



CSCI 5408

Dr. Saurabh Dey
saurabh.dey@dal.ca



A perspective view of a tunnel formed by a repeating pattern of binary digits (0s and 1s). The tunnel curves slightly to the right and recedes into the distance, creating a sense of depth. The binary code is rendered in a light blue color against a dark background.

Why Data?

JAN
2023

ESSENTIAL DIGITAL HEADLINES

OVERVIEW OF THE ADOPTION AND USE OF CONNECTED DEVICES AND SERVICES



GLOBAL OVERVIEW

TOTAL
POPULATION



we
are
social

8.01
BILLION

URBANISATION

57.2%

UNIQUE MOBILE
PHONE USERS



Meltwater

5.44
BILLION

vs. POPULATION

68.0%

INTERNET
USERS



K
KEPIOS

5.16
BILLION

vs. POPULATION

64.4%

ACTIVE SOCIAL
MEDIA USERS



4.76
BILLION

vs. POPULATION

59.4%

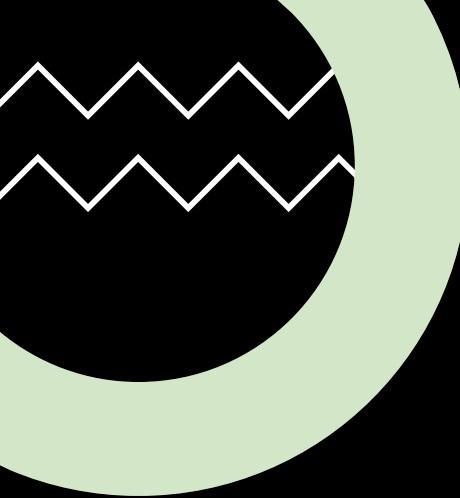
10

SOURCES: UNITED NATIONS; GOVERNMENT BODIES; GSMA INTELLIGENCE; ITU; WORLD BANK; EUROSTAT; CNNIC; APJI; IAMAI & KANTAR; CIA WORLD FACTBOOK; COMPANY ADVERTISING RESOURCES AND EARNINGS REPORTS; OCDH; BETA RESEARCH CENTER; KEPIOS ANALYSIS. ADVISORY: SOCIAL MEDIA USERS MAY NOT REPRESENT UNIQUE INDIVIDUALS. COMPARABILITY: SIGNIFICANT REVISIONS TO SOURCE DATA, INCLUDING COMPREHENSIVE REVISIONS TO POPULATION DATA. FIGURES ARE NOT COMPARABLE WITH PREVIOUS REPORTS. ALL FIGURES USE THE LATEST AVAILABLE DATA, BUT SOME SOURCE DATA MAY NOT HAVE BEEN UPDATED IN THE PAST YEAR. SEE NOTES ON DATA FOR FULL DETAILS.

we
are.
social

Meltwater

Image Source: <https://datareportal.com/reports/digital-2023-global-overview-report>



This course focuses on

- How to identify source
- How to gather data
- How to process data
- How to store data
- How to analyze data
 - :
 - :



