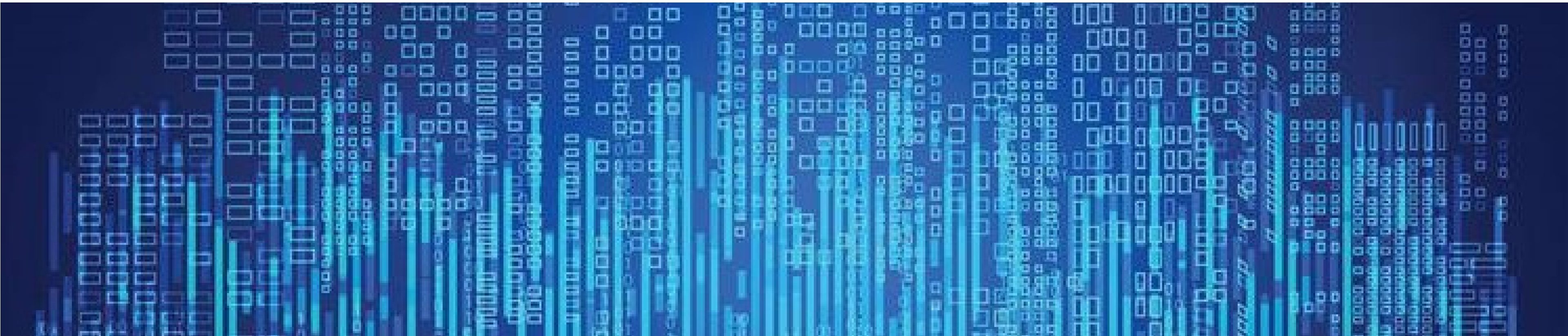


CLASS INF-305

PENDAHULUAN: DATA MINING

Prof. Dr. Taufik Fuadi Abidin, S.Si., M.Tech
Alim Misbullah, S.Si., M.S.

Department of Informatics - FMIPA -Universitas Syiah Kuala

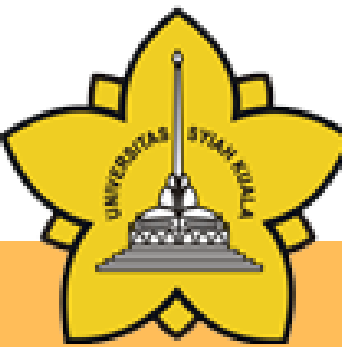




Taufik Fuadi Abidin






**Profesor at the Department of Informatics and
Chairman of the Institute for Research and Community Services
Universitas Syiah Kuala**

