานข้อมูลเพื่อใช้ในการตรวจสอบและแก้ไขข้อมูลให้ถูกต้องได้ เช่น อาจจะกำหนดได้ว่าอายุของพนักงานในบริษัทจะไม่ต่ำกว่า 18 ปี และจะไม่เกิน 60 ปี ดังนั้นหากมีการป้อนข้อมูลที่แสดงอายุเป็น 70 ปี ระบบจัดการฐานข้อมูลก็จะไม่ยอมหรืออาจกำหนดได้ว่า ชื่อของพนักงานที่จะปรากฏในฐานข้อมูล จะต้องเป็นชื่อที่ปรากฏในตารางแสดงประวัติพนักงานเท่านั้น ซึ่งเป็นตารางที่เก็บประวัติพนักงานที่บริษัทว่าจ้างอยู่ในปัจจุบัน (ตัวอย่างนี้เป็นกรณีของฐานข้อมูลเชิงสัมพันธ์) หรืออาจกำหนดได้ว่าหากพนักงานคนใดถูกลบออกจากตารางแสดงประวัติ ก็ให้ลบส่วนอื่น ๆ ที่เกี่ยวข้องกับพนักงานคนนั้นออกไปจากฐานข้อมูลด้วย เป็นต้น

4. สามารถสร้างหรือกำหนดระบบความปลอดภัยได้ จากตัวอย่างที่แสดง จะเห็นได้ว่าผู้บริหารฐานข้อมูลเป็นผู้ที่สามารถกำหนดได้ว่า จะให้ผู้ใช้คนใดเห็นข้อมูลในส่วนใดได้บ้าง โดยการสร้างวิวให้ผู้ใช้ เฉพาะในส่วนที่เขาจะมีสิทธิรับรู้เท่านั้น นอกจากนี้หากระบบจัดการฐานข้อมูลที่ใช้ เป็นตัวที่มีความสามารถมากพอ ก็จะสามารถกำหนดละเอียดลงไปได้อีกว่า สำหรับผู้ใช้ต่าง ๆที่ใช้วิวเดียวกันนี้ ผู้ใช้คนใดบ้างที่จะมีสิทธิในการแก้ไขข้อมูล หรืออาจจะมีสิทธิได้แค่เรียกดูข้อมูลได้เพียงอย่างเดียว และในระบบจัดการฐานข้อมูลที่มีระบบความปลอดภัยที่ดี จะให้ผู้ใช้หรือโปรแกรมเข้าถึงข้อมูลได้ ก็ต่อเมื่อติดต่อผ่านระบบจัดการฐานข้อมูลเท่านั้น คือเมื่อผู้ใช้ป้อนข้อมูลเข้ามา ระบบจัดการฐานข้อมูลก็จะต้องทำการเข้ารหัสข้อมูลก่อนที่จะนำไปเก็บภายในฐานข้อมูล เมื่อมีการเรียกใช้ข้อมูล จึงทำการถอดรหัสแล้วค่อยนำเสนอให้ผู้ใช้

5. เกิดความอิสระของข้อมูล คือเมื่อมีการเปลี่ยนแปลงโครงสร้างของข้อมูลในระดับล่าง จะต้องไม่มีผลกระทบต่อโครงสร้างของข้อมูลในระดับบน ซึ่งความอิสระของข้อมูลนี้ รวมไปถึงในกรณีที่สามารถกำหนดได้ว่า ข้อมูลที่ปรากฏในระดับต่าง ๆ กัน สามารถมีรูปแบบที่ต่างกันได้ตามต้องการ เช่น ในระดับหลักการ

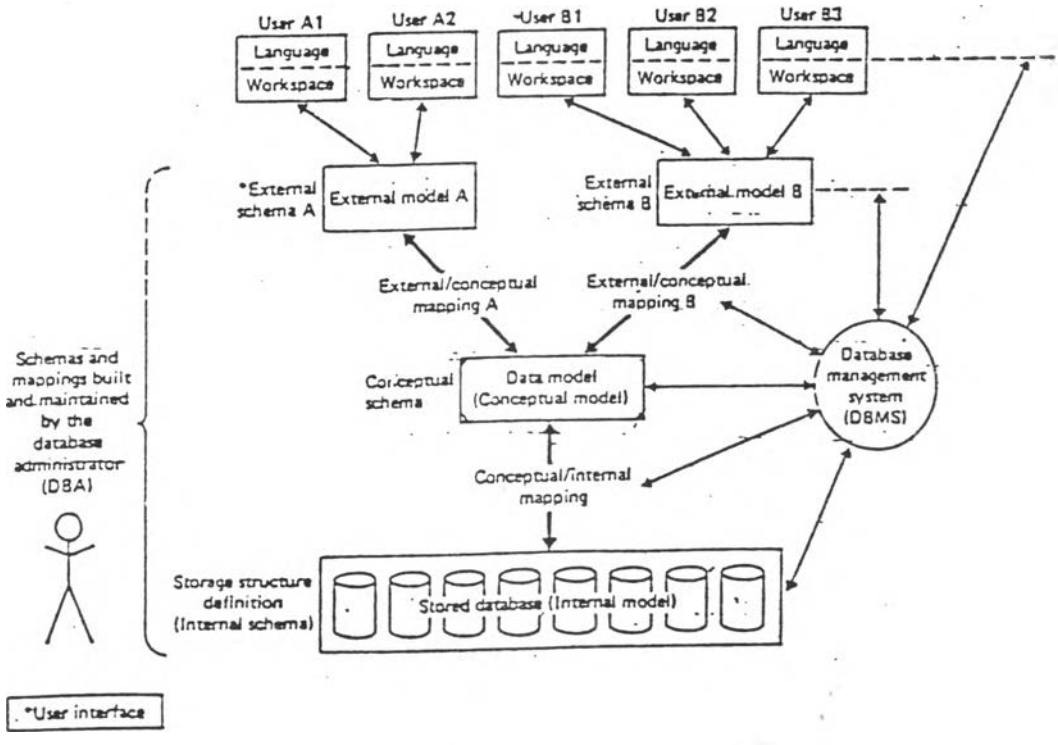
เงินเดือนของพนักงานอาจอยู่ในรูปของเลขฐานสอง แต่เมื่อสร้างวิวให้กับผู้ใช้คนหนึ่ง เงินเดือนที่ปรากฏอาจจะอยู่ในรูปของเลขฐานสิบ และในวิวของผู้ใช้อีกคนอาจเห็นเงินเดือนในรูปของเลขฐานสอง ซึ่งการแปลงข้อมูลนี้ก็จะเป็นหน้าที่ของระบบจัดการฐานข้อมูล โดยมีผู้บริหารฐานข้อมูลเป็นผู้กำหนด

<u>สถาปัตยกรรมของระบบฐานข้อมูล</u> ระบบฐานข้อมูลจะถูกแบ่งออกเป็น 3 ระดับคือ

1. ระดับกายภาพ (Internal or Physical Level) เป็นระดับที่ต่ำที่สุด อันได้แก่ ระดับของการจัดเก็บข้อมูลในสื่อบันทึกข้อมูล ถูกกำหนดโดยระบบบริหารฐานข้อมูล

2. ระดับหลักการ (Conceptual Level or Schema) เป็นระดับที่อยู่ตรงกลางระหว่างระดับในสุดและระดับนอกสุด เป็นการกำหนดรูปแบบของฐานข้อมูล อันได้แก่ ชนิด ลักษณะ และความสัมพันธ์เชิงตรรกของข้อมูลทั้งหมดของฐานข้อมูล ภาพที่เห็นในระดับนี้เรียกว่าภาพระดับหลักการ (Conceptual View) นอกจากนี้ยังกำหนดขอบเขตของข้อมูลที่ผู้ใช้ในระดับนอกสุดควรจะเห็น ภาพระดับหลักการนี้กำหนดโดยผู้บริหารฐานข้อมูล

3. ระดับนอกสุด (External Level or Subschema) เป็นระดับที่ใกล้กับผู้ใช้มากที่สุด คือในระดับนี้ภาพของฐานข้อมูลที่ผู้ใช้เห็นจะเป็นเพียงส่วนที่เกี่ยวข้องกับผู้ใช้เท่านั้น ภาพที่เห็นนี้บางครั้งก็เรียกว่าภาพนอกสุด (External View) ผู้ใช้จะสามารถเรียกค้นหรือแก้ไข (หากทำได้) ก็เฉพาะในส่วนที่มองเห็นนี้เท่านั้น ถ้าภาพระดับหลักการถือเป็นภาพรวมของฐานข้อมูล ภาพนอกสุดก็จัดเป็นภาพย่อยของฐานข้อมูล เช่น ถ้าแผนที่ถนนหนทางทั้งหมดในกรุงเทพมหานครเป็นภาพรวมแล้ว การวาดเฉพาะเส้นทางระหว่างบ้านกับที่ทำงานก็คือภาพย่อย



รูป 3.1 ภาพรวมของระบบฐานข้อมูล

รูป 3.1 เป็นภาพรวมของระบบฐานข้อมูล โดยในรูปจะเห็นว่าระบบฐานข้อมูลมี 3 ระดับ โดยผู้ใช้จะยุ่งเกี่ยวเฉพาะระดับนอกสุดและการที่แบ่งให้เป็น 3 ระดับเช่นนี้ ก็เพื่อจะให้เกิดความอิสระของข้อมูล คือเมื่อมีการเปลี่ยนแปลงโครงสร้างของข้อมูลในระดับล่าง จะต้องไม่มีผลกระทบต่อโครงสร้างของข้อมูลในระดับบน โดยมีระบบจัดการฐานข้อมูลเป็นผู้ทำหน้าที่ในการแบ่งระดับข้อมูลให้เป็น 3 ระดับ นอกจากนี้ผู้ใช้หรือโปรแกรมต่าง ๆ ที่ติดต่อกับฐานข้อมูล จะต้องติดต่อโดยผ่านระบบจัดการฐานข้อมูลเสมอ และจะใช้ภาษาฐานข้อมูลในการติดต่อกับระบบฐานข้อมูลดังกล่าว จะมีผู้บริหารฐานข้อมูลเป็นผู้ออกแบบฐานข้อมูล และดูแลให้ฐานข้อมูลมีรูปแบบที่เหมาะสมกับการใช้งานและถูกต้องเสมอ

