บทที่ 2

แนวความคิดและทฤษฎีพื้นฐาน

ข้อมูลและสารสนเทศ (สุโขทัยธรรมาธิราช, 2534)

ข้อมูล (data) หมายถึง ข้อเท็จจริงต่าง ๆ ที่มีอยู่ในธรรมชาติ เป็นกลุ่มสัญลักษณ์แทนปริมาณหรือการกระทำต่าง ๆ ที่ยังไม่ผ่านการประมวลผล ข้อมูลอาจจะอยู่ในรูปของตัวเลข ตัวหนังสือ และท้ายที่สุดข้อมูลก็คือ วัตถุดิบของสารสนเทศ

สารสนเทศ (information) ได้แก่ ข้อมูลต่าง ๆ ที่ได้รับการประมวลผลแล้วด้วยวิธีการต่าง ๆ เป็นความรู้ที่ต้องการสำหรับใช้ทำประโยชน์ เป็นส่วนผลลัพธ์หรือเอาท์พุตของระบบการประมวลผลข้อมูล เป็นสิ่งซึ่งสื่อความหมายให้ผู้รับเข้าใจ และสามารถนำไปกระทำกิจกรรมใดกิจกรรมหนึ่งโดยเฉพาะได้ หรือเพื่อเป็นการย้ำความเข้าใจที่มีอยู่แล้วให้มีมากยิ่งขึ้น

ความสัมพันธ์ของข้อมูล สารสนเทศ และระบบสารสนเทศ (สุโขทัยธรรมาธิราช, 2534)

ระบบสารสนเทศจะประกอบไปด้วยส่วนนำเข้า ส่วนกระบวนการหรือส่วนประมวลผล และส่วนผลลัพธ์ ซึ่งตัวข้อมูลจะเป็นวัตถุดิบของระบบในส่วนนำเข้าเพื่อประมวลผลข้อมูลด้วยวิธีการต่าง ๆ กัน และได้สารสนเทศเป็นผลลัพธ์ของระบบสารสนเทศ ระบบสารสนเทศในปัจจุบันนี้ได้นำเอาเครื่องคอมพิวเตอร์เข้ามาใช้ เพื่อให้การประมวลผลทำได้รวดเร็วขึ้น และเพื่อให้สารสนเทศที่ได้มีคุณภาพที่น่าเชื่อถือมากขึ้น

เนื่องจากสารสนเทศเป็นสิ่งสำคัญ ดังนั้นผู้ที่จะผลิตสารสนเทศเพื่อสื่อความหมายให้ได้ครบถ้วน ทำให้ผู้รับสารสนเทศแต่ละคนสามารถทำความเข้าใจได้ใกล้เคียงกัน จะต้องคำนึงถึงปัจจัยที่สำคัญต่อไปนี้ด้วย ได้แก่ สิ่งที่ใช้อ้างอิง ทัศนคติของผู้รับ ความตั้งใจในการที่จะทำความเข้าใจ และการได้ยินหรือได้เห็นที่แตกต่างกัน เป็นต้น

คุณสมบัติของสารสนเทศ (สุโขทัยธรรมาธิราช, 2534)

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารสนเทศที่จัดว่าเป็นสารสนเทศที่ดี ควรจะมีคุณสมบัติที่สำคัญ คือ ถูกต้องทันต่อการใช้งาน ครบถ้วนสมบูรณ์ กะทัดรัด และตรงกับความต้องการ ซึ่งคุณสมบัติเหล่านี้จะเป็นตัวบ่งบอกถึงคุณภาพและคุณค่าของสารสนเทศนั้น ในการนำสารสนเทศไปใช้งานครั้งหนึ่ง ๆ สารสนเทศควรจะมีคุณสมบัติดังกล่าวนี้เป็นปริมาณที่มากหรือน้อยเพียงใดนั้น ขึ้นอยู่กับความต้องการของผู้ใช้ ซึ่งเขาจะเป็นผู้พิจารณาถึงความคุ้มค่าของสารสนเทศที่ต้องการกับต้นทุนหรือค่าใช้จ่ายที่จะได้มาซึ่งสารสนเทศนั้น