www.fsd.unsyiah.ac.id/tfa
taufik.abidin@unsyiah.ac.id





Source: <https://datafloq.com/read/business-benefits-of-data-mining/6210>

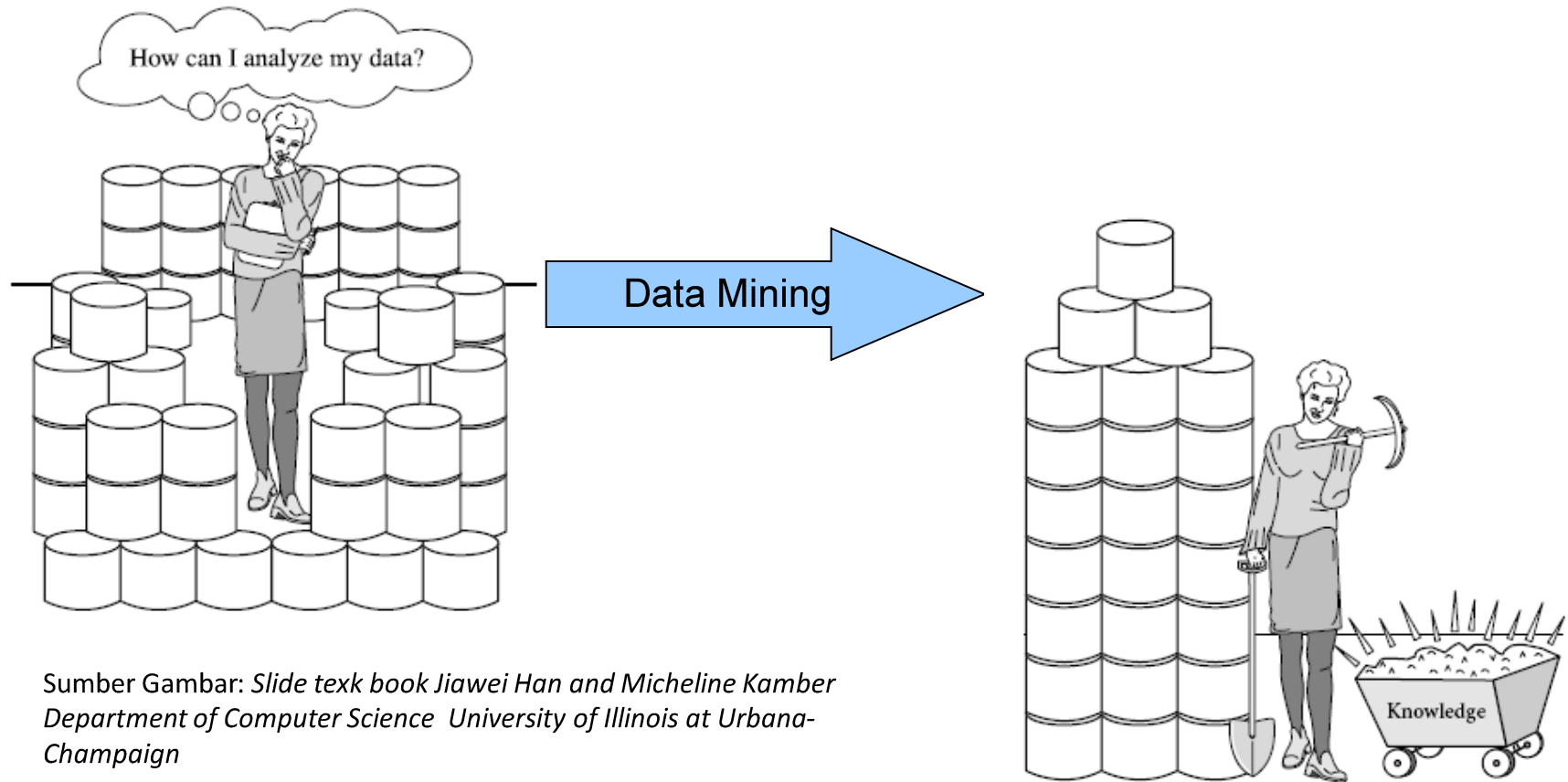
-  Pendahuluan Data Mining
-  Mengapa Diperlukan
-  Jenis Data
-  Fungsionalitas Data Mining
-  Metode-Metode yang Populer
-  Isu-Isu Utama



Data Mining

- Data Mining berkaitan dengan pengolahan data dalam jumlah besar. Saat ini, data telah mencapai ukuran terabyte sampai dengan petabyte
- Pengumpulan data saat ini sudah dapat dilakukan secara otomatis diantaranya data image (satelit), database, web, bioinformatics, remote sensing, simulasi, kamera digital, dan berita online
- Informasi dalam data-data itu perlu diolah dan dianalisa agar bermanfaat dan dapat dijadikan acuan dalam pengambilan keputusan

