

บทที่ 10 การประมวลผลข้อมูล

โดย

ร.ศ. พรรณีภา ไพบูลย์นิมิตร

การประมวลผล (Processing)

- การดำเนินการกับข้อมูลดิบ (Raw Data) เพื่อให้ได้ข้อมูลสารสนเทศ (Information)



ข้อมูลดิบ

- ข้อมูลที่รวบรวมมาจากแหล่งข้อมูลต่าง ๆ
- ตัวเลข ตัวอักษร รูปภาพ เสียง อุณหภูมิ ความเร็ว
- คะแนนสอบของนักศึกษา 1 กลุ่ม

ข้อมูลสารสนเทศ

- ข้อมูลที่ผ่านการประมวลผลแล้ว สามารถนำไปใช้ประโยชน์ได้
- ค่าเฉลี่ยคะแนนสอบของนักศึกษา 1 กลุ่ม

ประเภทของการประมวลผล

- **การประมวลผลด้วยมือ (Manual Data Processing)**
- **การประมวลผลอัตโนมัติ (Automatic Data Processing)**
 - **Online Processing**
 - **Batch Processing**

Online Processing

- **Real-time Processing**
- ระบบสำรองที่นั่งในเครื่องบิน
- **Time Sharing**
- เป็นการประมวลผลแบบโดยตรงหรือสุ่ม (Random)
- สื่อที่มีการเข้าถึงแบบโดยตรง เช่น ฮาร์ดดิสก์

Online Processing (ต่อ)

- ข้อมูลทันต่อเหตุการณ์
- สามารถทำรายงานได้ทันที
- I/O Device ต้องเชื่อมต่อกับ CPU ตลอดเวลา

Batch Processing

- **ไฟล์หลัก (Master Files)**
 - **ไฟล์รายการข้อมูลสินค้า**
- **ไฟล์รายการเปลี่ยนแปลง (Transaction Files)**
 - **รายการขายสินค้า**
 - **รายการซื้อสินค้า**

Batch Processing (ต่อ)

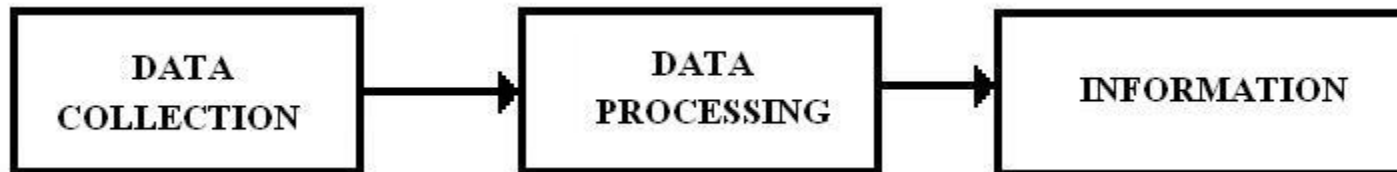
- ประมวลผลตามรายการเปลี่ยนแปลงทีละรายการ
- ประมวลผลทุกวันหรือทุกเดือน
- เป็นการประมวลผลแบบลำดับ (Sequential)
- สื่อที่มีการเข้าถึงแบบลำดับ เช่น เทปแม่เหล็ก
- ค่าใช้จ่ายต่ำ

Batch Processing (ต่อ)

- ข้อมูลที่ใหม่ทันต่อเหตุการณ์
- รายงานล่าช้าไม่ทันต่อเหตุการณ์
- I/O Device ไม่จำเป็นต้องเชื่อมต่อกับ CPU ตลอดเวลา

การประมวลผลด้วยคอมพิวเตอร์

- การเก็บข้อมูล (Data Collection)
- การประมวลผลข้อมูล (Data Processing)
- ผลลัพธ์ (Output / Information)



การเก็บข้อมูล

- วิธีการเก็บข้อมูล
- ลักษณะของข้อมูลที่ดี
- การเข้ารหัสข้อมูล
- การตรวจสอบความถูกต้องของข้อมูล
- การเก็บข้อมูลในสื่อบันทึกข้อมูล