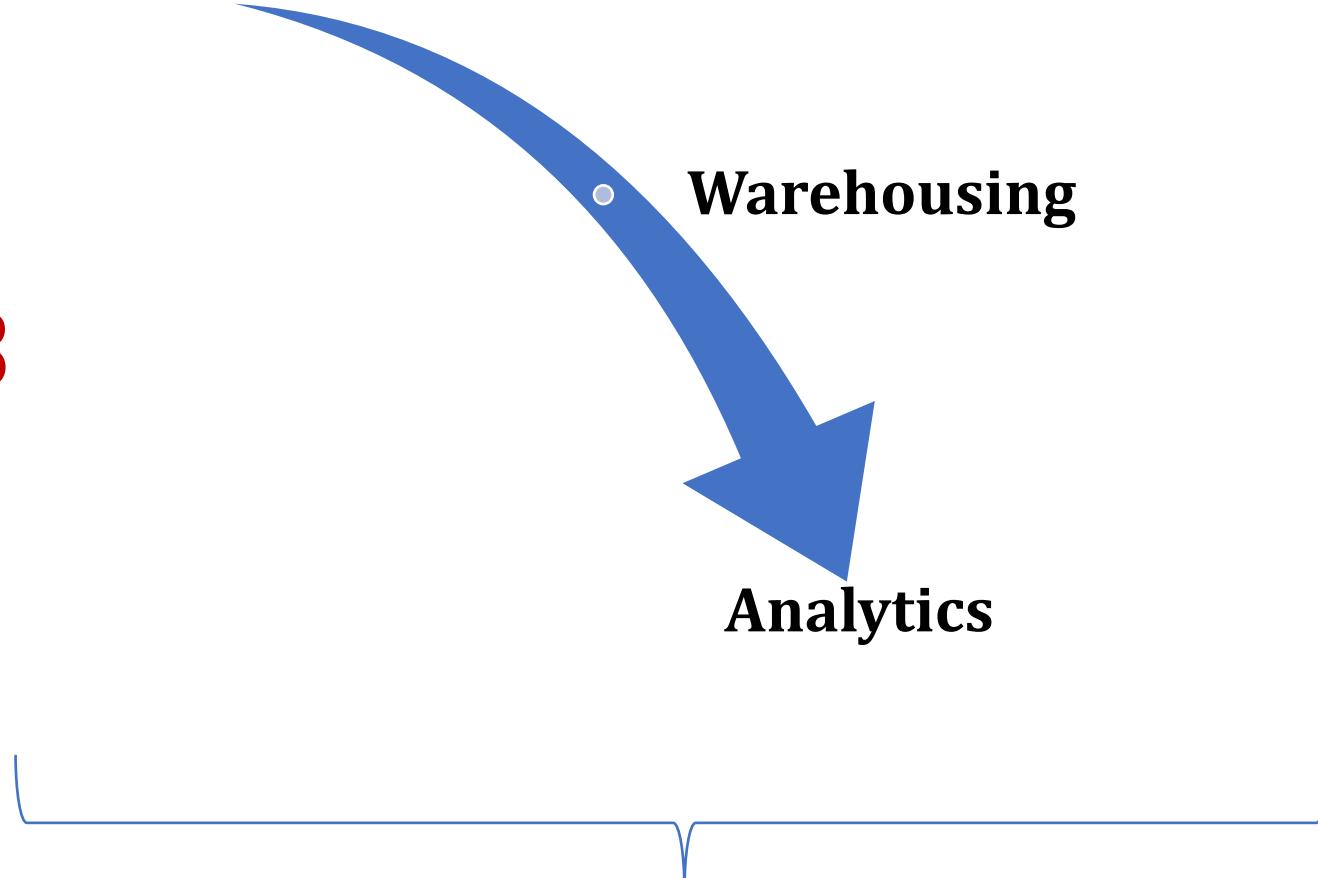
CSCI 5408

Database

Warehousing

Analytics

Data



Questions on Data?

01

CAN WE MANAGE
EVEN IF DATA IS
INCORRECT?

02

HOW TO SELECT A
SOURCE FOR DATA
COLLECTION?

03

WHEN CAN WE BE
SURE THAT THE
DATA WE GATHERED
IS CORRECT AND
RELEVANT?

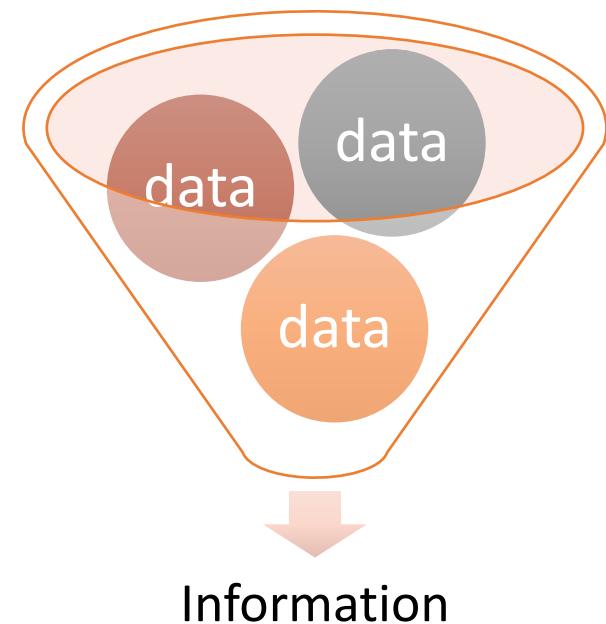
Basic Concepts

Data consists of raw facts

- Not yet processed to reveal meaning to the end user
- Building blocks of information

