

# Big Data & Cloud Ecosystems

ניהול וכריית נתוני עתק בסביבת ענן

Lesson 1



יוסי זגורי

[Yossi.Zaguri@gmail.com](mailto:Yossi.Zaguri@gmail.com)

052-4668866

# מטרות הקורס

- הכרת מושגים, תהליכים וטכנולוגיות לטיפול בנתוני עתק
  - KDD
  - אתגרי Big Data
  - ארכיטקטורה וגישות לאחסון ואחזור מאסות נתונים
- הקניית יכולות שימוש בתשתיות מחשוב ענן
  - שירותי תוכנה וספקים מרכזיים
  - הגדרת סביבה בענן לעיבוד נתוני עתק
- הבנת האתגרים ויישום אלגוריתמים נבחרים בלמידת מכונה
  - גישות חישוביות לכריית נתונים
  - מימוש משימות ניתוח והפקת ידע

# מושגי יסוד



If you wish to converse with me,  
define your terms.

~ Voltaire

AZ QUOTES

# רקע ותזכורת

- על ההבדלים בין עיצוב וארכיטקטורה.

- עקרון SOC ומודולציה.

- ניהול Dependencies.

- מהי תוכנה ומהו שירות תוכנה.



# רקע ותזכורת

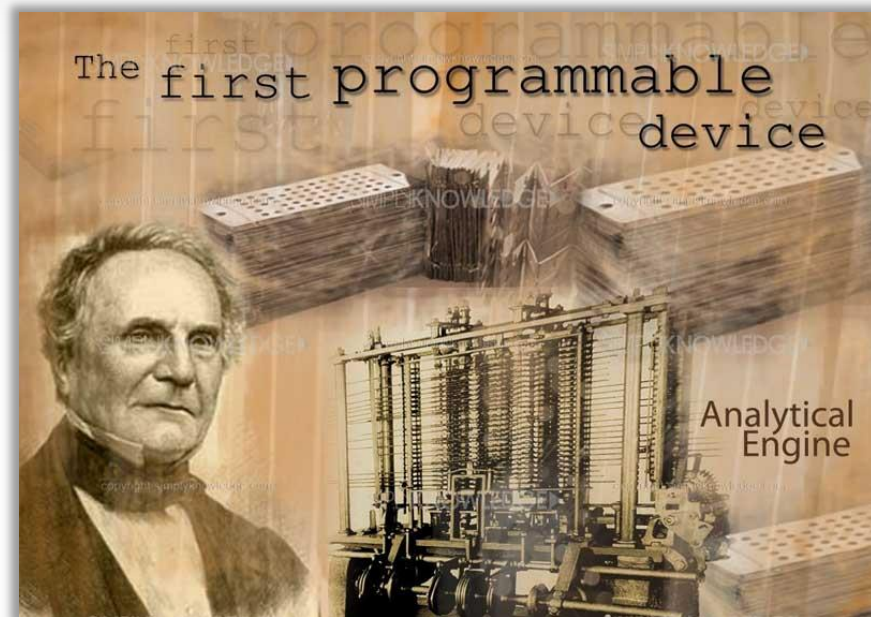
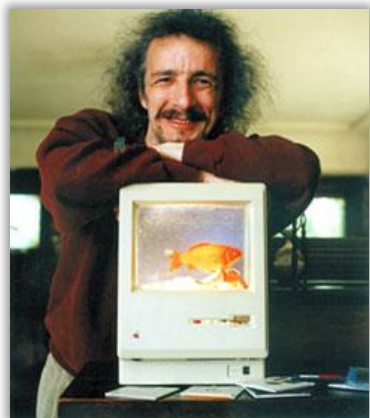
- ביזור ומשמעויותיו.
- המעבר ממודל Scale Up ל- Scale Out.
- תהליכי פיתוח תוכנה, DevOps ופעילויות CI/CD
- מה משמעות המונח "וירטואלי" ?