ภาษาฐานข้อมูล (Database Language)

โดยทั่วไปแล้ว ระบบจัดการฐานข้อมูลซึ่งเป็นสิ่งที่คั่นกลางระหว่างฐานข้อมูลและผู้ใช้ จะมีภาษาสำหรับสร้างฐานข้อมูลสำหรับให้ผู้ใช้เรียกค้นข้อมูล และจัดการด้านอื่น ๆ ในทางทฤษฎีแล้วภาษาฐานข้อมูลนี้จะประกอบด้วย 2 ส่วนคือ

1. ส่วนนิยามข้อมูล (Data Definition Language : DDL)

เป็นส่วนที่ใช้ในการสร้างฐานข้อมูลทั้ง 3 ระดับ หรือหากมองเป็นระบบฐานข้อมูลเชิงสัมพันธ์ ภาษานิยามข้อมูลนี้จะใช้สำหรับสร้างตารางหลัก (Base Table) วิว (View) และบอกถึงลักษณะการเก็บข้อมูลจริง นอกจากนี้ภาษานิยามข้อมูลยังกำหนดกฎเกณฑ์ เพื่อให้ระบบจัดการฐานข้อมูลใช้สำหรับตรวจสอบความถูกต้องของข้อมูลอีกด้วย

2. ส่วนจัดการข้อมูล (Data Manipulation Language : DML) เป็นส่วนที่ใช้ในการค้นหาข้อมูล (retrieve), เพิ่มข้อมูล (insert), แก้ไขข้อมูล (update) และลบข้อมูล ดังนั้นการใช้ฐานข้อมูลในระดับของผู้ใช้นั้นก็สามารถกระทำได้โดยการใช้ส่วนนี้นั่นเองโดยทั่วไปแล้วแบ่งได้เป็น 2 ประเภทคือ

ก. แบบต้องกำหนดวิธีการ (Procedural Language) แบบนี้เป็นหน้าที่ของผู้ใช้ที่จะต้องระบุสิ่งที่ต้องการ และวิธีการในการให้ได้มาซึ่งผลอันนั้นด้วย

ข. แบบไม่ต้องกำหนดวิธีการ (Nonprocedural Language) แบบนี้ผู้ใช้เพียงแต่ระบุสิ่งที่ต้องการ โดยไม่ต้องบอกว่าทำอย่างไร

โดยทั่วไปแล้วภาษาเอสคิวแอล (SQL : Structured Query Language) ซึ่งเป็นภาษาที่ใช้ในการติดต่อกับระบบจัดการฐานข้อมูล นอกจากจะประกอบด้วยส่วนจัดการและนิยามข้อมูลแล้วยังมีส่วนควบคุมข้อมูล(Data Control Language : DCL) เพื่อใช้ควบคุมและป้องกันการเข้าถึงข้อมูลที่เป็นส่วนตัวด้วย

<u>ส่วนประกอบของระบบฐานข้อมูล</u> มี 4 ประการคือ

1. ข้อมูล
2. ฮาร์ดแวร์ (Hardware)
3. ซอฟท์แวร์ (Software) โดยมีระบบจัดการฐานข้อมูล(Database Management System : DBMS) เป็นซอฟท์แวร์ที่ทำหน้าที่จัดการกับฐานข้อมูลตามที่ผู้ใช้ต้องการ
4. ผู้ใช้ข้อมูล (User) แบ่งออกเป็น
   ก. โปรแกรมเมอร์งานประยุกต์(Application Programmer)
   ข. ผู้ใช้งานจริง (End User)
   ค. ผู้บริหารฐานข้อมูล (Database Administrator : DBA)

มีหน้าที่คอยดูแลและจัดวางผังโครงสร้างต่างๆเพื่อที่จะให้ระบบฐานข้อมูลทำงานได้

ขั้นตอนการออกแบบฐานข้อมูล (Tsai,1988)

การออกแบบฐานข้อมูลของระบบงานใดระบบงานหนึ่ง เพื่อใช้งานประกอบด้วยหลักการออกแบบใหญ่ ๆ 2 หลักคือ

1. การออกแบบฐานข้อมูลแบบตรรก(Logical Database Design) เป็นการรวบรวมความต้องการของผู้ใช้ ข้อมูลทั้งหมดและนำมาออกแบบโดยคำนึงถึงความสัมพันธ์ข้อมูลเป็นหลัก ไม่สนใจถึงอุปกรณ์ฮาร์ดแวร์ โปรแกรมระบบ หรือโปรแกรมการจัดการฐานข้อมูล ดังนั้นการออกแบบฐานข้อมูลแบบตรรกจะเป็นอิสระจากอุปกรณ์ (Hardware Independent) ผลที่ได้จากการออกแบบจะเป็นโมเดลข้อมูลเชิงตรรก (Logical Data Modeling : LDM) ของระบบงาน

จากรูปที่ 3.2 แสดงให้ทราบถึงขั้นตอนการออกแบบฐานข้อมูล โดยการออกแบบฐานข้อมูลแบบตรรก จะครอบคลุมขั้นตอนที่ 1 ถึงขั้นตอนที่ 5

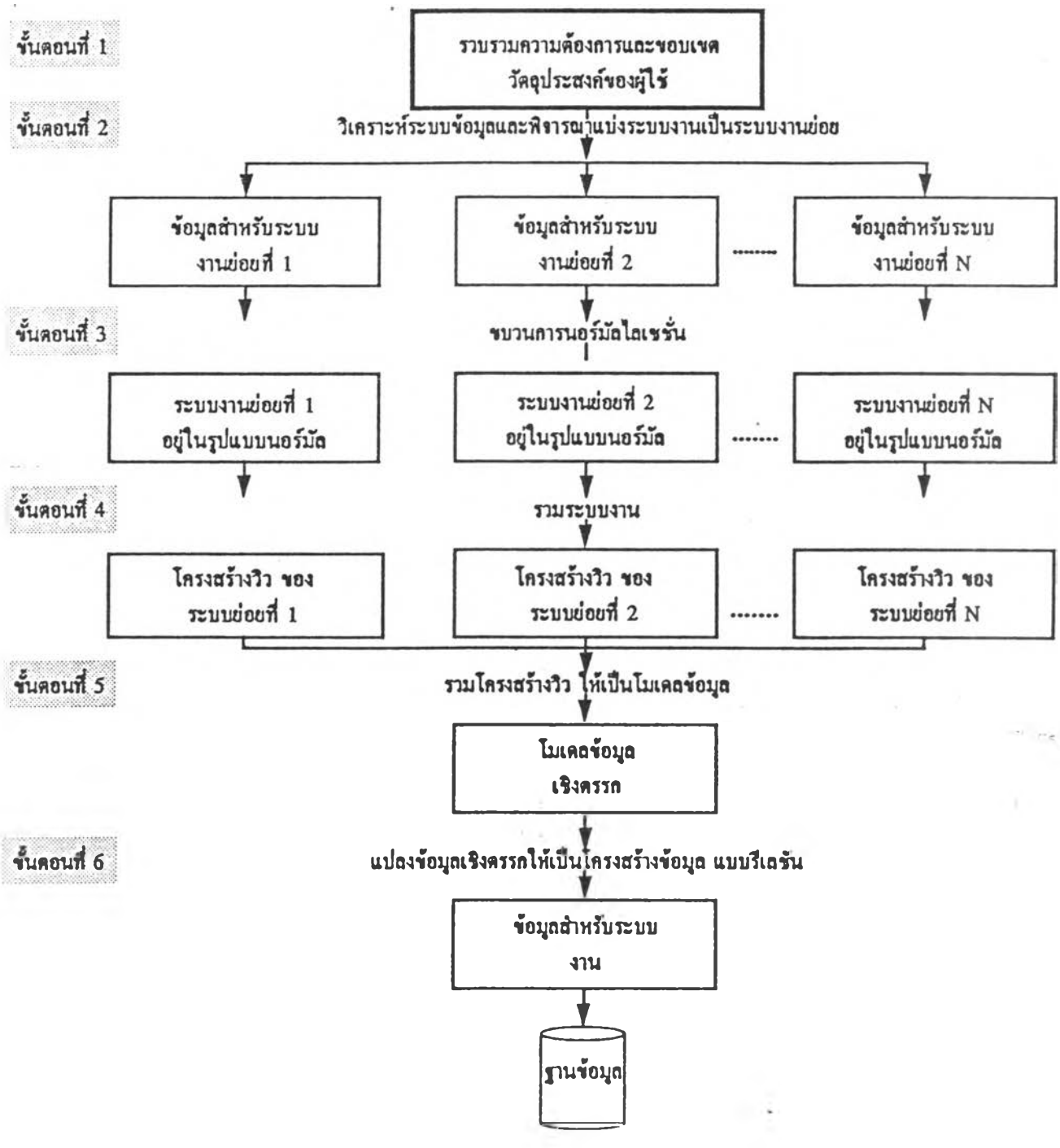
ขั้นตอนที่ 1 เป็นการรวบรวมความต้องการ (Requirement) ของผู้ใช้ซึ่งความต้องการเหล่านี้จะรวมถึงความต้องการในปัจจุบันของระบบงาน และความต้องการในอนาคตของระบบงานที่มีโอกาสเกิดขึ้น ผู้รวบรวมข้อมูล (Data Administrator ,Modeler) จะรวบรวมข้อมูลปัจจุบันและข้อมูลอนาคตที่คาดคะเนไว้ เช่น ข้อมูลที่เจ้าหน้าที่ตรวจสอบคอมพิวเตอร์ (Computer Auditor) จำเป็นต้องใช้ในการตรวจสอบการทำงานของโปรแกรมระบบงานต่าง ๆ ฯลฯ ซึ่งข้อมูลที่รวบรวมจะอยู่ในลักษณะที่แบ่งย่อยไม่ได้ (Atomic Data)