Ilustrasi Data Mining

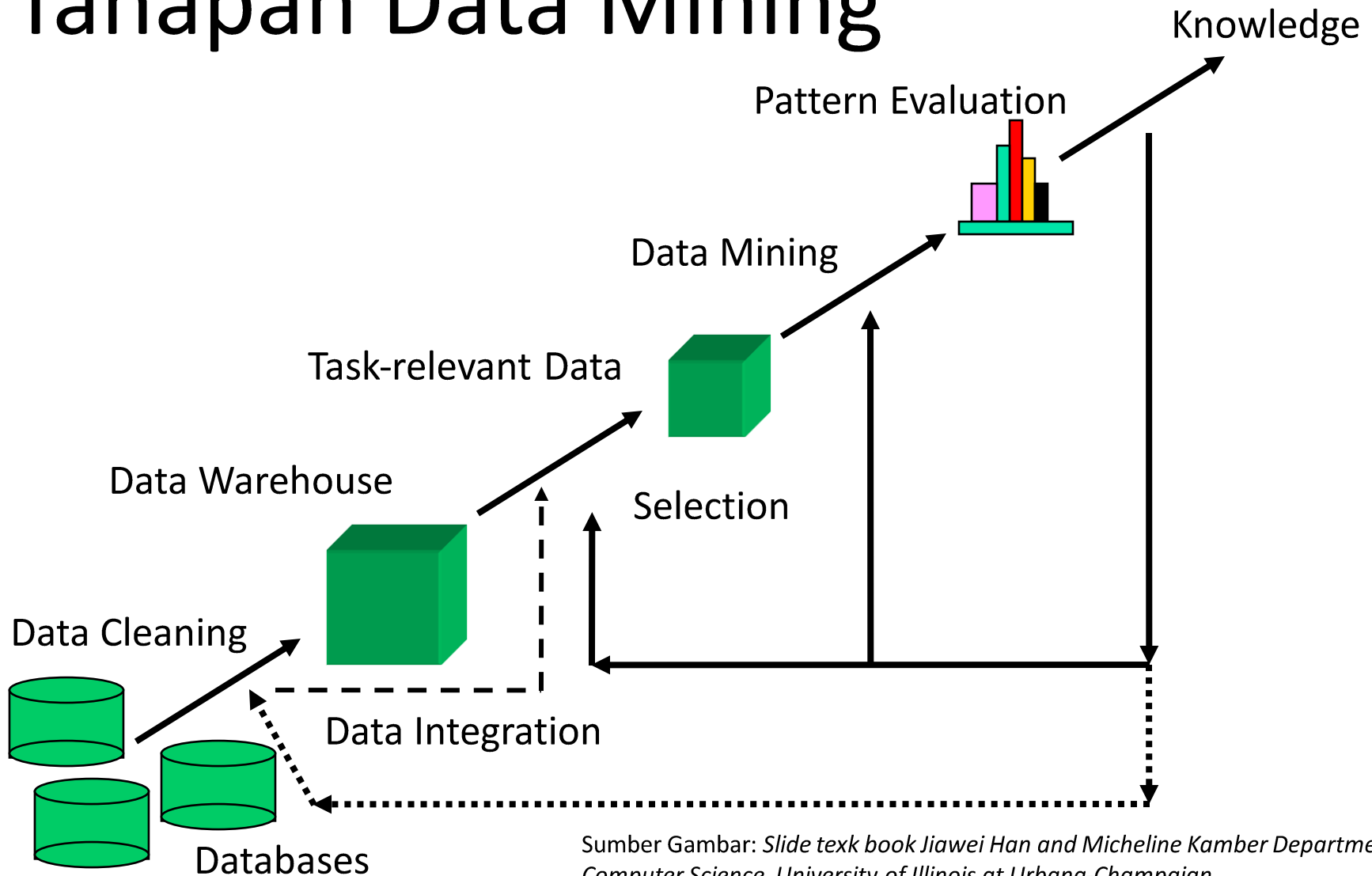


Sumber Gambar: Slide text book Jiawei Han and Micheline Kamber
Department of Computer Science University of Illinois at Urbana-
Champaign

Definisi Data Mining

- Data Mining juga dikenal dengan sebutan *Knowledge Discovery from Data* adalah proses mengekstrak pola (*pattern*) atau informasi (*knowledge*) dari sejumlah data berukuran besar
- Menurut Jiawei Hans, alternatif nama Data Mining adalah: *knowledge extraction, data/pattern analysis, data archeology, information harvesting, business intelligence*

