

CME2202  
Data Organization & Management

***Lecture #1: Welcome to Class***

**Asst.Prof.Dr. Özlem AKTAŞ**  
2023 Spring

# Teaching Team

Asst.Prof.Dr. Özlem AKTAŞ

ozlem@cs.deu.edu.tr

Res.Asst.Dr. Göksu TÜYSÜZOĞLU

goksu@cs.deu.edu.tr

Res.Asst.Dr. İbrahim Atakan KUBİLAY

atakan@cs.deu.edu.tr



# Objectives

The main objectives of this course are to discuss and teach

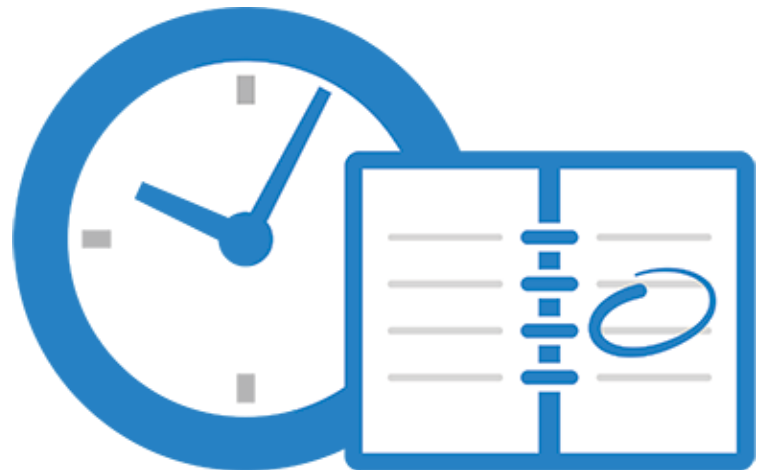
- Data, information and Knowledge triple,
- Data and Information representation,
- Organization and processing of files (file processing techniques),
- Physical characteristics of storage media,
- Indexing structures for fast access to data,
- Introduction to DBMS and SQL.

# Prerequisites

1. Students who wants to take this course must have taken and pass the course CME2201 Data structures and Algorithms. Some of the techniques will be discussed based on the fundamental data structure.
2. Assignments requires coding in C and Python programming language. A short tutorial will be given in lab sessions for both programming language.

# Weekly Schedule

- Lecture teoric session
  - Monday  
1. and 2. Groups: 13:00 – 15:00
- Lab practice session
  - Wednesday  
1. Group: 09:00 – 10:30  
2. Group: 10:30 – 12:00



# Assignments

- During the course term, 2 assignments will be given.
- All the assignments must be meticulously documented and must be returned in time.
- No late assignments will be accepted!
- However, assignments must be completed and submitted to take final exam.

# Academic dishonesty

- You can discuss the implementations and details of the assignments with your friends but always remember that assignments must be your **individual** work, otherwise stated.
- Cheating is **strictly prohibited !!!**
- If any cheating occurs, both of the works/assignments **will be graded with zero**. If it repeats department administration will be informed.

# Grading Policy

Midterm	20%
Assignments	30%
Final Exam	50%





# Weekly Plan

Topic	
1	Welcome to DOM
2	Data-Information-Knowledge
3	Data Representation
4	XML Technologies
5	Storage devices & Basic File Ops
6	Binary Search, Simple Indexing
7	B+tree
8	Multidimensional Indexing
9	Hashing
10	External Sorting
11	Introduction to DBMS
12	Data Normalizations & SQL
13	Database Design

# How to succeed in this course

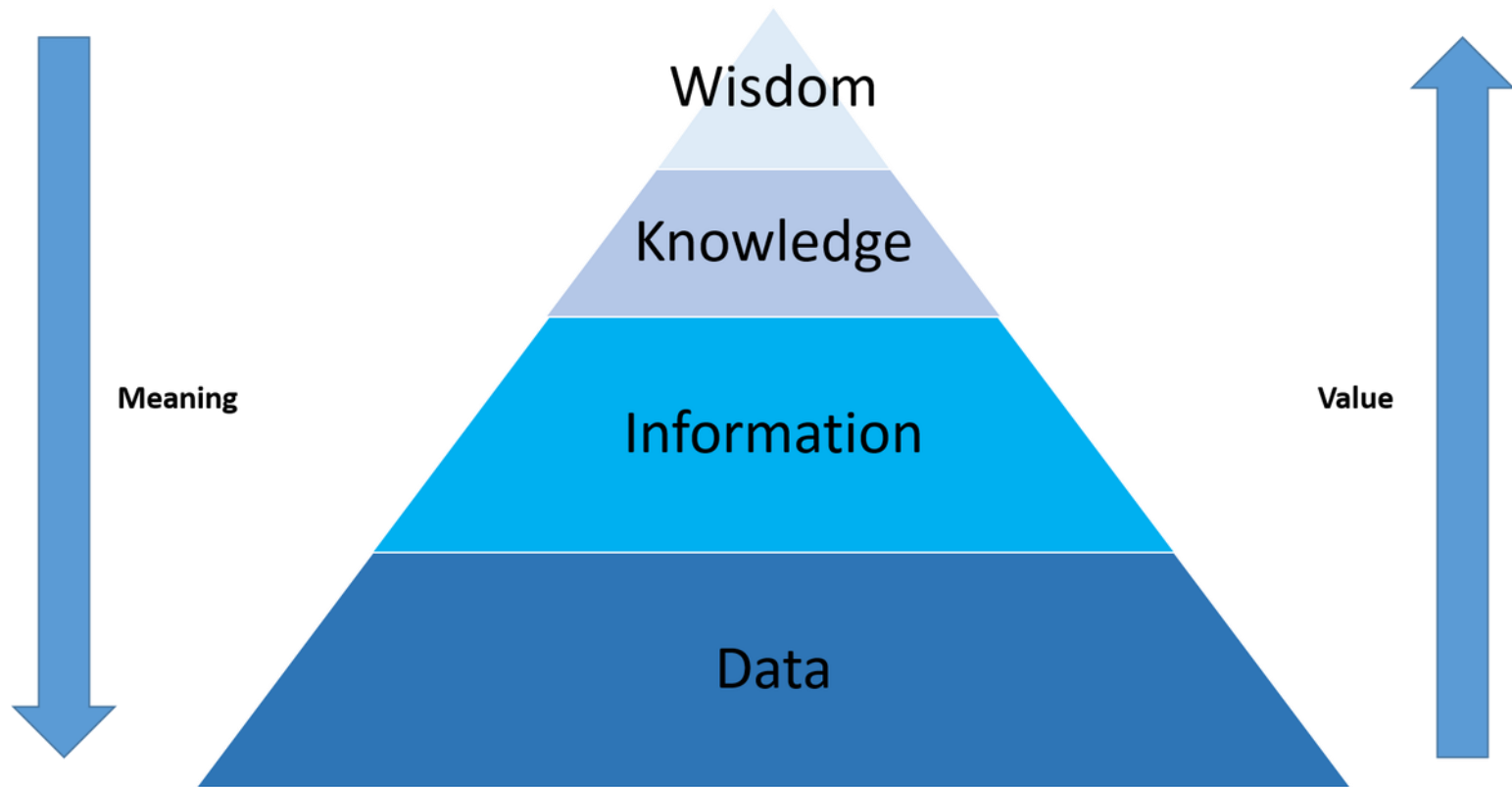
- Attend and **participate** classes. Ask question whenever you find something unclear.
- **Don't miss** the lab session.
- Do the **assignments** in time. Don't wait till the last night to start before submission!!
- Ask **help** whenever you don't understand something if you studied enough.

