

## เทคนิคในการสร้างคลังข้อมูล

- การเคลื่อนที่ของข้อมูลในคลังข้อมูล

ข้อมูลที่จัดเก็บภายในคลังข้อมูลมีการเคลื่อนที่ของข้อมูล (Information flow)

5 ประเภท ดังนี้

1. Inflow คือการนำข้อมูลจากฐานข้อมูลอื่นเข้าสู่คลังข้อมูลทั้งฐานข้อมูลภายในและภายนอกองค์กร โดยในขั้นนี้อาจมีการเปลี่ยนแปลงโครงสร้างข้อมูล การทำ denormalize การลบหรือการเพิ่มฟิลด์เพื่อให้ข้อมูลทั้งหมดอยู่ในเนื้อหาที่สนใจเดียวกัน ในขั้นตอนนี้อาจใช้เครื่องมือที่เรียกว่า data warehouse tool

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- การเคลื่อนที่ของข้อมูลในคลังข้อมูล
2. Upflow เมื่อข้อมูลที่เราต้องการอยู่ในคลังข้อมูลแล้ว ในบางครั้งอาจต้องมีการเพิ่มคุณค่าให้กับข้อมูลด้วยเพื่อให้ข้อมูลอยู่ในรูปแบบที่เป็นประโยชน์มากที่สุดต่อการนำเครื่องมือมาใช้ ซึ่งได้แก่การจัดกลุ่มข้อมูลหาค่าทางสถิติที่ซับซ้อน จัดข้อมูลให้อยู่ในรูปแบบหรือเทมเพลต มาตรฐาน

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- การเคลื่อนที่ของข้อมูลในคลังข้อมูล
3. Downflow เป็นขั้นตอนของการปรับปรุงเปลี่ยนแปลงข้อมูลเก่าและไม่อยู่ในเนื้อหาที่องค์กรสนใจออกไปจากคลังข้อมูลขององค์กร
  4. Outflow เป็นขั้นตอนที่ผู้ใช้เรียกใช้ข้อมูลในคลังข้อมูลผ่านเครื่องมือต่างๆ โดยการเรียกใช้อาจมีเพียงขอเรียกเป็นครั้งคราวเป็นประจำทุกวัน/เดือน หรือแม้กระทั่งต้องการแบบทันที

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- การเคลื่อนที่ของข้อมูลในคลังข้อมูล
5. Metaflow ข้อมูลที่จัดเก็บในคลังข้อมูลจะถูกทำข้อมูลไว้อีกชุดหนึ่ง เป็นแหล่งที่มาของข้อมูลนั้น หรือแม่กระทั่งที่อยู่ของข้อมูลนั้นในคลังข้อมูลและข้อมูลอื่นที่เกี่ยวข้อง

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล

วิธีการนี้ถูกเสนอโดย Kimball ในปี 1996 เรียกว่าระเบียบวิธี 9 ขั้น หรือ Nine-Step Methodology โดยวิธีการนี้เริ่มจากการออกแบบจากส่วนย่อยที่แสดงถึงแต่ละระบบงานขององค์กร หรือเรียกอีกอย่างหนึ่งว่าดาต้ามาร์ท (data mart) โดยเมื่อออกแบบแต่ละส่วนสำเร็จแล้ว จึงนำมารวมกันเป็นคลังข้อมูล ขององค์กรในขั้นสุดท้าย ซึ่งขั้นตอนทั้ง 9 ขั้นตอน มีรายละเอียดดังนี้

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล
  1. กำหนดดาต้ามาร์ท คือการเลือกที่จะสร้างดาต้ามาร์ทของระบบงานใดบ้าง และระบบงานใดเป็นระบบงานแรกโดยองค์กรจะต้องสร้าง E-R model ที่รวมระบบงานทุกระบบขององค์กรไว้ แสดงการเชื่อมโยงของแต่ละระบบงานอย่างชัดเจน