ขั้นตอนที่ 2 เป็นการรวบรวมคุณสมบัติข้อมูลของฐานข้อมูล และทำการวิเคราะห์ระบบข้อมูล โดยหลังจากวิเคราะห์แล้ว อาจแบ่งระบบงานเป็นระบบงานย่อย คุณสมบัติของข้อมูลที่รวบรวมจะประกอบด้วย

ก. ความหมายของข้อมูล (Data Meaning)

ข. ค่าความจริงของข้อมูล (Data Facts)

ในขั้นตอนนี้จะพบจำนวนของข้อมูลมีมากสำหรับระบบงาน เพื่อขจัดปัญหาในการบริหารข้อมูลที่มีมากเกินไปที่จะพบในขั้นตอนต่อไป จึงมีการแบ่งระบบงานเป็นระบบงานย่อยโดยรวมข้อมูลของงานประยุกต์ (Applications)ที่ใช้เข้า



รูป 3.2 ขั้นตอนการออกแบบฐานข้อมูล

อยู่ในระบบงานย่อยเดียวกัน แต่ก็มีข้อมูลบางส่วนที่ใช้ร่วมกัน สำหรับระบบงานย่อยหลาย ๆ ระบบงานย่อยซึ่งจะมีการขจัดในขั้นตอนรวมโมเดลระบบงานย่อยเป็นโมเดลระบบงานภายหลัง ในการวิเคราะห์ข้อมูลและแสดงคุณสมบัติข้อมูลจะใช้วิธีเทคนิคโมเดลข้อมูล (Data Modeling Technique) ซึ่งจะกล่าวต่อไป ผลจากการวิเคราะห์จะได้โมเดลของระบบงานย่อยของแต่ละระบบงานย่อยอย่างคร่าว ๆ เท่านั้น

ขั้นตอนที่ 3 เป็นการพิจารณาข้อมูลในแต่ละระบบงานย่อย เพื่อให้ได้เอนติตีและความสัมพันธ์ระหว่างเอนติตี มีการนำขบวนการนอร์มัลไลเซชันมาลดความซ้ำซ้อนข้อมูล และจัดข้อมูลให้อยู่ในรูปแบบนอร์มัล

ขั้นตอนที่ 4 เป็นการสร้างโมเดลข้อมูลสำหรับแต่ละระบบงานย่อย ซึ่งโมเดลข้อมูลนั้นเกิดจากการนำเอนติตีของแต่ละระบบงานย่อยมาเชื่อมโยงเข้าด้วยกัน

ขั้นดอนที่ 5 เป็นการนำโมเดลข้อมูลของแต่ละระบบงานย่อยมารวมกันเป็นโมเดลข้อมูล โมเดลข้อมูลเชิงตรรกของระบบงานด้วยการ

ก. รวมความสัมพันธ์ หรือยุบความสัมพันธ์หลาย ๆ ความสัมพันธ์ที่ใช้แสดงเอนติตีให้เป็นหนึ่งความสัมพันธ์

ข. เชื่อมเอนติตีในแต่ละระบบงานย่อย ให้เป็นโมเดลข้อมูลเชิงตรรกของระบบงาน

2. การออกแบบฐานข้อมูลแบบกายภาพ (Physical Database Design)

การออกแบบฐานข้อมูลแบบกายภาพอาศัยรายละเอียด คือโมเดลข้อมูลเชิงตรรก ซึ่งเกิดจากหลักการแรกเป็นพื้นฐานในการดำเนินงานต่อไป โดยมีขบวนการแปลง (Translation Process) จากโมเดลข้อมูลเชิงตรรกเป็นโมเดลข้อมูลเชิงกายภาพ

รูป 3.2 แสดงให้ทราบถึงขั้นตอนการออกแบบฐานข้อมูล โจยการออกแบบฐานข้อมูลแบบกายภาพ จะครอบคลุมขั้นดอนที่ 6 ซึ่งในบทนี้จะกล่าวถึงการออกแบบฐานข้อมูลกายภาพแบบรีเลชันนัล

ขั้นตอนที่ 6 เป็นการนำโมเดลข้อมูลเชิงตรรกของระบบงานรวม มาดัดแปลงเพื่อให้เหมาะสมกับระบบการจัดการฐานข้อมูลที่จะนำมาทำงานร่วมกัน ผลจากขั้นตอนนี้จะได้โมเดลข้อมูลเชิงกายภาพแบบรีเลชันนัล

การออกแบบฐานข้อมูลโดยพิจารณาจากทิศทางการออกแบบ สามารถแบ่งได้ 2 วิธีคือ

ก. การออกแบบจากระดับบนไปสู่ระดับล่าง (Top-Down Design)

เป็นการวิเคราะห์ข้อมูลของทั้งองค์กร โดยพิจารณาความหมายของข้อมูลในระดับบน ไม่เน้นการประยุกต์ใช้งานทำให้ง่ายต่อการรวมโมเดลข้อมูลแบบเบ็ดเสร็จ เรียกลักษณะนี้ว่าเป็นการทำโมเดลของสารสนเทศ (Information Modeling)

ข. การออกแบบจากระดับล่างไปสู่ระดับบน (Bottom-Up Design)

เป็นการวิเคราะห์ข้อมูลของระบบงาน โดยวิเคราะห์ความต้องการของผู้ใช้แต่ละคนหรือกลุ่ม และจำลองวิวของผู้ใช้แยกกัน จากนั้นจึงนำมารวมกันเป็นฐานข้อมูลของทั้งระบบ ซึ่งวิธีนี้ใช้บ่อยในองค์กรที่ระบบงานเดิมมีการประยุกต์ใช้งานอยู่แล้ว ทำให้โมเดลข้อมูลที่ได้มีขอบเขตชัดเจน และถูกต้องกับความเป็นจริงมาก แต่ประสบปัญหาเกี่ยวกับจุดเชื่อมโยงระหว่างวิวของผู้ใช้เพื่อให้เป็นโมเดลข้อมูลแบบเบ็ดเสร็จ ทำให้โมเดลข้อมูลที่ได้ไม่สามารถแสดงกิจกรรมทางธุรกิจทั้งหมดขององค์กรอย่างสมบูรณ์ เรียกลักษณะนี้เป็นการรวบรวมวิว (View Integration)

การเลือกระบบจัดการฐานข้อมูล

การเลือกนี้จะต้องคำนึงถึงทางด้านเทคนิคและด้านเศรษฐกิจ ตลอดจนโครงสร้างในการเก็บข้อมูล ทิศทางการเข้าถึงข้อมูลที่ฐานข้อมูลชนิดนั้นจะทำได้ ภาษาระดับสูงที่ใช้ได้

การพิจารณาทางด้านเทคนิค ขอกล่าวจากฐานข้อมูลแต่ละประเภทดังนี้

1. แบบเชิงลำดับชั้น

เป็นฐานข้อมูลในรุ่นแรก จะเก็บความสัมพันธ์ของข้อมูลเป็นแบบต้นไม้

(tree) โดยมีหลักการว่า ข้อมูลจากแม่ (parent)ชี้ไปยังข้อมูลในลูก (child) ผู้ออกแบบจะต้องกำหนดแม่และลูก ตั้งแต่การสร้างว่าข้อมูลอะไรเกี่ยวข้องกัน และมีข้อจำกัดว่า ใน 1 ลูก จะเกี่ยวข้องกับ 1 แม่เท่านั้น ฐานข้อมูลแบบนี้ ได้แก่ IMS/VS

ข้อดี : ความสัมพันธ์เป็นแบบง่าย ๆ ทำให้เข้าถึงข้อมูลได้ง่าย รวดเร็ว

ข้อเสีย : ถ้าเป็นงานที่มีความสัมพันธ์กันเป็นจำนวนมาก (M:N) จะไม่เหมาะสมกับโครงสร้างแบบต้นไม้ และถ้าต้องการเปลี่ยนความสัมพันธ์ของข้อมูลภายหลังจากใช้งานไปแล้วทำได้ยาก

2. แบบข่ายงาน

ข้อเสียของแบบเชิงลำดับชั้นถูกแก้ด้วยแบบข่ายงาน คือ 1 เมมเบอร์ (member/child) จะเกี่ยวข้องกับโอนเนอร์ (owner/parent) มากกว่า 1 ได้ แต่ก็ยังต้องกำหนดความสัมพันธ์ตั้งแต่เริ่มสร้างฐานข้อมูล ไม่สามารถเปลี่ยนแปลงในระหว่างการทำงานได้ ฐานข้อมูลแบบนี้ได้แก่ DBTG

ข้อดี : เหมาะสมกับระบบที่มีความสัมพันธ์ของข้อมูลเป็นจำนวนมาก (M:N)

ข้อเสีย : ค่อนข้างซับซ้อน โปรแกรมเมอร์ที่เขียนคำสั่งเพื่อดึงข้อมูลที่ต้องการจะต้องเข้าใจและทราบตำแหน่งของข้อมูล จึงยุ่งยากกว่าแบบอื่น