# נתונים, מידע וידע

“Errors using inadequate data are much less than those using no data at all.”

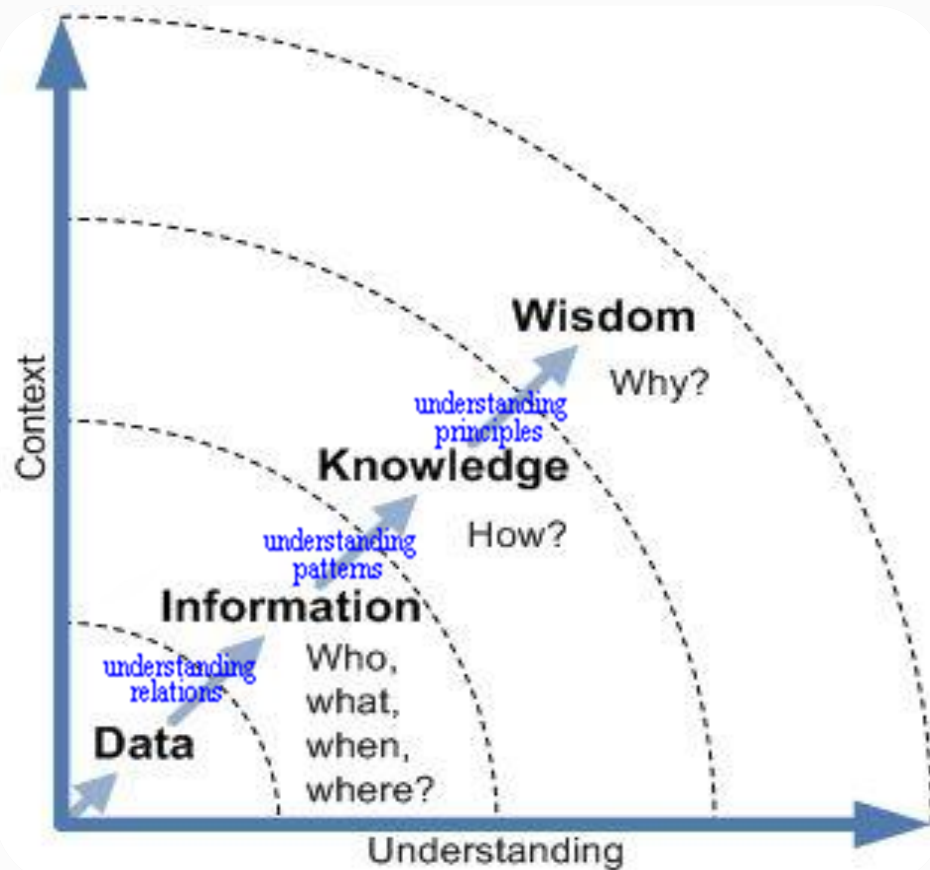
*Charles Babbag*



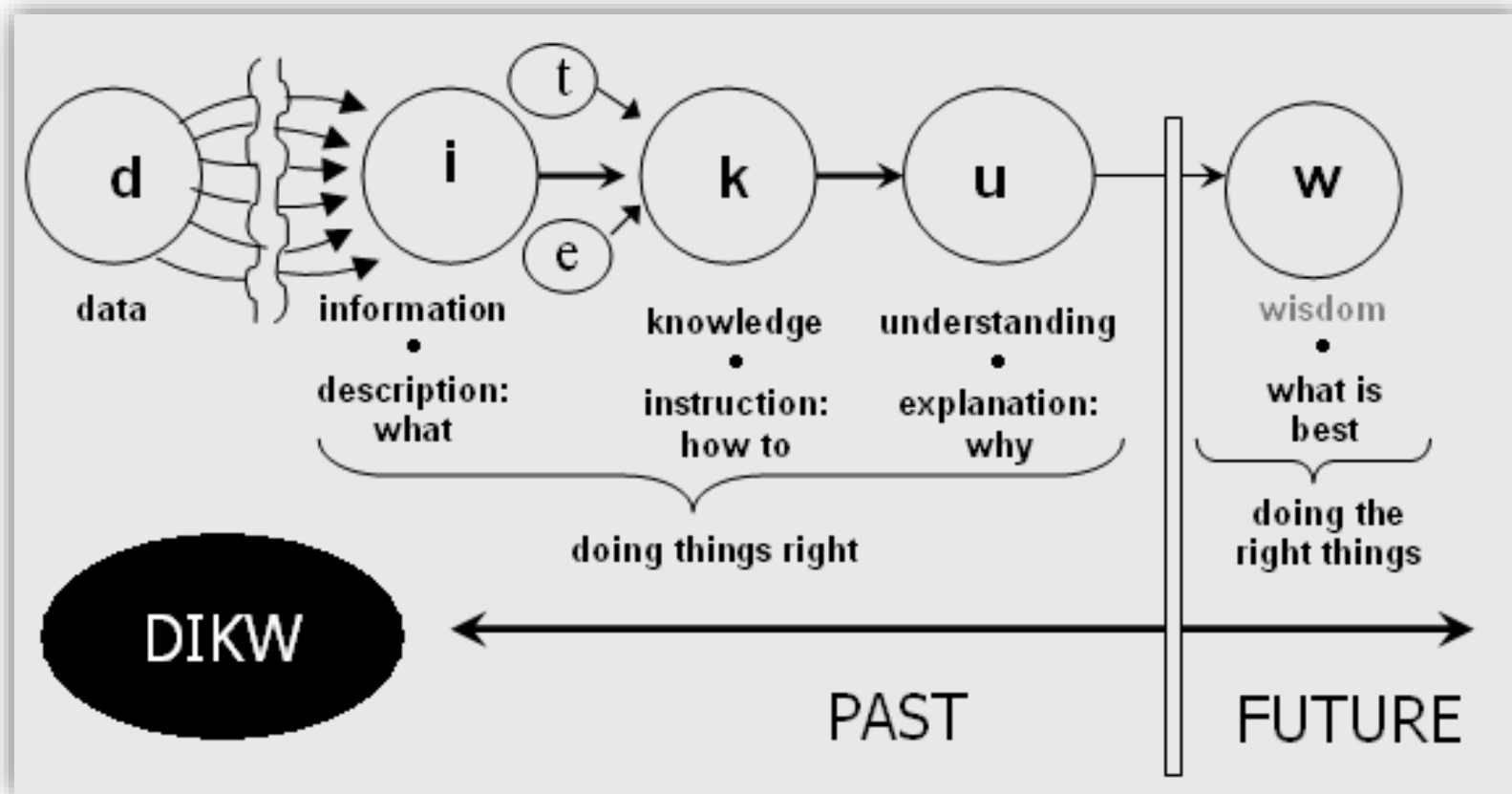
“Data is not information, information is not knowledge, knowledge is not understanding, understanding is not wisdom.”

*Clifford Stoll*

# מודל DIKW



# מודל DIKW

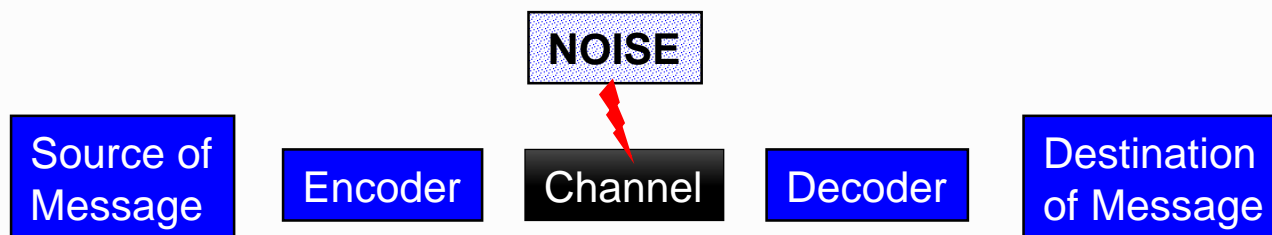




# Measuring Information

Information Theory is a branch of science that deals with the analysis of a communications system.