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล
  1. กำหนด dimension table ของดาต้ามาร์ท คือกำหนดเนื้อหาหลักที่ควรจะเป็นของดาต้ามาร์ท โดยการเลือกเอนทิตีหลักและกระบวนการที่เกี่ยวข้องกับเอนทิตีนั้นๆ ออกมาจาก E-R model ขององค์กร
  2. กำหนด fact table ของดาต้ามาร์ท คือกำหนดเนื้อหาหลักที่ควรจะเป็นของดาต้ามาร์ท โดยการเลือกเอนทิตีหลักและกระบวนการที่เกี่ยวข้องกับเอนทิตีนั้นๆ ออกมาจาก E-R model ขององค์กร

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล

3. กำหนดแอตทริบิวต์ที่จำเป็นในแต่ละ dimension table คือ การกำหนดแอตทริบิวต์ที่บอกหรืออธิบายรายละเอียดของ dimension ได้ ทั้งนี้แอตทริบิวต์ที่เป็น primary key ควรเป็นค่าที่คำนวณได้ กรณีที่มีค่าตัวมากกว่าหนึ่งค่าตัวมี dimension เหมือนกัน นั้นหมายถึงว่า แอตทริบิวต์ใน dimension นั้นจะต้องเหมือนกันทุกประการ



## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล
4. กำหนดแอตทริบิวต์ที่จำเป็นใน fact table โดยแอตทริบิวต์หลักใน fact table จะมาจาก primary key ในแต่ละ dimension table นอกจากนี้แล้ว ยังสามารถมีแอตทริบิวต์ที่จำเป็นอื่นๆ ประกอบอยู่ด้วย เช่น แอตทริบิวต์ที่ได้จากการคำนวณค่าเบื้องต้นที่จำเป็นสำหรับการคงอยู่ของแอตทริบิวต์อื่นใน fact table เรียกอีกอย่างหนึ่งว่า measure

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล
5. จัดเก็บค่าการคำนวณเบื้องต้นใน fact table คือการจัดเก็บที่ได้จากการคำนวณให้เป็นแอตทริบิวต์ใน fact table
  6. เขียนคำอธิบาย dimension table ทั้งนี้ก็เพื่อให้ผู้ใช้สามารถใช้งานตามาร্থได้อย่างมีประสิทธิภาพเพราะเกิดความเข้าใจอย่างดีในส่วนต่างๆ

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล
7. กำหนดระยะเวลาในการจัดเก็บข้อมูลในฐานข้อมูล โดยอาจจะเป็นการจัดเก็บเพียงช่วงระยะเวลา 1-2 ปี หรือนานกว่านั้นขึ้นอยู่กับความต้องการขององค์กร

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล
8. การติดตามปัญหาการเปลี่ยนแปลงของ dimension อย่างช้าๆ คือ การเปลี่ยนเอาแอตทริบิวต์ของ dimension table เก่ามาใช้ แล้วส่งผลกระทบต่อข้อมูลปัจจุบันของ dimension table โดยสามารถแบ่งประเภทของปัญหาที่เกิดขึ้นได้ 3 ประเภท ดังนี้
- 1 เกิดการเขียนทับข้อมูลใหม่โดยข้อมูลเก่า
  - 2 เกิดเรคอร์ดใหม่ๆ ขึ้นใน dimension
  - 3 เกิดเรคอร์ดที่มีทั้งค่าเก่าและใหม่ปนกันไป

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- วิธีการออกแบบฐานข้อมูลสำหรับคลังข้อมูล
9. กำหนดวิธีเป็นการออกแบบด้านกายภาพเพื่อให้ผู้ใช้เกิดความสะดวกในการทำงานและสามารถทำงานได้อย่างมีประสิทธิภาพ

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- การแปลงข้อมูลเข้าสู่ตลาดมาร์ท Extraction Transformation and Loading (ETL)

