Introduction to Data Processing

Michael J. Watts

http://mike.watts.net.nz

What's it all about?

- Every information system has to deal with data
- Any non-trivial system will have to collect, process and present data
- So, we need to understand
  - Sources and types of data
  - How data is represented and stored
  - The theory of information systems
  - How data is analysed and transformed
  - How data is presented to users

Sources and Types of Data

- Some issues in this section are:
  - Where does data come from?
  - What processes do the data represent?
  - What issues are involved in collecting data?
- Why study it?
  - Understanding the problem is crucial to developing a system to represent it
  - If data is collected improperly it is useless

Sources and Types of Data

- Topics we will cover in this section
  - Quantitative and qualitative data
  - Macromolecular and physical data
  - Image and sound data

Data Representation and Storage

- Some issues in this section are:
  - How is data represented and stored?
  - What data structures are appropriate for particular applications?
- Why study it?
  - In implementing an information system, it is necessary to store data
  - The correct choice will make implementing the system much easier

Data Representation and Storage

- Topics we will cover in this section:
  - Data structures
  - Metadata

Theory of Information Systems

- Why study it?
  - All practical methods are based on theory
  - An understanding of the theory will assist in making the appropriate choices when creating an information system
  - An understanding of theory can stop you doing something stupid

Theory of Information Systems

- Topics we will cover in this section
  - Probability
  - Information theory
  - Data compression
  - Measurement theory
  - Formal language theory
  - Computer theory

Data Analysis and Transformation

- Some topics in this section are:
  - What is data analysis?
  - What is data transformation?
  - What methods are appropriate for different data sets?
- Why study it?
  - Analysis of data aids understanding the problem
  - Transformation of data can ease analysis

Data Analysis and Transformation

- Topics we will cover in this section
  - Statistical methods
  - Feature selection and extraction
  - Clustering
  - Linear & non-linear transforms

Presentation of Data and Information

- Some issues in this section are:
  - How do you visualise data?
  - How do you design user interfaces?
  - How do you communicate data to the user?
- Why study it?
  - Data becomes information when it is presented to a user
  - Inappropriate presentation of data can have fatal consequences

Presentation of Data and Information

- Topics covered in this section are:
  - Cognitive factors of HCI
  - Designing UI and communicating data
  - Visualisation of data in 2 & 3 dimensions
  - Usability evaluation