Tahapan Data Mining



Sumber Gambar: Slide text book Jiawei Han and Micheline Kamber Department of Computer Science University of Illinois at Urbana-Champaign

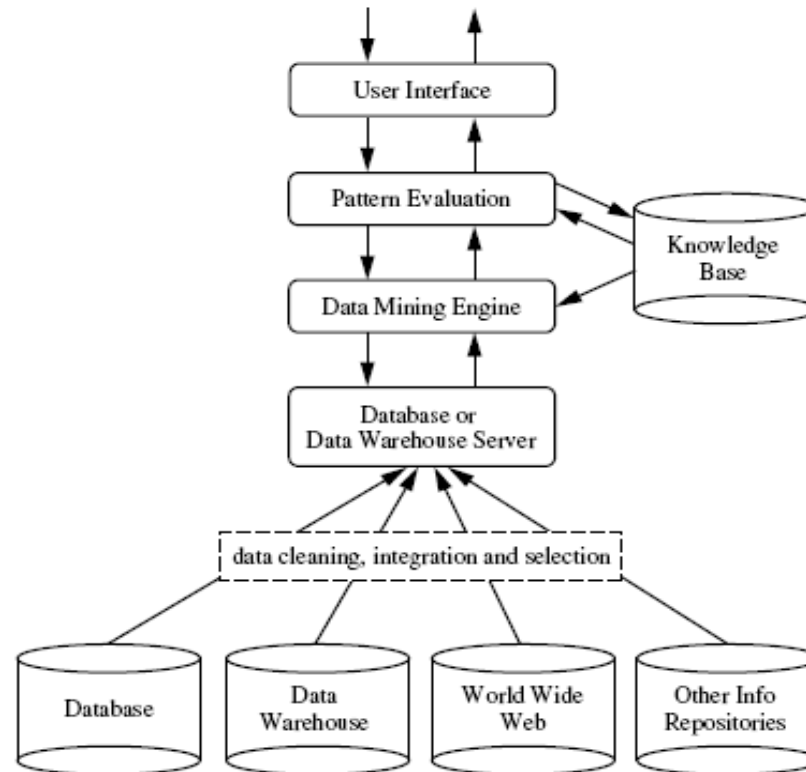
Tahapan Data Mining

- Pembersihan Data (*Data Cleaning*) yaitu proses menghilangkan noise dan data yang tidak konsisten
- Integrasi Data adalah proses menggabungkan data dari berbagai sumber membentuk sebuah data warehouse
- Seleksi Data adalah proses pengumpulan data-data yang relevan untuk dianalisa
- Transformasi Data adalah perubahan data menjadi dalam format yang sesuai
- Evaluasi pola untuk mengidentifikasi pola yang menarik