นอกจากคุณสมบัติของสารสนเทศดังกล่าวแล้ว ยังมีคุณสมบัติที่แอบแฝงของสารสนเทศอีกบางลักษณะที่สัมพันธ์กับระบบสารสนเทศ และวิธีการดำเนินงานของระบบสารสนเทศ คุณสมบัติเหล่านี้จะมีความสำคัญแตกต่างกันไปตามลักษณะงานเฉพาะอย่าง คุณสมบัติเหล่านี้ ได้แก่ ความละเอียดแม่นยำ คุณสมบัติเชิงปริมาณ ความยอมรับได้ การใช้ได้ง่าย ความไม่ลำเอียง ชัดเจน

ขั้นตอนในการประมวลผลข้อมูลเพื่อสารสนเทศ (สุโขทัยธรรมาธิราช, 2534)

ขั้นตอนในการประมวลผลข้อมูลเพื่อสารสนเทศ สามารถจัดแบ่งการปฏิบัติการแยกตามส่วนต่าง ๆ ของระบบสารสนเทศได้ดังนี้

ก. การปฏิบัติการในส่วนนำเข้า

1. การเก็บรวบรวมข้อมูล เป็นการจัดทำข้อมูล หรือเก็บรวบรวมข้อมูลให้อยู่ในรูปแบบบางอย่างสำหรับการประมวลผล

2. การตรวจสอบความถูกต้องของข้อมูล วิธีการนี้เป็นการตรวจสอบให้แน่ใจว่าข้อมูลที่เก็บรวบรวมมาหรือบันทึกไว้นั้นถูกต้อง

ข. การปฏิบัติการในส่วนประมวลผล

1. การแบ่งประเภท เป็นการจัดแบ่งประเภทข้อมูลที่ได้มาออกเป็นหมวดหมู่หรือเป็นกลุ่ม ซึ่งมีความหมายต่อผู้ใช้

2. การจัดเรียงลำดับ เป็นการจัดเรียงข้อมูลตามลำดับที่กำหนดไว้

3. การคำนวณ เป็นการคำนวณทางคณิตศาสตร์หรือทางตรรก

4. การสรุป เป็นการจัดรวมข้อมูลเข้าด้วยกัน หรือแบ่งกลุ่มข้อมูลและรวมยอดแต่ละกลุ่ม ซึ่งมีด้วยกัน 2 ลักษณะ คือ ลักษณะแรกเป็นการคำนวณทางคณิตศาสตร์ ลักษณะที่สองเป็นการสรุปตามเงื่อนไขทางตรรก

ค. การปฏิบัติการในส่วนผลลัพธ์

1. การแสดงผล เป็นวิธีการย้ายข้อมูลจากที่หนึ่งไปยังอีกที่หนึ่ง

2. การเก็บรักษาข้อมูล เป็นการเก็บข้อมูลไว้ในสื่อเก็บข้อมูลบางชนิด

3. การนำข้อมูลที่เก็บมาใช้ เป็นวิธีค้นหาข้อมูลจากสื่อเก็บข้อมูลที่เก็บข้อมูลนั้นไว้ออกมาใช้งาน

4. การคัดลอกข้อมูล เป็นการคัดลอกข้อมูลจากแฟ้มข้อมูลหนึ่งไปยังอีกแฟ้มข้อมูลหนึ่ง หรือเป็นการคัดลอกข้อมูลจากสื่อหนึ่งไปเก็บไว้ในอีกสื่อหนึ่ง

แบบจำลองข้อมูล (Data Model)

