

# Chapter 1 Introduction

13/7/64

นาย ปิณฑิตา เกษมศักดิ์ศิริ  
623020760 - 0

## Data and Information Systems (DAIS) Course Structures at CS/UIUC

- Coverage: Database, data mining, text information systems, Web and bioinformatics
- Data mining
  - Intro. to data warehousing and mining (CS412)
  - Data mining: Principles and algorithms (CS512)
- Database Systems:
  - Intro. to database systems (CS411)
  - Advanced database systems (CS511)
- Text information systems
  - Text information system (CS410)
  - Advanced text information systems (CS510)

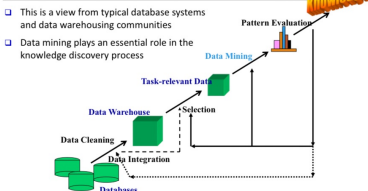
## Why Data Mining?

- The Explosive Growth of Data: from terabytes to petabytes
  - Data collection and data availability
  - Automated data collection tools, database systems, Web, computerized society
- Major sources of abundant data
  - Business: Web, e-commerce, transactions, stocks, ...
  - Science: Remote sensing, bioinformatics, scientific simulation, ...
  - Society and everyone: news, digital cameras, YouTube
- We are drowning in data, but starving for knowledge!
- "Necessity is the mother of invention"—Data mining—Automated analysis of massive data sets

## \* Why Data Mining?

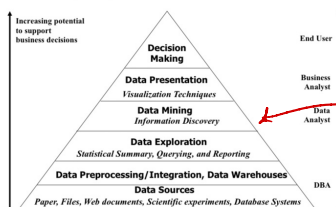
- the explosive Growth of data: from terabytes to petabytes.  
= มีพหุคูณจาก terabytes to petabytes
- มีแนวโน้มรวมข้อมูล ทล: การมีอยู่ใช้งานของข้อมูล

## Knowledge Discovery (KDD) Process



\* มีที่มาจาก: ขบวนการจากข้อมูลทั่วไป ทล: คลังข้อมูล  
การหาข้อสรุปเชิงสถิติจากข้อมูลในคลังข้อมูล  
พบความรู้

## Data Mining in Business Intelligence



ผู้ให้ข้อมูลทั้งหมดที่ถูกประมวลให้แล้วมา Analysis ต่อ  
งานงานขึ้น สืบต่อไป ทล: เสนอ Presence งาน

## Data Mining Functions

ข้อมูลทล: มีสร้าง คลังข้อมูล  
Ex.1 ทล: คลังข้อมูล ทล: พบบ  
การรวม ทล: โมเดลข้อมูล  
ทล: วิเคราะห์

- (1) Generalization
- (2) Pattern Discovery
- (3) Classification