Tahapan Data Mining

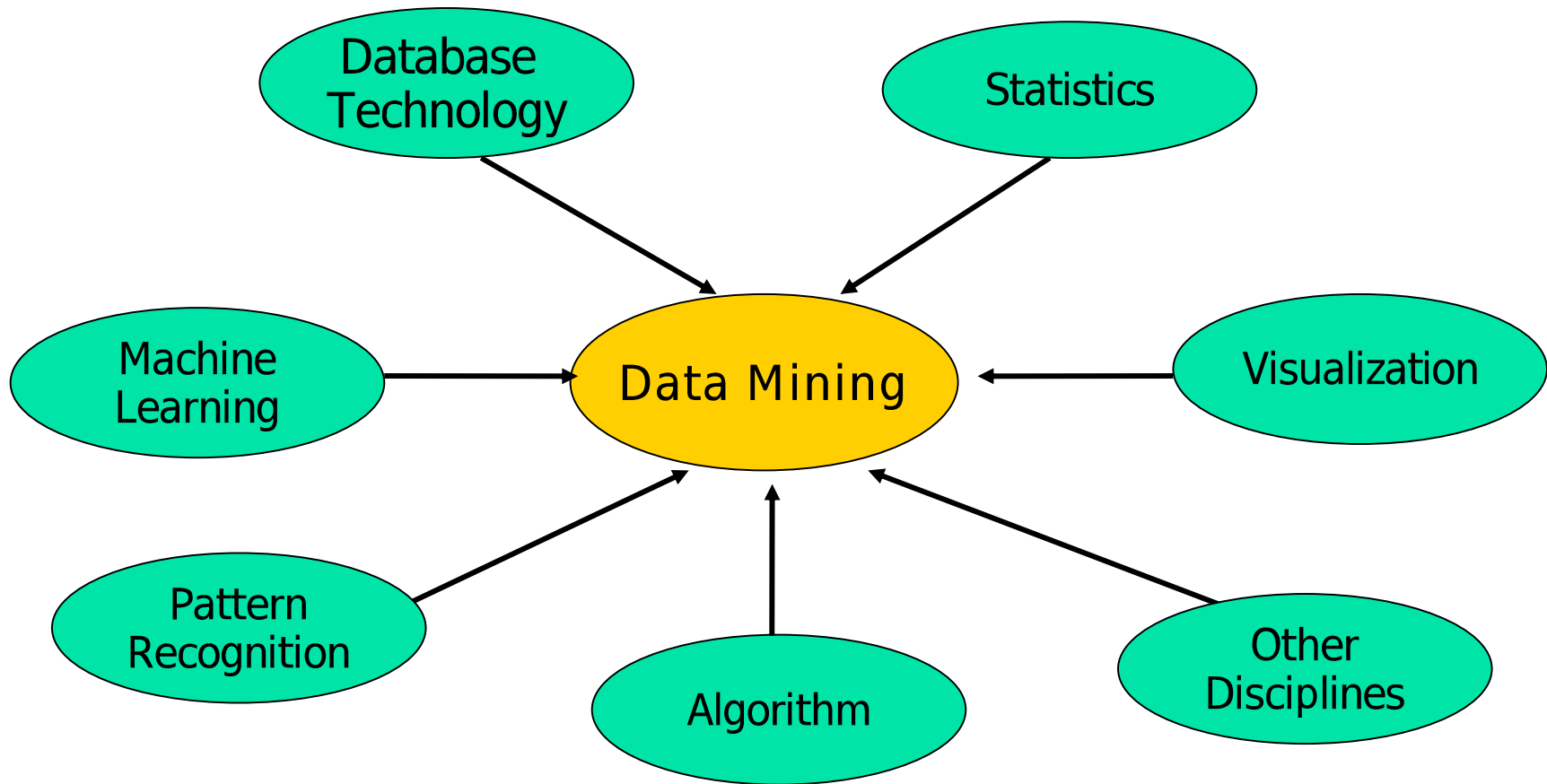
- Data Mining adalah penerapan metode yang pintar untuk mengekstrak pattern atau pola yang menarik
- Presentasi Informasi adalah visualisasi informasi untuk pengguna

Data Mining Sistem



Sumber Gambar: Slide text book Jiawei Han and Micheline Kamber Department of Computer Science University of Illinois at Urbana-Champaign

Melibatkan Banyak Disiplin



Sumber Gambar: Slide text book Jiawei Han and Micheline Kamber
Department of Computer Science University of Illinois at Urbana-
Champaign

Mining Berbagai Jenis Data

- Streams data dan data sensor
- Data time-series data, data temporal
- Data berstruktur, graphs, data jaringan sosial dan data yang terkoneksi (Web)
- Database dengan berbagai data (heterogeneous database)
- Data spatial dan data spatiotemporal
- Database multimedia database
- Text database dan WWW

Fungsionalitas Data Mining

- Klasifikasi dan Prediksi
- Clustering (Pengelompokan)
- Pola yang sering muncul (*frequent pattern*)
- ARM (Association Rule *Mining*)
- Pendeteksian Pencilan (*outlier*)
- Analisis Trend
- Analisa berdasarkan kesamaan, misalnya mendeteksi kesamaan halaman WWW

Questions & Discussion





TERIMA KASIH

—