การวิเคราะห์ระบบเพื่อให้ได้แฟ้มข้อมูลต่าง ๆ ที่ต้องใช้ในระบบนั้น โครงสร้างของแฟ้มข้อมูลที่ได้ อาจไม่เหมาะสมกับการเปลี่ยนแปลงที่เกิดขึ้นในอนาคต เมื่อเกิดความต้องการใหม่ ๆ ที่เกี่ยวข้องกับข้อมูลชุดเดียวกัน จึงจำเป็นที่จะต้องมีโครงสร้างข้อมูลที่สามารถรองรับงานที่มีอยู่เดิม รวมทั้งงานที่อาจจะเกิดขึ้นในอนาคตซึ่งมีความเกี่ยวข้องกับข้อมูลชุดนั้น โครงสร้างข้อมูลนี้ก็คือแบบจำลองข้อมูล ซึ่งต้องได้รับการออกแบบเป็นอย่างดี สามารถเก็บข้อมูลต่าง ๆ ที่มีผลกระทบต่อระบบเพื่อให้ได้ข้อมูลอ้างอิงต่าง ๆ ที่ใช้ร่วมกัน (Korth and Silberschatz, 1986)

แบบจำลองข้อมูลประกอบด้วยเอนทิตี (Entity) และความสัมพันธ์ (Relationship) ระหว่างเอนทิตี ใช้เป็นเครื่องมือเพื่อช่วยอธิบายข้อมูล ความสัมพันธ์ระหว่างข้อมูล และกฎเกณฑ์ข้อบังคับต่าง ๆ ของข้อมูล ซึ่งจะช่วยอำนวยความสะดวกในการติดต่อระหว่างนักพัฒนาระบบและผู้ใช้ขั้นสุดท้าย (End User) นอกจากนั้นยังสามารถใช้แก้ปัญหาที่อาจจะเกิดขึ้นในภายหลัง เช่น การมีข้อมูลซ้ำซ้อน ความไม่ตรงกันของข้อมูลในแต่ละหน่วยงาน ความปลอดภัยของข้อมูล เป็นต้น (Korth and Silberschatz, 1986)

ก. ประเภทของแบบจำลองข้อมูล

แบ่งตามแนวความคิดที่ใช้อธิบายโครงสร้างฐานข้อมูล (Database Structure) ได้เป็น 3 ประเภท (Elmasri and Navathe, 1989) คือ

1. แบบจำลองข้อมูลระดับสูง (High-Level หรือ Conceptual Data Model)

ให้แนวความคิดที่ใกล้เคียงกับวิธีที่ผู้ใช้งานมองเห็นข้อมูล โดยไม่ต้องเกี่ยวข้องกับข้อมูลที่เก็บจริง เป็นอิสระจากระบบจัดการฐานข้อมูล (DBMS) และอุปกรณ์คอมพิวเตอร์ (Elmasri and Navathe, 1989)

2. แบบจำลองข้อมูลระดับต่ำ (Low-Level หรือ Physical Data Model)

แสดงรายละเอียดว่าข้อมูลถูกเก็บในคอมพิวเตอร์อย่างไร (Elmasri and Navathe, 1989)

3. แบบจำลองข้อมูลระดับกลาง (Implementation Data Model)

อยู่ระหว่างแบบจำลองข้อมูลระดับสูง (High-Level Data Model) และแบบจำลองข้อมูลระดับต่ำ (Low-Level Data Model) ให้แนวความคิดที่เข้าใจได้โดยผู้ใช้งาน แต่ไม่รวมถึงวิธีการเก็บข้อมูลในระบบคอมพิวเตอร์ สามารถติดตั้งบนระบบคอมพิวเตอร์ได้โดยตรง ได้แก่ แบบจำลองข้อมูลเชิงสัมพันธ์ (Relational Data Model) แบบจำลองข้อมูลเครือข่าย (Network Data Model) หรือ แบบจำลองข้อมูลลำดับชั้น (Hierachical Data Model) (Elmasri and Navathe, 1989)

ข. แบบจำลองข้อมูลความสัมพันธ์-เอนทิตี (Entity-Relationship Data Model)

เป็นแบบจำลองข้อมูลระดับสูง (Conceptual Data Model) โดยเป็นการรวบรวมข้อมูลต่าง ๆ ของระบบ แสดงในรูปของเอนทิตี (Entity) ลักษณะประจำ (Attribute) และความสัมพันธ์ระหว่างเอนทิตี โดยที่เอนทิตี (Entity) คือสิ่งต่าง ๆ ในระบบ อาจเป็นสิ่งที่มีอยู่จริงทางกายภาพ (Physical Existance) เช่น พนักงาน สินค้า หนังสือ หรือสิ่งที่มีอยู่ทางความคิด (Conceptual Existance) เช่น บริษัท รายวิชา เป็นต้น (Elmasri and Navathe, 1989)

