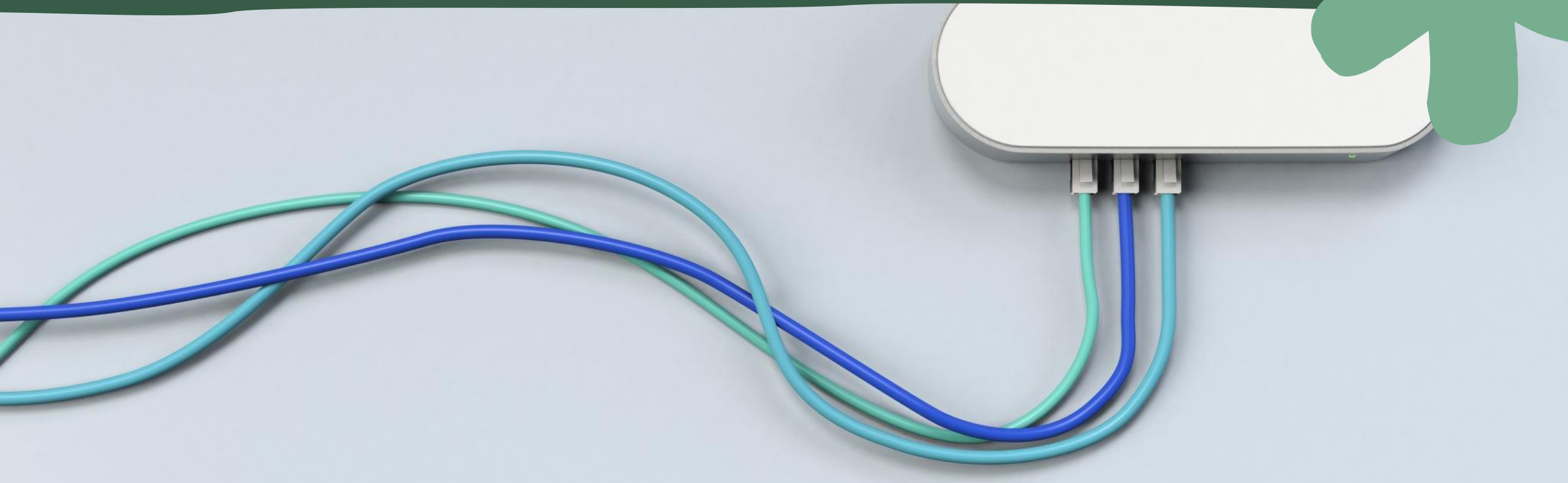
11. ฐานข้อมูลเบื้องต้น

01418111: วิทยาการคอมพิวเตอร์เบื้องต้น (INTRODUCTION OF COMPUTER SCIENCE)



อ.พิชยภรณ์ พงศกรรังศิลป์

วัตถุประสงค์

- เข้าใจ DBMS และ อธิบายองค์ประกอบของ DBMS ได้
- เข้าใจสถาปัตยกรรมของ DBMS และระดับของ DBMS
- เข้าใจแนวคิดของการดำเนินการของฐานข้อมูลสัมพันธ์
- ใช้ Structured Query Language (SQL) เพื่อนิยามความสัมพันธ์ได้

Database

ฐานข้อมูล คือ ฐานข้อมูลเป็นพื้นที่จัดเก็บข้อมูลที่ถูกออกแบบมาเพื่อการจัดเก็บข้อมูลอย่างมีประสิทธิภาพ สามารถให้ผู้ใช้จัดการข้อมูล การเรียกใช้ และการเข้าถึงข้อมูลในลักษณะต่างๆ และรวมทั้งการบำรุงรักษาข้อมูลได้ ฐานข้อมูลมีหลายประเภทตามความ เหมาะสมสำหรับความต้องการของแต่ละอุตสาหกรรม

Database (ต่อ)

ฐานข้อมูลมีต้นกำเนิดมาจากการค้นคว้าวิจัยในเชิงของวิทยาการคอมพิวเตอร์ เป็นพื้นที่จัดเก็บข้อมูลที่ถูกออกแบบมาเพื่อการจัดเก็บข้อมูลอย่างมีประสิทธิภาพ สามารถให้ผู้ใช้จัดการข้อมูล การเรียกใช้ และการเข้าถึงข้อมูลในลักษณะต่างๆ รวมทั้งบำรุงรักษาข้อมูลได้ ฐานข้อมูลมีหลายประเภทตามความเหมาะสมสำหรับความต้องการของแต่ละอุตสาหกรรม มีความสามารถในการจัดเก็บข้อมูลประเภทแบบไบนารีไฟล์ เอกสาร รูปภาพ วีดีโอ ข้อมูลเชิงสัมพันธ์ ข้อมูลเชิงมิติ ข้อมูลทรานแซคชัน ข้อมูลเชิงวิเคราะห์ ข้อมูลภูมิศาสตร์และข้อมูลประเภทอื่นๆ

Database (ต่อ)

ข้อมูลสามารถจัดเก็บได้หลากหลายรูปแบบได้แก่ แบบตาราง แบบลำดับชั้น และแบบกราฟ กรณีที่ข้อมูลถูกจัดเก็บในรูปแบบของตารางจะถูกเรียกว่า **ฐานข้อมูลเชิงสัมพันธ์ (Relational database)** เมื่อเป็นการจัดเก็บข้อมูลแบบโครงฐานแบบทรี (Tree) จะถูกเรียกว่า ฐานข้อมูลแบบลำดับชั้น (Hierarchical database) ข้อมูลถูกจัดเก็บในรูปแบบของกราฟความสัมพันธ์ระหว่างออบเจ็กต์จะหมายถึง ฐานข้อมูลแบบเครือ ข่าย (Network database)

Database (ต่อ)

ระบบจัดการฐานข้อมูล (DBMS) เป็นเครื่องมือที่ทำให้สามารถจัดเก็บข้อมูล และอำนวยความสะดวกในการ เข้าถึง และการเรียกใช้ข้อมูล โดยข้อมูลที่จัดเก็บในฐานข้อมูลจะต้องมีความสอดคล้องกันของข้อมูล ซึ่งปกติระบบฐานข้อมูลจะมีความสามารถในการรองรับผู้ใช้งานหลายคนเข้ามาทำงานพร้อมกันได้ โดยระบบฐานข้อมูลจะต้องมีระบบการป้องกัน เช่น การเพิ่ม การปรับปรุง และการลบข้อมูลตัวเดียวกันได้โดยไม่มีผลกระทบต่อผู้ใช้งานอื่น หมายความว่าผู้ใช้งานแต่ละคนจะไม่ทำให้เกิดความไม่สอดคล้องกันของข้อมูล ข้อมูลจะไม่สูญหายโดยไม่ตั้งใจระหว่างการดำเนินการ นอกจากนี้จำเป็นต้องมีเครื่องมือที่ใช้ติดต่อกับระบบฐานข้อมูล (Interface) แบบมาตรฐานสำหรับการเข้าถึงข้อมูล เครื่องมือสำหรับการสำรองข้อมูล การคืนกลับข้อมูล และกู้คืน และระบบฐานข้อมูลที่ดีควรมีวิธีการจัดการกับสิ่งที่นอกเหนือจากความคาดหมายอื่นๆ เช่น การรองรับการทำงานกับข้อมูลจำนวนมหาศาล และผู้ใช้จำนวนมาก