CME2202  
Data Organization & Management

# ***Lecture #2: Data, Information, Knowledge, Wisdom (DIKW) Model***

Asst.Prof.Dr. Özlem AKTAŞ  
2023 Spring

# Data, Information, Knowledge, Wisdom (DIKW) Model



- We frequently hear the words **Data**, **Information** and **Knowledge** used as if they are the same thing.
- **Data** represents unorganized and unprocessed facts.
  - Usually data is static in nature.
  - It can represent a set of discrete facts about events.
  - Data is a prerequisite to information.
  - An organization sometimes has to decide on the nature and volume of data that is required for creating the necessary information.

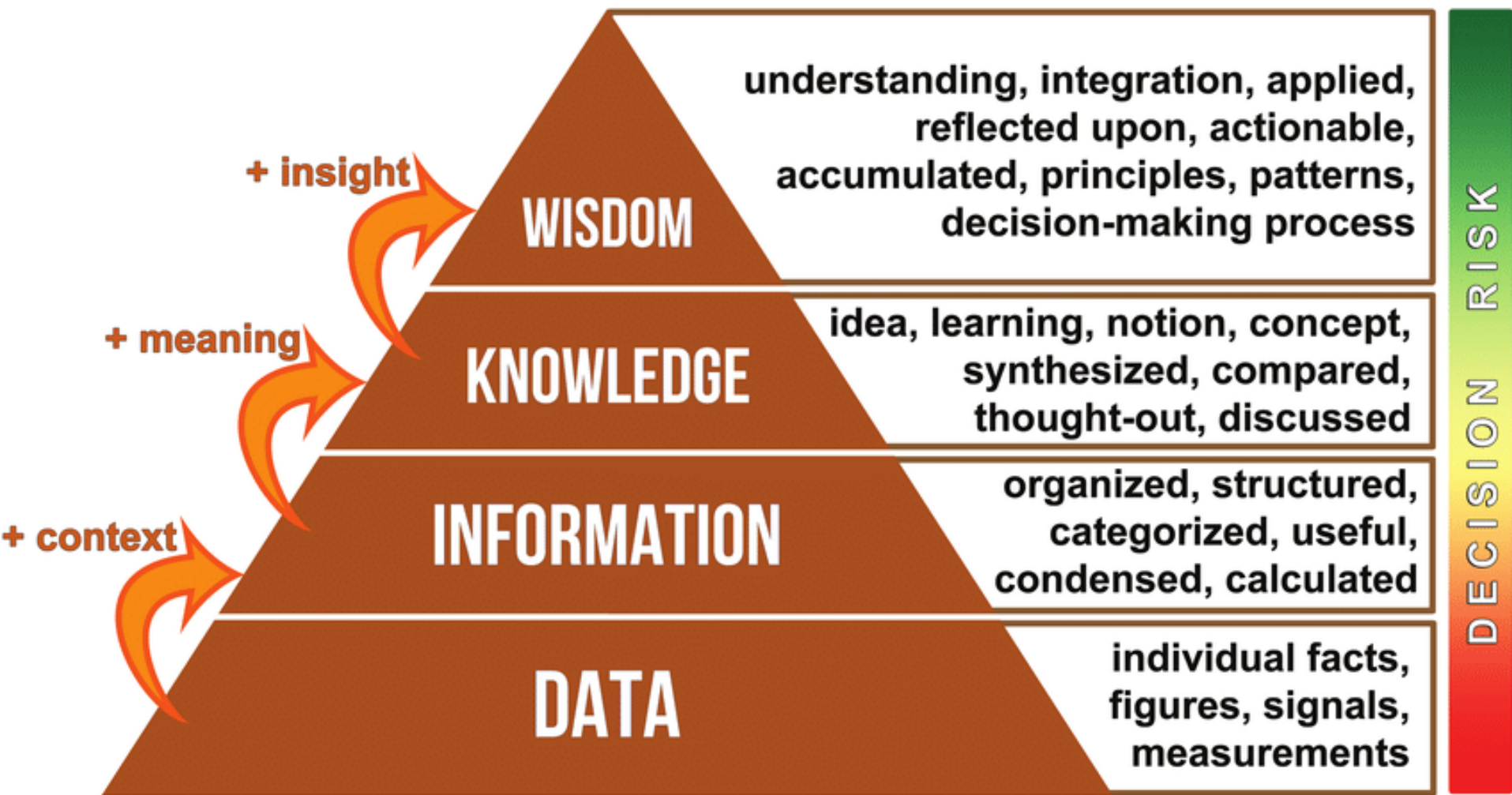
- **Information**

- Information can be considered as an aggregation of data (processed data) which makes decision making easier.
- Information has usually got some meaning and purpose.

