

การจัดการข้อมูลและสารสนเทศ

Information Sharing

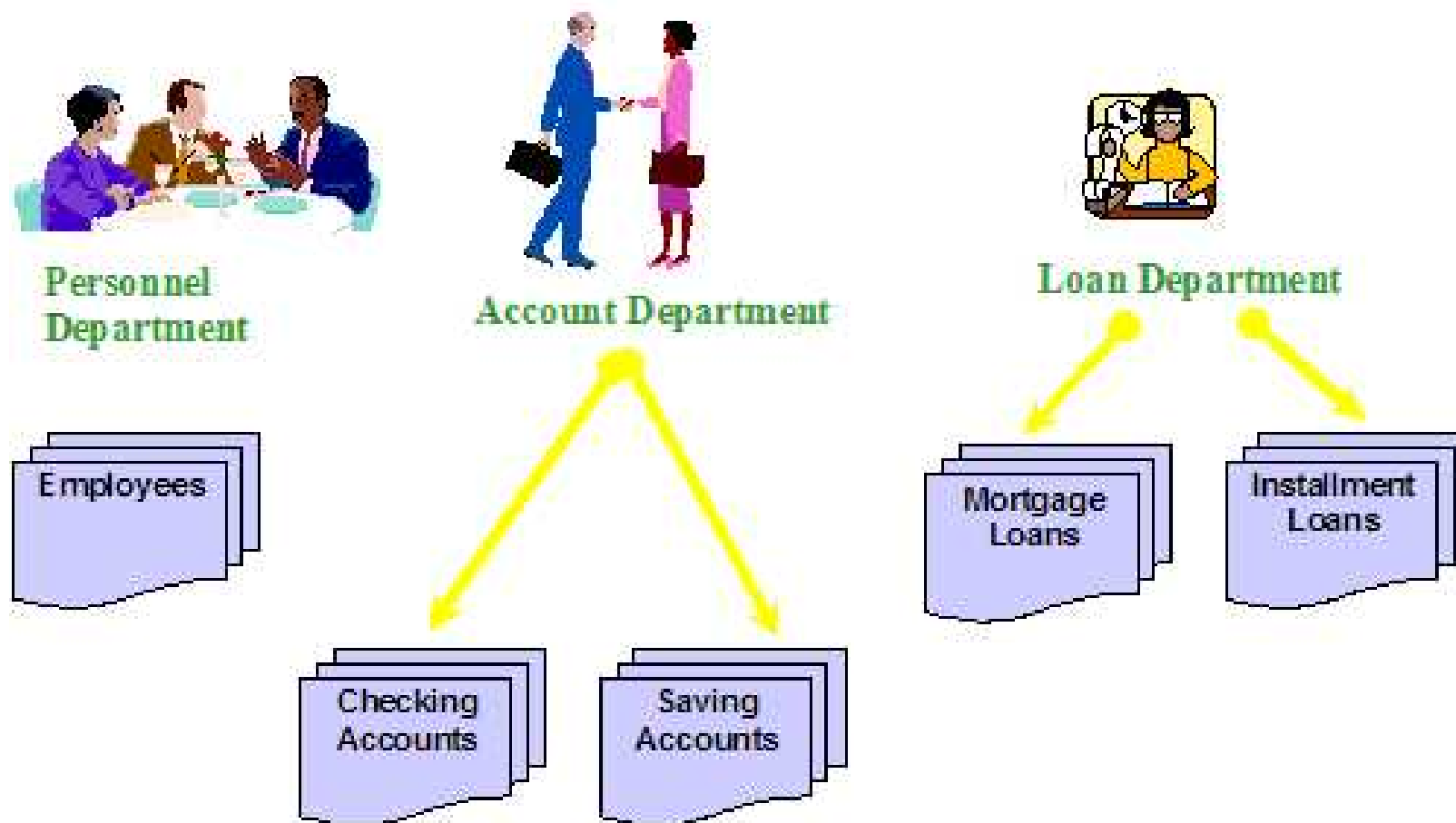
ทัศนวรรณ ศูนย์กลาง
ภาควิชาคอมพิวเตอร์ คณะวิทยาศาสตร์