Database (ต่อ)

ระบบจัดการฐานข้อมูล (DBMS) ทำหน้าที่ในการกำหนด สร้าง และดูแลรักษาฐานข้อมูล นอกจากนี้ระบบจัดการฐานข้อมูลยังอำนวยความสะดวกให้ผู้ใช้สามารถควบคุมการใช้ข้อมูลในฐานข้อมูลได้อีกด้วย ระบบจัดการฐานข้อมูลประกอบด้วยองค์ประกอบหลัก 5 ส่วนคือ

1. ฮาร์ดแวร์ (hardware)
2. ซอฟต์แวร์ (software)
3. ข้อมูล (data)
4. ผู้ใช้ (users)
5. กระบวนการ (procedures)

Database (ต่อ)

ผู้ผลิตซอฟต์แวร์ระบบฐานข้อมูลส่วนใหญ่ได้มีการพัฒนาระบบการจัดการฐานข้อมูลในลักษณะที่เป็นฐานข้อมูลเชิงสัมพันธ์ (Relational DataBase Management System: RDBMS) ซึ่งปัจจุบันได้กลายมา เป็นหลักสำคัญในการรองรับข้อมูลขององค์กรต่างๆ และทำงานร่วมกันกับแอพพลิเคชั่นในหลากหลายอุตสาหกรรม เช่น งานธนาคาร งานขนส่ง งานสาธารณสุขและอื่นๆ อีกทั้งแอพพลิเคชั่นที่ทำงานบนเว็บที่มีปริมาณเพิ่มขึ้นทั้งในด้านของจำนวนผู้ใช้และขอบเขตของการใช้งาน เช่น ในกรณีของงานธุรกรรมอิเล็กทรอนิกส์เพื่อรองรับข้อมูลของการทำธุรกิจและการค้าแบบออนไลน์

Database (ต่อ)

1. **ฮาร์ดแวร์ (hardware)** เป็นอุปกรณ์ที่ใช้สำหรับจัดเก็บข้อมูลประกอบด้วย 2 ส่วนหลักๆ คือ หน่วยความจำสำรอง (Storage) ที่ใช้เก็บข้อมูล และหน่วยความจำหลัก (Main Memory) เป็นส่วนที่ใช้ในการประมวลผล โดยเตรียมข้อมูลที่ต้องใช้ในการประมวลผลมารอ เพื่อให้การประมวลผลข้อมูลมีความเร็ว
2. **ซอฟต์แวร์ (Software)** ซึ่งโปรแกรมที่ใช้ติดต่อกับฐานข้อมูลเรียกว่า DBMS (Database Management System) มีหน้าที่ในการจัดการและควบคุมความถูกต้อง ความสัมพันธ์ และความซ้ำซ้อนในฐานข้อมูล

Database (ต่อ)

3. **ข้อมูล (Data)** ข้อมูลถือได้ว่าเป็นทรัพย์สินที่สำคัญส่วนหนึ่งของธุรกิจ ที่ใช้สำหรับการบริหารและการจัดการภายใน องค์กร โดยข้อมูลอาจจะมีการรวบรวมจากแหล่งข้อมูลต่างๆ เช่น การจัดเก็บและรวบรวมข้อมูลของผู้บริโภคที่มีการใช้บัตรเครดิตเพื่อหาแบบแผนการใช้บัตรเครดิตของลูกค้า ข้อมูลขององค์กรจึงมีความสำคัญมาก และต้องการความ เสถียรภาพ ความปลอดภัย อีกทั้งยังต้องการโปรแกรมที่มีความสามารถในการเรียกใช้งานข้อมูล จัดเก็บและประมวลผลข้อมูลได้อย่างรวดเร็ว คำตอบของความต้องการเหล่านี้คือฐานข้อมูลที่มีความเสถียรภาพและเชื่อถือได้

Database (ต่อ)

4. **ผู้ใช้ (User)** คำว่า “ผู้ใช้” ตามนัยของ DBMS จะมีความหมายที่กว้าง เราสามารถแบ่งผู้ใช้ออกได้เป็น 2 กลุ่มหลักคือผู้ใช้ที่เป็นคนหรือที่เรียกว่า end users กับผู้ใช้ที่เป็นโปรแกรมประยุกต์หรือที่เรียกว่า application program

   - **End Users** คือ คนที่สามารถเข้าถึงข้อมูลและนำข้อมูลออกมาใช้ แบ่งออกเป็น 2 กลุ่ม คือ ผู้บริหารฐานข้อมูล (Database Administrator : DBA) กับ ผู้ใช้ทั่วไป โดย DBA มีสิทธิ์สูงสุดในการเข้าถึงและควบคุมฐานข้อมูล เป็นผู้กำหนดสิทธิ์ในการใช้ฐานข้อมูลทั้งหมด
   - **Application Programs** เป็นโปรแกรมที่ใช้ในการเข้าถึงและประมวลผลข้อมูลในฐานข้อมูล เช่นโปรแกรมประมวลผลเงินเดือน โปรแกรมคิดคะแนนเฉลี่ยของนักศึกษาในแต่ละภาคการศึกษา

Database (ต่อ)

5. **กระบวนการ (procedures)** กระบวนการในการจัดการฐานข้อมูล โดยปกติจะทำผ่านซอฟต์แวร์ โดย DBA สามารถสร้าง นำเข้า ปรับปรุง ลบ ข้อมูลได้ โดยอาศัยหลักการการบริหารจัดการฐานข้อมูลในรูปแบบต่างๆ

วัฒนาการของฐานข้อมูล

ในปี ค.ศ. 1960 ระบบแบบเครือข่ายและระบบแบบลำดับชั้นเป็นเทคโนโลยีที่ล้ำสมัยในขณะนั้นถูกนำมาใช้ทำธุรกรรมแบบอัตโนมัติของธนาคาร ด้านบัญชี ด้านการสั่งซื้อ และมีระบบการประมวลผลโดยใช้คอมพิวเตอร์แม่ข่ายที่เรียกว่า mainframe ระบบที่ใช้งานดังกล่าวเป็นพื้นฐานของระบบเบื้องต้น

ในปี 1970 มีแนวความคิดใหม่ตามเอกสารที่ชื่อว่า “A relational model of data for large shared data banks” โดยช่วยให้การจัดการฐานข้อมูลง่ายขึ้น เป็นงระบบการจัดการฐานข้อมูลด้วยแบบจำลองเชิงสัมพันธ์ที่ให้ความเป็นอิสระของข้อมูล ปัจจุบันระบบจัดการฐานข้อมูลเชิงสัมพันธ์เป็นระบบที่ถูกใช้มากที่สุดในการจัดการฐานข้อมูล เช่น Oracle, Microsoft SQL Server