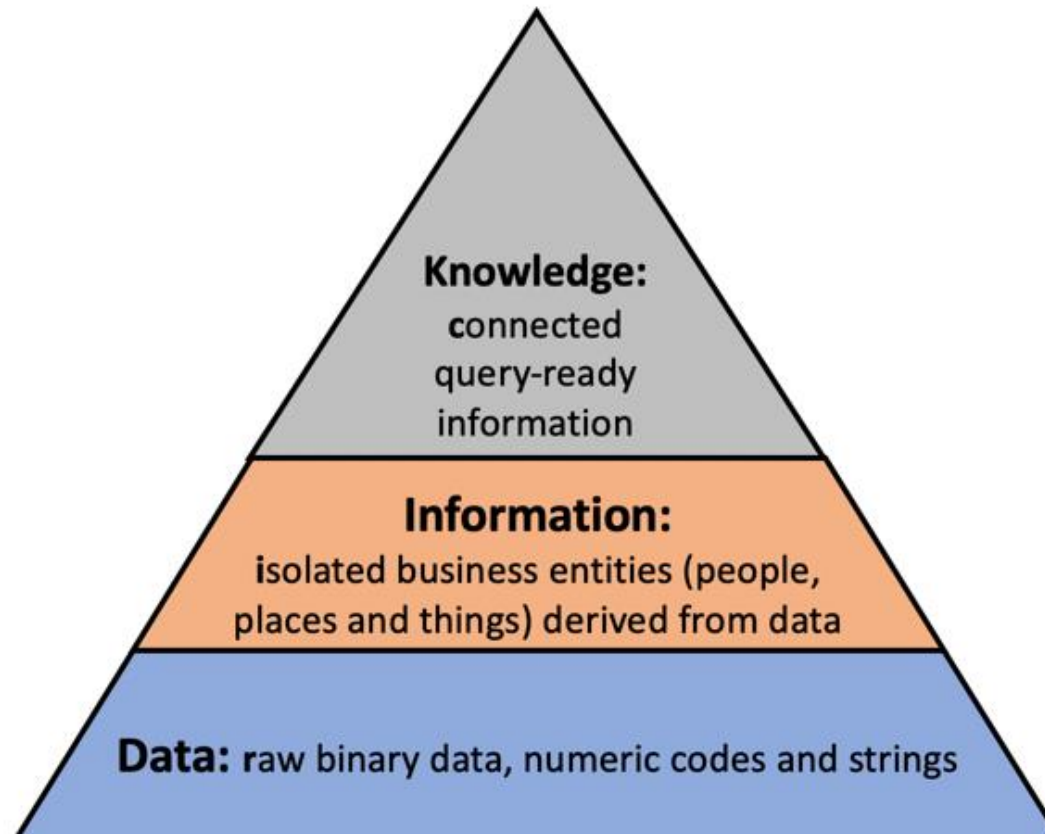
- **Knowledge**

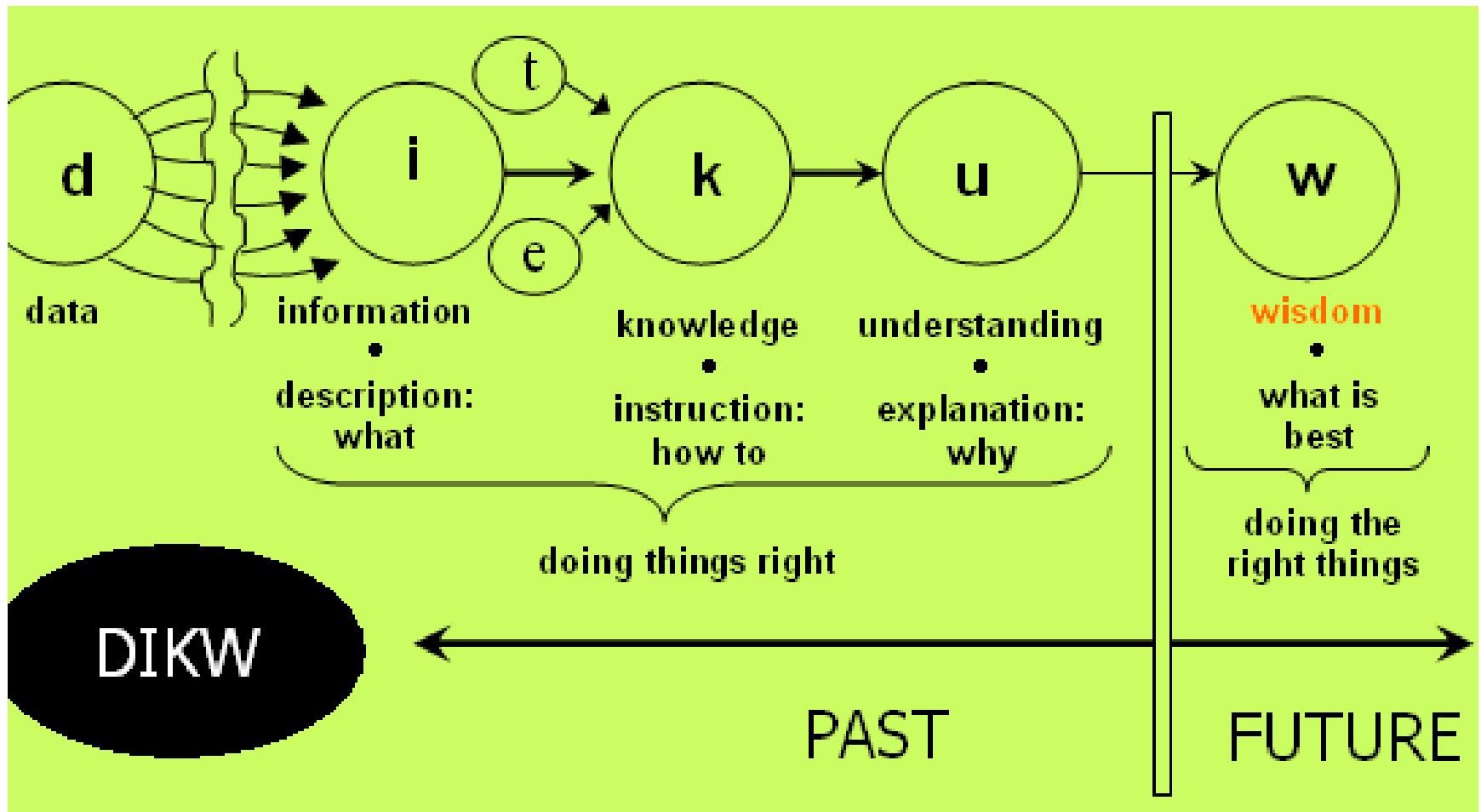
- By knowledge we mean *human understanding of a subject matter that has been acquired through proper study and experience.*
- Knowledge is usually based on learning, thinking, and proper understanding of the problem area.
- Knowledge is not information and information is not data.
- Data → Information → Knowledge

- Knowledge is derived from information in the same way information is derived from data.
- We can view it as an understanding of information based on its perceived importance or relevance to a problem area.
- It can be considered as the integration of human perceptive processes that helps them to draw meaningful conclusions.