เนื้อหาวันนี้

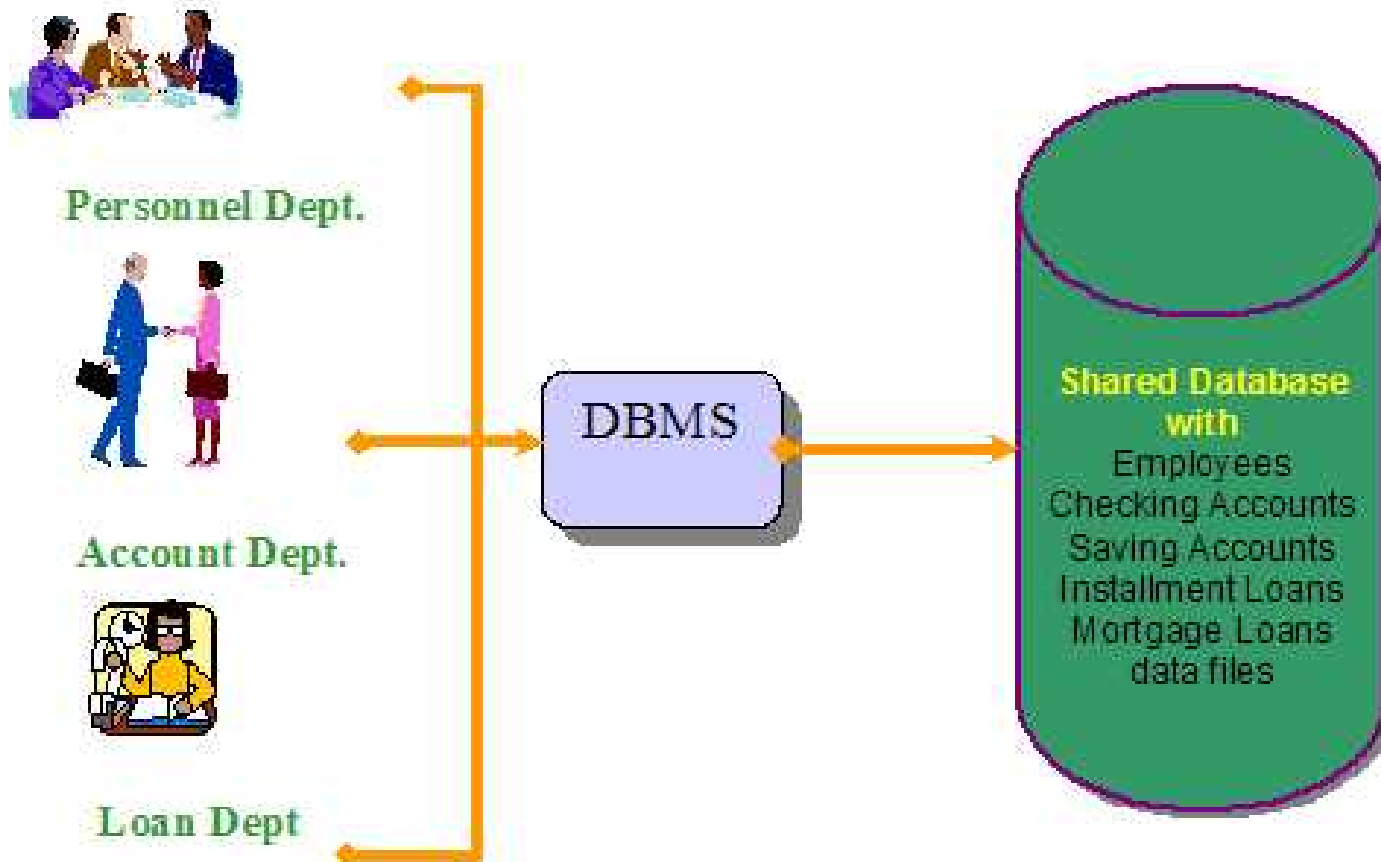
○ การจัดการข้อมูลและสารสนเทศ (Information sharing)

- ระบบแฟ้มข้อมูล
- ระบบฐานข้อมูล
- การประมวลผลข้อมูล
- คลังข้อมูล

ระบบเพิ่มข้อมูล



ระบบฐานข้อมูล



แนวคิดของการใช้ฐานข้อมูล



ทำไมต้องมีฐานข้อมูล

Item	File	Database
Impact of changes to data format on program	Large	Small
Redundancy of data	Data is sometimes redundant on a task-by-task basis	No redundancy
Consistency between related data	Hard to maintain	Can be maintained
Sharing of data between tasks	Sharing is difficult	Sharing is easy
Data backup	Complicated	Simple and easy

ฐานข้อมูล

- ฐานข้อมูล (Database)
 - การจัดเก็บรวบรวมข้อมูลให้เป็นระเบียบเพื่อสามารถนำข้อมูลไปใช้งานต่อได้และง่ายต่อการค้นคืน
- แบบจำลองข้อมูล (Database model)
 - การออกแบบรูปแบบการจัดเก็บข้อมูล
- วัตถุประสงค์ของการทำ Database model
 - รวบรวมข้อมูลที่กระจัดกระจายให้เป็นหมวดหมู่ ง่ายต่อการบริหารจัดการ ทำให้สามารถใช้ข้อมูลที่มีอยู่ เพื่อประกอบการวิเคราะห์ ประเมินผลในทางธุรกิจได้