วัฒนาการของฐานข้อมูล

วิวัฒนาการต่อมาของระบบจัดการฐานข้อมูลคือ การขยายความสามารถ ภาษา SQL ที่ประดิษฐ์โดยไอบีเอ็ม ตั้งแต่ต้นปี ค.ศ. 1970 ได้มีการปรับปรุงอย่างต่อเนื่อง ถึงแม้ภาษา SQL จะมีประสิทธิภาพสูงและผู้ใช้สามารถสร้างภาษา SQL เพิ่มเติมได้เองแล้ว

การพัฒนาในลำดับต่อไปเป็นการเน้นเรื่องของการรวบรวมข้อมูล (Data Integration) ซึ่งทุกวันนี้หลายองค์กรจำเป็นต้องมีการแลกเปลี่ยนข้อมูลซึ่งกันและกัน ซึ่ง eXtensible Markup Language (XML) เป็นเทคโนโลยีหนึ่งที่ได้รับความสนใจเพื่อตอบวัตถุประสงค์ดังกล่าว ปัจจุบันเทคโนโลยีที่กำลังได้รับความสนใจคือ Cloud Computing

Information Models

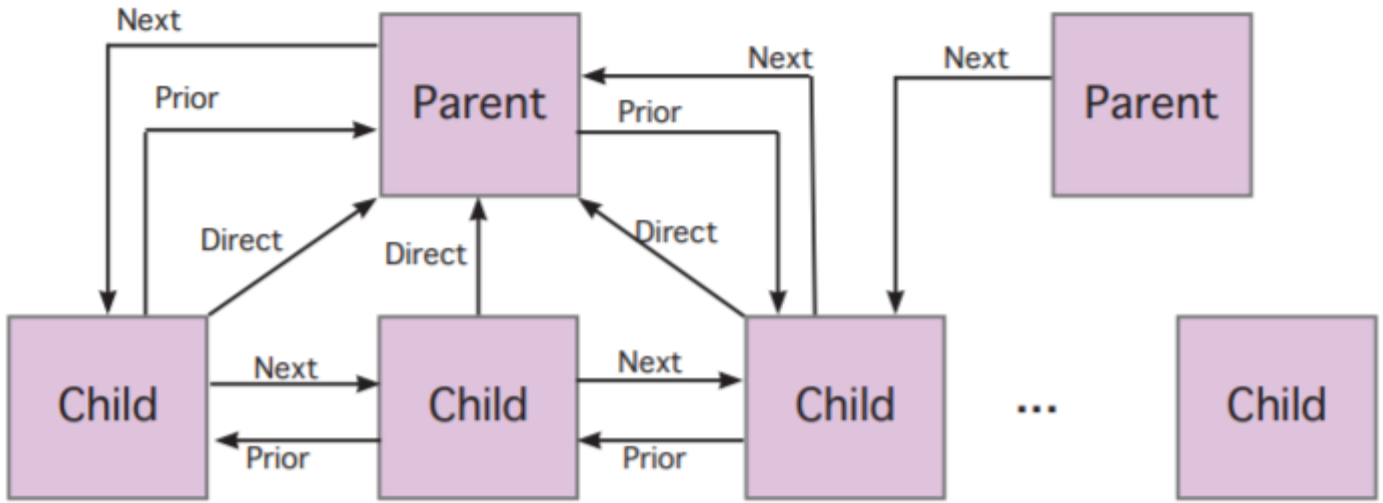
รูปแบบข้อมูล (Information Models) เป็นการนำเสนอแบบจำลองโครงสร้างของข้อมูล โดยมีการนำเสนอเป็นลักษณะเอนทิตี (entity) ที่มีคุณสมบัติ (properties) ของตัวเอง มีความสัมพันธ์ (relationship) ระหว่างเอนทิตี และสิ่งที่สามารถกระทำ ได้กับเอนทิตี โดยเอนทิตีจะถูกสร้างเป็น**แบบจำลอง**ที่มาจากความต้องการของธุรกิจ และปัญหาที่เกิดในโลกของความเป็นจริง (Real-world problem) เช่น เอนทิตีของระบบการเรียกเก็บเงิน แรงจูงใจที่อยู่เบื้องหลังเป็นแนวคิดที่ต้องการสร้างรูปแบบเพื่ออธิบายปัญหาให้เกิดความกระจ่างและสามารถนำไปพัฒนาได้จริงโดยระบบซอฟต์แวร์ การสร้างรูปแบบดังกล่าวนั้นเป็นการแปลงรูปแบบข้อมูลที่เรียกว่าการทำแบบจำลองข้อมูล ซึ่งอาจจะใช้แบบจำลองเชิงวัตถุ (Unified Modeling Language - UML) แบบจำลองอีอาร์ (Entity Relationship Models : E-R model) หรือ XML schemas

Information Models (ต่อ)

การสร้าง**แบบจำลองข้อมูล**นั้นมีความสำคัญโดยต้องคำนึงถึงความยืดหยุ่นที่ต้องรองรับการเปลี่ยนแปลงในอนาคตและไม่ให้เกิดผลกระทบที่จะเกิดขึ้นต่อการทำงาน แบบจำลองที่สร้างขึ้นต้องเข้ากันได้และต้องรองรับส่วนต่อขยายในอนาคต รูปแบบข้อมูล และแบบจำลองข้อมูลนั้นมีความต่างกันเนื่องจากตอบวัตถุประสงค์คนละอย่างกัน จุดประสงค์หลักของรูปแบบข้อมูล นั้นก็คือการจัดการแบบจำลองในระดับแนวคิด (Conceptual level) ไม่ยึดติดกับวิธีการพัฒนาหรือโปรโตคอลในการส่งข้อมูล ระดับของรายละเอียดที่อธิบาย รูปแบบข้อมูลนั้นจะขึ้นอยู่กับความต้องการของนักออกแบบ