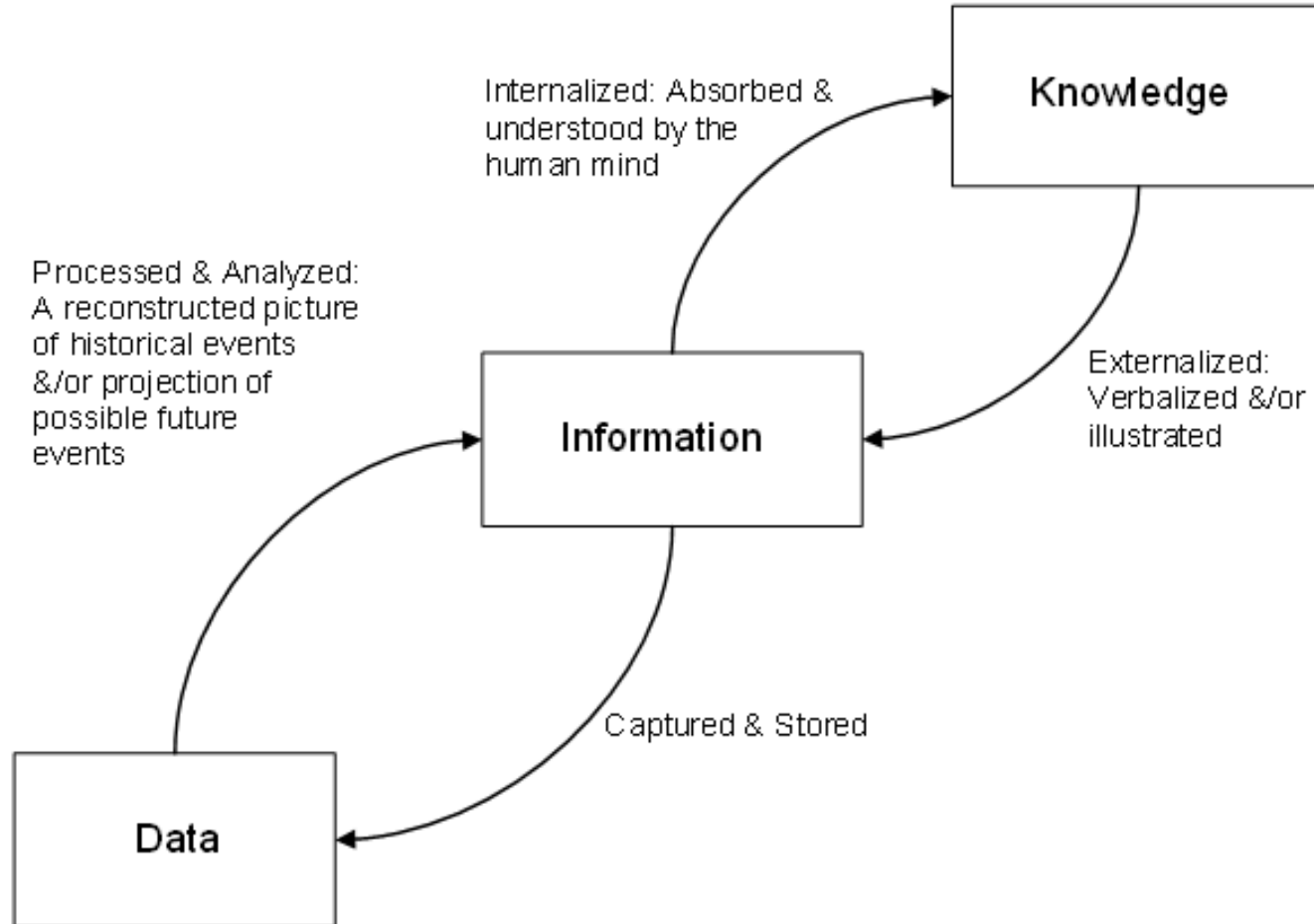








# Relationships Amongst Knowledge, Information, and Data

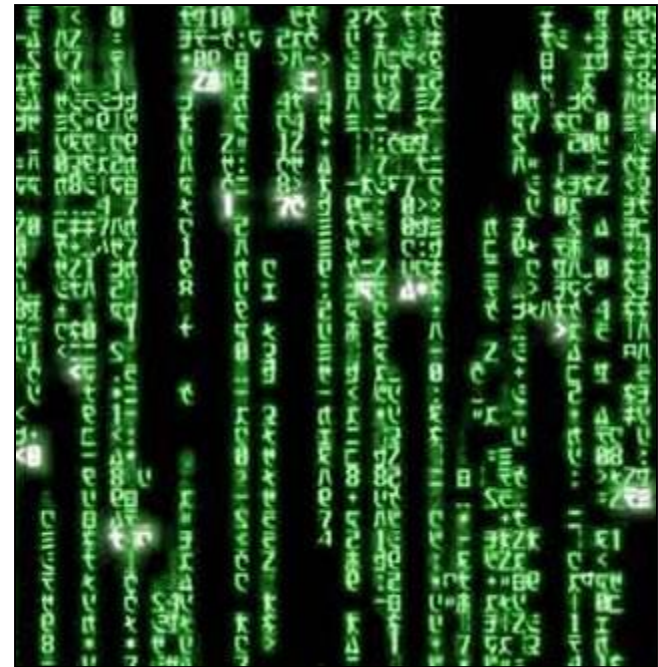


# Definitions

- Data management
- Information management
- Knowledge management

# Data

- Data **are** raw facts and figures that on their own have no meaning
- These can be any alphanumeric characters i.e. text, numbers, symbols



# Data Examples

- Yes, Yes, No, Yes, No, Yes, No, Yes
- 42, 63, 96, 74, 56, 86
- 111192, 111234
- None of the above data sets have any meaning until they are given a **CONTEXT** and **PROCESSED** into a useable form

# Data Into Information

- To achieve its aims the organisation will need to **process** data into information.
- Data needs to be turned into meaningful information and presented in its most useful format
- Data must be processed in a **context** in order to give it meaning

