

แนวคิดเกี่ยวกับข้อมูลและ ฐานข้อมูล

204204 Database Design & Development
บรรยายครั้งที่ 1

วัตถุประสงค์รายครั้ง

นักศึกษาสามารถ

- อธิบายวิวัฒนาการของระบบสารสนเทศได้
- บอกความหมายของระบบแฟ้มข้อมูล และฐานข้อมูลได้
- อธิบายความถูกต้องและความเป็นอิสระของข้อมูลได้อย่างถูกต้อง
- บอกคุณลักษณะที่ดี องค์ประกอบ และบุคคลที่เกี่ยวข้องกับฐานข้อมูลได้อย่างถูกต้อง

หัวข้อ

บรรยาย

- Information คืออะไร? ต่างจาก Data อย่างไร?
- วิวัฒนาการของระบบสารสนเทศ
- ระบบแฟ้มข้อมูล
- ฐานข้อมูล
- ความถูกต้องและความเป็นอิสระของข้อมูล
- คุณลักษณะที่ดีของฐานข้อมูล
- องค์ประกอบของระบบฐานข้อมูล
- บุคคลที่เกี่ยวข้องกับฐานข้อมูล

หัวข้อ

ปฏิบัติ

- การใช้โปรแกรม Microsoft Access 2000

Information คืออะไร? ต่างจาก Data อย่างไร?

- Data หรือ ข้อมูล หมายถึงสิ่งที่เก็บรวบรวมเพื่อนำไปประมวลผลต่อไป
- Information หรือ ข้อมูลสารสนเทศ หมายถึงผลได้จากการประมวลผลข้อมูล

DATA → PROCESS → INFORMATION

- Information System คือระบบที่ประมวลผลข้อมูลดิบ (Data) เพื่อให้ได้มาซึ่งข้อมูลที่ก่อให้เกิดประโยชน์หรือข้อมูลสารสนเทศ (Information)
- การประมวลผลข้อมูลอาจใช้คน หรือใช้คอมพิวเตอร์ ที่เรียกว่า Computer Information System (CIS)

ตัวอย่างข้อมูลดิบ

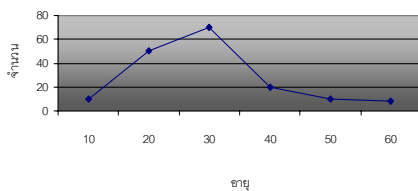
รหัสลูกค้า	ชื่อลูกค้า	รหัสอัลบั้ม	ชื่ออัลบั้ม	จำนวน	วันที่ซื้อ
1100	สมศักดิ์	622	เพลงร็อก	2	01/02/96
2200	วิไล	633	เพลงร็อก Vol 2	3	10/02/96
4600	อริญญา	855	เพลงไทยสากล	1	07/03/96
3500	วิรัตน์	553	เพลงบรรเลง	2	10/03/96
6700	วิल्ली	855	เพลงไทยสากล	2	10/03/96
4300	รงชัย	996	เพลงฮาร์ตร็อก	1	11/03/96
8700	ใหม่	750	เพลงคันทรี่	4	01/04/96
4500	มนต์สิทธิ์	553	เพลงบรรเลง	2	12/04/96
6800	เจ	996	เพลงฮาร์ตร็อก	3	14/05/96

ตัวอย่างสารสนเทศที่แสดงด้วยตาราง

รหัสลูกค้า	ชื่อลูกค้า	รหัสอัลบั้ม	ชื่ออัลบั้ม	จำนวน	วันที่ซื้อ
4300	รงชัย	996	เพลงฮาร์ตร็อกเพลงร็อก	1	11/03/96
1100	สมศักดิ์	622	เพลงร็อก Vol 2	2	01/02/96
2200	วิไล	855	เพลงฮาร์ตร็อก	3	10/02/96
6800	เจ	996		3	14/05/96

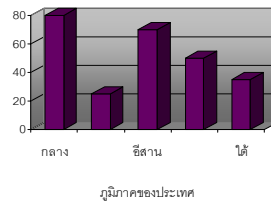
ตัวอย่างสารสนเทศที่แสดงด้วยกราฟเส้น

กราฟแสดงจำนวนการซื้ออัลบั้มเพลงร็อก จำแนกตามอายุ



ตัวอย่างสารสนเทศที่แสดงด้วยกราฟแท่ง

กราฟแสดงการซื้ออัลบั้มเพลงร็อก จำแนกตามภูมิภาค



- ระบบสารสนเทศ เป็นระบบจัดการข้อมูลสารสนเทศต่าง ๆ เพื่อให้เกิดประโยชน์ต่อระบบอื่น ๆ ที่ต้องใช้ข้อมูลเหล่านั้น
- หน้าที่หลักของระบบสารสนเทศก็คือการแปลงข้อมูลจำนวนมากให้เป็นสารสนเทศที่มีประโยชน์