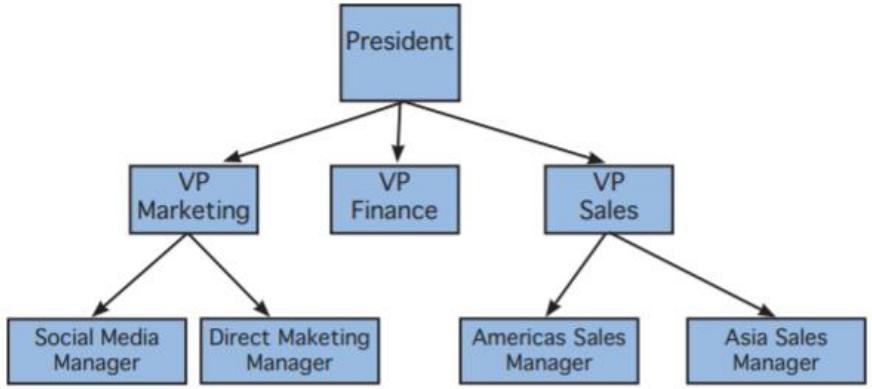
Data Models

1. **แบบจำลองเครือข่าย (Network model)** ใช้โครงสร้างแผนภูมิ เรคอร์ดเหล่านั้นสามารถระบุได้ด้วยคีย์ กลุ่มของชนิดของเรคอร์ดจะประกอบกันเป็นเครือข่ายหรือ ฐานข้อมูล โดยที่ตัวลูก (child) หนึ่งตัวสามารถมาจากตัวแม่ (parent) หลายตัวและเชื่อมต่อกันด้วยพอยน์เตอร์ไปยังตัวถัดไปและเชื่อมไปยังตัวก่อนหน้าโดยตรง



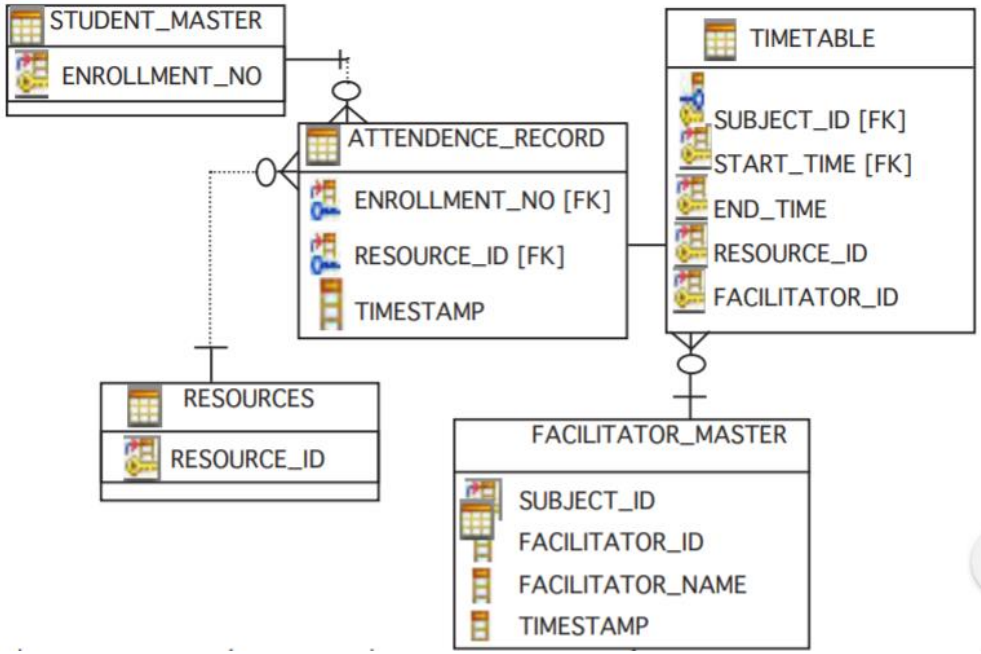
Data Models (ต่อ)

2. **แบบจำลองลำดับ (Hierarchical Model)** ชั้นมีโครงสร้างการจัดการโดยโครงสร้างแบบต้นไม้ รากของต้นไม้ (root) จะทำหน้าที่เป็นโหนดแม่ ซึ่งจะประกอบด้วยโหนดลูก ในแบบจำลองลำดับชั้นจะจัดเก็บชื่อฟิลด์ซึ่งเกี่ยวข้องกับชนิดข้อมูลเรียกว่าชนิดเรคอร์ด ชนิดของเรคอร์ดจะต้องอธิบายข้อมูลที่จัดเก็บข้างในว่าเป็นข้อมูลอะไร บางฟิลด์ในชนิดเรคอร์ดถูกใช้เป็นคีย์ ระบบการจัดการฐานข้อมูลแบบลำ ดับชั้นระบบแรกคือ IMS (Information Management System) นำเสนอโดยไอบีเอ็ม ในปี ค.ศ. 1968



Data Models (ต่อ)

3. **แบบจำลองเชิงสัมพันธ์ (Relational model)** แบบจำลองเชิงสัมพันธ์เป็นแบบจำลองที่เข้าใจง่ายและมีรูปแบบที่เป็นระบบ มีรากฐานมาจากคณิตศาสตร์บนทฤษฎีและการพิสูจน์แคลคูลัส แบบจำลองเชิงสัมพันธ์เป็นแบบจำลองที่ถูกใช้มากที่สุดในฐานข้อมูลปัจจุบัน ด้วยโครงสร้างอย่างง่าย ทำให้ข้อมูลเชิงตรรกะมีความเป็นอิสระ

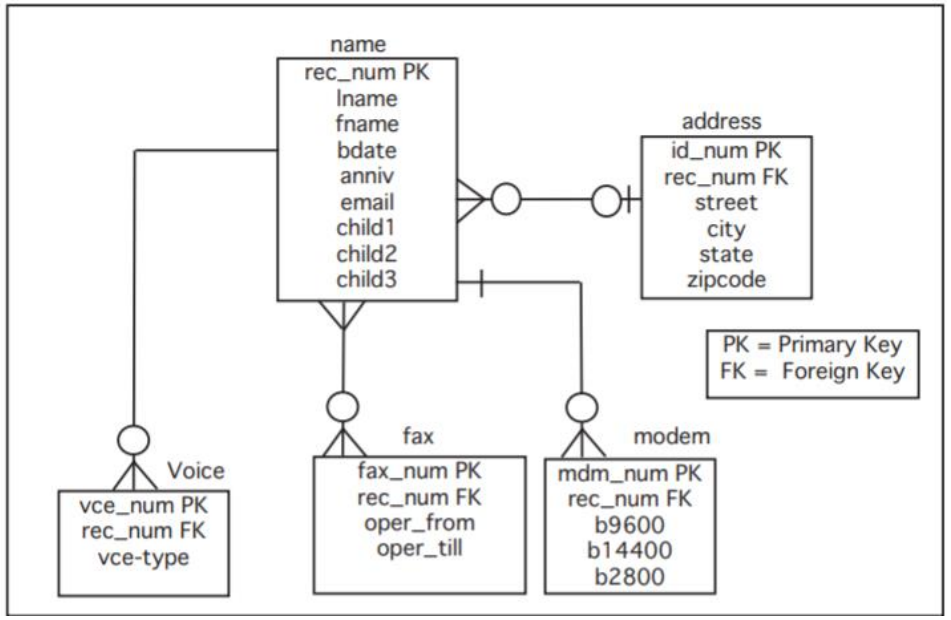


Data Models (ต่อ)