วิธีการเก็บข้อมูล

- สังเกต (Observation)
- วัดค่า (Measurement)
- เอกสาร (Document)
- สัมภาษณ์ (Interview)

วิธีการเก็บข้อมูล

- **แบบสอบถาม (Questionnaire)**

- **Manual**

- **Internet**

แบบสอบถาม

ชื่อ	นาย ▼	<input type="text"/>	นามสกุล	<input type="text"/>
อายุ	ต่ำกว่า 20 ▼		ระดับการศึกษา	ต่ำกว่า ปวช. ▼
ตำแหน่ง	<input type="text"/>		อีเมล	<input type="text"/>
หน่วยงาน	<input type="text"/>		ประเภทหน่วยงาน	หน่วยงานราชการ ▼
ที่อยู่เลขที่	<input type="text"/>		ถนน	<input type="text"/>
แขวง/ตำบล	<input type="text"/>		เขต/อำเภอ	<input type="text"/>
Province	กรุงเทพมหานคร ▼		รหัสไปรษณีย์	<input type="text"/>
โทรศัพท์	<input type="text"/>		โทรสาร	<input type="text"/>

ลักษณะของข้อมูลที่ดี

- ถูกต้อง (Accuracy)
- ปัจจุบัน (Recency)
- สมบูรณ์ (Completeness)

การเข้ารหัสข้อมูล

- ประหยัดเนื้อที่ในการเก็บข้อมูล
- มีความเร็วในการค้นหาสูง
- ใช้ตัวเลขหรือรหัสแทนข้อความ
- การปรับอัตราส่วนของจำนวน

1 กรุงเทพ
2 เชียงใหม่
3 ลำพูน
....

หน่วย =*1000 บาท

การเข้ารหัสข้อมูล

ข้อมูลดิบ

นาย	สมชาย	คนไทย	เชียงใหม่	12,000 บาท
-----	-------	-------	-----------	------------

เข้ารหัส

ข้อมูลผ่านการเข้ารหัส

M	สมชาย	คนไทย	2	12
---	-------	-------	---	----

M แทน นาย
N แทน นางสาว
O แทน นาง

1 กรุงเทพฯ
2 เชียงใหม่
3 ลำพูน
....

หน่วย = *1000 บาท

การตรวจสอบความถูกต้องของข้อมูล

- **Edit** การแก้ไข
- **Verify** การทวนสอบ
- **Validate** ความสมเหตุสมผล
 - **Range Check** เช่น อายุ เป็น -10 ไม่ได้
 - **Relation Check** เช่น เป็นผู้หญิง คำนำหน้านามเป็นนาย ไม่ได้

การเก็บข้อมูลในสื่อบันทึกข้อมูล

- การเปลี่ยนสื่อบันทึกข้อมูล
- ข้อมูลจริง
- ข้อมูลในสื่อบันทึกข้อมูล

ข้อมูลจริง

ชื่อ	แผนก	ชั่วโมง	อัตรา
นายสมชาย คนไทย	คอมพิวเตอร์	12	60
นางสาวแพร มีบุญ	บัญชี	20	50
นางศศิวิมล คนดี	ประชาสัมพันธ์	5	80
...

M แทน นาย
N แทน นางสาว
O แทน นาง

1 คอมพิวเตอร์
2 บัญชี
3 ประชาสัมพันธ์
....

ข้อมูลในสื่อบันทึกข้อมูล

M สมชาย คนไทย	1	12	60
N แพร มีบุญ	2	20	50
O ศศิวิมล คนดี	3	5	80
...

ระเบียน (Record)

แอททริบิวต์ (Attribute)

การประมวลผลข้อมูล

- **คำนวณ (Calculating)**
- **เรียงลำดับ (Sorting)**
- **ปรับปรุง (Updating)**
- **ค้นหา (Searching)**
- **รายงาน (Reporting)**
- **แบ่งประเภท (Classification)**
- **สรุป (Summarizing)**

คำนวณ

- นำกลุ่มของแอททริบิวต์ที่มีอยู่มาคำนวณหาค่า
- ค่าเฉลี่ย (Mean)
- มัธยฐาน (Median)
- ฐานนิยม (Mode)

ปรับปรุง

- นำแอททริบิวต์ที่มีอยู่มาคำนวณ แล้วแทนที่
- การปรับค่าอัตราเงินเดือนของพนักงาน
- การปรับค่าเงินสุทธิของบัญชีธนาคาร

รายงาน - แสดงรายการทั้งหมด

ชื่อ	แผนก	อัตรา	ชั่วโมง	ค่าจ้าง
นายสมชาย คนไทย	คอมพิวเตอร์	60	12	720
นางสาวแพร มีบุญ	บัญชี	50	20	1,000
นางศศิวิมล คนดี	ประชาสัมพันธ์	80	5	400
...