3. แบบเชิงสัมพันธ์

เกิดขึ้นในช่วงหลัง เป็นฐานข้อมูลที่ได้รับการพัฒนาโดยที่ความสัมพันธ์ของข้อมูลสามารถกำหนดขึ้นมาเมื่อใดก็ได้ที่ผู้ใช้ต้องการ มีลักษณะคล้ายคลึงกับการเก็บเป็นระเบียน และเขตข้อมูล โดยเรียกระเบียนว่าทูเปิล (tuple) และเรียกเขตข้อมูลว่าแอททริบิวต์ (attribute) ทำให้เข้าใจง่าย ฐานข้อมูลแบบนี้ได้แก่ System R, DB2, SQL/DS, INGRES, ORACLE

ข้อดี : การออกแบบและเปลี่ยนแปลงความสัมพันธ์ทำได้ง่าย และอาจประยุกต์เป็นแบบเชิงลำดับชั้นหรือข่ายงานก็ได้

ข้อเสีย : ถ้าข้อมูลมีความสัมพันธ์กันเป็นจำนวนมาก จะทำให้การค้นหาข้อมูลช้ากว่าแบบอื่น

การพิจารณาทางด้านเศรษฐกิจ เป็นการพิจารณาถึงค่าใช้จ่ายที่เกี่ยวข้องเมื่อเลือกระบบจัดการฐานข้อมูลชนิดนั้น ค่าใช้จ่ายต่าง ๆ ได้แก่

ก. ราคาของซอฟท์แวร์ การเลือกต้องพิจารณาหน้าจอ วิธีการดึงข้อมูล การสำรองข้อมูลประกอบการพิจารณาราคา

ข. ค่าบำรุงรักษาระบบ

ค. ราคาของฮาร์ดแวร์ ต้องพิจารณาร่วมกับชนิดของเทอร์มินอล หน่วยความจำหลัก หน่วยความจำสำรองที่จะได้มา

ง. ค่าใช้จ่ายในการสร้างระบบใหม่ และการเปลี่ยนจากระบบเก่าเป็นระบบใหม่

จ. ค่าจ้างบุคลากร รวมถึงค่าฝึกหัดต่าง ๆ

<u>การออกแบบฐานข้อมูลทางกายภาพ</u> มีสิ่งที่ต้องคำนึงถึงดังนี้

ก. เวลาในการตอบสนอง (Response Time) คือช่วงเวลาตั้งแต่การส่งงานเข้าไปถึงเมื่อได้รับผลลัพธ์ที่ต้องการออกมา

ข. การใช้ที่ว่าง (Space Utilization) จำนวนที่ว่างของหน่วยความจำที่จะถูกใช้ โดยแฟ้มข้อมูลของฐานข้อมูลและโครงสร้างทิศทางการเข้าถึงข้อมูล

ค. งานที่ได้ออกมา (Transaction Throughput) จะคิดเป็นค่าเฉลี่ย คำนวณจากจำนวนงานที่สามารถประมวลผลได้โดยระบบจัดการฐานข้อมูลต่อหนึ่งหน่วยเวลา

<u>การควบคุมความปลอดภัย</u> (Security Control)

การควบคุมความปลอดภัย หมายถึงการป้องกันไม่ให้ทรัพยากรในระบบคอมพิวเตอร์ รวมทั้งข้อมูลเกิดความเสียหายขึ้นอันอาจเกิดจากเจตนาหรือไม่ก็ตาม การควบคุมความปลอดภัย เป็นการเพิ่มความมั่นใจว่าระบบมีมาตรฐาน และประสิทธิภาพสูงขึ้น ด้วยการป้องกันสาเหตุที่ก่อให้เกิดความเสียหายกับอุปกรณ์คอมพิวเตอร์ ซอฟต์แวร์ต่าง ๆ ประกอบด้วยเทคนิค 2 เทคนิคดังนี้

1. เทคนิคความปลอดภัยทางกายภาพ (Physical Security Techniques)

เป็นการพิจารณาถึงความปลอดภัยของอุปกรณ์ ที่ตั้งอุปกรณ์ ฯลฯ เพื่อไม่

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ให้ความเสียหายเกิดขึ้น ด้วยการคำนึงถึง

1.1 การควบคุมการเข้าถึงอุปกรณ์ต่าง ๆ

เป็นพื้นฐานของระบบความปลอดภัย มีหลักว่าความเสียหายจะมีโอกาสเกิดขึ้นน้อยมาก เมื่อมีการตรวจสอบอย่างเข้มงวดในการเข้าถึงอุปกรณ์ เช่น การอนุญาตเฉพาะบุคลากรที่ทำงานเกี่ยวข้องกับคอมพิวเตอร์เท่านั้นสามารถเข้า-ออกหน่วยงานคอมพิวเตอร์ได้ ด้วยการตรวจสอบบัตรอนุญาตเข้า-ออก เป็นต้น

1.2 สถานที่ตั้ง

ตำแหน่งที่ตั้งหน่วยงานคอมพิวเตอร์ มีสำคัญในการวางแผนความปลอดภัย เช่น พิจารณา

ก. สถานที่ตั้งระบบคอมพิวเตอร์ แยกเป็นอิสระจากสถานที่ทำงานของหน่วยงานอื่นๆ

ข. เส้นทางและที่ตั้งของสายไฟ สายสื่อสารต่างๆที่ใช้

ค. สถานที่เก็บสื่อกลางอุปกรณ์ที่ใช้ในการสำรองข้อมูล

1.3 การป้องกันอุปกรณ์ (Physical Protection)

เป็นการพิจารณาถึงการป้องกันอุปกรณ์อื่น ๆ นอกจากที่กล่าวมาแล้ว เช่น ระบบดับเพลิงต่างๆในกรณีเกิดอุบัติเหตุไฟไหม้ ระบบไฟสำรองในกรณีไฟดับ มาตรการต่างๆที่เกิดจากภัยพิบัติทางธรรมชาติ

2. เทคนิคดำเนินการความปลอดภัย (Procedural Security Technique) เป็นการพิจารณาในลักษณะการใช้งานประยุกต์ต่าง ๆ โดยคำนึงถึง

2.1 การกำหนดรหัสผ่านแก่ผู้มีสิทธิ์ และให้เข้าถึงข้อมูลได้เฉพาะส่วนที่มีสิทธิ์เท่านั้น

2.2 ควบคุมการเข้าถึงรายละเอียดของเอนติตี รีเลชันชิปและแอททริบิวต์ซึ่งในส่วนนี้ ภาษารุ่นที่ 4 จะมีคำสั่งสร้างวิว (View) เป็นส่วนช่วยให้การเข้าถึงระดับของคอลัมน์และแถวของตารางได้ อีกทั้งกำหนดการทำงานต่าง ๆ เช่น สามารถเลือกดู (SELECT) แก้ไข (UPDATE) เพิ่ม (INSERT) ภายในตารางนั้นได้

I17125632

เนื่องจากงานประมาณราคาระบบส่งไฟฟ้า เป็นลักษณะงานที่ข้อมูลมีความสัมพันธ์กับหลายหน่วยงาน อาจมีการเปลี่ยนแปลงความสัมพันธ์ของข้อมูลเพื่อให้การใช้งานเป็นไปตามความต้องการของผู้ใช้ในภายหลัง ดังนั้นเพื่อให้การแก้ไขปรับปรุง และบำรุงรักษาข้อมูลเป็นไปได้โดยง่าย สามารถพัฒนาโปรแกรมได้โดยไม่มีผลกระทบต่อข้อมูล ผู้ใช้ทั่วไปที่ไม่ใช่โปรแกรมเมอร์ สามารถเรียกดูข้อมูลได้รวดเร็ว เพราะมีภาษาระดับง่ายสำหรับผู้ใช้โดยเฉพาะ อีกทั้งการใช้ข้อมูลก็ไม่ได้ต้องการผลตอบสนองในทันที จากเหตุผลดังกล่าวจึงเลือกออกแบบและพัฒนางานประมาณราคาระบบส่งไฟฟ้าบนฐานข้อมูลแบบเชิงสัมพันธ์

ขั้นตอนการออกแบบระบบฐานข้อมูล สำหรับงานประมาณราคาระบบส่งไฟฟ้าที่สำคัญมี 2 ประการคือ

1. โมเดลข้อมูลเชิงตรรก (Logical Data Modeling : LDM)

2. ฐานข้อมูลแบบเชิงสัมพันธ์ (Relational Database Design : RDD)

1. <u>โมเดลข้อมูลเชิงตรรก</u> (Flemming and Von Halle, 1989)

เป็นโมเดลที่ใช้ในขั้นการออกแบบฐานข้อมูลเชิงมโนภาพ โดยแทนโครงสร้างและลักษณะของข้อมูลที่ต้องการ ให้อยู่ในรูปแบบที่พร้อมที่จะสร้างเป็นฐานข้อมูลแบบอื่น ๆ ต่อไป แบบจำลองข้อมูลเชิงตรรกนี้ จะทำให้ได้ฐานข้อมูลที่ถูกต้องไม่ขัดแย้งกัน สามารถใช้ข้อมูลร่วมกันได้ และสามารถปรับปรุงแก้ไขข้อมูลได้ง่าย

การสร้างแบบจำลองข้อมูลเชิงตรรกมีขั้นตอนต่าง ๆ ดังนี้

LDM 1 : กำหนดเอนติตีหลัก

เอนติตี เปรียบเสมือนกับเป็นคำนามอันได้แก่ บุคคล สถานที่และสิ่งของ ซึ่งอาจเป็นสิ่งที่มีตัวตน เช่น อาจารย์ นิสิต หรือเป็นนามธรรม เช่น โครงการ อัตราแลกเปลี่ยนเงินตรา เป็นต้น