4. **แบบจำลองอีอาร์ (Entity-Relationship diagram)** ในกลางทศวรรษที่ 1970 Peter Chen ได้นำเสนอแบบจำลองอีอาร์ซึ่งเป็นอีกทางเลือกหนึ่งในการทำรูปแบบความสัมพันธ์ ได้เสนอแนวคิดของฐานข้อมูลที่เป็นกลุ่มของอินสแตนซ์ของเอนทิตี เอนทิตีเป็นออบเจ็กต์ที่เป็นอิสระจากเอนทิตีอื่นในฐานข้อมูล เอนทิตีมีแอตทริบิวต์ซึ่งแอตทริบิวต์แสดงถึงคุณสมบัติหรือลักษณะของเอนทิตี แอตทริบิวต์หนึ่งตัวหรือหลายตัวถูกกำหนดให้เป็นคีย์ สุดท้ายเอนทิตีสามารถมีความสัมพันธ์ระหว่างกันเป็นแบบ หนึ่งต่อหนึ่ง (1 to 1) หนึ่งต่อกลุ่ม (1 to many) กลุ่มต่อกลุ่ม (many to many)

Data Models (ต่อ)

แบบจำลองอีอาร์ (Entity-Relationship diagram) เอนทิตีแทนด้วยรูปสี่เหลี่ยม ในภาพก็คือเอนทิตี name, address, voice, fax, และ modem ภายในเอนทิตีจะประกอบด้วยคุณสมบัติเรียกว่าแอตทริบิวต์ ตัวอย่าง เอนทิตี voice ประกอบด้วย แอตทริบิวต์ vce_num rec_num และ vce-type PK เป็นสัญลักษณ์แทนคีย์หลัก FK เป็นสัญลักษณ์แทนคีย์นอก



Data Models (ต่อ)

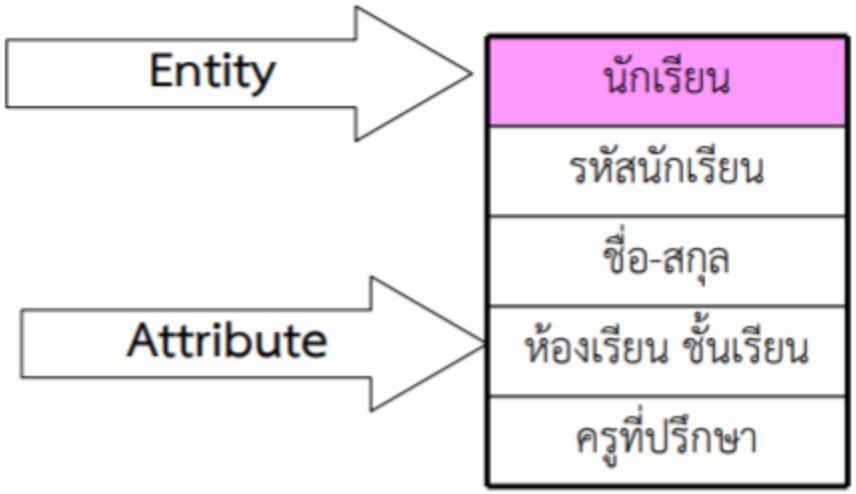
5. **แบบจำลองความสัมพันธ์เชิงวัตถุ (Object-relational model)** แบบจำลองความสัมพันธ์เชิงวัตถุมีความคล้ายคลึงกับแบบจำลองเชิงสัมพันธ์ แต่จะมองเอนทิตีเป็นเชิงวัตถุและมีความสัมพันธ์ที่เป็นแบบคุณสมบัติการสืบทอด คุณลักษณะและข้อดีของแบบจำลองความสัมพันธ์แบบวัตถุคือ
   - ผู้ใช้สามารถสร้างชนิดของข้อมูลขึ้นมาเอง
   - สนับสนุนคุณสมบัติการสืบทอดของออบเจ็กต์
   - สนับสนุนส่วนขยายของออบเจ็กต์ เช่น สนับสนุนการจัดเก็บข้อมูลแผนที่

แนวคิดของแบบจำลองข้อมูลเชิงสัมพันธ์

การสร้างแบบจำลองข้อมูลเชิงสัมพันธ์มีองค์ประกอบที่สำคัญหลายส่วนด้วยกันเช่น แอตทริบิวต์ รีเลชัน candidate key, primary key, foreign key

แนวคิดของแบบจำลองข้อมูลเชิงสัมพันธ์ (ต่อ)

- **Entity** คือ ชื่อของสิ่งใดสิ่งหนึ่งที่ต้องการจัดเก็บข้อมูลไว้ เช่น Entity นักเรียน
- **Attributes** คือ รายละเอียดของข้อมูลที่แสดงรายละเอียดและคุณสมบัติของ Entity หนึ่งๆ
- **แถว (row)** ในตารางข้อมูล หรือ เรคอร์ดในแฟ้มข้อมูล



แนวคิดของแบบจำลองข้อมูลเชิงสัมพันธ์ (ต่อ)

- **รีเลชัน (Relations)** เป็นส่วนหลักของฐานข้อมูลเชิงสัมพันธ์ ประกอบด้วยส่วนหัว (heading) และส่วนตัวตาราง (body) หรือรายละเอียดที่เก็บข้อมูลในรีเลชัน ส่วนหัวจะประกอบด้วยเซตของแอตทริบิวต์ สามารถสรุปคุณสมบัติของรีเลชันได้ 4 ข้อดังนี้
  - ข้อมูล (Record/Tuple) ต้องไม่ซ้ำกันในรีเลชัน
  - แต่ละข้อมูล (Record/Tuple) ไม่มีการเรียงลำดับ (บนลงล่าง)
  - แอตทริบิวต์ไม่มีการเรียงลำดับ (ซ้ายไปขวา)
  - ค่าแอตทริบิวต์ทั้งหมดเป็นค่า Atomic นั่นคือ ค่าของข้อมูลที่ปรากฏอยู่ในตารางจะเป็นค่า ๆ เดียว เป็นลิสต์ของค่าหลาย ๆ ค่าไม่ได้

แนวคิดของแบบจำลองข้อมูลเชิงสัมพันธ์ (ต่อ)

- **โครงร่าง (Schemas)** โครงร่างของฐานข้อมูลเป็นรูปแบบมาตรฐานที่อธิบายถึงรายละเอียดของฐานข้อมูลเชิงสัมพันธ์และความสัมพันธ์ระหว่างรีเลชันที่อยู่ภายใน
- **คีย์ (Keys)** แบบจำลองข้อมูลเชิงสัมพันธ์ใช้คีย์เพื่อบ่งบอกถึงความแตกต่างระหว่างทูเปิลของรีเลชัน คีย์ถูกใช้สำหรับบังคับใช้เป็นกฎ และ/หรือ การควบคุมข้อมูลในฐานข้อมูล การควบคุมข้อมูลดังกล่าวใช้จัดการความสอดคล้องกันและความถูกต้องของข้อมูล ระบบจัดการฐานข้อมูล (DBMS) อนุญาตให้มีการใช้คีย์และเป็นส่วนที่ระบบจัดการฐานข้อมูลเชิงสัมพันธ์ใช้ในการตรวจสอบและรักษาความสอดคล้องกันและความถูกต้องของข้อมูลในฐานข้อมูล