การแปลงข้อมูลเป็นสิ่งสำคัญมากสำหรับการสร้างคลังข้อมูล จะแตกต่างกันไปตามคลังข้อมูลของแต่ละองค์กร โดยการแปลงข้อมูลหมายถึงตั้งแต่การวิเคราะห์แหล่งข้อมูล กำหนดการส่งข้อมูลรวบรวมหรือสร้างข้อมูลภายนอก วางแผนและสร้างรูทีนของการแปลงข้อมูล และตรวจสอบความถูกต้องของข้อมูลที่ได้สามารถสรุปเป็นขั้นตอนได้ดังนี้

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- การแปลงข้อมูลเข้าสู่ดาต้ามาร์ท Extraction Transformation and Loading (ETL)
  1. วิเคราะห์แหล่งข้อมูล เช่น ปริมาณของข้อมูล จำนวนและชนิดของการเข้าถึงแหล่งข้อมูล แพลตฟอร์มและภาษาโปรแกรมที่ใช้ เป็นต้น
  2. ย้ายข้อมูลที่ต้องการจากระบบเดิมมาไว้ในบริเวณที่ใช้ปรับแต่งข้อมูล หรือเรียกบริเวณนี้ว่า staging area เพื่อนำมาเลือกเฉพาะส่วนที่ต้องการแปลงข้อมูลและตรวจสอบความถูกต้อง หรือการทำความสะอาดข้อมูล

## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- การแปลงข้อมูลเข้าสู่ดาต้ามาร์ท Extraction Transformation and Loading (ETL)
3. กำหนด primary key ของ fact table และ dimension table และกำหนด foreign key ระหว่าง fact table กับ dimension table
  4. ย้ายข้อมูลที่ทำความสะอาดแล้วจาก staging area ลงสู่ เซิร์ฟเวอร์ของดาต้ามาร์ท



## เทคนิคในการสร้างคลังข้อมูล (ต่อ)

- การแปลงข้อมูลเข้าสู่ดาต้ามาร์ท Extraction Transformation and Loading (ETL)
- 5. สร้าง metadata ของแต่ละดาต้ามาร์ท โดยเก็บรายละเอียดของข้อมูลการอัปเดตและส่งออกไว้ในดาต้ามาร์ท
- 6. ตรวจสอบความถูกต้องของข้อมูล ซึ่งจะต้องกระทำตลอดทั้งกระบวนการแปลงข้อมูล

## Data Warehousing (การคลังข้อมูล)

การคลังข้อมูล คือ การออกแบบ และสร้างโครงสร้างของข้อมูล  
ในคลังข้อมูล รวมถึงวิธีการได้มาของข้อมูลวิธีการสร้างผลลัพธ์  
จากข้อมูลที่มีวิธีการดูแลรักษา และวิธีการปรับปรุง  
ประสิทธิภาพ

## กระบวนการคลังข้อมูล

- แนวคิดการออกแบบและสร้างโครงสร้างของข้อมูล
- กระบวนการในการย้ายข้อมูล
- วิธีการเพื่อให้ได้มาซึ่งข้อมูล
- วิธีการวิเคราะห์ข้อมูล
- วิธีการจัดการกับข้อมูลที่ล้าสมัย

## กระบวนการจัดเก็บข้อมูล(Data Storing)

- กระบวนการจัดเก็บข้อมูล หมายถึง กระบวนการต่างๆที่ดำเนินการ เพื่อให้การจัดเก็บข้อมูลเป็นไปอย่างถูกต้อง สามารถบอกเล่าถึงธุรกิจขององค์กรได้ถูกต้องแม่นยำและมีประสิทธิภาพ

## กระบวนการเคลื่อนย้ายข้อมูล(Data Propagation)

- เป็นกระบวนการในการเคลื่อนย้ายข้อมูลจากองค์ประกอบหนึ่งไปยังอีกองค์ประกอบหนึ่ง เทคนิคที่ใช้กันในคลังข้อมูลเรียกว่า ELT (Extract Transform Load)

## กระบวนการเตรียมการเพื่อใช้ข้อมูล (Data Provisioning)

- เป็นกระบวนการเพื่อการออกแบบโครงสร้างข้อมูลใน Data Mart และการใช้ข้อมูลจาก Data Mart