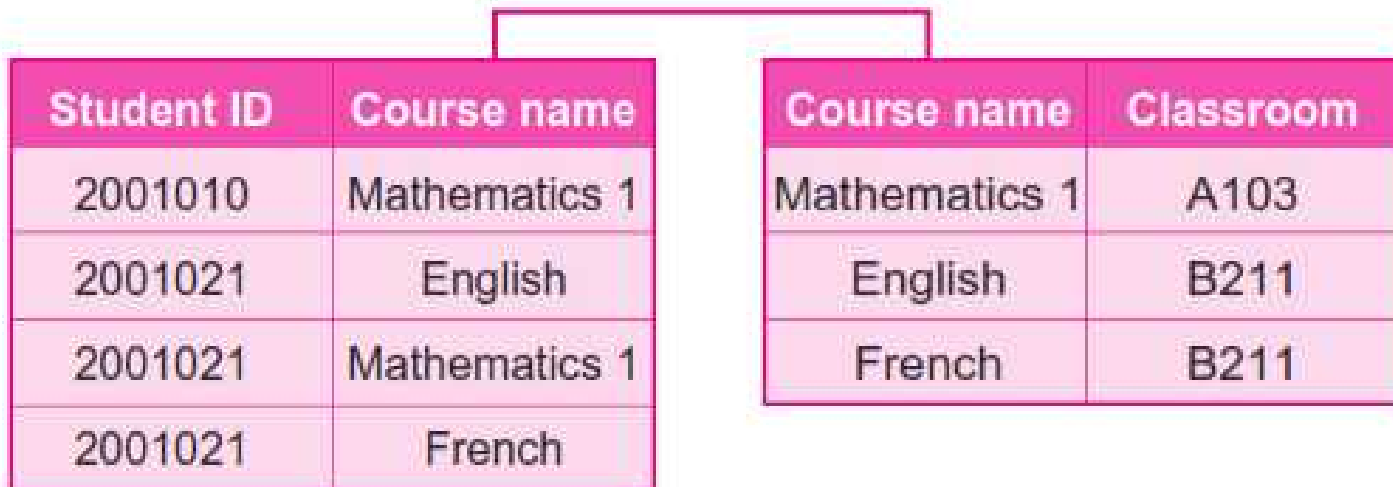
Database model

ประเภทของฐานข้อมูล แบ่งเป็นหลายประเภท ตัวอย่างเช่น

- ฐานข้อมูลเชิงสัมพันธ์ (relational database)
- ฐานข้อมูลแบบลำดับชั้น (hierarchical database)
- ฐานข้อมูลแบบเครือข่าย (network database)
- ฐานข้อมูลเชิงวัตถุ (object-oriented database)

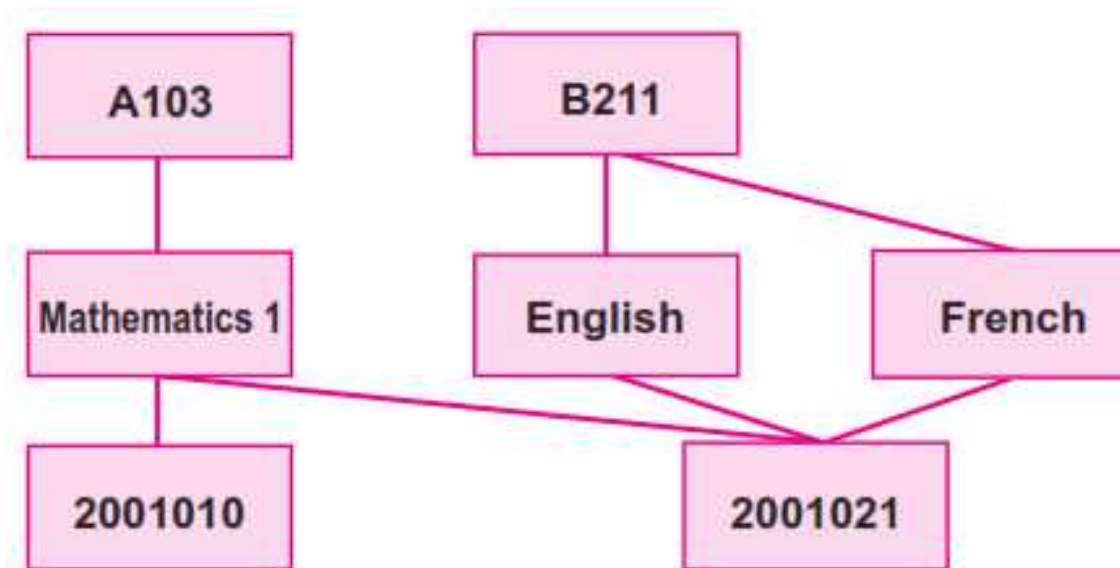
Database model

- Relational database เก็บข้อมูลแบบตาราง แต่ละตารางมีความสัมพันธ์กันโดยมีคอล์มน์ที่เหมือนกัน
- ง่ายในการทำความเข้าใจสำหรับผู้ใช้
- ง่ายในดึงข้อมูลที่ผู้ใช้งานต้องการ