- ลักษณะประจำ (Attribute) คือ คุณสมบัติของเอนทิตี
- ความสัมพันธ์ (Relationship) คือ ความสัมพันธ์ระหว่างเอนทิตี

แบบจำลองข้อมูลความสัมพันธ์-เอนทิตี ถูกเสนอครั้งแรกโดยเชน (Chen) เป็นเครื่องมือที่ช่วยในการติดต่อระหว่างผู้ออกแบบระบบและผู้ใช้ขั้นสุดท้าย เนื่องจากความง่ายต่อการเข้าใจและความสะดวกในการแทนข้อมูล แผนภาพที่ใช้การแสดงข้อมูลและความสัมพันธ์ระหว่างข้อมูล ทำให้เห็นภาพได้ชัดเจนและรัดกุมกว่าข้อมูลในรูปแบบข้อความ สามารถสื่อความหมายได้ง่าย โดยไม่จำเป็นต้องมีความรู้พื้นฐานในเรื่องแบบจำลองข้อมูลและการออกแบบฐานข้อมูล (Elmasri and Navathe, 1989)

ค. แบบจำลองข้อมูลเชิงสัมพันธ์ (Relational Data Model)

แบบจำลองข้อมูลเชิงสัมพันธ์ เป็นแนวความคิดที่แสดงให้เห็นถึงข้อมูล ที่ถูกมองเห็นโดยผู้ใช้งาน โดยจะแสดงรายละเอียดของข้อมูลที่ปรากฏแก่ผู้ใช้ การดำเนินการกับข้อมูลและกฎเกณฑ์ต่าง ๆ (Fleming and Halle, 1989)

1. ส่วนประกอบของแบบจำลองข้อมูลเชิงสัมพันธ์ (Fleming and Halle, 1989)

ก) โครงสร้างข้อมูล (Data Structure) เป็นข้อมูลที่มองเห็นโดยผู้ใช้งาน อยู่ในลักษณะของตารางความสัมพันธ์ (Relation)

ข) การดำเนินการกับข้อมูลในตารางความสัมพันธ์ (Data Manipulation)

ค) กฎและข้อบังคับสำหรับข้อมูลในตารางความสัมพันธ์ (Data Integrity)

2. ข้อดีของแบบจำลองข้อมูลเชิงสัมพันธ์ (Fleming and Halle, 1989)

ก) แบบจำลองข้อมูล ที่เสนอต่อผู้ใช้งานอยู่ในรูปแบบที่เข้าใจง่าย การทำงานเพื่อตอบสนองต่อความต้องการของผู้ใช้จะเกี่ยวข้องเฉพาะข้อมูล ไม่ต้องคำนึงถึงความซับซ้อนของอุปกรณ์คอมพิวเตอร์

ข) แบบจำลองข้อมูล แยกโครงสร้างหน่วยเก็บข้อมูล (Storage Structure) และวิธีเข้าถึงข้อมูล (Access Strategy) ออกจากตัวเชื่อมประสานกับผู้ใช้ (User Interface) ซึ่งแบบจำลองข้อมูลเชิงสัมพันธ์มีข้อดีในแง่นี้เหนือกว่าแบบจำลองข้อมูลอื่น ๆ

ค) แบบจำลองข้อมูลเชิงสัมพันธ์ มีพื้นฐานบนทฤษฎีทางคณิตศาสตร์ที่พัฒนามาเป็นอย่างดี และวิธีการออกแบบฐานข้อมูลโดยการทำให้เป็นบรรทัดฐาน (Normalization) ทำให้ได้แบบจำลองข้อมูลที่มีพื้นฐานดี

ระบบฐานข้อมูล (Database System)