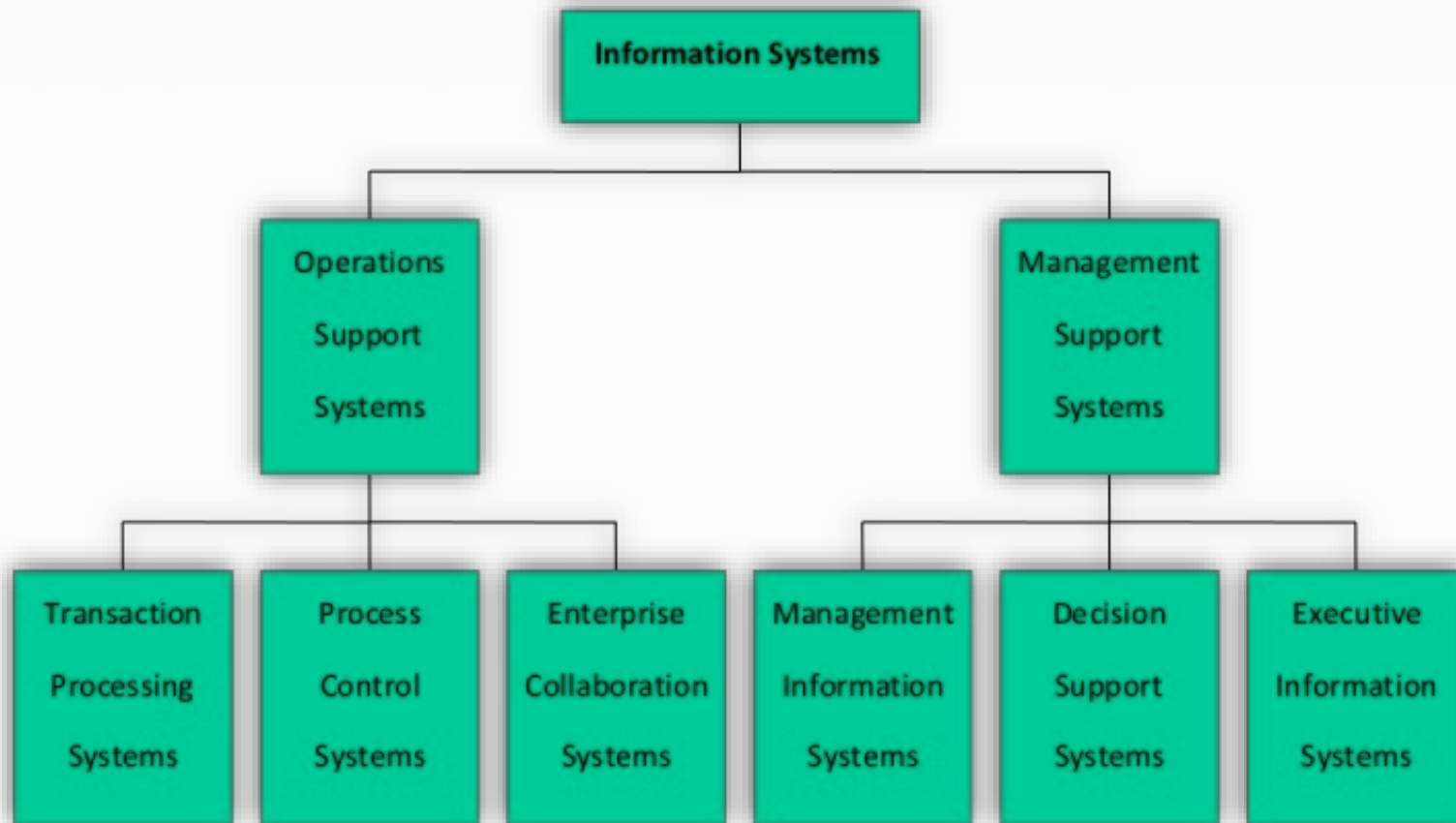
Channels can vary from a file to network protocol.