Data หรือ ข้อมูลดิบ

Information หรือ สารสนเทศ
ได้จาก อัตรา x ชั่วโมง

สรุป- คำนวณและแสดงผลลัพธ์แบบรวมทุกรายการ

ชื่อ	แผนก	อัตรา	ชั่วโมง	ค่าจ้าง
นายสมชาย คนไทย	คอมพิวเตอร์	60	12	720
นางสาวแพรว มีบุญ	บัญชี	50	20	1,000
นางศศิวิมล คนดี	ประชาสัมพันธ์	80	5	400
...
รวมรายจ่ายทั้งหมด ข้อมูลเป็นประโยชน์ต่อเจ้าของกิจการ (คนจ่ายเงิน)			72	4,170

สรุป - แสดงข้อมูลสรุปแบบกลุ่ม

แผนก	อัตรา	ชั่วโมง	ค่าจ้าง
คอมพิวเตอร์	60	22	1,320
บัญชี	50	33	1,650
ประชาสัมพันธ์	80	15	1,200
		70	4,170

สรุปค่าจ้างแยกตามแผนก
เป็นประโยชน์ในการลงทุนขยายกิจการ

เรียงลำดับ

- ตัวเลข เช่น คะแนนสอบ
- ข้อความ เช่น ชื่อ
- รหัส เช่น รหัสสินค้า

ค้นหา

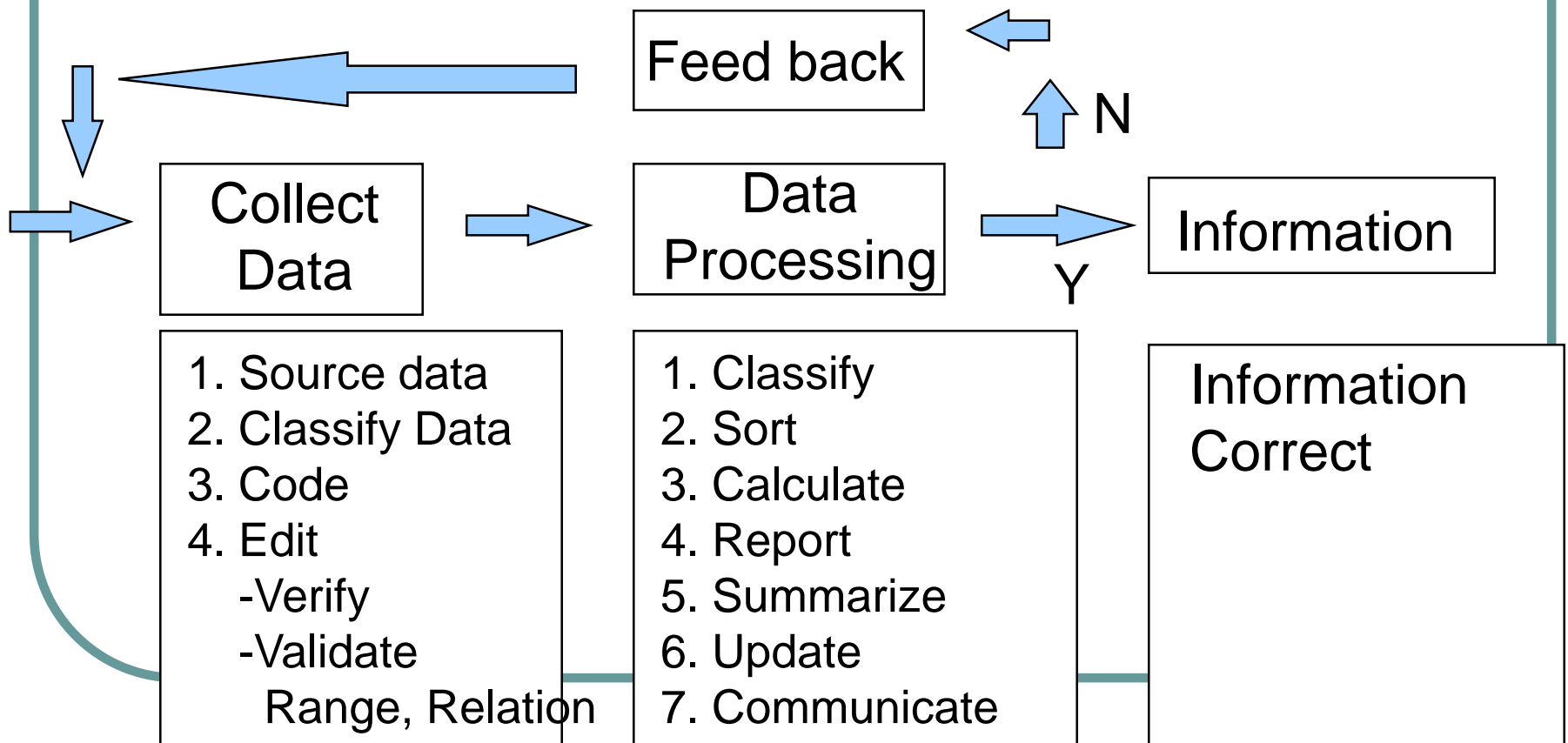
- ค้นหาระเบียบที่ต้องการจากเงื่อนไขที่กำหนด
- ค้นหานักเรียนที่ GPA < 1.75
- ค้นหาพนักงานที่ เงินเดือน ≤ 5,000 หรือ อยู่ฝ่ายบัญชี
- ค้นหาสินค้าที่ ราคา = 100 และ เป็นอาหาร