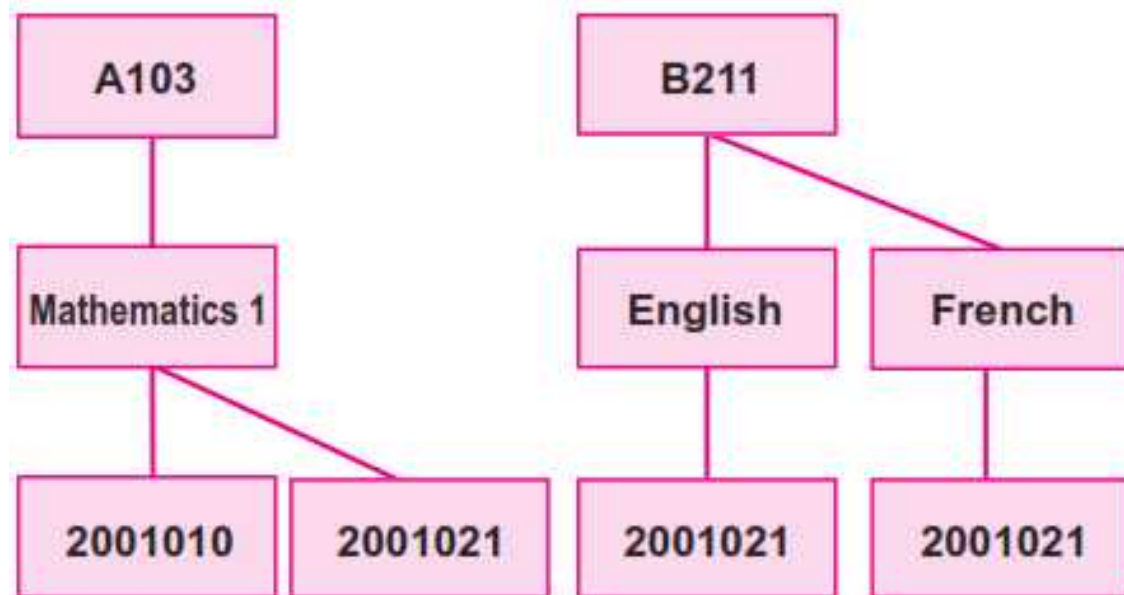
Database model

- Network database ข้อมูลแต่ละตัวเชื่อมโยงกันแบบโครงข่าย อาจมีความสัมพันธ์กันแบบ many-to-many



Database model

- Hierarchical database ข้อมูลแต่ละตัวถูกแบ่งออกเป็นชั้นๆ อาจมีความสัมพันธ์กันแบบ one-to-many