## กระบวนการบริหารจัดการข้อมูล (Housekeeping)

- กระบวนการเพื่อจัดเตรียมพื้นที่ให้เพียงพอต่อการใช้งานตามช่วงเวลาที่เหมาะสม รวมถึงการจัดการข้อมูลที่ล้าสมัย

## กระบวนการเพื่อการควบคุมการทำงาน (Controlling)

- ระบบคลังข้อมูล จำเป็นต้องมีการควบคุมการทำงาน เพื่อที่จะใช้ข้อมูลประเภทพิเศษที่เรียกว่า Metadata

กระบวนการเตรียมข้อมูลสำหรับ

คลังข้อมูล

Data Preparation



## Data Preparation

- คือ กระบวนการสำหรับการนำข้อมูลเข้ามาสู่คลังข้อมูล และการตรวจสอบคุณภาพของข้อมูล

## Data Preparation

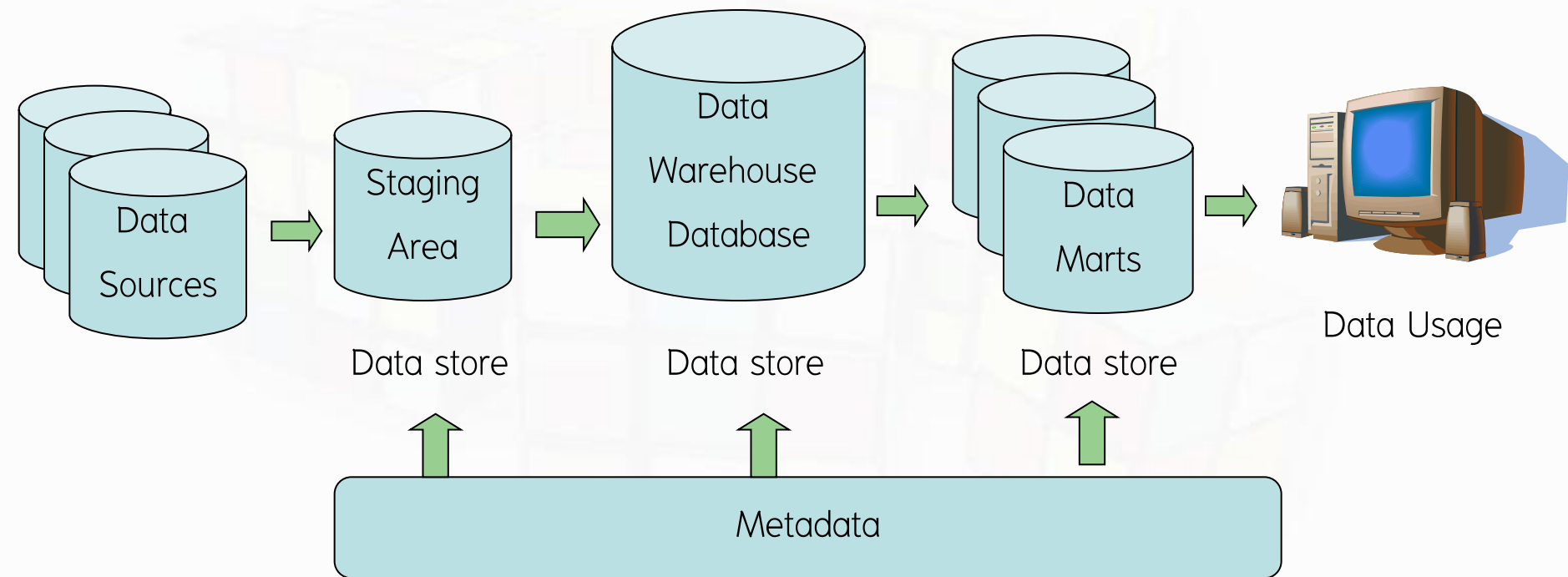
- กระบวนการ Data Preparation แบ่งออกเป็น 2 ประเภท คือ
  1. กระบวนการเพื่อการนำข้อมูลเข้าสู่คลังข้อมูล (Data Propagation)
  2. กระบวนการเพื่อการตรวจสอบคุณภาพของข้อมูล (Data Validation)