# Information

- Data that has been processed within a context to give it meaning

OR

- Data that has been processed into a form that gives it meaning

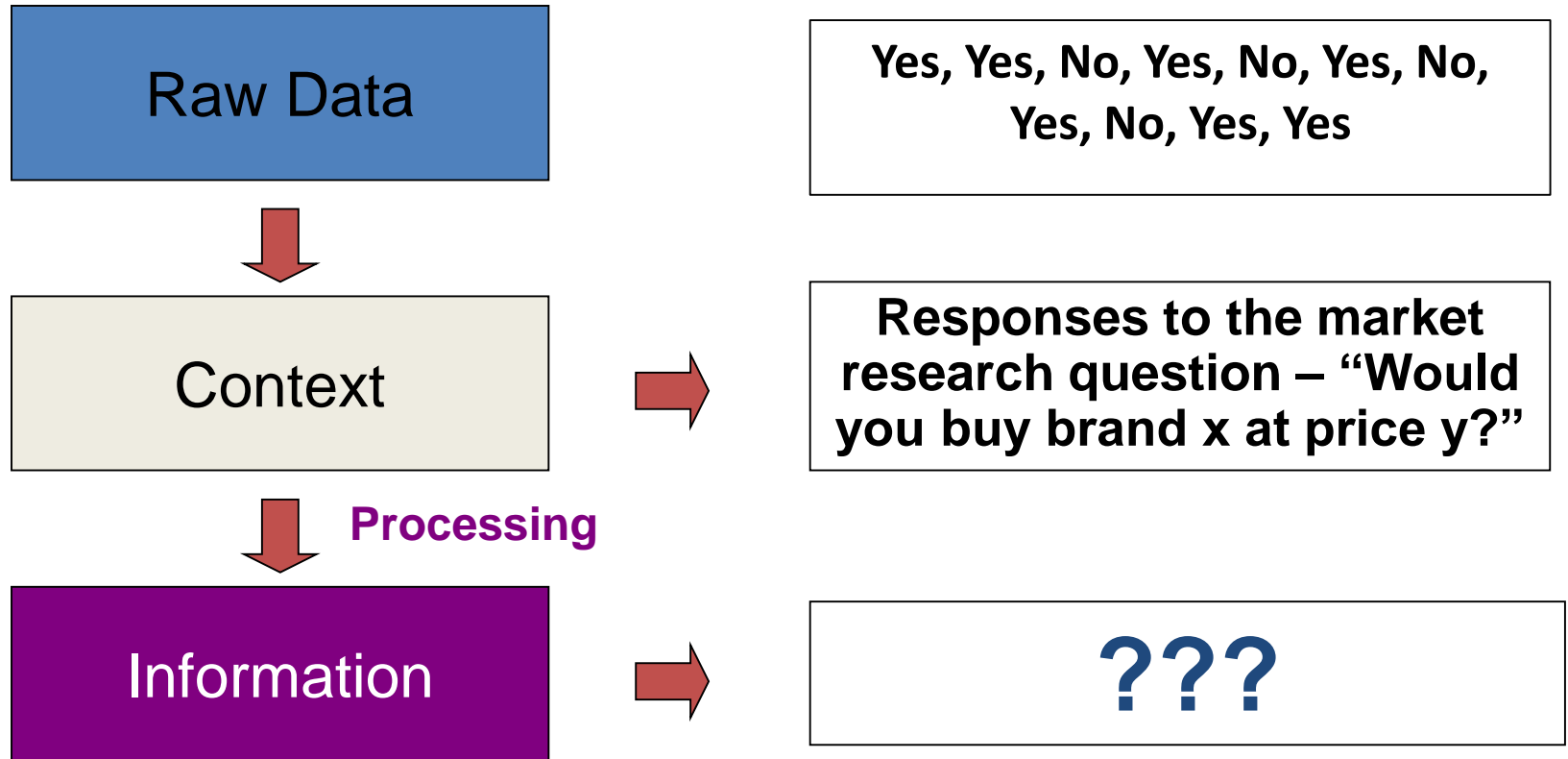


# Examples

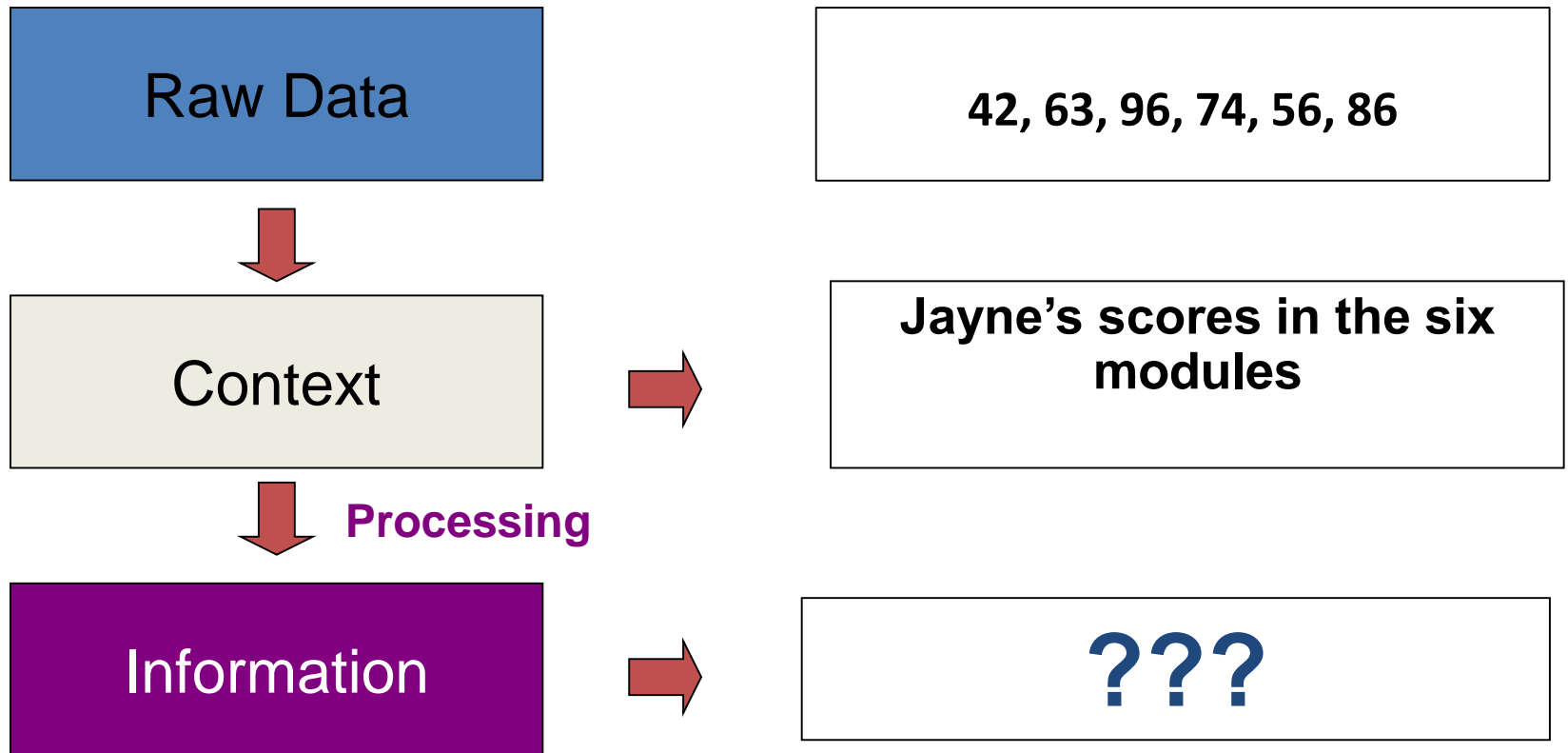
- In the next 3 examples explain how the data could be processed to give it meaning
- What information can then be derived from the data?