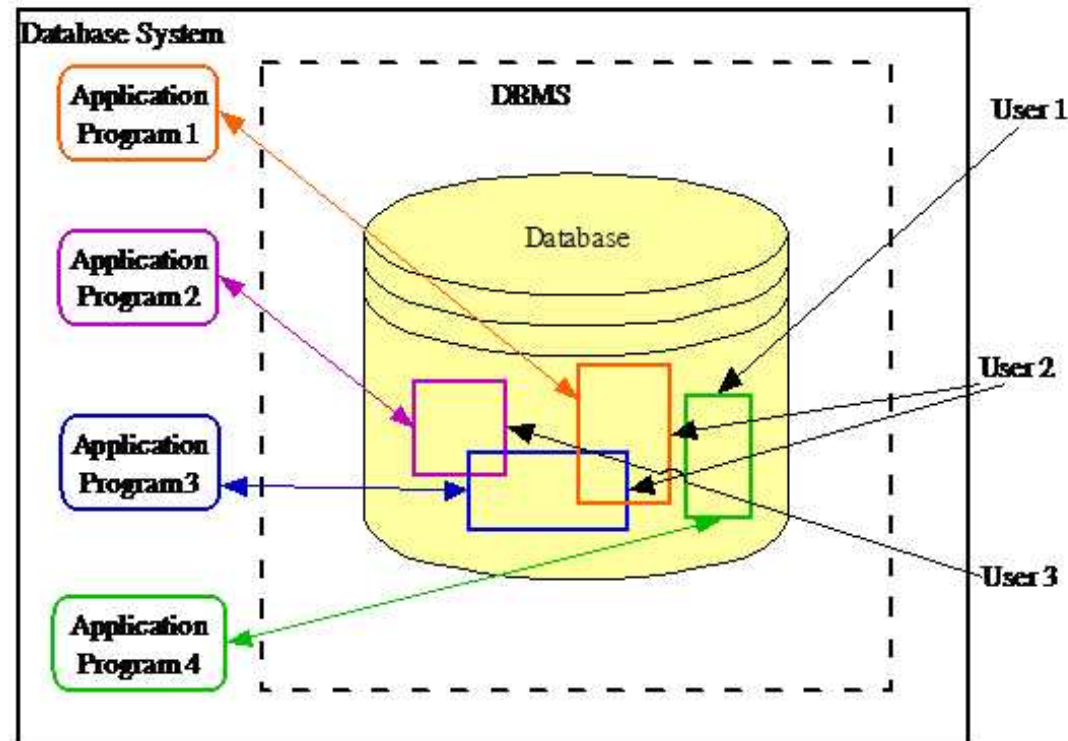


Database model

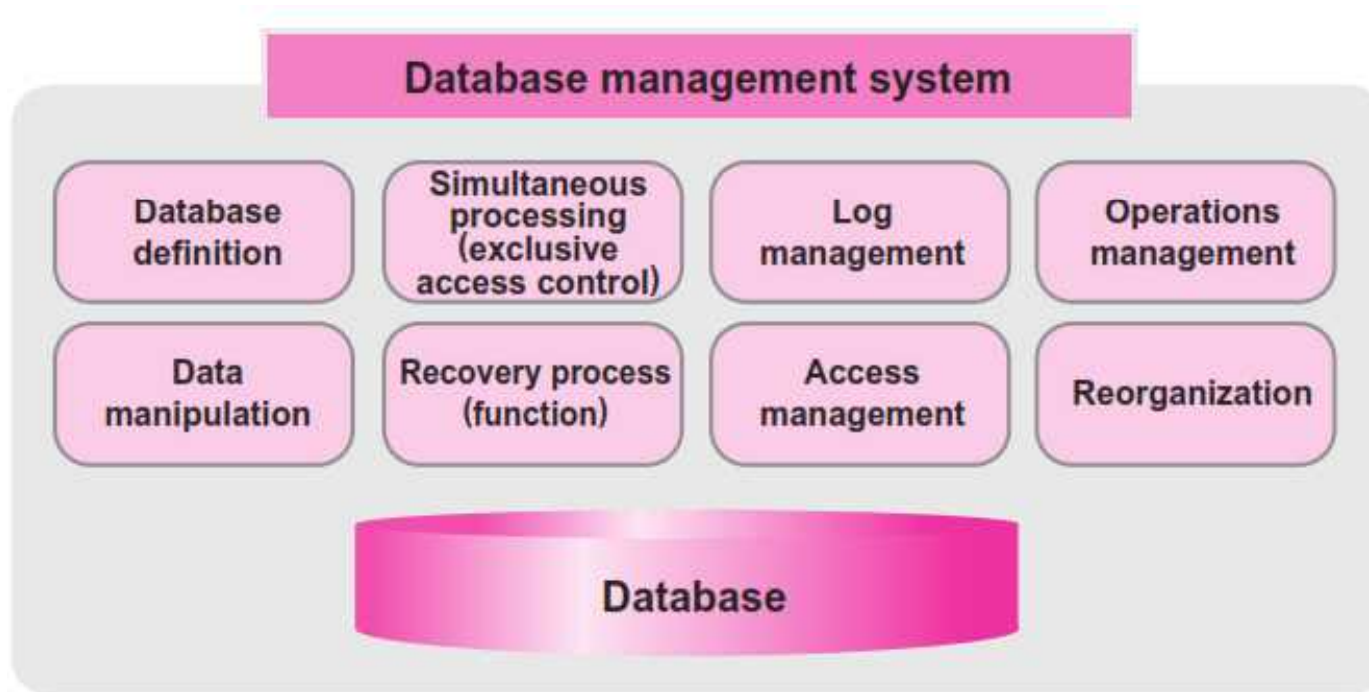
- Object-oriented database เหมาะสำหรับจัดเก็บข้อมูลที่ไม่มีโครงสร้าง เช่น ข้อมูล ภาพ เสียง และวีดิทัศน์

เครื่องมือสำหรับจัดการฐานข้อมูล

- Database Management System – DBMS
 - ระบบที่ใช้สร้างและบริหารจัดการฐานข้อมูลเพื่อช่วยให้ผู้ใช้งานมีความสะดวกในการใช้งาน



หน้าที่หลักของ DBMS



หน้าที่หลักของ DBMS

- database definition นิยามโครงสร้าง DB เช่น table, item, index
- data manipulation เช่น ค้นหา เพิ่ม ลบ แก้ไขข้อมูลใน DB
- Simultaneous processing จัดการ data consistency เมื่อมีผู้ใช้หลายคน
- Recovery process กู้คืนข้อมูลหรือ restore ข้อมูลไปยังจุดก่อนมีปัญห
- Log management เก็บและจัดการ log file เพื่อใช้สำหรับ recovery
- Access control กำหนดสิทธิการใช้งาน
- Operation management ทำ backup หรือ restore
- Reorganization จัดเรียงข้อมูลที่เกิด fragmentation จากการเพิ่มลบข้อมูล

ประโยชน์ของการใช้ DBMS

- รองรับทำงานใช้เวลาน้อยในการค้นหา สืบเปลี่ยน คำนวณค่าของข้อมูล
- สามารถทำงานออกมาได้หลากหลายรูปแบบ
- สามารถรองรับผู้ใช้หลายๆ คนได้พร้อมๆ กันในเวลาเดียวกัน
- หากเกิดความเสียหายกับ Hardware และ Software อาจไม่ส่งผลต่อข้อมูลที่ทำกรจัดเก็บ

การออกแบบฐานข้อมูล

- ขั้นตอนการออกแบบ
 - Data analysis ต้องตรวจสอบข้อมูลที่นำมาสร้างฐานข้อมูล โดยแบ่งออกเป็นส่วนๆ ประกอบด้วย หัวข้อ รายละเอียดของหัวข้อ ความถี่ในการใช้งาน
 - Data design ในกรณีของ Relational database จะนำข้อมูลที่ได้มาออกแบบในรูปแบบของตาราง
 - Data normalization เพื่อไม่ให้ข้อมูลซ้ำซ้อนกัน จำเป็นต้องออกแบบตารางให้เหมาะสม

ส่วนประกอบของตาราง

- แถว (Row) เรียกรายละเอียดในแถวว่า Record
- คอลัมน์ (Column) เรียกรายละเอียดในคอลัมน์ว่า Field
- คีย์หลัก (Primary key) สามารถใช้ field ที่เป็น primary key เป็นตัวแทนของตารางนั้นได้ ข้อมูลที่เป็น primary key ในตารางใดๆ จะต้องเป็นข้อมูลที่ไม่มีค่าซ้ำกันเลยในตารางนั้น

Customer table

Customer code		Item name (field name)	
Customer code	Customer name	Phone number	Location
A-1	Nanboku Denki	03-3592-123X	Tokyo
B-1	Nihon Kogyo	06-6967-123X	Osaka
A-20	Iroha Denshi	078-927-123X	Hyogo

Primary key: Identifies each record in the table
For example, if the customer code is specified, the corresponding customer name can be identified.