แนวคิดของแบบจำลองข้อมูลเชิงสัมพันธ์ (ต่อ)

- **คีย์ (Keys)** สามารถระบุคีย์ตามประเภทต่างๆ ดังนี้
  - Primary Key ใช้แสดงความเป็นหนึ่งเดียวของแต่ละทูเปิลในรีเลชัน เป็นคีย์ที่กำหนดขึ้น โดยข้อมูลต้องไม่ซ้ำและไม่เป็นค่าว่าง
  - Candidate Keys คือสิ่งที่สามารถใช้ระบุทูเปิลที่ไม่ซ้ำ กันในรีเลชัน ตามคำจำกัดความ รีเลชันหนึ่งมี Candidate Keys อย่างน้อย 1 ตัว
  - Foreign keys เป็นคีย์ที่ใช้เชื่อมตารางเข้าด้วยกัน โดยความสัมพันธ์ของ Foreign-to-primary-key การอ้างอิงรีเลชันหนึ่งไปยังอีกรีเลชันหนึ่ง เหมือนเป็นกาวที่ยึดโยงรีเลชันต่างๆ ไว้ด้วยกันในฐานข้อมูล

ซอฟต์แวร์ระบบจัดการฐานข้อมูล

ซอฟต์แวร์ที่ใช้ในการสร้าง จัดเก็บ เข้าถึง ค้นคืน จัดการแฟ้มต่างๆ และบำรุงรักษาฐานข้อมูล รวมถึงช่วยจัดทำแบบฟอร์มนำเข้าข้อมูล และออกรายงานที่เป็นสิ่งออก ระบบจัดการฐานข้อมูลเป็นส่วนต่อประสานระหว่างฐานข้อมูล ผู้ใช้ และระบบประยุกต์ต่างๆ ทำให้การใช้ฐานข้อมูลมีประสิทธิภาพมากขึ้น ซึ่งผู้ใช้สามารถใช้ข้อมูลในฐานข้อมูลผ่านระบบจัดการฐานข้อมูลได้โดยตรงผ่านโปรแกรมประยุกต์ส่วนหน้า (front-end application) หรือใช้ข้อมูลในฐานข้อมูลผ่านระบบประยุกต์ที่เรียกใช้ระบบจัดการฐานข้อมูลอีกทอดหนึ่งที่เรียกว่า โปรแกรมประยุกต์ส่วนหลัง (back-end application) ก็ได้ เช่น โปรแกรมประยุกต์บนเว็บที่มีการเรียกใช้ระบบจัดการฐานข้อมูลอยู่เบื้องหลัง

ซอฟต์แวร์ระบบจัดการฐานข้อมูล

ส่วนประกอบของซอฟต์แวร์ระบบจัดการฐานข้อมูลที่สำคัญ ได้แก่

- ส่วนเครื่องประมวลผลฐานข้อมูล (database engine)
- ส่วนการนิยามข้อมูล (data definition)
- ส่วนการจัดการข้อมูล (data manipulation)
- ส่วนการสร้างระบบประยุกต์ (application generation)
- ส่วนการบริหารข้อมูล (data administration)

ซอฟต์แวร์ระบบจัดการฐานข้อมูล

- **ส่วนเครื่องประมวลผลฐานข้อมูล (database engine)**

รับผิดชอบการจัดเก็บข้อมูล การจัดการ และการค้นคืนข้อมูล โต้ตอบกับส่วนประกอบอื่นๆ ของระบบจัดการฐานข้อมูล เพื่อแปลงคำขอเชิงตรรกะจากผู้ใช้ ให้เป็นการดำเนินการทางกายภาพทำให้ผู้ใช้ไม่จำเป็นต้องทราบว่า ข้อมูลมีโครงสร้างทางกายภาพหรือถูกเก็บอย่างไรบนสื่อหน่วยเก็บ หรือกระบวนการใดที่ระบบจัดการฐานข้อมูลใช้ เพื่อจัดการกับข้อมูล แต่ผู้ใช้สามารถอัปเดตข้อมูลด้วยการเพิ่ม ลบ และแก้ไขข้อมูลได้ง่าย

ซอฟต์แวร์ระบบจัดการฐานข้อมูล

- **ส่วนการนิยามข้อมูล (data definition)**

ส่วนการนิยามข้อมูลใช้ในการสร้างและจัดการพจนานุกรมข้อมูล และกำหนดหรือเปลี่ยนแปลงโครงสร้างฐานข้อมูล พจนานุกรมข้อมูล (data dictionary) จะรวมคำอธิบายโดยละเอียดเกี่ยวกับตารางและนิยามเกี่ยวกับฟิลด์ต่างๆ การนิยามข้อมูลมีการกำหนดกฎเพื่อบูรณภาพของข้อมูล (data integrity) เช่น การกำหนดกฎเกี่ยวกับ Primary Key และ Foreign keys การอ้างอิง (referential integrity) ที่เป็นการจับคู่ระหว่าง Primary Key และ Foreign keys สำหรับโครงสร้างฐานข้อมูล (database structure) ถูกแสดงในรูปเค้าร่างฐานข้อมูล (database schema) ที่เป็นคำอธิบายเกี่ยวกับฐานข้อมูลทั้งหมด ซึ่งกำหนดโครงสร้างเชิงตรรกะและโครงสร้างทางกายภาพของฐานข้อมูล ระบุถึงตารางและคุณลักษณะในแต่ละตาราง และความสัมพันธ์ระหว่างคุณลักษณะในตารางต่างๆ

ซอฟต์แวร์ระบบจัดการฐานข้อมูล

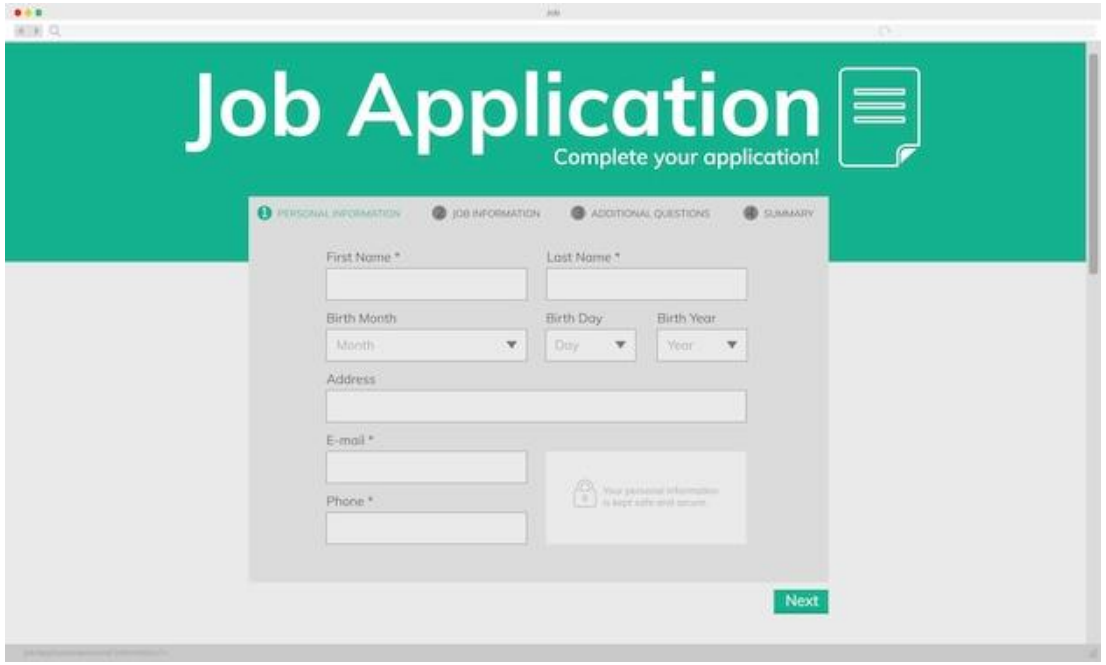
- **ส่วนการจัดการข้อมูล (data manipulation)**