วิวัฒนาการของระบบสารสนเทศ

- วิธีเก็บรวบรวมข้อมูล
 - หนังสือนำ
 - ไบรอน
 - กระดาษ
 - Electronic file
 - Database Management System (DBMS)
- การติดต่อสื่อสารในรูปแบบสื่อประสม (Multimedia)

- ระบบสารสนเทศได้อยู่ทุกหนทุกแห่ง และมีส่วนเกี่ยวข้องกับชีวิตประจำวันของเรามากมาย
- ระบบสารสนเทศรุ่นใหม่ ก็คือการนำคอมพิวเตอร์มาช่วยในการเก็บข้อมูลและประมวลผล
- หัวใจหลักของการจัดการระบบสารสนเทศรุ่นใหม่ก็คือระบบการจัดการข้อมูลอย่างมีประสิทธิภาพ การเรียนรู้เกี่ยวกับฐานข้อมูลจึงเป็นสิ่งจำเป็น

ลักษณะของสารสนเทศที่ดี

- ความเป็นปัจจุบัน (Current)
 - ข้อมูลปรับเปลี่ยนไปเรื่อย ๆ มีความทันสมัย เช่น เกรดนักศึกษา เป็นต้น
- ทันเวลา (Timely)
 - มีคุณค่าทางเวลามากเกี่ยวข้องกับ ถ้าไม่ได้สารสนเทศในเวลาที่ต้องการ อาจเกิดการสูญเสียโอกาสได้
- ความเกี่ยวข้อง (Relevant)
 - ข้อที่ได้นี้ต้องมีคุณสมบัติ ถูกต้อง

ลักษณะของสารสนเทศที่ดี

- ความคงที่ (Consistent)
 - ข้อมูลที่เก็บไว้หลาย ๆ ที่ อาจไม่ตรงกัน ขัดแย้งกัน สารสนเทศที่ดีต้องไม่มีความขัดแย้งกัน หรือขัดแย้งกันน้อยที่สุด
- นำเสนอรูปแบบที่มีประโยชน์ (Present in usable form)
 - มีรูปแบบในการนำเสนอที่เข้าใจง่าย เหมาะสม

ระบบเพิ่มข้อมูล (Electronic file or Files system)

- เพิ่มข้อมูลประกอบด้วย
 - กลุ่มของระเบียบ (Records) ที่มีรูปแบบ (Format) เหมือนกัน
 - แต่ละระเบียบจะเก็บข้อมูลที่แทนที่ของอย่างหนึ่ง (An instance or occurrence) ในกลุ่มของของที่จัดเก็บ
- ปัญหาของระบบเพิ่มข้อมูล
 - Data redundancy
 - Data inconsistency
 - Data anomaly

Data redundancy

- ได้แก่อข้อมูลซ้ำซ้อนกัน ข้อมูลชุดเดียวกันถูกจัดเก็บใน 2 แห่งหรือมากกว่า
- fact หมายถึง ความสัมพันธ์ระหว่างข้อมูลที่เป็นไปได้หรือเป็นจริง เช่น
 - ลำไย สำเร็จการศึกษาสาขาเทคโนโลยีสารสนเทศ
 - ลำไย เป็นโปรแกรมเมอร์
 - ลำไย มีงานอดิเรกในการร้องเพลง
 ข้อความทั้งหมดที่ยกมานี้ ไม่ถือว่าซ้ำซ้อนกัน เพราะเป็นคนละ fact กัน

Data redundancy

- Redundancy หมายถึง fact ที่ปรากฏมากกว่า 1 ครั้ง หรือถูกเก็บไว้มากกว่า 1 ครั้ง
- ยกตัวอย่างเช่น เก็บชื่อ ตำแหน่ง และแผนกไว้ในตารางพนักงานและตารางโครงการถือว่าซ้ำซ้อนกัน

- ฐานข้อมูลแบบตารางจะมีความซ้ำซ้อนของข้อมูลเกิดขึ้นได้ 3 แบบ ดังนี้

- ความซ้ำซ้อนระหว่างตาราง (Inter relation redundancy)
- ความซ้ำซ้อนภายในตารางเดียวกัน (Intra relation redundancy)
- ความซ้ำซ้อนบน row เดียวกัน (Intra row redundancy)