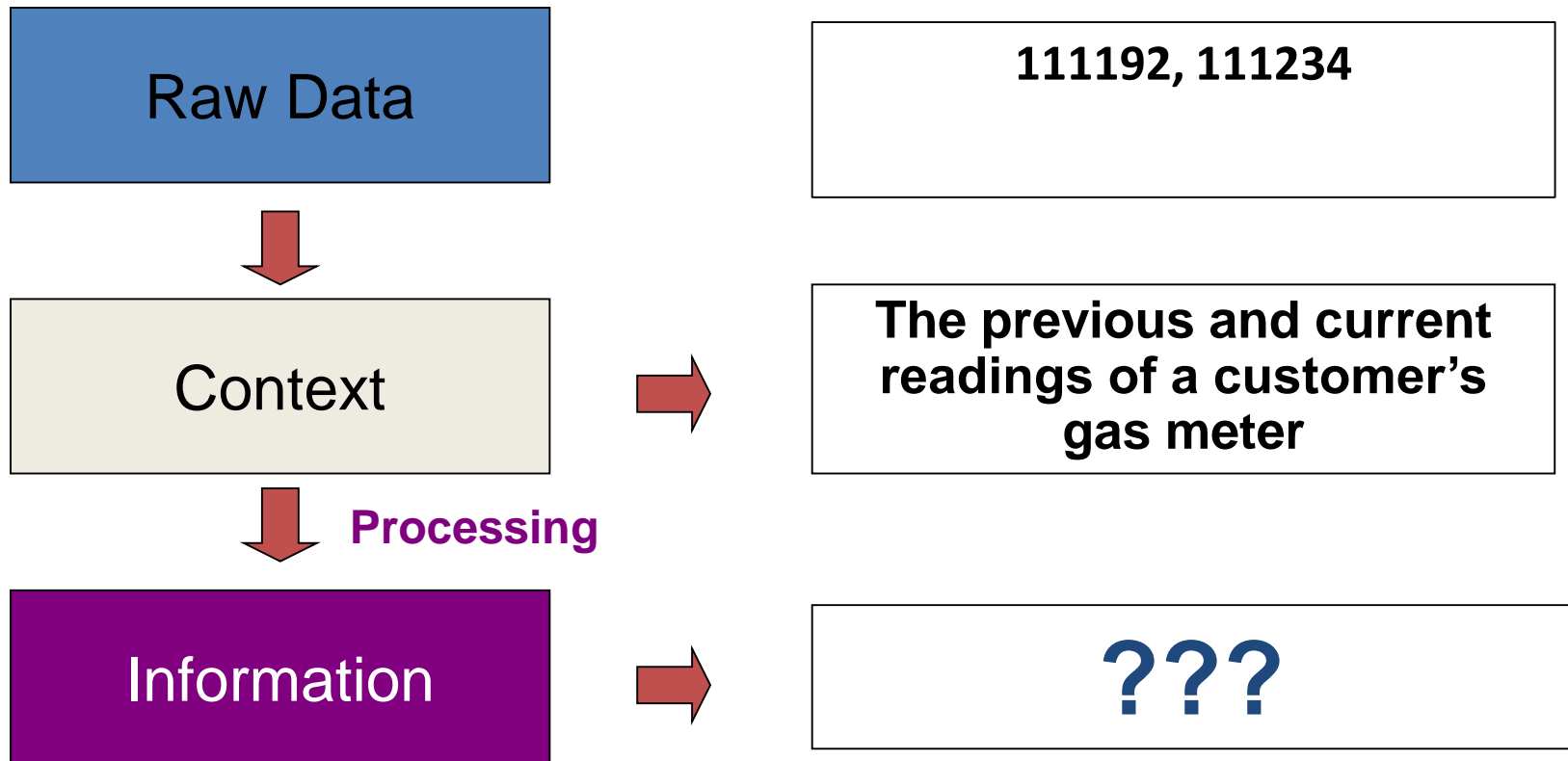
# Example #1



# Example #2



# Example 3



# Knowledge

- Knowledge is the understanding of rules needed to interpret information

“...the capability of understanding the relationship between pieces of information and what to actually do with the information”

# Knowledge

- Data and information deal with facts and figures
- Knowing what to do with them requires **knowledge**
- **Knowledge = information + rules**
- Rules tell us the likely effect of something
- For example: you are more likely to pass your A level IF you do your coursework and revise for your exam!

# Knowledge Examples

- Using the 3 previous examples:
  - A Marketing Manager could use this information to decide whether or not to raise or lower price  $y$
  - Jayne's teacher could analyse the results to determine whether it would be worth her re-sitting a module
  - Looking at the pattern of the customer's previous gas bills may identify that the figure is abnormally low and they are fiddling the gas meter!!!

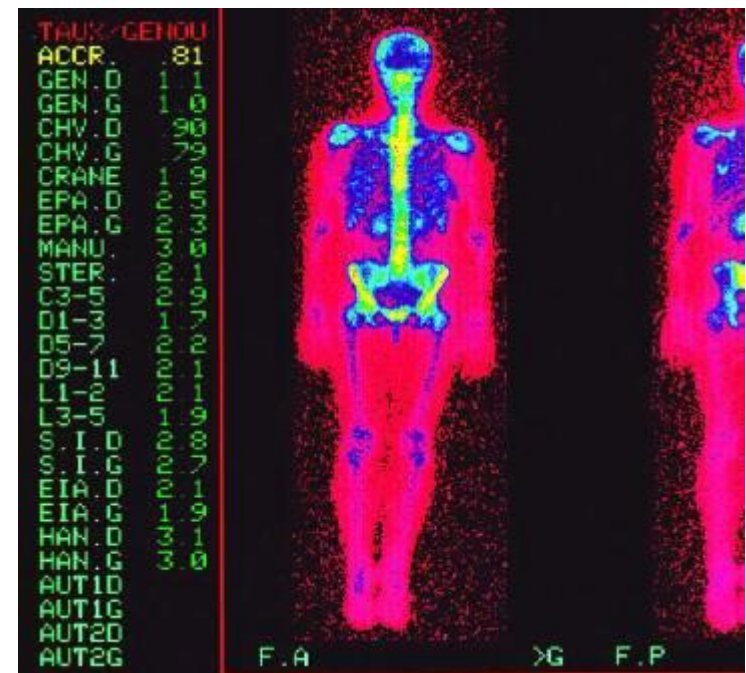
# Knowledge Workers

- Knowledge workers have specialist knowledge that makes them “experts”
  - Based on formal and informal rules they have learned through training and experience
- Examples include doctors, managers, librarians, scientists...