แบ่งประเภท หรือจัดกลุ่ม

- แบ่งระเบียบทั้งหมดออกเป็นกลุ่ม โดยที่ระเบียบในกลุ่มเดียวกันต้องเป็นประเภทเดียวกัน
- แบ่งพนักงานออกเป็นแผนก
- แบ่งนักศึกษาออกเป็นคณะหรือสาขา
- แบ่งสินค้าออกเป็นสินค้านำราคาแพงหรือราคาถูก

ภาพวงจรการประมวลผลด้วยคอมพิวเตอร์

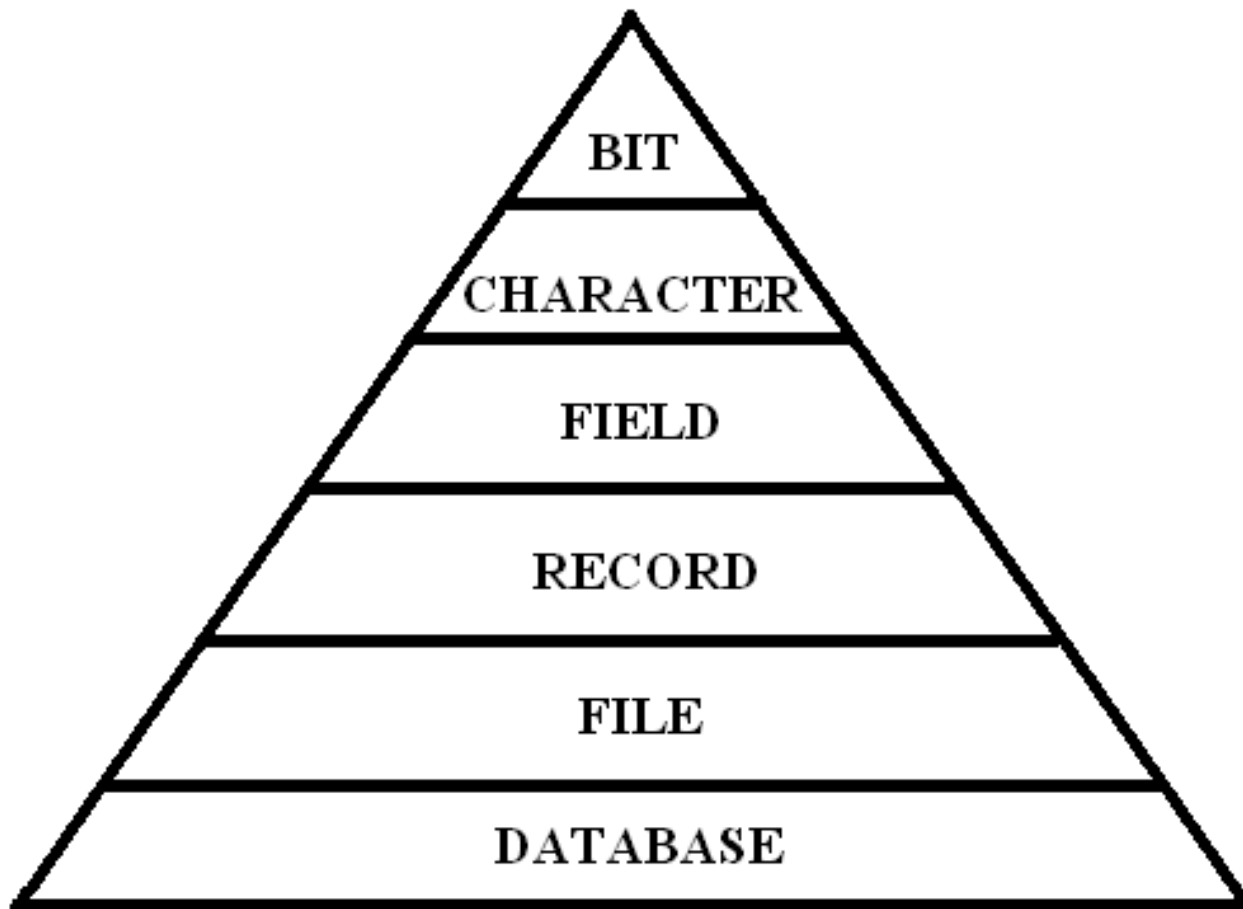
● Processing



องค์ประกอบข้อมูล (Data Organization)

- บิต (Bit)
- อักขระ (Character)
- ขอบเขตข้อมูล (Field)
- ระเบียบ (Record)
- แฟ้มข้อมูล (File)
- ฐานข้อมูล (Database)

องค์ประกอบข้อมูล (ต่อ)



บิต

- ตัวเลข 0 และ 1
- หน่วยพื้นฐานที่จัดเก็บในคอมพิวเตอร์

อักขระ

- กลุ่มของบิต เช่น 011101011
- 1 อักขระ นับเป็น 1 ไบต์
- โดยทั่วไป 1 ไบต์ (อักขระ) มี 8 บิต
- เช่น 01101011 แทน F เป็นต้น