## Data Preparation

- กระบวนการ Data Preparation เป็นกระบวนการที่ทำหน้าที่แปลงและเคลื่อนย้ายข้อมูลจากที่หนึ่งไปยังอีกที่หนึ่งในคลังข้อมูล จะมีการเคลื่อนย้ายข้อมูล
  1. การแปลงและเคลื่อนย้ายข้อมูลจาก Data sources ไปยัง Data Warehouse Database
  2. การแปลงและเคลื่อนย้ายข้อมูลจาก Data Warehouse ไปยัง Data Marts

# Data Warehousing and Applications

Lect. Rattanawadee Panthong  
School of ICT



รูปรายละเอียดและลำดับการเกิดของกิจกรรมกระบวนการคลังข้อมูล

## Data Validation

- Data Validation เป็นกระบวนการเพื่อการตรวจสอบและตรวจรับคุณภาพของข้อมูล
  1. มีความถูกต้องของข้อมูล (Data Correctness)
  2. มีความสอดคล้องกันของข้อมูล (Data Consistency)
  3. มีความบริสุทธิ์สะอาดของข้อมูล (Data Cleanliness)

## Staging Areas

- Data Model เน้นให้สามารถสืบถึงเรื่องราวขององค์กรโดยผ่านการ
- นำเสนอความสัมพันธ์ระหว่างข้อมูลต่างๆ
- Data Acquisition File เน้นการสะดวกต่อการจัดส่ง ไม่ได้เน้นที่การ
- แสดงความเป็นจริงของข้อมูลเท่ากับ Data Model

## Staging Areas

- ภายใน Data Staging Area ประกอบด้วย การตรวจสอบความถูกต้องของข้อมูลในเบื้องต้น สำรองข้อมูล
- ข้อมูลจะถูกส่งเข้า Data Acquisition Subsystem ก่อนเป็นจุดแรก
- จากนั้นก็จะเข้าสู่ Data Staging Area เป็นจุดถัดไป

## Staging Areas

- Data Staging Area เป็นส่วนที่ติดต่อกับทั้ง Data Acquisition และ Data Warehouse Database
- การตรวจสอบความถูกต้องสอดคล้องกันระหว่างข้อมูลที่ได้รับมา กับข้อมูลที่มีอยู่ใน Data Warehouse Database สามารถทำได้ในภายใน Data Staging Area ในทางกลับกัน Data Acquisition System สามารถทำได้เพียงการตรวจสอบความถูกต้องของข้อมูลได้เบื้องต้นเท่านั้น ไม่สามารถตรวจสอบใน Warehouse database ได้





## Staging Areas

Data Acquisition → Data Staging Area → Data Warehouse Database



## กิจกรรมต่างๆที่เกิดขึ้นบน Data Staging Area

- การตรวจสอบความถูกต้องของข้อมูลก่อนการ Load เข้าสู่ Data Warehouse Database
- ความถูกต้องในแง่ของการมีค่าของข้อมูล (Data Consistency) ความถูกต้องในแง่ของค่าต่างๆที่เป็นไปได้ของข้อมูล (Possible Values) ความถูกต้องในแง่ของความสัมพันธ์ของข้อมูล (Data Relationship)
- การทำหน้าที่ Temporary Backup

## ความถูกต้องในแง่ของการมีค่าของข้อมูล (Data Consistency)

- Data Staging Area จะทำหน้าที่ตรวจสอบว่าข้อมูลที่เข้ามา มี Cardinality ของ Field ต่างๆ ตรงตามที่กำหนดไว้ใน Data Model หรือไม่
- พิจารณา Mandatory หรือ Optional

## ความถูกต้องในแง่ของค่าต่างๆที่เป็นไปได้ของ ข้อมูล (Possible Values)

- ความถูกต้องในแง่ของค่าต่างๆที่เป็นไปได้ของข้อมูล
- (Possible Values)
- ค่าที่เป็นไปได้ของข้อมูล (Possible Values) คือ ข้อจำกัดของค่าที่จะมี อยู่ใน Field ใด Field หนึ่งของข้อมูล