ฐานข้อมูล (Database) คือ กลุ่มของข้อมูลที่มีความสัมพันธ์กัน ข้อมูลในที่นี้หมายถึงข้อเท็จจริง ที่สามารถบันทึกได้และมีความหมายแน่ชัด เช่น ชื่อ หมายเลขโทรศัพท์ ที่อยู่ เป็นต้น (Elmasri and Navathe, 1989)

ระบบจัดการฐานข้อมูล (Data Base Management System : DBMS) คือ กลุ่มของโปรแกรมที่ช่วยให้ผู้ใช้งานสามารถสร้างและบำรุงรักษาฐานข้อมูล (Elmasri and Navathe, 1989)

ระบบฐานข้อมูล (Database System) คือ การประกอบกันของฐานข้อมูลและโปรแกรม (Elmasri and Navathe, 1989)

ก. ข้อดีของฐานข้อมูล (Date, 1986)

1. ช่วยลดความซ้ำซ้อนของข้อมูล
2. หลีกเลี่ยงการเกิดความไม่ต้องกัน (Inconsistency) ของข้อมูล
3. สามารถใช้ข้อมูลร่วมกันได้
4. ทำให้เกิดมาตรฐานของข้อมูล
5. มีการกำหนดมาตรการต่าง ๆ เพื่อความปลอดภัยของข้อมูล

ข. สถาปัตยกรรมของระบบจัดการฐานข้อมูล (Elmasri and Navathe, 1989)

1. ระดับภายใน (Internal Level) ใช้เค้าร่างภายใน (Internal Schema) อธิบายโครงสร้างหน่วยเก็บข้อมูล โดยใช้แบบจำลองข้อมูลกายภาพ (Physical Data Model)

2. ระดับแนวคิด (Conceptual Level) ใช้เค้าร่างเชิงแนวคิด (Conceptual Schema) อธิบายโครงสร้างของฐานข้อมูลทั้งหมดแต่จะไม่แสดงรายละเอียดของโครงสร้างหน่วยเก็บข้อมูล

3. ระดับภายนอก (External หรือ View Level) ประกอบด้วยเค้าร่างภายนอก (External Schema) หรือ ภาพระดับผู้ใช้งาน (User View) ของระบบทั้งหมด โดยที่แต่ละภาพระดับผู้ใช้งาน (User View) จะเป็นภาพที่อธิบายบางส่วนของฐานข้อมูลซึ่งผู้ใช้งานกลุ่มใดกลุ่มหนึ่งสนใจ โดยจะไม่แสดงส่วนอื่น ๆ ที่ไม่เกี่ยวข้องกับผู้ใช้กลุ่มนั้น

ผู้ใช้งานจะอ้างถึงเฉพาะเค้าร่างภายนอก (External Schema) ของตนเอง ระบบจัดการฐานข้อมูลจะทำการเปลี่ยนความต้องการที่ระบุในเค้าร่างภายนอก (External Schema) ให้อยู่ในรูปของเค้าร่างเชิงแนวคิด (Conceptual Schema) จากนั้นจึงเปลี่ยนเป็นเค้าร่างภายใน (Internal Schema) เพื่อประมวลผลข้อมูลที่เก็บอยู่จริง

ค. ความไม่พึ่งพิงของข้อมูล (Data Independence)

ความไม่พึ่งพิงของข้อมูล (Data Independence) คือ ความสามารถในการเปลี่ยนแปลงเค้าร่าง (Schema) ที่ระดับหนึ่ง โดยไม่มีผลกระทบกับเค้าร่าง (Schema) ในระดับที่สูงกว่า มี 2 ประเภท (Elmasri and Navathe, 1989) คือ

1. ความไม่พึ่งพิงในเชิงตรรก (Logical Data Independence) การเปลี่ยนเค้าร่างเชิงแนวคิด (Conceptual Schema) ทำได้โดยที่เค้าร่างภายนอก (External Schema หรือ Application Program) ไม่ต้องเปลี่ยนตาม (Elmasri and Navathe, 1989)

2. ความไม่พึ่งพิงในเชิงกายภาพ (Physical Data Independence) การเปลี่ยนแปลงเค้าร่างภายใน (Internal Schema) สามารถทำได้โดยเค้าร่างเชิงแนวคิด (Conceptual Schema) ไม่ต้องเปลี่ยนตาม (Elmasri and Navathe, 1989)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