ตัวอย่างของ DBMS

- ฐานข้อมูลขนาดใหญ่ เช่น Oracle, SQL Server, DB2
- ฐานข้อมูลขนาดกลาง เช่น MySQL, PostgreSQL
- ฐานข้อมูลขนาดเล็ก เช่น Microsoft Access

ชนิดของฐานข้อมูล

จำแนกตามชนิดของการใช้งาน แบ่งได้ 5 ประเภท

- ฐานข้อมูลส่วนบุคคล (individual database)
- ฐานข้อมูลขององค์กร (company database/shared database)
- ฐานข้อมูลชนิดกระจาย (distributed database)
- ฐานข้อมูลชนิดมีเจ้าของเฉพาะ (proprietary database)
- ฐานข้อมูลเว็บ (web database)

ฐานข้อมูลส่วนบุคคล

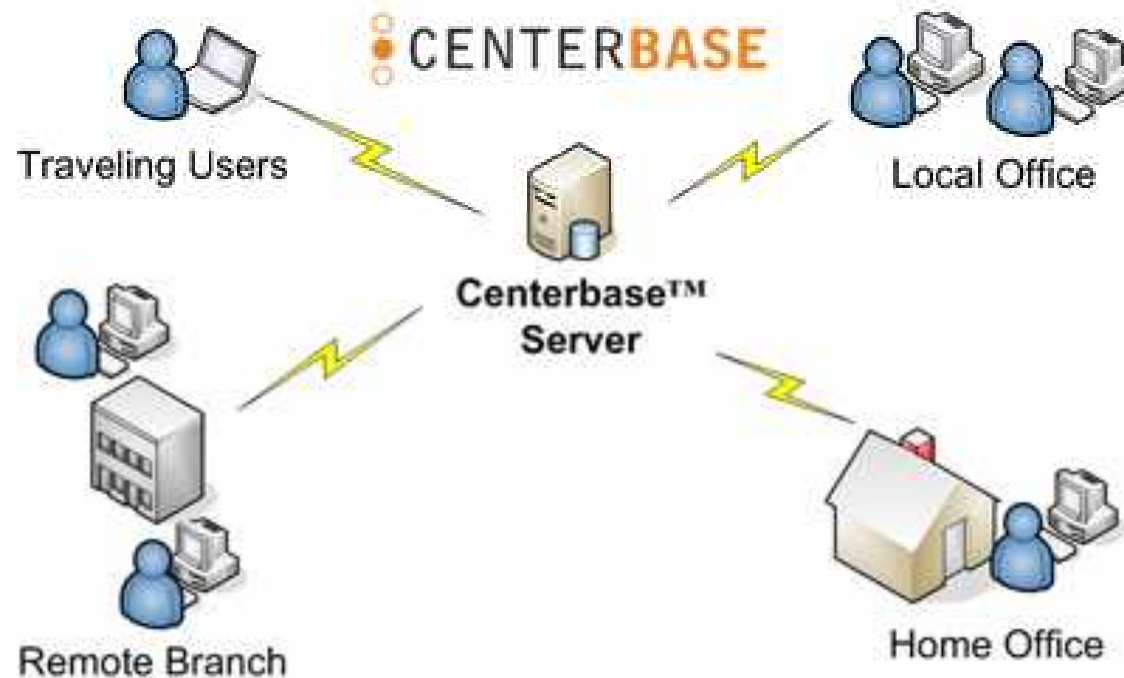
- จัดเก็บข้อมูลในฮาร์ดดิสก์
 - ภายในเครื่อง
 - ภายในเครือข่ายระยะใกล้
- มักใช้ในงานประจำ
 - ฝ่ายขายใช้เก็บข้อมูลลูกค้า ดูรายชื่อลูกค้าได้
 - ผู้จัดการฝ่ายขายสามารถค้นหาพนักงานที่ทำยอดขายสูงสุดได้
 - ผู้บริหารฝ่ายโฆษณาสามารถคิดค่าโฆษณาจากลูกค้า

ฐานข้อมูลขององค์กร

- จัดเก็บข้อมูลไว้ที่เครื่องคอมพิวเตอร์แบบเมนเฟรม
- จัดการฐานข้อมูลโดยผู้บริหารฐานข้อมูล
- ผู้ต้องการเรียกใช้ข้อมูลจะใช้เครื่องเชื่อมต่อเป็นระบบเครือข่ายระยะไกลหรือใกล้ก็ได้
- แบ่งเป็น 2 ประเภทย่อย
 - ฐานข้อมูลสำหรับการปฏิบัติงานทั่วไป (common operational database) ข้อมูลเกิดจากการทำงานภายในองค์กร
 - ฐานข้อมูลที่ขึ้นกับผู้ใช้ (common user database) ข้อมูลเกิดจากการทำงานและได้จากฐานข้อมูลชนิดมีเจ้าของเฉพาะ เพื่อสร้างสารสนเทศสำหรับผู้บริหาร