# Expert Systems

- Because many rules are based on probabilities computers can be programmed with “subject knowledge” to mimic the role of experts
- One of the most common uses of expert systems is in medicine
  - The ONCOLOG system shown here analyses patient data to provide a reference for doctors, and help for the choice, prescription and follow-up of chemotherapy



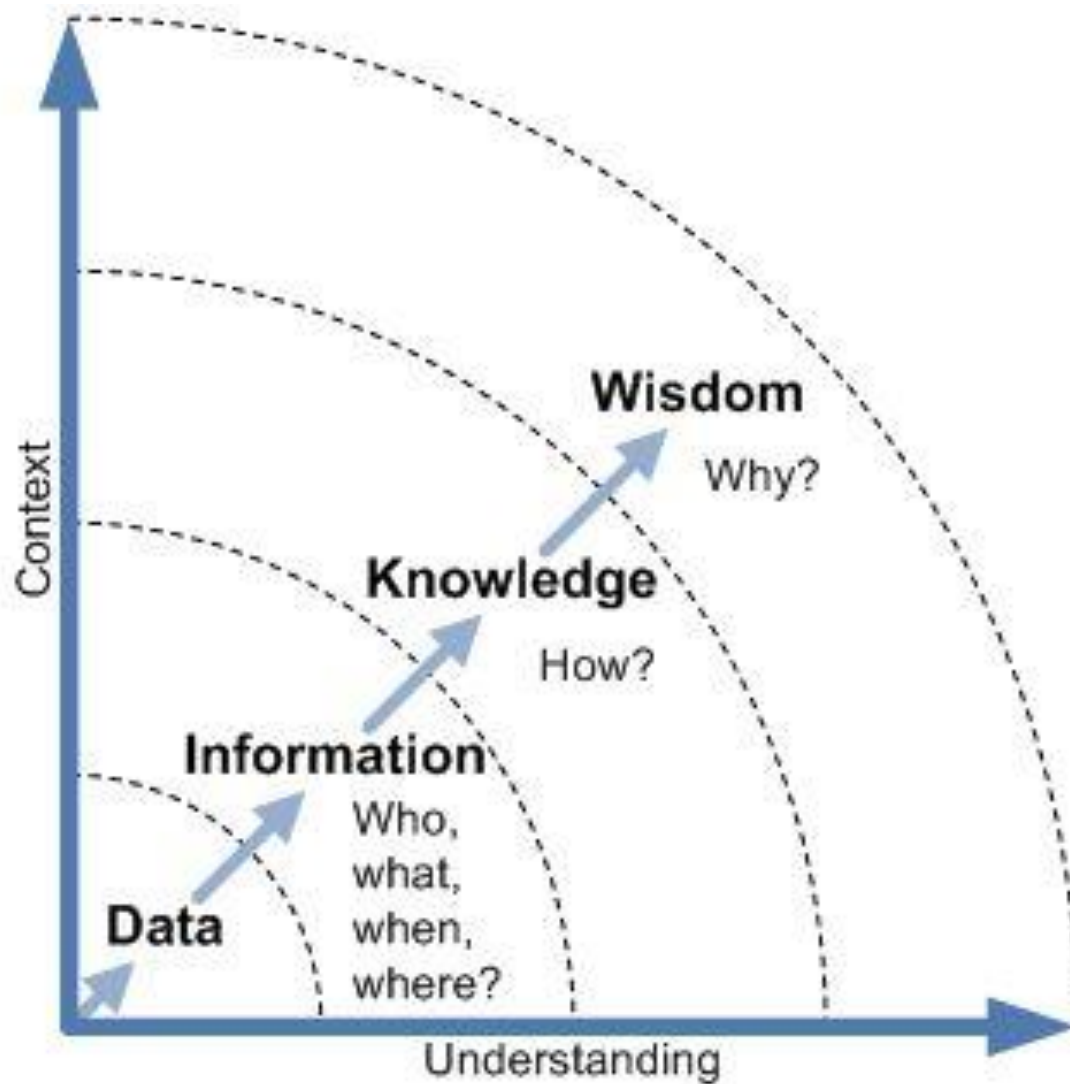
# Summary



**Data** – raw facts and figures

**Information** – data that has been processed (in a context) to give it meaning

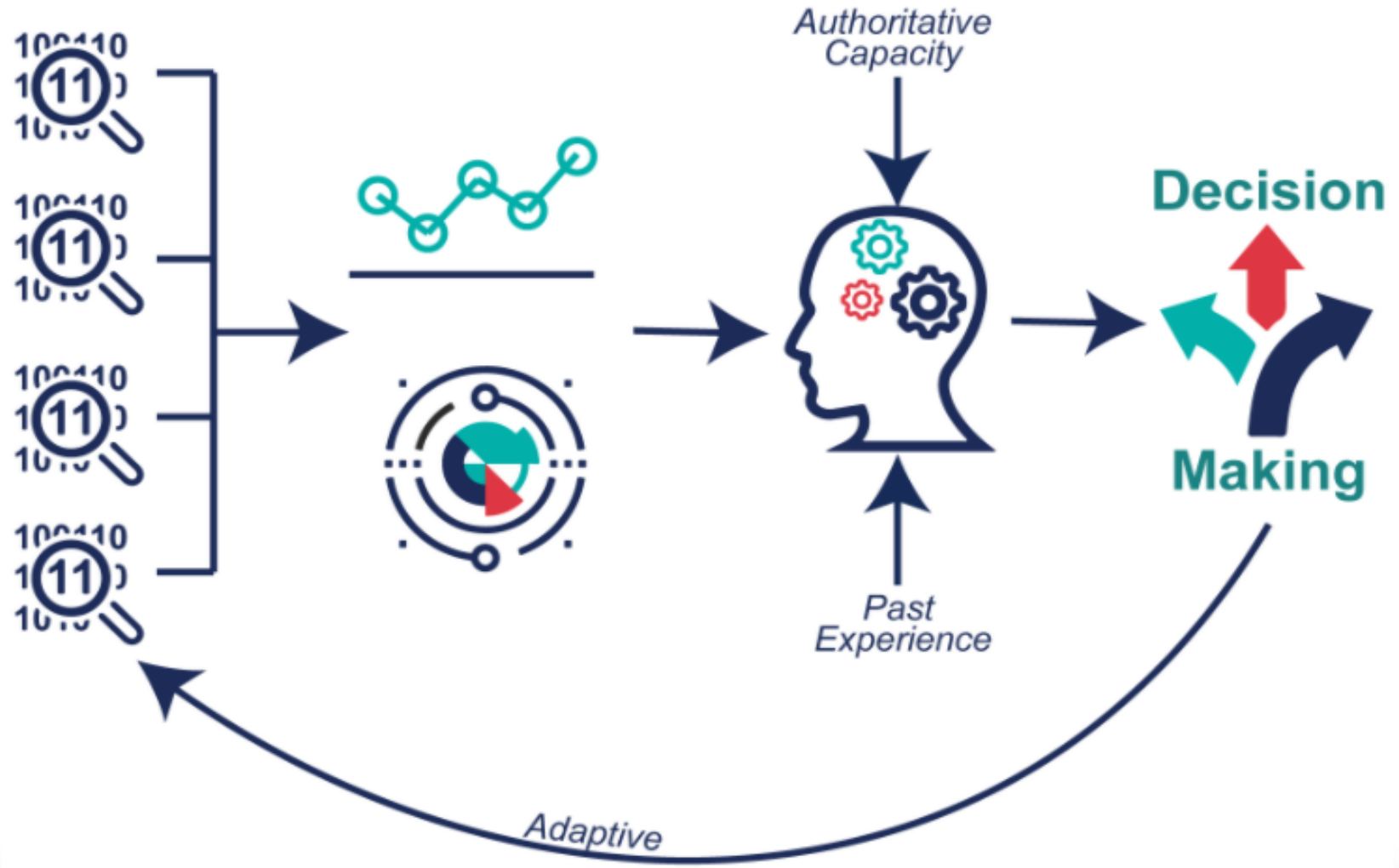




**Data**  
*Raw*

**Information**  
*Processed*

**Knowledge**  
*Actionable*



# Data Organization & Management

*“Data is stored in files”*

# Data Structures vs File Structures

- Both involve

Representation of Data

+

Operations for accessing data

- Difference:
  - **Data structures** deal with data **in the main memory**
  - **File structures** deal with the data **in the secondary storage**

# Goal of the File structures

- Performance
  - Time
    - Minimize the number of **hops** in order to get desired information
    - Group related information so that we are likely to get everything we need with fewer hops.
  - Memory
    - Balance the memory size and the time
- How to improve performance
  - Use the right file structure
    - Understand the advantages disadvantages of alternative methods