ขอบเขตข้อมูล

- กลุ่มของอักขระ
- จำนวน 1,024
- ข้อความ Cat
- อักขระเดียว A

ระเบียบ

- กลุ่มของขอบเขตข้อมูลที่เกี่ยวข้องกัน
- นายพร มีบุญ แผนกบัญชี 20 ชั่วโมง 50 บาท

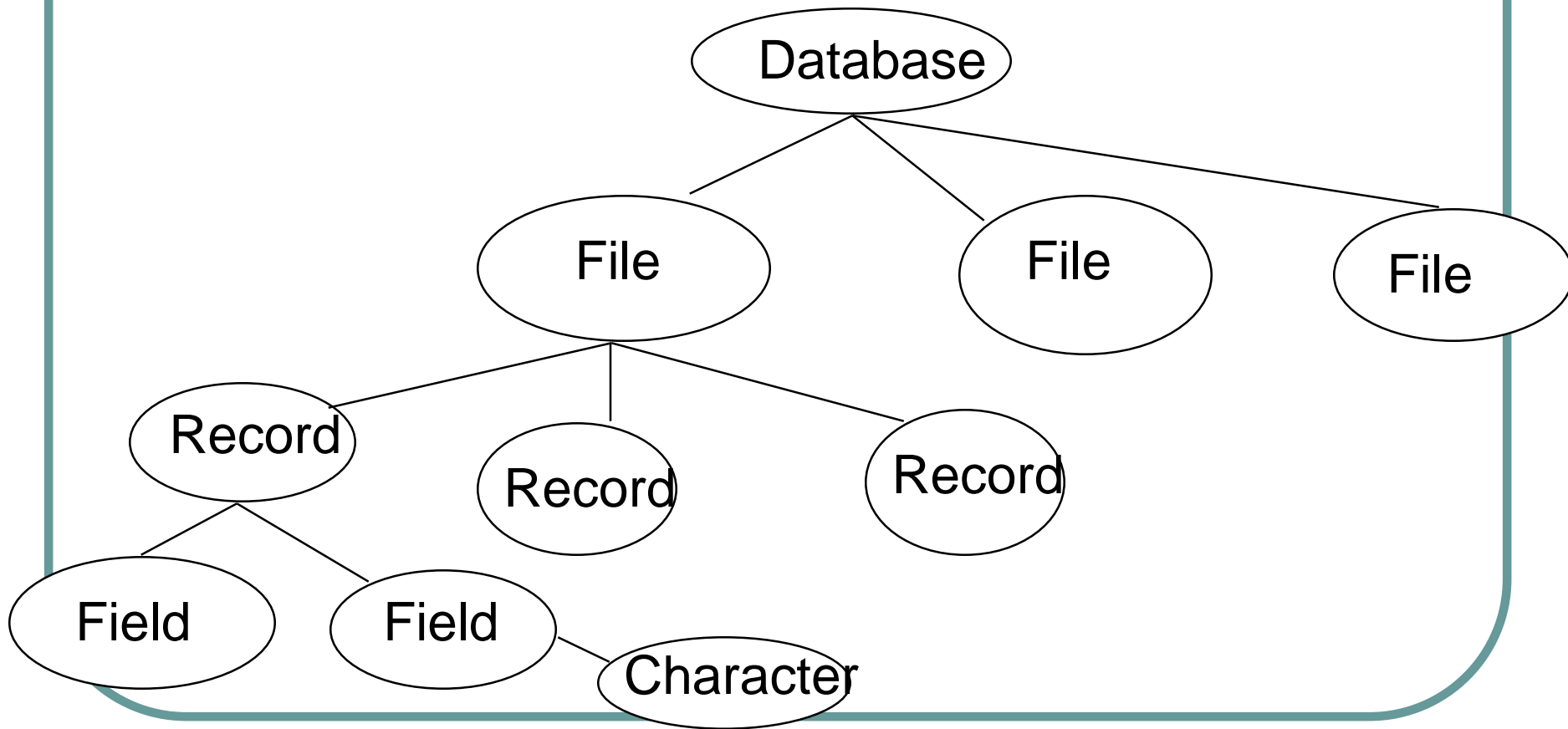
เพิ่มข้อมูล

กลุ่มของระเบียบที่มีรูปแบบเดียวกัน

เพิ่ม/ File

ชื่อ	แผนก	ชั่วโมง	อัตรา
นายพร มีบุญ Record	คอมพิวเตอร์	20 เรคคอร์ด	50
นางสาวศศิวิมล คนดี	บัญชี	5	80
นาง ปณิตดา ดวงคำ	ประชาสัมพันธ์	17	90
...	ฟิลด์/Field

- DATABASE



ฐานข้อมูล

● กลุ่มของแฟ้มข้อมูลที่สัมพันธ์กัน

ชื่อ	แผนก	ชั่วโมง	อัตรา
นายพร มีบุญ	คอมพิวเตอร์	20	50 File A
นางสาวศศิวิมล คนดี	บัญชี	5	80

แผนก	เวลาเริ่มงาน	จำนวนพนักงาน
คอมพิวเตอร์	12:00	52 File B
บัญชี	9:00	23

ลักษณะฐานข้อมูล

- ข้อมูลมีโครงสร้างที่แน่นอน
- มีชุดเดียวใช้ได้ทั้งหน่วยงาน
- เป็นอิสระจากโปรแกรม
- มีภาษาสอบถาม (Query Language)