## ความถูกต้องในแง่ของค่าต่างๆที่เป็นไปได้ของ ข้อมูล (Possible Values)

แบ่งออกได้เป็น 2 ประเภท

- Universal Possible Value มีอยู่หลายชนิด เช่น วันที่ เดือน สกุลเงิน ประเทศ
- System-Based Values เป็นค่าที่เป็นไปได้ต่างๆที่ระบบจำเป็นต้องมี เช่น
- ชนิดสินค้าต่างๆที่จำหน่ายในร้านค้า อาจจะมีต่างๆกันไป ตามแต่ละองค์กร

## ความถูกต้องในแง่ของค่าต่างๆที่เป็นไปได้ของ ข้อมูล (Possible Values)

- Data Staging Area จะมีหน้าที่ในการตรวจสอบค่าของข้อมูลที่ได้รับมาจาก Data Acquisition Subsystem ว่ามีค่าที่ตรงกับค่าใดที่ระบุไว้ใน Possible Value หรือไม่

## ความถูกต้องในแง่ของความสัมพันธ์ของข้อมูล (Data Relationship)

- ต้องทำหน้าที่ตรวจสอบความถูกต้องในความสัมพันธ์ เช่น ข้อมูลตัวหนึ่งมีความสัมพันธ์กับข้อมูลอีกตัวหนึ่งอย่างไรบ้าง ตรงกับที่ได้ทำในแบบจำลองหรือไม่
- ยกตัวอย่างความสัมพันธ์แบบ Foreign key เป็นต้น

## Staging area (ต่อ)

- Staging area เป็นกระบวนการในการเตรียมข้อมูลก่อนที่จะนำเข้า warehouse data repository
- Staging area หรือ ETL เป็นกระบวนการที่ประกอบด้วยขั้นตอนย่อย ดังนี้
  1. Extraction = การดึงข้อมูลจากฐานข้อมูลหรือไฟล์ข้อมูลในระบบ โดยใช้เครื่องมือในการดึงข้อมูลอย่างง่าย คือ SQL



## Staging area (ต่อ)

2. Transformation = การแปลงข้อมูลหลายระบบให้อยู่ใน format ของ data warehouse ก่อนข้อมูลจะถูก load เข้า data warehouse ซึ่งขั้นตอนนี้จะรวมถึงกระบวนการทำงานดังนี้
  - 2.1 Cleaning = การทำให้ข้อมูลเข้าถูกต้อง
  - 2.2 Standardizing = การทำให้ข้อมูลอยู่ใน format ของ data warehouse ให้เป็นรูปแบบเดียวกัน ไม่ว่าจะเป็นอย่างใด หน่วยที่ใช้วัด

## Staging area (ต่อ)

2.3 Reformatting = การทำให้ข้อมูลจากภายนอกองค์กรที่อยู่ในรูปแบบอื่นๆ เข้าอยู่ในรูปแบบที่เก็บใน warehouse repository

2.4 Summarizing = การสรุปสารสนเทศได้หลายรูปแบบ

3. Loading = การโหลดข้อมูลเก็บในหัวข้อต่างๆ เข้า data warehouse

## การทำหน้าที่ Temporary Backup

- เมื่อข้อมูลมาถึงคลังข้อมูล จะมีการรอนำเข้าข้อมูล ETL (Extract – Transform – Load ) กระบวนการนี้ใช้เวลาาน ยืงหากมีระบบใหญ่ Data Staging Area จะทำหน้าที่เป็นส่วนสำรองข้อมูลชั่วคราว (Temporary Backup) เพื่อสำรองข้อมูลขณะที่กระบวนการนำเข้า

## การทำหน้าที่ Temporary Backup

- ข้อมูลดำเนินอยู่ถ้าหากเกิดความผิดพลาดในกระบวนการนำเข้าข้อมูล ระบบจะเริ่มต้นกระบวนการนำเข้าใหม่ด้วยข้อมูลที่สำรองไว้และเมื่อกระบวนการเสร็จสิ้น ข้อมูลสำรองจะถูกกำจัดออกจาก Data Staging Area