ใช้เพื่อเพิ่ม ลบ แก้ไข และดึงข้อมูลจากฐานข้อมูล โดยมีคำสั่งในการจัดการทั่วไปเพื่อการสร้าง view ที่แสดงข้อมูลบางอย่างในฐานข้อมูล ซึ่งข้อมูลมาจากตารางหรือฟิลด์ที่ถูกสอบถามหรือคิวรี (query) ขึ้นมาแสดง การจัดการข้อมูลสามารถทำได้ด้วยภาษาสอบถามเชิง โครงสร้างหรือเอสคิวแอลในกลุ่มภาษาจัดการข้อมูล (data manipulation language: DML) เช่น คำสั่ง SELECT, INSERT, UPDATE, DELETE เป็นต้น

ซอฟต์แวร์ระบบจัดการฐานข้อมูล

- **ส่วนการสร้างระบบประยุกต์ (application generation)**

ถูกใช้ในการสร้างโปรแกรมประยุกต์ส่วนหน้าที่โต้ตอบกับผู้ใช้ได้โดยตรง เช่น การจัดทำหน้าจอหรือแบบฟอร์มนำเข้าข้อมูล



ซอฟต์แวร์ระบบจัดการฐานข้อมูล

- **ส่วนการบริหารข้อมูล (data administration)**
  - การนำเข้าและส่งออกข้อมูล (import/ export data) ที่ช่วยนำเข้าข้อมูล และส่งออกข้อมูลในรูปแบบต่างๆ ได้ การทำซ้ำข้อมูล (data replication) ที่ช่วยจัดการสำเนาข้อมูลเดียวกันที่มีหลายชุด อยู่ในหลายที่ เพื่อการเพิ่มประสิทธิภาพในการเข้าถึงข้อมูล
  - การสำรองข้อมูล (backup) ที่เป็นการทำสำเนาฐานข้อมูล
  - การกู้ข้อมูล (recovery) ที่คืนค่าฐานข้อมูลไปยังสถานะที่ถูกต้องหรือสถานะก่อนหน้าจากสถานะที่ไม่ถูกต้อง
  - การรักษาความมั่นคง (security) ที่เป็นการรักษาความมั่นคงพื้นฐานที่ฐานข้อมูลก่อน ป้องกันการเข้าถึงฐานข้อมูลโดยไม่ได้รับอนุญาต ไม่ว่าจะโดยเจตนาหรือไม่ เช่น การให้สิทธิ์ การระบุบุคคล การเข้ารหัส
  - การจัดการการอัปเดตข้อมูลพร้อมกัน (concurrent update) ที่รับรองความถูกต้อง

ซอฟต์แวร์ระบบจัดการฐานข้อมูล

ระบบจัดการฐานข้อมูลมีผู้ดูแลฐานข้อมูล ทำหน้าที่ออกแบบและจัดการฐานข้อมูล กำหนดมาตรการรักษาความมั่นคง พัฒนาขั้นตอนการกู้ข้อมูล ประเมินประสิทธิภาพของฐานข้อมูล และปรับแต่งฐานข้อมูลเพื่อเพิ่มประสิทธิภาพ ระบบจัดการฐานข้อมูลคงความเชื่อถือได้และความถูกต้องของข้อมูล ด้วยการรักษาคุณสมบัติหลักของรายการเปลี่ยนแปลงในฐานข้อมูล (ACID properties)

- Atomicity รายการเปลี่ยนแปลงหรือธุรกรรม ต้องได้รับการปฏิบัติแบบครบถ้วน เช่น การโอนเงินที่ยอดเงินทั้ง 2 บัญชี ต้องเปลี่ยนแปลงอย่างถูกต้อง
- Consistency ฐานข้อมูลจะต้องอยู่ในสถานะที่สอดคล้องกันหลังจากการทำรายการเปลี่ยนแปลง
- Isolation การกระทำกับข้อมูลที่แยกจากกัน ทำให้การทำรายการแต่ละครั้งไม่เกิดการแทรกแซงกัน
- Durability ฐานข้อมูลควรมีความทนทานเพียงพอที่จะรองรับการปรับปรุงข้อมูลให้เป็นปัจจุบัน

SQL (Structured Query Language)

เป็นภาษามาตรฐานในการเข้าถึง ฐานข้อมูล สามารถใช้งานภาษา SQL ได้จากโปรแกรมต่างๆ ที่ต้องทำการกับระบบฐานข้อมูล เช่น ใช้ SQL ในการทำการดึงข้อมูล (Retrieve Data) จากฐานข้อมูล และ SQL เป็นมาตรฐานกลางที่ใช้ในระบบฐานข้อมูลต่างๆ

การดำเนินความสัมพันธ์

การดำเนินการ Insert

COURSES

No	Course-Name	Unit
CIS15	Intro to C	5
CIS17	Intro to Java	5
CIS19	UNIX	4
CIS51	Networking	5

→ **Insert** →

No	Course-Name	Unit
CIS15	Intro to C	5
CIS17	Intro to Java	5
CIS19	UNIX	4
CIS51	Networking	5
CIS52	*TCP/IP Protocols*	*6*

การดำเนินความสัมพันธ์ (ต่อ)

การดำเนินการ Delete

COURSES

No	Course-Name	Unit
CIS15	Intro to C	5
CIS17	Intro to Java	5
CIS19	UNIX	4
CIS51	Networking	5
CIS52	TCP/IP Protocols	6

→ **Delete** →

No	Course-Name	Unit
CIS15	Intro to C	5
CIS17	Intro to Java	5
CIS51	Networking	5
CIS52	TCP/IP Protocols	6

การดำเนินความสัมพันธ์ (ต่อ)