Information results from processing raw data

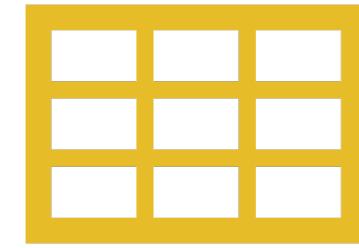
- Requires context
- Should be accurate, relevant, and timely



- 1. Who decides the context?**
- 2. Can correct data generate incorrect information**

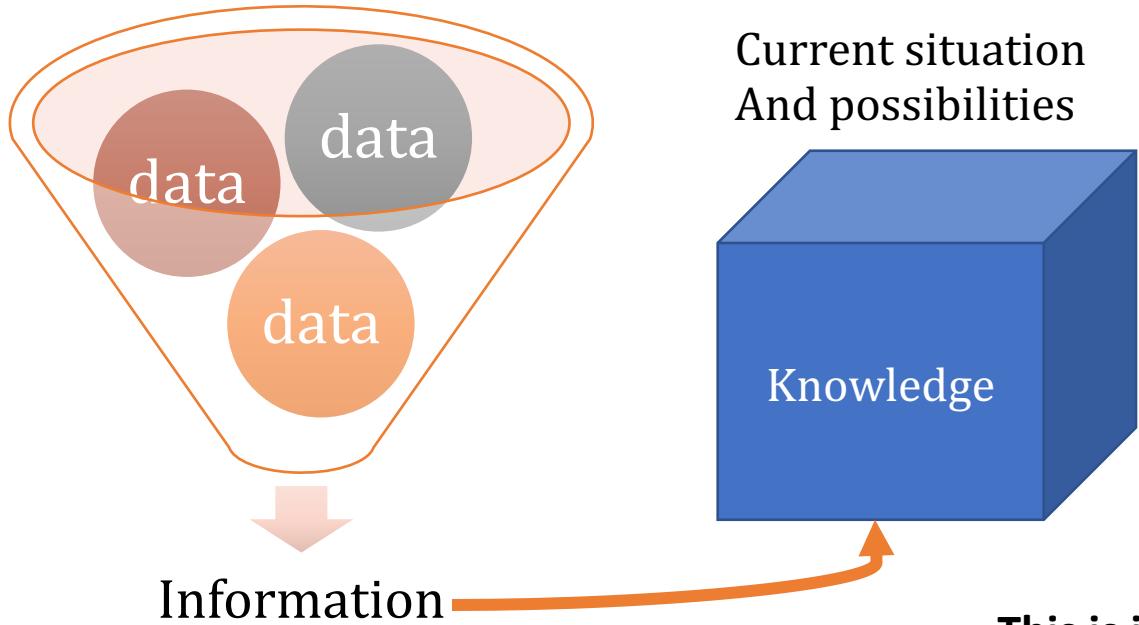


Information needs some context



Can you give an example when raw data is considered as information?

Knowledge and Wisdom



This is just an example; Please do not take it seriously 😊

There will be power outage in Halifax for longer period.

Data: Number of People will be affected/ area/ weather predictions

Information: Percentage of people and residence affected

Knowledge: Area that needs emergency backup or quick restore

Wisdom: We need to research for alternative power source

End of Lecture Questions

1. Will data reach a point, when it will not be possible to handle?
2. Can you think of any scenario when data can be considered information in its raw form?
3. Can we consider the term knowledge as storage and wisdom as analysis?
4. Will you consider File System for your business or personal use?