LDM 2 : กำหนดความสัมพันธ์ระหว่างเอนติตี

ความสัมพันธ์คือลักษณะของการเกี่ยวพันกันระหว่างเอนติตีโดยรีเลชันชิป ประกอบด้วยทิศทางความสัมพันธ์ (Direction) และ สัดส่วนความสัมพันธ์ (Cardinality Ratio) ของสมาชิกในเอนติตีแม่ และเอนติตีลูก ความสัมพันธ์มีอยู่ 3 แบบคือ

1. ความสัมพันธ์ที่มีอยู่จริง (Existence Relationships) เช่น ผู้เข้าประกวดราคาเสนอราคาโครงการระบบส่งไฟฟ้า ดังนั้นความสัมพันธ์ระหว่างเอนติตีผู้เข้าประกวดราคาและเอนติตีโครงการระบบส่งไฟฟ้า คือ การเสนอราคาเป็นความสัมพันธ์ที่มีอยู่จริง

2. ความสัมพันธ์ตามหน้าที่ (Functional Relationships) เช่น คณะกรรมการที่ได้รับมอบหมาย ให้ทำหน้าที่คำนวณหาราคากลางของงานประกวดราคา ดังนั้นความสัมพันธ์ระหว่างเอนติตีคณะกรรมการที่ได้รับการมอบหมาย และเอนติตีงานประกวดราคา คือการคำนวณหาราคากลางเป็นความสัมพันธ์ตามหน้าที่

3. ความสัมพันธ์ตามเหตุการณ์ที่เกิดขึ้น (Event Relationships) เช่น เมื่อโครงการระบบส่งไฟฟ้าใดได้รับการอนุมัติ ก็จะมีการดำเนินการจัดซื้ออุปกรณ์ ดังนั้นความสัมพันธ์ระหว่างเอนติตีโครงการที่ได้รับการอนุมัติ และเอนติตีอุปกรณ์ คือการจัดซื้อ เป็นความสัมพันธ์ตามเหตุการณ์ที่เกิดขึ้น

<u>ความสัมพันธ์ (Relationships)</u>

ทิศทางของความสัมพันธ์ จะถูกกำหนดโดยใช้ลูกศรชี้จากเอนติตีแม่ไปยังเอนติตีลูกเสมอ โดยอาศัยสัดส่วนความสัมพันธ์ สามารถแบ่งความสัมพันธ์ได้ 3 ลักษณะดังนี้

1. ความสัมพันธ์แบบหนึ่งต่อหนึ่ง (1:1 or One-to-One Relationships) แต่ละค่าของเอนติตีแม่มีความสัมพันธ์กับเอนติตีลูกได้อย่างมากที่สุด 1 ค่า

เช่น อัตราเงินเฟ้อในแต่ละปีงบประมาณจะมีเพียงค่าเดียวเท่านั้น

หรือ อัตราเงินเฟ้อ <-----> ปีงบประมาณ

2. ความสัมพันธ์แบบหนึ่งต่อกลุ่ม (1:N or One-to-Many

Relationships) แต่ละค่าของเอนติตีแม่มีความสัมพันธ์กับเอนติตีลูกได้หลายค่า แต่ในทางกลับกันแต่ละค่าของเอนติตีลูกมีความสัมพันธ์กับเอนติตีแม่ได้เพียงค่าเดียว เช่น ผู้เข้าประกวดราคาโครงการจะมีสิทธิ์ชนะการประกวดราคาได้หลายโครงการ แต่การประกวดราคาโครงการหนึ่งจะมีผู้ชนะการประกวดราคาได้เพียงรายเดียวเท่านั้น

หรือ ผู้ชนะการประกวดราคา <-----›› โครงการ

3. ความสัมพันธ์แบบกลุ่มต่อกลุ่ม (M:N or Many-to-Many Relationships) แต่ละค่าของเอนติตีแม่มีความสัมพันธ์กับเอนติตีลูกได้หลายค่า และในทางกลับกัน แต่ละค่าของเอนติตีลูกมีความสัมพันธ์กับเอนติตีแม่ได้หลายค่า เช่น งานระบบส่งไฟฟ้างานหนึ่งจะมีการจัดซื้ออุปกรณ์หลายประเภท และอุปกรณ์แต่ละประเภทจะถูกซื้อนำไปใช้ในหลายงานได้

หรือ งานระบบส่งไฟฟ้า ‹‹-----›› อุปกรณ์

LDM 3 : กำหนดคีย์หลักและคีย์รอง (Primary and Alternate Keys)

ขั้นตอนนี้เป็นการเพิ่มข้อมูลที่เรียกว่า แอททริบิวต์ ลงในทุกเอนติตีโดยที่แอททริบิวต์ก็คือข้อมูลที่แสดงลักษณะและคุณสมบัติของเอนติตี เช่น แอททริบิวต์ของบุคคล ก็อาจจะได้แก่ สีตา สีผม อายุ เพศ ชื่อ ฯลฯ เปรียบได้กับฟิลด์ (field) ของแฟ้มข้อมูลนั่นเอง

คีย์หลัก จะประกอบด้วยค่าของแอททริบิวต์ 1 ตัว หรือมากกว่า 1 ตัวก็ได้ ที่สามารถใช้เป็นตัวเจาะจงบอกว่าต้องการอ้างถึงทูเปิลใด เช่น รหัสของโครงการ เป็นต้น

คีย์รอง จะมีคุณสมบัติเหมือนคีย์หลักทุกประการ แต่ไม่ได้ถูกเลือกให้เป็นคีย์หลัก

คีย์หลักของเอนติตีที่เป็นซับไทป์นั้น จะต้องเหมือนกับคีย์หลักของซูเปอร์ไทป์เอนติตี

LDM 4 : กำหนดคีย์นอก (Foreign Keys)

คีย์นอกคือแอททริบิวต์ของเอนติตีหนึ่งที่ซ้ำกับแอททริบิวต์ในอีกเอนติตีหนึ่ง

เป็นตัวทำให้เกิดความสัมพันธ์ระหว่างเอนติตี และใช้เป็นส่วนช่วยควบคุมความถูกต้องของข้อมูลด้วย โดยคีย์นอกจะอยู่ในเอนติตีลูก และมีค่าเท่ากับคีย์หลักของเอนติตีแม่

LDM 5 : กำหนดกฎการจัดการกับข้อมูล (Key Business Rules)

เป็นสิ่งที่ใช้ตรวจสอบเพื่อให้แน่ใจในความถูกต้อง และสอดคล้องกันของข้อมูลมี 3 เรื่องดังนี้

5.1 โดเมน (Domain)

5.2 กฎการเพิ่มและลบข้อมูล (Insert and Delete Rule)

5.3 ทริกเกอร์ดำเนินการ (Trigger Operation)

5.1 โดเมน หมายถึงกรอบของค่าต่าง ๆ ที่เป็นไปได้ของแอททริบิวต์ เช่น กรอบของแอททริบิวต์วันที่ จะกินความเฉพาะค่าวันที่ที่เป็นไปได้ เช่น จำนวนวันอยู่ระหว่าง 1 ถึง 31 มีเดือนอยู่ระหว่าง 1 ถึง 12 หรือโดเมนของชื่อต้องเป็นตัวอักษรเท่านั้น เป็นต้น

5.2 กฎการเพิ่มและลบข้อมูล

5.2.1 กฎการเพิ่มข้อมูล เป็นการกำหนดเงื่อนไขที่ต้องตรวจสอบสำหรับเอนติตีแม่ ในการเพิ่มคีย์นอกในเอนติตีลูก โดยทั่วไปมี 6 ลักษณะคือ

ก. การเพิ่มแบบขึ้นต่อกัน (Dependent) อนุญาติให้เพิ่มข้อมูลในเอนติตีลูกได้ ถ้ามีข้อมูลที่สัมพันธ์กันในเอนติตีแม่

ข. การเพิ่มแบบอัตโนมัติ (Automatic) ถ้าตรวจสอบพบว่าไม่มีข้อมูลที่สัมพันธ์กันในเอนติตีแม่ จะเพิ่มค่าข้อมูลในเอนติตีแม่ให้เองทันที

ค. เปลี่ยนเป็นค่าว่าง (Nullify) ถ้าตรวจสอบพบว่าไม่มีข้อมูลที่สัมพันธ์กันในเอนติตีแม่จะเปลี่ยนค่าคีย์นอกเป็นค่าว่าง โดยที่ค่าว่าง ก็คือค่าที่ไม่ทราบแน่ชัด หรือค่าที่ไม่เหมาะสม กล่าวอีกนัยหนึ่งก็คือ ค่าว่างจะเป็นค่าที่ไม่อยู่ในกรอบของโดเมนนั่นเอง

ง. เปลี่ยนเป็นค่าที่กำหนด (Default) กรณีที่ไม่มีข้อมูลในเอนติตีแม่ จะเปลี่ยนค่าคีย์นอกเป็นค่าที่กำหนดไว้

จ. แบบมีเงื่อนไข (Customized) จะเพิ่มข้อมูลในเอนติตี

ลูกได้ ก็ต่อเมื่อเงื่อนไขที่กำหนดไว้ถูกตรวจสอบว่าเป็นจริง

ฉ. ไม่ตรวจสอบ (No Effect) สามารถเพิ่มข้อมูลในเอนติตีลูกได้โดยไม่ต้องตรวจสอบจากเอนติตีแม่ว่ามีค่าตรงกับเอนติตีแม่หรือไม่

5.2.2 กฎการลบข้อมูล เป็นการกำหนดเงื่อนไขในการลบคีย์หลักในเอนติตีแม่ที่ถูกอ้างอิงถึงโดยคีย์นอกในเอนติตีลูก โดยทั่วไปมี 6 ลักษณะคือ

ก. การลบแบบมีข้อจำกัด (Restricted) จะลบข้อมูลในเอนติตีแม่ได้ ก็ต่อเมื่อไม่มีข้อมูลที่สัมพันธ์กันในเอนติตีลูกเท่านั้น