## การทำหน้าที่ Temporary Backup

- การแบ่ง Data Area และ Back up Area อาจจะเป็น Disk ตัวเดียวกันหรือแยกกันคนละตัวก็ได้
- แต่เพื่อความปลอดภัยมากที่สุด ควรมีการแยกออกเป็น Disk คนละตัวมากกว่า เพราะมีโอกาสที่จะเกิดความผิดพลาดขึ้นกับ Disk ทั้งหมดได้ซึ่งอาจจะทำให้ Data Area และ Back up Area เสียหายทั้งคู่ หากอยู่ใน Disk เดียวกัน

## แนวทางและทางเลือกสำหรับการออกแบบ Data Staging Area กับกระบวนการ ETL

กระบวนการ ETL ประกอบไปด้วย 3 กระบวนการคือ

- Extract คือกระบวนการในการดึงข้อมูลออกจาก Source
- Transform คือกระบวนการแปลงข้อมูลจากโครงสร้างเดิมที่กำหนดไว้ใน Source ให้อยู่ในรูปแบบโครงสร้างตามที่ได้กำหนดใน Destination

## แนวทางและทางเลือกสำหรับการออกแบบ

### Data Staging Area กับกระบวนการ ETL

กระบวนการ ETL ประกอบไปด้วย 3 กระบวนการคือ

- Load คือ การนำข้อมูลที่เปลี่ยนแปลงรูปแบบแล้วไปเก็บไว้ใน Destination (Data Warehouse)

## แนวทางและทางเลือกสำหรับการออกแบบ Data Staging Area กับกระบวนการ ETL

กระบวนการ ETL มีสองแนวทางให้เลือก

- แนวทางแรกคือ การออกแบบโครงสร้างข้อมูลให้มีลักษณะแบบเดียวกันกับโครงสร้างข้อมูลของ Data Acquisition System



## แนวทางและทางเลือกสำหรับการออกแบบ

### Data Staging Area กับกระบวนการ ETL

กระบวนการ ETL มีสองแนวทางให้เลือก

- แนวทางที่สองคือ การออกแบบโครงสร้างข้อมูลให้มีลักษณะแบบเดียวกับ Data Warehouse Database การออกแบบของ Data Staging Area ทั้งสองแบบจะส่งผลกระทบต่อการออกแบบกระบวนการทำงานของ ETL

## แนวทางและทางเลือกสำหรับการออกแบบ Data Staging Area กับกระบวนการ ETL

- ทั้งสองแนวทางจะเกิด ETL ขึ้นสองครั้งถ้าหากออกแบบ Data Staging Area ให้รับโครงสร้างข้อมูลแบบ Data Acquisition System ระบบจะต้องเสียเวลามากขึ้นในการย้ายข้อมูลจาก Data Staging Area เข้าไปยัง Data Warehouse Database

## แนวทางและทางเลือกสำหรับการออกแบบ Data Staging Area กับกระบวนการ ETL

- ถ้าหากออกแบบ Data Staging Area ให้รับโครงสร้างข้อมูลแบบ Data Warehouse Database ระบบจะต้องเสียเวลามากขึ้นในการย้ายข้อมูลจาก Data Acquisition System เข้าไปยัง Data Staging Area

## แนวทางและทางเลือกสำหรับการออกแบบ Data Staging Area กับกระบวนการ ETL

- แต่หากระบบที่มีข้อมูลเข้าที่มีปริมาณมากและความถี่สูง แต่ไม่มีความซับซ้อนในการตรวจสอบมากนัก ควรมีการออกแบบโครงสร้างข้อมูลให้มีลักษณะแบบเดียวกับ Data Acquisition System เพื่อลดปัญหาการรอเข้ามาของข้อมูล ซึ่งก่อให้เกิดปัญหาการเกิดคอขวดในการเข้ามาของข้อมูลได้