ฐานข้อมูลขององค์กร

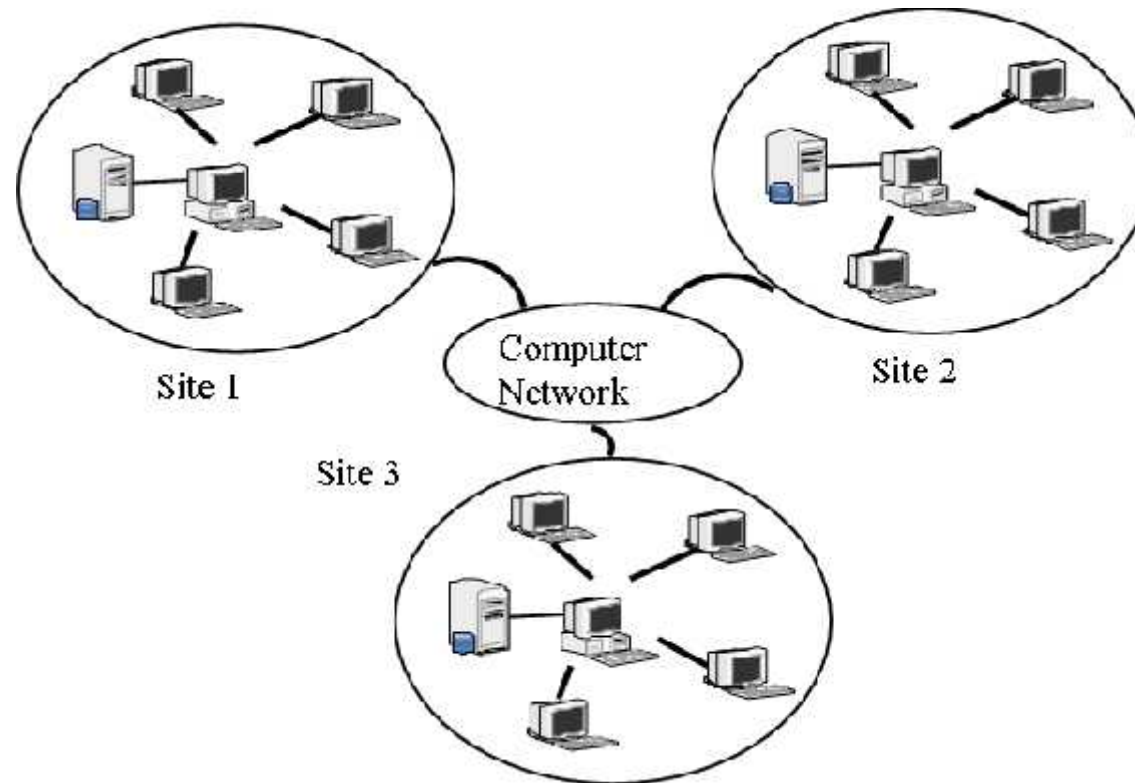
- ฐานข้อมูลแบบรวมศูนย์ (centralized database)



source: http://www.centerbase.com/inside.aspx?module_name=system

ฐานข้อมูลชนิดกระจาย

- จัดเก็บข้อมูลในฐานข้อมูลซึ่งไม่ได้เก็บไว้ ณ สถานที่เดียว แต่จะเก็บแยกกันหลายที่



Source: <http://cnx.org/content/m28156/latest/>

ฐานข้อมูลชนิดมีเจ้าของเฉพาะ

- ฐานข้อมูลขนาดใหญ่ มีหน่วยงานพัฒนาขึ้นโดยมีเนื้อหาเฉพาะ
- ให้บริการกับบริษัทหรือองค์กรที่ต้องการข้อมูลจากภายนอก
- ผู้ขอใช้ต้องชำระค่าบริการ อาจเรียกว่า ธนาคารข้อมูล
- ตัวอย่างเช่น
 - เครดิตบูโร ข้อมูลประวัติการชำระสินเชื่อ และการชำระบัตรเครดิตของบุคคลจากสถาบันการเงินหลายๆ แห่ง
 - IEEE Xplorer เป็นเครื่องมือที่ใช้สืบค้นเอกสารสิ่งพิมพ์ทางด้านอิเล็กทรอนิกส์

ฐานข้อมูลเว็บ

- มีลักษณะเช่นเดียวกับฐานข้อมูลชนิดอื่นๆ แต่ผู้ใช้สามารถติดต่อผ่านเครือข่ายอินเทอร์เน็ตได้
- นิยมใช้บริการสืบค้นข้อมูลของเว็บไซต์