- ข้อเสีย ของความซ้ำซ้อน คือ

- เปลืองเนื้อที่ในการจัดเก็บ
- เพิ่มค่าใช้จ่ายในการนำเข้าข้อมูล
- จะต้องตามแก้ไขข้อมูลทุกที่ เมื่อข้อมูลมีการเปลี่ยนแปลง
- การเชื่อมตารางจะใช้ทรัพยากรมาก และช้า

- ข้อดี

- การทำ query report จะเร็วขึ้น

Data inconsistency

- เป็นผลมาจาก Data redundancy คือ ข้อมูลชุดเดียวกันมีค่าต่างกัน ทำให้ไม่ทราบว่าข้อมูลชุดใดคือข้อมูลที่ถูกต้อง

Data anomaly

- เป็นผลมาจาก Data redundancy เช่นกัน ทำให้การเพิ่ม ลบ หรือเปลี่ยนแปลงข้อมูลชุดเดียวกันในแฟ้มข้อมูลต่าง ๆ ไม่ครบถ้วน ซึ่งเกิดใน 3 ลักษณะ ดังนี้
- Modification anomaly
- Insertion anomaly
- Deletion anomaly

Modification anomaly

- เป็นการเปลี่ยนแปลงค่าของข้อมูลในแฟ้มข้อมูลต่าง ๆ ที่สัมพันธ์กันไม่ครบถ้วน เช่น เปลี่ยนชื่อ “ดวงใจ” ไปเป็น “ดวงพร” ในแฟ้มข้อมูล Employee โดยไม่เปลี่ยนชื่อของพนักงานคนเดียวกันในแฟ้ม Salesman จะส่งผลให้ “ดวงพร” และ “ดวงใจ” ไม่ใช่คนเดียวกันเมื่อทำการเรียกข้อมูลมาดู เป็นต้น

Insertion anomaly

- เป็นการกำหนดข้อมูลเพิ่มเติมให้กับแฟ้มข้อมูลต่าง ๆ ที่สัมพันธ์กันไม่ครบถ้วน เช่น มีการเพิ่มข้อมูลพนักงานขายคนใหม่ชื่อ “สมบูรณ์” เฉพาะแฟ้มข้อมูล Salesman โดยไม่ได้เพิ่มข้อมูลลงในแฟ้มข้อมูล Employee เป็นต้น

Deletion anomaly

- เป็นการลบข้อมูลจากเพิ่มข้อมูลต่าง ๆ ที่สัมพันธ์กัน ไม่ครบถ้วน เช่น กรณีที่พนักงานชื่อ “กนกวรรณ” ลาออกแล้ว มีการลบข้อมูลของพนักงานคนนั้นเฉพาะเพิ่มข้อมูลพนักงาน โดยไม่ได้ลบข้อมูลของพนักงานคนอื่นเดียวกันในเพิ่มข้อมูลพนักงานขาย เป็นต้น

ฐานข้อมูล (Database system)

- ฐานข้อมูล หมายถึง ที่เก็บข้อมูลและความสัมพันธ์ระหว่างข้อมูลเหล่านั้น (A collection of data and relationships)
- จะเกี่ยวข้องกับ logical file มากกว่า physical file
- ยกตัวอย่างเช่น
 - การออกแบบฐานข้อมูลแบบ relational model โครงสร้างที่เห็นจะเป็นตาราง
 - ฐานข้อมูลที่มีโครงสร้างแบบ hierarchical model ผู้ใช้จะมองเห็นเป็น tree
 - ฐานข้อมูลที่มีโครงสร้างแบบ network model ผู้ใช้จะมองเห็นเป็น link list

ความถูกต้องของข้อมูล (Data integrity)

- Static integrity (State of transition)
 - เป็นความถูกต้องของเนื้อข้อมูล เช่น ผู้หญิงลาบวชไม่ได้ ผู้ชายลาคลอดไม่ได้ เป็นต้น
 - Dynamic integrity (State of transition)
 - เป็นความถูกต้องของลำดับการแก้ไข
- ความถูกต้องของข้อมูลจะถูกบังคับโดย integrity rule หรือ integrity constraints และไม่ควรถูกจัดการโดยโปรแกรม แต่ควรถูกจัดการโดยระบบจัดการฐานข้อมูล