# Metrics used to measure efficiency and effectiveness of a File structure

- Time complexities,
- Space complexities,
- Simplicity,
- Reliability,
- Scalability,
- Programmability, and
- Maintainability.



## Metrics used to measure efficiency and effectiveness of a File structure

- The file structures involve two domains: **hardware** and **software**.
  - Hardware - physical characteristics of the storage medium.
  - Software - data structures and algorithms to deal with these structures.
- Hardware + SW (data structures and algorithms) are used to predict the efficiency of file operations.

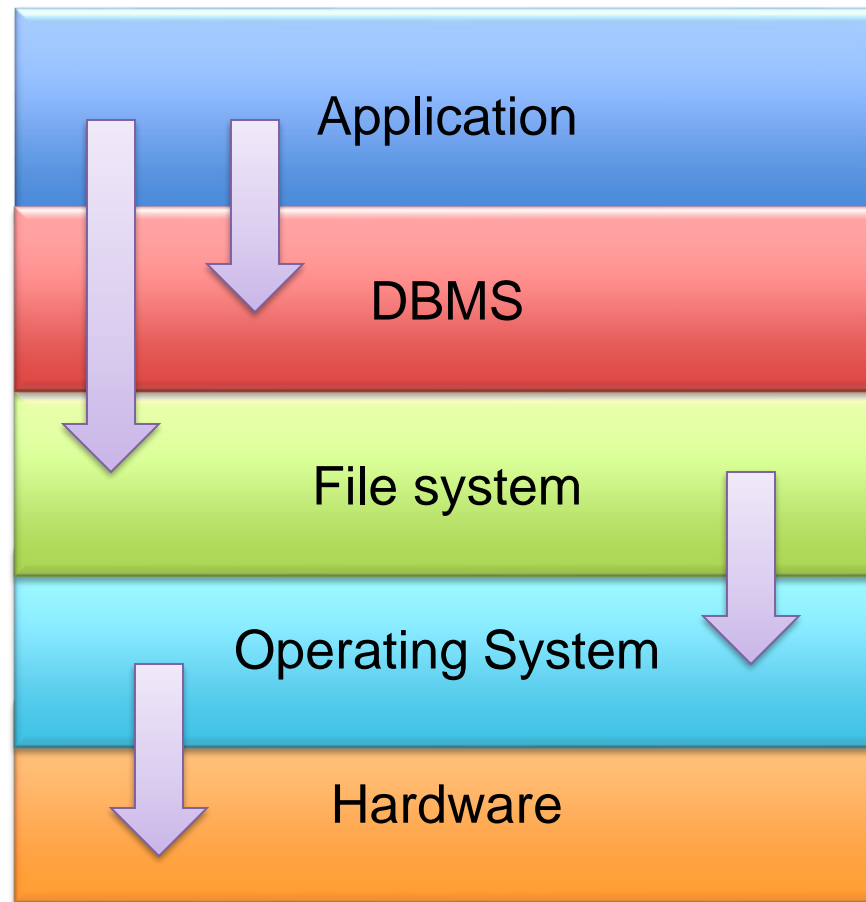
# Basic File operations

- Search for a particular data in a file,
- Add a certain data item,
- Remove a certain item,
- Order the data items according to a certain criterion,
- Merge of files,
- Creation of new files from existing file(s).
- create, open, and close operations which have implications in the operating system.

# Basic File Processing Operations

- Opening
- Closing
- Reading
- Writing
- Seeking
- Updating
- Compacting

# Where do File System fits in Computer System?



# File structures versus DBMS

- According to Alan Tharp, “file structures is used to process data in physical level, DBMS is used to manage data in a logical level”
- According to Raghu Ramakrishnan, “DBMS is a piece of software designed to make data maintenance easier, safer, and more reliable”.
- Thus, file structure is a pre-requisite to DBMSs.