การดำเนินการ Update

COURSES

No	Course-Name	Unit
CIS15	Intro to C	5
CIS17	Intro to Java	5
CIS19	UNIX	4
CIS51	Networking	5
CIS52	TCP/IP Protocols	6

→ **Update** →

No	Course-Name	Unit
CIS15	Intro to C	5
CIS17	Intro to Java	5
CIS19	UNIX	4
CIS51	Networking	**6**
CIS52	TCP/IP Protocols	6

การดำเนินความสัมพันธ์ (ต่อ)

การดำเนินการ Select

COURSES

No	Course-Name	Unit
CIS15	Intro to C	5
CIS17	Intro to Java	5
CIS19	UNIX	4
CIS51	Networking	5
CIS52	TCP/IP Protocols	6

→ **Select** →

No	Course-Name	Unit
CIS15	Intro to C	5
CIS17	Intro to Java	5
CIS51	Networking	5

การดำเนินความสัมพันธ์ (ต่อ)

การดำเนินการ Join

CIS15-Roster

Student-ID	F-Name	L-Name
145-67-6754	John	Brown
232-56-5690	George	Yellow
345-89-6580	Anne	Green
459-98-6789	Ted	Purple

CIS52-Roster

Student-ID	F-Name	L-Name
342-88-9999	Rich	White
145-67-6754	John	Brown
232-56-5690	George	Yellow

Union

Student-ID	F-Name	L-Name
145-67-6754	John	Brown
232-56-5690	George	Yellow
345-89-6580	Anne	Green
459-98-6789	Ted	Purple
342-88-9999	Rich	White

การดำเนินความสัมพันธ์ (ต่อ)

การดำเนินการ Union

CIS15-Roster

Student-ID	F-Name	L-Name
145-67-6754	John	Brown
232-56-5690	George	Yellow
345-89-6580	Anne	Green
459-98-6789	Ted	Purple

CIS52-Roster

Student-ID	F-Name	L-Name
342-88-9999	Rich	White
145-67-6754	John	Brown
232-56-5690	George	Yellow

Union

Student-ID	F-Name	L-Name
145-67-6754	John	Brown
232-56-5690	George	Yellow
345-89-6580	Anne	Green
459-98-6789	Ted	Purple
342-88-9999	Rich	White

การดำเนินความสัมพันธ์ (ต่อ)

การดำเนินการ Intersection

CIS15-Roster

Student-ID	F-Name	L-Name
145-67-6754	John	Brown
232-56-5690	George	Yellow
345-89-6580	Anne	Green
459-98-6789	Ted	Purple

CIS52-Roster

Student-ID	F-Name	L-Name
342-88-9999	Rich	White
145-67-6754	John	Brown
232-56-5690	George	Yellow

Intersection

Student-ID	F-Name	L-Name
145-67-6754	John	Brown
232-56-5690	George	Yellow

การดำเนินความสัมพันธ์ (ต่อ)

การดำเนินการ Difference

CIS15-Roster

Student-ID	F-Name	L-Name
145-67-6754	John	Brown
232-56-5690	George	Yellow
345-89-6580	Anne	Green
459-98-6789	Ted	Purple

CIS52-Roster

Student-ID	F-Name	L-Name
342-88-9999	Rich	White
145-67-6754	John	Brown
232-56-5690	George	Yellow

Difference

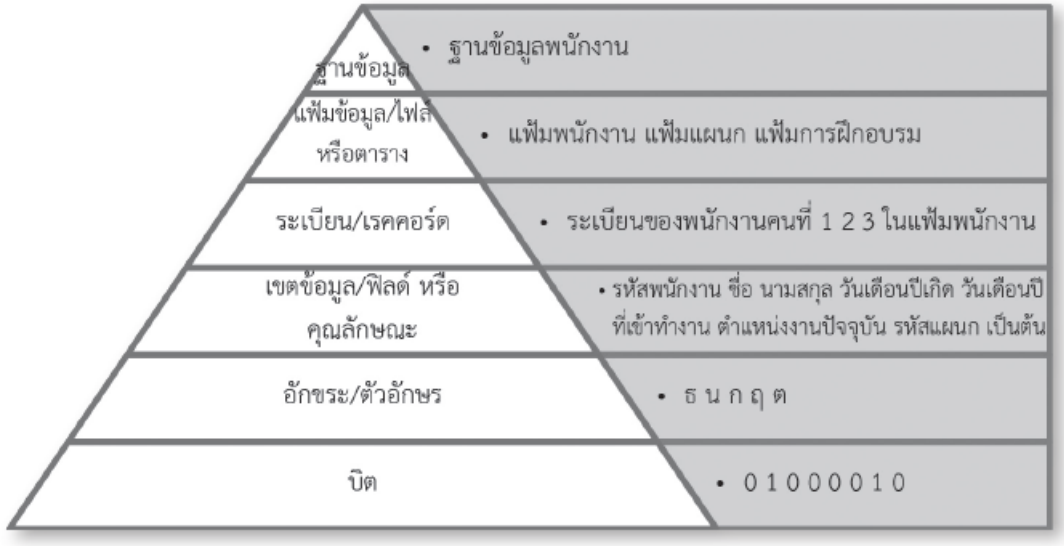
Student-ID	F-Name	L-Name
345-89-6580	Anne	Green
459-98-6789	Ted	Purple

ตัวอย่าง SQL

- **Select** name **From** Student **Where** PGP > 3.00
- **Select** name **From** Student **Where** Department = “Computer Science”
- **Select** name **From** Student **Where** Province = “Bangkok”
- **Select** name, OrderNo **From** Customer, Order **Where** Customer.ID = Order.CustID
- **Insert into** Courses values (“CT105”, “Introduction to Computer Science”,3)
- **Delete from** Courses **Where** No = “CT105”

ประโยชน์ของฐานข้อมูล

1. ลดความซ้ำซ้อนของข้อมูล (Data Redundancy)
2. หลีกเลี่ยงความขัดแย้งของข้อมูลได้ (Data Inconsistency)
3. สามารถใช้ข้อมูลร่วมกันในองค์กร
4. รักษาความถูกต้องและเชื่อถือได้ของข้อมูล
5. กำหนดความเป็นมาตรฐานเดียวกันได้
6. สามารถกำหนดระบบความปลอดภัยให้กับข้อมูลได้
7. ความเป็นอิสระของข้อมูล

ประโยชน์ของฐานข้อมูล

ข้อมูลที่ถูกจัดเก็บมักเป็นข้อมูลที่มีความสำคัญต่อการบรรลุเป้าหมายองค์กร สนับสนุนการดำเนินกระบวนการทางธุรกิจ หรือเป็นสิ่งที่ธุรกิจให้ความสนใจ และจะถูกนำไปใช้ในระบบสารสนเทศทางธุรกิจประเภทต่างๆ ฐานข้อมูลมีทั้งขนาดเล็กและขนาดใหญ่ ฐานข้อมูลขนาดใหญ่ เช่น ฐานข้อมูลประชากร เป็นต้น