ความเป็นอิสระของข้อมูล (Data independence)

- ความเป็นอิสระทางกายภาพ (Physical Data Independence)
 - การเปลี่ยนแปลงโครงสร้างข้อมูลระดับต่าง (Physical structure) จะไม่มีผลกระทบต่อโปรแกรม
 - ความเป็นอิสระทางตรรกภาพ (Logical Data Independence)
 - การเปลี่ยนแปลงโครงสร้างข้อมูลระดับกลางหรือระดับหลักการ (conceptual level) จะไม่มีผลกระทบต่อโปรแกรม
- ความเป็นอิสระของข้อมูลทำให้โปรแกรมสามารถเรียกใช้ข้อมูลได้ด้วยภาษาต่างกัน และยังสามารถเห็นข้อมูลได้หลายรูปแบบแตกต่างกัน บทที่ 3

คุณลักษณะที่ดีของฐานข้อมูล

- ลดความซ้ำซ้อนของข้อมูลให้เหลือน้อยที่สุด (Minimum Redundancy)
- ความถูกต้องสูงสุด (Maximum Integrity : Correctness)
- มีความเป็นอิสระของข้อมูล (Data Independence)
- มีระบบความปลอดภัยของข้อมูลสูง (High Degree of Data Security)
- การควบคุมจะอยู่ที่ส่วนกลาง (Logically Centralized Control)

องค์ประกอบของระบบฐานข้อมูล

- Data
 - ข้อมูลที่ถูกเก็บไว้ในระบบฐานข้อมูล รวมถึงความสัมพันธ์ระหว่างข้อมูลด้วย
- Hardware
 - คอมพิวเตอร์ อุปกรณ์เก็บข้อมูล และอุปกรณ์อื่น ๆ
- Software
 - โปรแกรมที่จัดการเกี่ยวกับฐานข้อมูล
- User
 - บุคคลที่เกี่ยวข้องกับฐานข้อมูล

บุคคลที่เกี่ยวข้องกับฐานข้อมูล

- ผู้บริหารฐานข้อมูล (Database Administrator หรือ DBA)
 - ทำหน้าที่ควบคุมและบริหารทรัพยากรฐานข้อมูล
 - ประยุกต์ใช้ฐานข้อมูลโดยความร่วมมือจากผู้เกี่ยวข้อง
 - สามารถประสานงานกับผู้บริหารระดับสูง ผู้ใช้แผนกต่าง ๆ และเจ้าหน้าที่ฝ่ายปฏิบัติการ
 - ควรมีความรู้ทั้งหลักการบริหารและเทคนิคของระบบจัดการฐานข้อมูล
 - หน้าที่
 - การออกแบบฐานข้อมูล
 - การปฏิบัติงานกับ DBMS

บุคคลที่เกี่ยวข้องกับฐานข้อมูล

- ผู้บริหารฐานข้อมูล (Database Administrator หรือ DBA)
- นักวิเคราะห์และออกแบบ (System Analyst)
- โปรแกรมเมอร์ (Programmer)
- ผู้ใช้ (End User)

ข้อเสียของระบบฐานข้อมูล

- มักมีขนาดใหญ่และราคาแพง
- ต้องอาศัยผู้ดูแลที่มีความรู้ความเข้าใจเทคโนโลยีระบบการจัดการฐานข้อมูล
- ถ้าระบบเสียจะทำให้มีผลต่อผู้ใช้หลายคน
- ความเป็นเจ้าของข้อมูลลดลง

หัวข้อ

บรรยาย

- Information คืออะไร? ต่างจาก Data อย่างไร?
- วิวัฒนาการของระบบสารสนเทศ
- ระบบเพิ่มข้อมูล
- ฐานข้อมูล
- ความถูกต้องและความเป็นอิสระของข้อมูล
- คุณลักษณะที่ดีของฐานข้อมูล
- องค์ประกอบของระบบฐานข้อมูล
- บุคคลที่เกี่ยวข้องกับฐานข้อมูล

วัตถุประสงค์รายครั้ง

นักศึกษาสามารถ

- อธิบายวิวัฒนาการของระบบสารสนเทศได้
- บอกความหมายของระบบเพิ่มข้อมูล และฐานข้อมูลได้
- อธิบายความถูกต้องและความเป็นอิสระของข้อมูลได้อย่างถูกต้อง
- บอกคุณลักษณะที่ดี องค์ประกอบ และบุคคลที่เกี่ยวข้องกับฐานข้อมูลได้อย่างถูกต้อง