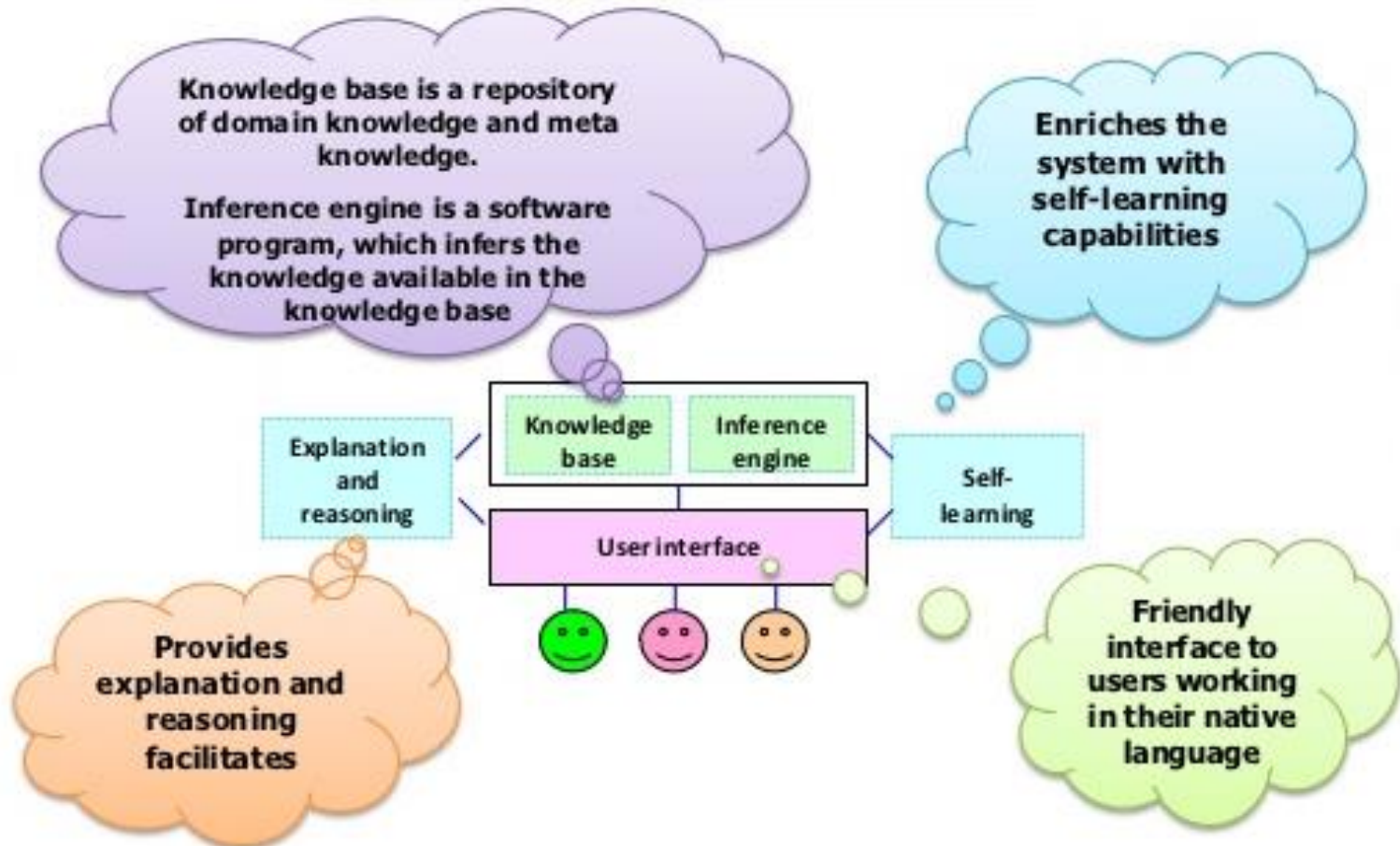
**Claude Shannon** Published a landmark paper in 1948 that was the beginning of the branch of information theory.