ข. กระทำแบบเป็นทอด ๆ (Cascade) จะลบข้อมูลในเอนติตีแม่และจะลบข้อมูลที่สัมพันธ์กันในเอนติตีลูกออกด้วย

ค. เปลี่ยนเป็นค่าว่าง การลบข้อมูลจะทำการเปลี่ยนข้อมูลของคีย์นอกที่อ้างอิงถึงให้เป็นค่าว่าง

ง. เปลี่ยนเป็นค่าที่กำหนด การลบข้อมูล จะทำการเปลี่ยนข้อมูลของคีย์นอกที่อ้างอิงถึงให้เป็นค่าที่กำหนดไว้

จ. แบบมีเงื่อนไข การลบข้อมูลจะกระทำได้ก็ต่อเมื่อเป็นไปตามเงื่อนไขที่กำหนดไว้

ฉ. ไม่ตรวจสอบ การลบข้อมูลไม่ต้องทำการตรวจสอบเงื่อนไขใด ๆ ทั้งสิ้น

5.3 ทริกเกอร์ดำเนินการ เป็นข้อกำหนดถึงผลกระทบที่ต้องดำเนินการกับเอนติตีอื่น หรือแอททริบิวต์อื่นในเอนติตีเดียวกัน เมื่อมีการเพิ่ม ลบ แก้ไข และเรียกใช้ข้อมูล

LDM 6 : เพิ่มแอททริบิวต์อื่น ๆ ลงในเอนติตี

หลังจากกำหนดคีย์หลัก คีย์รองและคีย์นอกให้กับแต่ละเอนติตีแล้ว ต่อไปก็จะเป็นการกำหนดแอททริบิวต์อื่น ๆ ที่เหลือ (nonkey attribute) ให้กับเอนติตี โดยมีข้อแม้ว่าต้องเป็นแอททริบิวต์ที่ขึ้นกับคีย์หลักเท่านั้น และยังคงทำให้คีย์หลักมีค่าไม่ซ้ำกันด้วย (unique)

ในกรณีที่แอททริบิวต์ใด สามารถหาได้จากสูตร หรือคำนวณได้จากแอททริบิวต์อื่น จะเรียกว่าเป็นดีไรฟ์แอททริบิวต์ (Derived Attribute) ให้

ระบุด้วย 'd' ในโมเดลข้อมูลด้วย กรณีเป็นแฟล็กใส่ไว้เพื่อระบุชับไทด์-ซุปเปอร์ไทด์ ให้ระบุด้วย 's' ในโมเดลเช่นกัน

LDM 7 : ตรวจสอบความถูกต้องของมุมมองผู้ใช้โดยใช้ทฤษฎีนอร์มัลไลเซชัน (Validate User Views Through Normalization Rules)

ข้อดีของการทำนอร์มัลไลเซชัน

ก. ลดเนื้อที่ที่ใช้เก็บข้อมูล

ข. ลดปัญหาเรื่องความขัดแย้งของข้อมูล

ค. ลดความซ้ำซ้อนของข้อมูล

นอร์มัลไลเซชันเป็นทฤษฎีที่ใช้วิเคราะห์และจัดโครงสร้างข้อมูลใหม่เพื่อลดความซ้ำซ้อนของโครงสร้างข้อมูล และได้โครงสร้างที่มีเสถียรภาพ โดยการออกแบบระบบให้อยู่ในรูปนอร์มัล (Normal Form) ก่อนที่จะกล่าวถึงรูปนอร์มัล ขอกล่าวถึงนิยามที่เกี่ยวข้องดังต่อไปนี้

1. ตัวเลือก (Determinant) คือแอททริบิวต์หรือกลุ่มของแอททริบิวต์ใด ๆ ที่สามารถเลือก (determine) แอททริบิวต์ตัวอื่น ๆ ได้

2. ฟังก์ชันการขึ้นต่อกัน (Functional Dependency)

นิยาม ถ้าให้ B และ A เป็นแอททริบิวต์ กล่าวได้ว่า B มีฟังก์ชันการขึ้นอยู่กับค่าของ A ก็ต่อเมื่อ สามารถใช้ A เป็นตัวเลือกค่าของ B ได้เพียงหนึ่งค่าเสมอ

สัญญลักษณ์ที่ใช้แทน B มีฟังก์ชันการขึ้นอยู่กับ A คือ A -> B

3. การขึ้นต่อกันอย่างสมบูรณ์ (Full Functional Dependency)

นิยาม ถ้าให้ B และ A เป็นแอททริบิวต์ กล่าวได้ว่า B มีฟังก์ชันการขึ้นอยู่กับค่าของ A อย่างสมบูรณ์ ก็ต่อเมื่อ B มีฟังก์ชันการขึ้นอยู่กับค่าของ A แต่ไม่มีคุณสมบัติฟังก์ชันการขึ้นอยู่กับแอททริบิวต์อื่นที่เป็นส่วนประกอบของ A

4. การขึ้นต่อกันเชิงกลุ่ม (Multivalued Dependency)

นิยาม ในเอนติตีที่ประกอบด้วยแอททริบิวต์ A B และ C กล่าวได้ว่าการขึ้นต่อกันเชิงกลุ่มระหว่าง B และ A โดยที่ B ขึ้นต่อ A คือค่าของ A หนึ่งค่าจะอิงกับกลุ่มของ B โดยการขึ้นต่อกันนี้จะเป็นอิสระกับค่าของ C

หลักการนอร์มัลไลเซชันโดยสังเขป คือการจัดโครงสร้างข้อมูลให้อยู่ในรูปนอร์มัล ซึ่งประกอบด้วยระดับต่าง ๆ ดังต่อไปนี้

รูปแบบนอร์มัลระดับที่ 1 (First Normal Form : 1NF)

นิยาม เอนติตีจะอยู่ในรูป 1NF ถ้าไม่มีกลุ่มของแอททริบิวต์ที่มีค่าซ้ำกัน

รูปแบบนอร์มัลระดับที่ 2 (Second Normal Form : 2NF)

นิยาม เอนติตีจะอยู่ในรูป 2NF เมื่อเอนติตีดังกล่าวอยู่ในรูป 1NF และไม่มีนันคีย์แอททริบิวต์ตัวใด ขึ้นกับส่วนใดส่วนหนึ่งของคีย์

รูปแบบนอร์มัลระดับที่ 3 (Third Normal Form : 3NF)

นิยาม เอนติตีจะอยู่ในรูป 3NF เมื่อเอนติตีดังกล่าวอยู่ในรูป 2NF และนำแอททริบิวต์อื่นที่ขึ้นกับนันคีย์แอททริบิวต์ออก

รูปนอร์มัลแบบบอยซ์/คอดด์ (Boyce/Codd Normal Form:BCNF)

นิยาม เอนติตีจะอยู่ในรูป BCNF ก็ต่อเมื่อเอนติตีดังกล่าวอยู่ในรูป 3NF และตัวเลือกทุกตัวต้องเป็นแคนดิเดตคีย์เท่านั้น

รูปแบบนอร์มัลระดับที่ 4 (Fourth Normal Form : 4NF)

นิยาม เอนติตีจะอยู่ในรูป 4NF ก็ต่อเมื่อเอนติตีดังกล่าวอยู่ในรูปBCNF และเป็นความสัมพันธ์ที่ไม่ขึ้นต่อกันเชิงกลุ่ม

รูปแบบนอร์มัลระดับที่ 5 (Fifth Normal Form : 5NF)

นิยาม เอนติตีจะอยู่ในรูป 5NF ก็ต่อเมื่อเอนติตีดังกล่าวอยู่ในรูป 4NF และไม่สามารถแยกเอนติตีดังกล่าวออกได้อีก

สำหรับการออกแบบโมเดลข้อมูลเชิงตรรก โดยการประยุกต์ใช้ทฤษฎีนี้ถึงแม้ว่าจะได้โครงสร้างที่ถูกขจัดความซ้ำซ้อนและมีเสถียรภาพ แต่เมื่อแปลงอยู่ในรูปฐานข้อมูลเชิงกายภาพ อาจมีปัญหาในเรื่องประสิทธิภาพในการประมวลผล จึงอาจต้องมีการปรับให้อยู่ในรูปแบบที่เหมาะสมในภายหลัง นอกจากนี้การที่จะให้แน่ใจว่าระบบที่ได้ถูกต้องตามความต้องการของผู้ใช้ การใช้ทฤษฎีนอร์มัลไลเซชันควรกระทำภายหลังจากการออกแบบผ่านขั้นตอน LDM 1 ถึง LDM 6 มาก่อน

ปกติการจัดโครงสร้างข้อมูลให้อยู่ในรูป 3NF ก็เพียงพอในการขจัดปัญหาความซ้ำซ้อนและความขัดแย้งกันของข้อมูล ส่วนการจัดให้อยู่ในรูปนอร์มัลระดับสูงขึ้นต้องพิจารณาถึงความจำเป็นของข้อมูล

LDM 8 : กำหนดโดเมนของแอททริบิวต์ พร้อมบันทึกรายละเอียดในพจนานุกรมข้อมูล (Data Dictionary) ซึ่งประกอบด้วยรายละเอียดดังนี้

ก. ชนิดของข้อมูล (Data Type) เช่น จำนวนเต็ม ทศนิยม ตัวอักษร วันที่

ข. รูปแบบข้อมูล (Format) เช่น dd/mm/yy (วันที่) nnn-nnnn (เบอร์โทรศัพท์)