<u>ฐานข้อมูลเชิงสัมพันธ์</u> (Relational Database)

ก. โครงสร้างของฐานข้อมูลเชิงสัมพันธ์

ฐานข้อมูลเชิงสัมพันธ์ (Relational Database) จะแทนข้อมูลในฐานข้อมูลในลักษณะของตารางความสัมพันธ์ (Relation) โดยที่แต่ละตารางความสัมพันธ์ (Relation) จะเสมือนเป็น 1 ตาราง และแถวต่าง ๆ ในตารางจะแสดงค่าของข้อมูลที่มีความสัมพันธ์กัน ชื่อของตาราง

(Table) และ ชื่อของสดมภ์ (Column) จะใช้แปลความหมายของค่าในแต่ละแถวของตารางและได้กำหนดคำเพื่อใช้อธิบายตารางดังนี้ (Elmasri and Navathe, 1989)

ตารางความสัมพันธ์ (Relation) หมายถึง ตาราง (Table)

ทูเพิล (Tuple) หมายถึง แถว (Row หรือ Record)

ลักษณะประจำ (Attribute) หมายถึง สดมภ์ (Column หรือ Field)

โดเมน (Domain) หมายถึง ค่าที่เป็นไปได้ของข้อมูลในแต่ละสดมภ์

ข. ทฤษฎีการทำให้เป็นบรรทัดฐาน (Normalization)

การทำให้เป็นบรรทัดฐาน (Normalization) เป็นทฤษฎีที่ใช้ทำการวิเคราะห์และแยกส่วนโครงสร้างของข้อมูลออกไปเป็นชุด ๆ ของความสัมพันธ์ ในทางปฏิบัติเป็นเทคนิคซึ่งประกอบด้วยหลายขั้นตอนซึ่งช่วยให้เกิดประโยชน์ต่าง ๆ เช่น ช่วยลดขนาดที่ว่างที่ใช้เก็บข้อมูล ช่วยลดความขัดแย้งกันภายในฐานข้อมูล ช่วยลดปัญหาที่เกิดขึ้นเมื่อทำการลบหรือปรับปรุงข้อมูล ทำให้โครงสร้างข้อมูลมีเสถียรภาพสูงสุด (Date, 1986)

1. รูปแบบบรรทัดฐานระดับที่ 1 (First Normal Form, 1NF)

ทุกสมาชิกในเอนทิตี (Entity) จะมีค่าลักษณะประจำ (Attribute) หนึ่ง ๆ ได้เพียงค่าเดียว หรือกล่าวคือจะมีกลุ่มของค่าลักษณะประจำ (Attribute) ในสมาชิกหนึ่งไม่ได้ (Date, 1986)

2. รูปแบบบรรทัดฐานระดับที่ 2 (Second Normal Form, 2NF)

ความสัมพันธ์จะมีคุณสมบัติเป็นรูปแบบบรรทัดฐานระดับที่ 2 เมื่อมีคุณสมบัติของรูปแบบบรรทัดฐานระดับที่ 1 และลักษณะประจำที่ไม่ใช่กุญแจ (Non-Key Attribute) ทุก ๆ ลักษณะประจำ (Attribute) ต้องขึ้นตรงกับกุญแจหลัก (Primary Key) ทั้งชุด (Date, 1986)

3. รูปแบบบรรทัดฐานระดับที่ 3 (Third Normal Form, 3NF)

ความสัมพันธ์จะมีคุณสมบัติเป็นรูปแบบบรรทัดฐานระดับที่ 3 เมื่อมีคุณสมบัติของรูปแบบบรรทัดฐานระดับที่ 2 และลักษณะประจำที่ไม่ใช่กุญแจ (Non-Key Attribute) ใด ๆ จะขึ้นกับลักษณะประจำ (Attribute) อื่น ๆ ที่ไม่ใช่กุญแจหลักหรือกลุ่มกุญแจหลัก (Primary Key) ไม่ได้ (Date, 1986)