In course context, we are interested in quantifying Information levels each data element may convey

# Information Systems



# Knowledge Based Systems



# Knowledge & AI

## Principle Of Rationality

If an agent has knowledge that one of its actions will lead to one of its goals, then the agent will select this action.

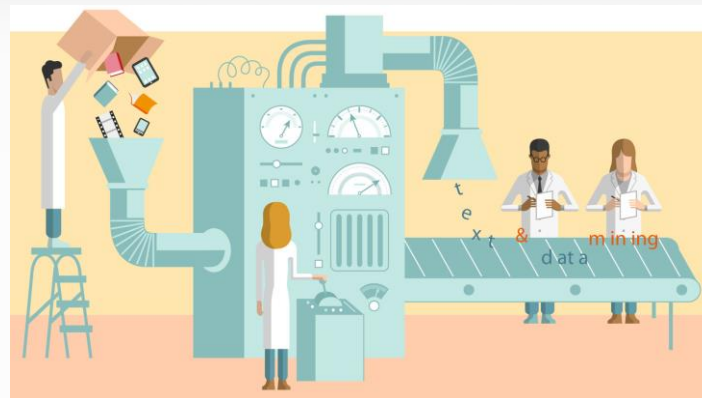
No guarantee that the knowledge is in fact sufficient to determine the right actions or even that it exists.

**Knowledge** makes the PR work as a *law of behavior*.

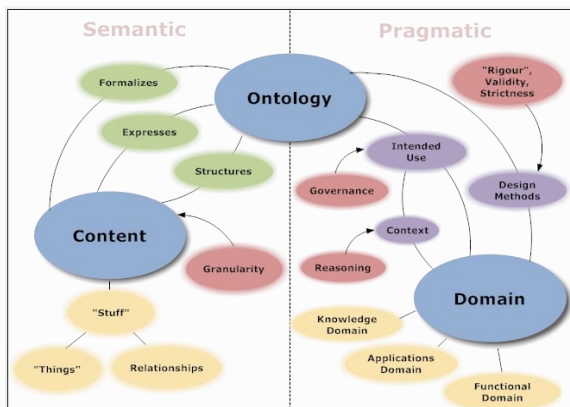
In this Context, **Knowledge** is whatever can be ascribed to the agent such that its behavior can be computed according to the PR.

In our course, we will reveal and represent **Knowledge** using Machine Learning methods, similar to those used in AI to make artificial agents *learn from experience* and improve performance over time, choosing actions according to the principle of rationality.

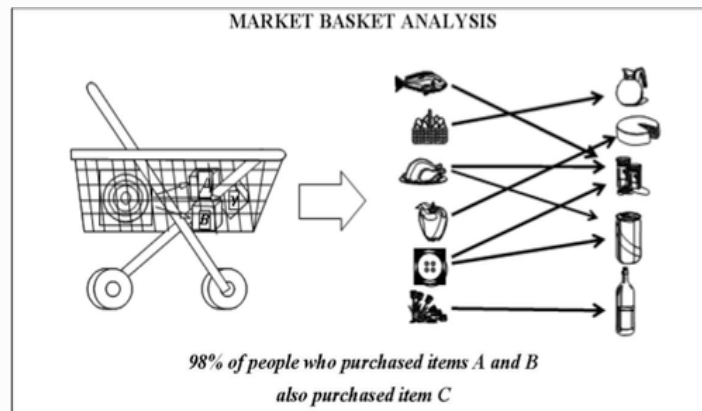
# Acquiring & Representing Knowledge



## Ontology



## Association Rules



## What is Knowledge Representation ?

# Ontology Example: Bitcoin