ค. ค่าที่ยอมรับได้ (Range) เช่น อายุพนักงานจะเป็นเลขจำนวนเต็ม 2 หลักเท่านั้น

ง. ความหมาย (Meaning) อธิบายความหมายของแอททริบิวต์นั้นว่าคืออะไร

จ. ความยาว (Length) เช่น 7 หลัก 20 ตัวอักษร

ฉ. มีค่าซ้ำกันได้หรือไม่ (Uniqueness)

ช. เป็นค่าว่างได้หรือไม่ (Null Support)

ได้มีการกำหนดกฎไว้เป็นพิเศษสำหรับโดเมนของแอททริบิวต์บางพวกต่อไปนี้

8.1 คีย์หลัก ต้องมีค่าไม่ซ้ำกัน และต้องไม่เป็นค่าว่าง

8.2 คีย์รอง ต้องมีค่าไม่ซ้ำกัน แต่เป็นค่าว่างได้

8.3 คีย์นอก ต้องสอดคล้องกับโดเมนของคีย์หลักในเอนติตีแม่

8.4 ดีไรฟ์แอททริบิวต์ ต้องมีค่าอยู่ในช่วงของผลลัพธ์ที่ได้จากการผ่านอัลกอริทึมของแอททริบิวต์ที่เป็นที่มา และมีชนิดของข้อมูลแบบเดียวกันด้วย

8.5 แอททริบิวต์ที่เป็นคีย์หลักของซับไทป์เอนติตี ต้องมีค่าเป็นสับเซท (subset) ของคีย์หลักของซูเปอร์ไทป์เอนติตี

LDM 9 : กำหนดทริกเกอร์ดำเนินการ

LDM 10 : รวมมุมมองของผู้ใช้เข้าด้วยกันเพื่อลดความซ้ำซ้อน ความไม่สอดคล้องของข้อมูล และเพิ่มความสัมพันธ์ระหว่างมุมมองของผู้ใช้ ซึ่งประกอบด้วยการรวมเอนติตี การรวมความสัมพันธ์ และการรวมแอททริบิวต์

ข้อควรพิจารณาในการรวมเอนติตี

1. การรวมเอนติตีที่มีคีย์หลักตัวเดียวกัน และค่าที่เป็นไปได้ของคีย์หลักเหมือนกัน จะต้องได้เอนติตีใหม่ที่มีแอททริบิวต์รวมของสองเอนติตีเดิม

2. ถ้าเอนติตีสองอันมีคีย์หลักเดียวกัน แต่ค่าที่เป็นไปได้ของคีย์หลักนั้นเป็นสับเซตของกัน จะรวมได้ในรูปของซูเปอร์ไทป์-ซับไทป์ โดยต้องตัดแอททริบิวต์ที่มีแล้วในซูเปอร์ไทป์เอนติตีออกจากซับไทป์เอนติตี

3. ถ้าเอนติตีสองอันมีคีย์หลักเดียวกัน แต่มีผลไปกำหนดแอททริบิวต์ที่ต่างกันบางตัว ให้กำหนดซูเปอร์ไทป์ขึ้นมาอันหนึ่งให้สัมพันธ์กับสองเอนติตีเดิม

4. การเชื่อมเอนติตีสองตัวที่มีคีย์หลักของตัวหนึ่ง เป็นคีย์รองของอีกตัว จะได้เอนติตีใหม่ที่มีคีย์หลักตามเอนติตีตัวหนึ่ง ส่วนคีย์หลักของเอนติตีอีกตัวจะกลายเป็นคีย์รองไป และมีแอททริบิวต์รวมระหว่างสองเอนติตีเดิม แล้วตัดแอททริบิวต์ที่ซ้ำซ้อนออกเสีย และต้องกำหนดกฎการแก้ไขให้ด้วยตามข้อบังคับเก่า เช่น ในกรณีคีย์หลักเดิมที่กลายมาเป็นคีย์รองในเอนติตีใหม่ก็ต้องยังคงห้ามเป็นค่าว่างด้วย

5. การรวมเอนติตีใด ๆ ก็ตาม ต้องไม่มีผลไปเปลี่ยนแปลงเอนติตีอื่นที่ไม่เกี่ยวข้อง

ข้อควรพิจารณาในการรวมรีเลชันชิป

1. ให้รวมรีเลชันชิประหว่างเอนติตีที่ให้ความหมายเหมือนกัน เข้าด้วยกันโดยถ้าผลทำให้เกิดเป็นรีเลชันชิปแบบ M:N จะต้องทำการแตกให้เป็นรีเลชันชิปแบบ 1:N สองอัน

2. การรวมรีเลชันชิปใด ๆ ต้องไม่ไปกระทบกับรีเลชันชิปอื่นที่ไม่ต้องการการเปลี่ยนแปลง นอกจากจะพิจารณาแล้วว่าควรตัดออกเนื่องจากซ้ำซ้อน หรืออาจเพิ่มขึ้นใหม่เพื่อความเหมาะสม

3. จากการรวมเอนติตีที่มีคีย์หลักเป็นคีย์รองของเอนติตีอีกตัว ให้ตรวจสอบคีย์นอกของเอนติตีอื่น ๆ ที่อ้างมาถึงว่าได้อ้างถึงคีย์หลักหรือคีย์รองของเอนติตีใหม่ที่ได้จากการรวมนั้น ถ้าอ้างถึงคีย์รองต้องทำการเปลี่ยนให้เป็นคีย์หลัก

4. เมื่อรวมมุมมองต่าง ๆ แล้ว ให้กำหนดกฎการแก้ไขคีย์สำหรับรีเลชันชิปใหม่ด้วย

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ข้อควรพิจารณาในการรวมแอททริบิวต์

1. ให้ทำการรวมแอททริบิวต์ที่มีความหมายเหมือนกัน ภายในเอนติตีเดียวกัน และรวมค่าที่เป็นไปได้รวมถึงกฎการจัดการเข้าด้วยกันด้วย โดยพิจารณาค่าที่เป็นไปได้ของแอททริบิวต์อื่นด้วยว่าเปลี่ยนไปหรือไม่

2. เมื่อรวมเอนติตีแล้วให้พิจารณาตัดแอททริบิวต์ที่เป็นคีไรพ์หรือแฟล็กที่ไม่จำเป็นทิ้งเสีย

3. หลังจากได้รวม ตัดหรือเพิ่มรีเลชันชิปแล้วให้ทำการนอร์มัลไลซ์อีกครั้งเพื่อตัดสิ่งที่ซ้ำซ้อนออกเสีย

LDM 11 : รวมโมเดลที่ได้เข้ากับโมเดลที่มีอยู่เดิม

LDM 12 : วิเคราะห์ความเสถียรภาพและการเติบโตในอนาคต การออกแบบโมเดลในขั้นตอนที่ผ่านมาจะพิจารณาข้อมูลที่เห็นได้ในปัจจุบันเป็นส่วนมาก ในขั้นสุดท้ายนี้จะพิจารณาสิ่งที่อาจเกิดขึ้นหรือเป็นไปได้ในอนาคต แล้วทำการดัดแปลงโมเดลไว้รองรับ หรือจดบันทึกไว้เพื่อพิจารณาในภายหลัง

2. <u>ฐานข้อมูลเชิงสัมพันธ์</u>

เป็นหลักเกณฑ์ที่ใช้ในการแปลงโมเดลข้อมูลเชิงตรรก เป็นโมเดลข้อมูลเชิงกายภาพ เพื่อให้สะดวกในการใช้งานจริงบนฐานข้อมูลเชิงสัมพันธ์ ซึ่งประกอบด้วยโครงสร้างข้อมูล (Data Structure) และความถูกต้องของข้อมูล (Data Integrity)

2.1 <u>โครงสร้างของฐานข้อมูลเชิงสัมพันธ์</u> ประกอบด้วย

2.1.1 คอลัมน์ คือคุณสมบัติของข้อมูลเมื่อเปรียบเทียบกับโมเดลข้อมูลเชิงตรรก ก็คือแอททริบิวต์ในเอนติตีนั่นเอง

2.1.2 แถว คือค่าของข้อมูลในตาราง

2.1.3 ตารางความสัมพันธ์ เป็นตาราง 2 มิติ ที่

ก. ต้องไม่มีคอลัมน์ที่ซ้ำกันในตาราง

ข. ค่าของข้อมูลในคอลัมน์จะต้องมีคุณสมบัติเหมือนกัน

ค. ข้อมูลในตาราง จะต้องไม่มีข้อมูลแถวใดที่ซ้ำกัน

ง. ลำดับของคอลัมน์ไม่มีความสำคัญ

จ. ลำดับข้อมูลแต่ละแถวไม่มีความสำคัญ

ฉ. ชื่อคอลัมน์ในตารางต้องไม่ซ้ำกัน

2.2 <u>ความถูกต้องของข้อมูล</u>

เมื่อกล่าวถึงความถูกต้องของข้อมูล โดยทั่วไปแล้วจะกินความหมายถึงกฎเกณฑ์ความถูกต้องอยู่ 3 กรณีคือ

2.2.1 ความถูกต้องของเอนติตี

ซึ่งกล่าวว่า 'แอททริบิวต์ทุกตัวที่เป็นส่วนของคีย์หลัก จะไม่อนุญาตให้มีค่าว่าง' เป็นกฎสำหรับการเพิ่ม ปรับปรุง และลบข้อมูลในตาราง

2.2.2 ความถูกต้องของการอ้างอิง

ซึ่งกล่าวว่า 'ถ้าเรามีตารางความสัมพันธ์ R2 ซึ่งมี FK เป็นคีย์นอกที่อ้างอิงถึงคีย์หลัก PK ในตารางความสัมพันธ์ R1 สำหรับทุก ๆ ค่าของ FK ใน R2 จะต้อง