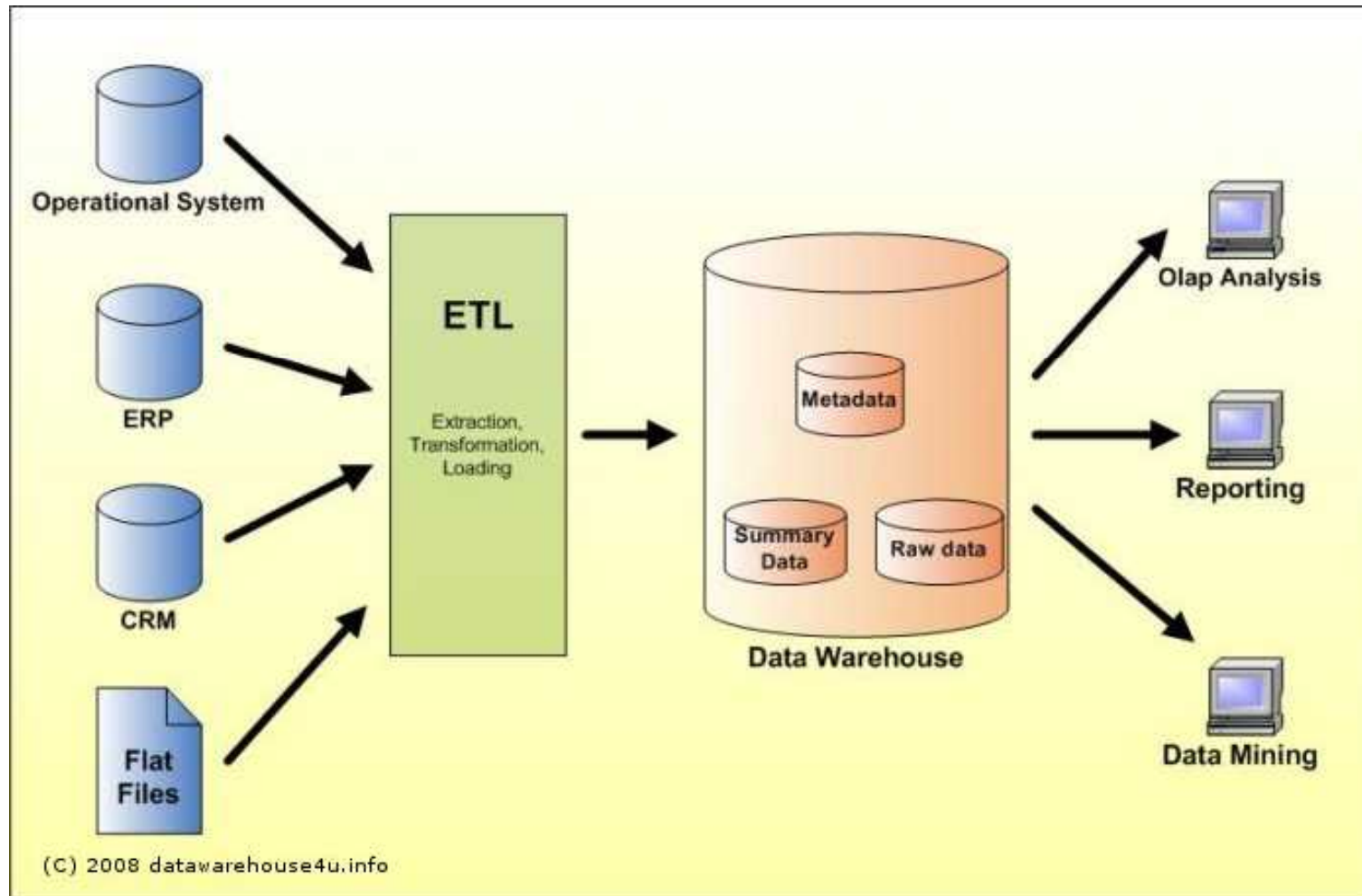
การประมวลผลข้อมูล

- การประมวลผลแบบกลุ่ม (batch processing)
 - ข้อมูลถูกจัดเก็บเอาไว้ระยะเวลาหนึ่ง จึงจะประมวลผลให้เสร็จในคราวเดียว
 - การคิดค่าโทรศัพท์ระบบรายเดือน
 - การตัดบัญชีจากการใช้บัตรเครดิต
- การประมวลผลแบบทันที (real-time processing)
 - การประมวลผลเกิดขึ้นทันที อาจจะเป็นการประมวลผลแบบออนไลน์ได้
 - การคิดค่าโทรศัพท์ระบบเติมเงิน
 - การถอนเงินผ่านเครื่องเอทีเอ็ม

คลังข้อมูล

- คลังข้อมูล (Data warehouse)
 - ที่เก็บข้อมูลขององค์กรที่ได้รับการออกแบบเพื่อช่วยในการตัดสินใจของผู้บริหาร
 - มีเครื่องมือสำหรับดำเนินการกับข้อมูล
 - ข้อมูลสำหรับการบริหาร โดยมากเป็นข้อมูลสรุป จากอดีต ปัจจุบัน หรือข้อมูลอ้างอิง
 - ข้อมูลอาจได้จากฐานข้อมูลในงานประจำวัน หรือจากภายนอกองค์กร

คลังข้อมูล



เป้าหมายการสร้างคลังข้อมูล

- เข้าถึงข้อมูลในองค์กรได้ ผู้บริหารสามารถเชื่อมต่อเข้าไปยังคลังข้อมูลได้ทันที
- มีเครื่องมือในการวิเคราะห์ข้อมูล และออกรายงานได้ง่าย
- ข้อมูลสามารถวิเคราะห์ได้โดยแบ่งหรือรวมข้อมูลมาวิเคราะห์ตามความต้องการ
- ข้อมูลมีความถูกต้องตรงกันหมด

ประโยชน์ของคลังข้อมูล

- ตอบสนองงานในรูปแบบการตัดสินใจ
- แยกฐานข้อมูลออกจากฐานข้อมูลที่ใช้งานประจำวัน
- เก็บข้อมูลในรูปแบบข้อมูลสรุป
- ข้อมูลสรุป เป็นข้อมูลที่มีความสำคัญต่อการตัดสินใจหรือใช้งานการบริหาร และการกำหนดแผนงานในอนาคต

เหมืองข้อมูล

- เทคนิคที่ใช้ค้นสารสนเทศที่มีความสัมพันธ์กัน
- มีการคำนึงถึงความปลอดภัยในการเข้าถึงฐานข้อมูล
- กระบวนการที่กระทำกับข้อมูล (โดยส่วนใหญ่จะมีจำนวนมาก) เพื่อค้นหา รูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น
- อาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง และหลักคณิตศาสตร์

เหมืองข้อมูล