ก. มีค่าเท่ากับ PKในแถวใดแถวหนึ่งในตารางความสัมพันธ์ R1

ข. มีค่าของแอททริบิวต์ทุกตัวใน FK เป็นค่าว่าง

ความสำคัญของกฎข้อนี้คือ เมื่อมีการอ้างอิงจากความสัมพันธ์ที่ 1 ไปยังความสัมพันธ์ที่ 2 ต้องมีค่าเสมอ

2.2.3 ความถูกต้องของโดเมน

หมายถึงกฎการรักษาความถูกต้องของคอลัมน์ทุกคอลัมน์ในตาราง รวมทั้งคีย์หลัก คีย์นอกและคอลัมน์ที่ไม่ใช่คีย์ ซึ่งกฎนี้ประกอบด้วย

ก. ชนิดของข้อมูล

ข. รูปแบบของข้อมูล

ค. ความยาวของข้อมูล

ง. ช่วงค่าของข้อมูล

จ. ค่าที่กำหนดไว้

ฉ. มีค่าซ้ำกันหรือค่าว่างได้หรือไม่

เวลาที่พูดถึงกฎความถูกต้องดังกล่าว เราต้องการที่จะประกันความถูกต้องนี้ในทุก ๆ สภาวะ และทุก ๆ ขณะของฐานข้อมูล นั่นก็หมายความว่าเมื่อมีการแก้ไขข้อมูลในระบบเกิดขึ้นเมื่อไร ก็เป็นหน้าที่ของระบบจัดการฐานข้อมูลที่จะต้องตรวจสอบอยู่เสมอว่า ยังคงมีความถูกต้องตามกฎเกณฑ์อยู่ และถ้าการแก้ไขข้อมูลครั้งใด จะยังผลให้ผิดกฎเกณฑ์ความถูกต้องไปแล้วก็ขึ้นอยู่กับการออกแบบของระบบจัดการฐานข้อมูลแต่ละตัวว่าจะแก้ไขสถานการณ์อย่างไร

วิธีการแก้ไขสถานการณ์แบ่งได้เป็น 3 วิธีคือ

วิธีที่ 1 : ระบบจะไม่ยอมรับการเปลี่ยนแปลงแก้ไขข้อมูลที่ทำให้เกิดปัญหานั้น โดยอาจจะแสดงข้อความออกมาบอกผู้ใช้

วิธีที่ 2 : ระบบจะอนุญาติให้การแก้ไขนั้น ๆ เกิดขึ้นได้ โดยระบบจะทำการปรับให้ผลลัพธ์มีความคงสภาพโดยอัตโนมัติ

วิธีที่ 3 : ระบบจะอนุญาติให้ผู้ใช้หรือผู้บริหารฐานข้อมูล สามารถตัดสินใจได้ในแต่ละคราวว่าจะยอมหรือไม่ยอมให้มีการแก้ไขข้อมูล

2.3 <u>ขั้นตอนการออกแบบฐานข้อมูลเชิงสัมพันธ์</u> มีขั้นตอนสำคัญดังนี้

RDD 1 : เปลี่ยนเอนติตีของโมเดลข้อมูลเชิงตรรก ให้อยู่ในรูปตาราง โดย 1 เอนติตี ต่อ 1 ตาราง

RDD 2 : แปลง 1แอททริบิวต์ของโมเดลข้อมูลเชิงตรรกเป็น 1คอลัมน์

RDD 3 : ปรับโครงสร้างของโมเดลที่ได้จาก RDD 1 และRDD 2 ให้เหมาะกับการติดตั้งระบบตามผลิตภัณฑ์ที่เลือกใช้

RDD 4 : ออกแบบกฎการจัดการข้อมูลกับเอนติตี

RDD 5 : ออกแบบกฎการจัดการข้อมูลกับความสัมพันธ์

RDD 6 : ออกแบบกฎการจัดการข้อมูลกับแอททริบิวต์ที่เหลือ

2.4 <u>ลักษณะของฐานข้อมูลเชิงสัมพันธ์ที่สมบูรณ์</u> (Hursch, 1991)

นักคณิตศาสตร์ชื่อคอดด์ (E.F.Codd) ได้เสนอรากฐานของระบบฐานข้อมูลเชิงสัมพันธ์ว่า คือการรวบรวมตารางแบบต่าง ๆ ที่สามารถกระทำการใด ๆ บนคณิตศาสตร์ของตรรกะได้ และตารางนี้จะต้องเป็นไปตามข้อบัญญัติดังนี้ (Fidelity Rules)

กฎข้อที่ 0 (Foundation Rule) : ระบบใด ๆ จะประกาศว่าเป็น

ระบบจัดการฐานข้อมูลเชิงสัมพันธ์จะต้องจัดการกับฐานข้อมูลทั้งหมด ด้วยความสามารถในลักษณะที่เป็นตาราง

กฎข้อที่ 1 (Information Rule) : ข้อมูลทั้งหมดในฐานข้อมูลเชิงสัมพันธ์จะต้องถูกแสดงในระดับตรรกะด้วยวิธีการเดียวเท่านั้น โดยลักษณะของตารางต่าง ๆ เช่น ประกอบด้วย ชื่อตาราง ชื่อคอลัมน์ ชนิดของข้อมูลในแต่ละคอลัมน์ เป็นต้น

กฎข้อที่ 2 (Guaranteed Access Rule) : ในแต่ละและทุก ๆ ค่าของข้อมูลในฐานข้อมูลเชิงสัมพันธ์ จะต้องรับประกันได้ว่าผู้ใช้จะสามารถเข้าถึงข้อมูลทุกตัวในทุก ๆ ตารางได้โดยการรวมกันของชื่อตาราง ค่าคีย์หลัก และชื่อคอลัมน์

กฎข้อที่ 3 (Missing Information Rule) : การแยกแยะข้อมูลที่ยังไม่พร้อมหรือยังไม่มีค่านั้น จะใช้ค่าว่างเป็นตัวบ่ง เพื่อให้เกิดความแตกต่างระหว่างสตริงที่ว่างหรือสตริงของค่าแบลงค์ หรือค่า 0

กฎข้อที่ 4 (System Catalog Rule) : โครงสร้างแคตตาลอกของระบบจะต้องมีสภาพของฐานข้อมูลเชิงสัมพันธ์ ดังนั้น ผู้ใช้จะสามารถเรียกดู หรือแก้ไขได้ด้วยภาษาเดียวกันกับการเรียกดูข้อมูลของระบบ

กฎข้อที่ 5 (Comprehensive Language Rule) : ไม่ว่าจะมีภาษาอยู่กี่ภาษาหรือโหมดในการติดต่ออยู่กี่โหมดที่ใช้ได้ แต่อย่างน้อยที่สุดจะมี 1 ภาษาที่สามารถนิยามโครงสร้างของระบบข้อมูล นิยามวิว เปลี่ยนแปลงแก้ไขข้อมูล ทั้งแบบการเขียนโปรแกรม หรือใช้คำสั่งแบบมีการประมวลผลโดยทันที ควบคุมความถูกต้องของข้อมูลและให้สิทธิการใช้งาน

กฎข้อที่ 6 (View Updatability Rule) : ระบบจัดการฐานข้อมูลจะต้องมีการแก้ไขข้อมูลผ่านทางวิวได้โดยสามารถตัดสินใจได้ว่า จะให้ผู้ใช้เพิ่มหรือลบแถวในตารางหรือแก้ไขคอลัมน์ใด ๆ โดยผ่านทางวิวได้เลยหรือไม่

กฎข้อที่ 7 (Set Level Updates Rule) : ระบบจัดการฐานข้อมูลควรจะมีภาษาที่สามารถให้ผู้ใช้สามารถเพิ่ม ลด และแก้ไขข้อมูลในหลาย ๆ แถวหรือหลาย ๆ คอลัมน์ได้

กฎข้อที่ 8 (Physical Data Independence Rule) : ระบบ

จัดการฐานข้อมูลจะต้องมีความเป็นอิสระของข้อมูลในระดับกายภาพ คือไม่ขึ้นกับฮาร์ดแวร์หรือวิธีจัดการกับข้อมูล (Access Method)

กฎข้อที่ 9 (Logical Data Independence Rule) : ระบบจัดการฐานข้อมูลจะต้องมีความเป็นอิสระของข้อมูลในระดับตรรก ไม่มีผลต่อการเปลี่ยนแปลงตารางหลัก

กฎข้อที่ 10 (Integrity Independence Rule) : ระบบจัดการฐานข้อมูลจะต้องมีความเป็นอิสระในการจัดเก็บข้อมูลเพื่อให้คงความถูกต้อง โดยไม่ขึ้นกับโปรแกรม

กฎข้อที่ 11 (Distribution Independence Rule) : ระบบจัดการฐานข้อมูลจะต้องไม่ขึ้นต่อการกระจายของข้อมูล คือผู้ใช้ไม่ต้องสนใจว่าข้อมูลจะอยู่ในส่วนใด หรือคอมพิวเตอร์เครื่องใดที่เชื่อมโยงกันเป็นระบบ

กฎข้อที่ 12 (Nonsubversion Rule) : ระบบจัดการฐานข้อมูลจะต้องไม่ยอมให้ภาษาอื่นใด สามารถทำลายกฎต่าง ๆ โดยเฉพาะอย่างยิ่งในกฎที่เกี่ยวกับความถูกต้องของระบบ หรือไม่ว่าจะเป็นเครื่องมือใด ๆ ก็ตาม

ปัจจุบันยังไม่มีผลิตภัณฑ์ใดที่มีคุณสมบัติครบถ้วนตามกฎดังกล่าว ดังนั้นการตัดสินใจเลือกใช้ระบบจัดการฐานข้อมูลตัวใด จึงขึ้นกับการลักษณะงานว่ามีความจำเป็นต้องปฏิบัติให้เป็นไปตามกฎข้อใด และกฎข้อใดสามารถละเลยได้