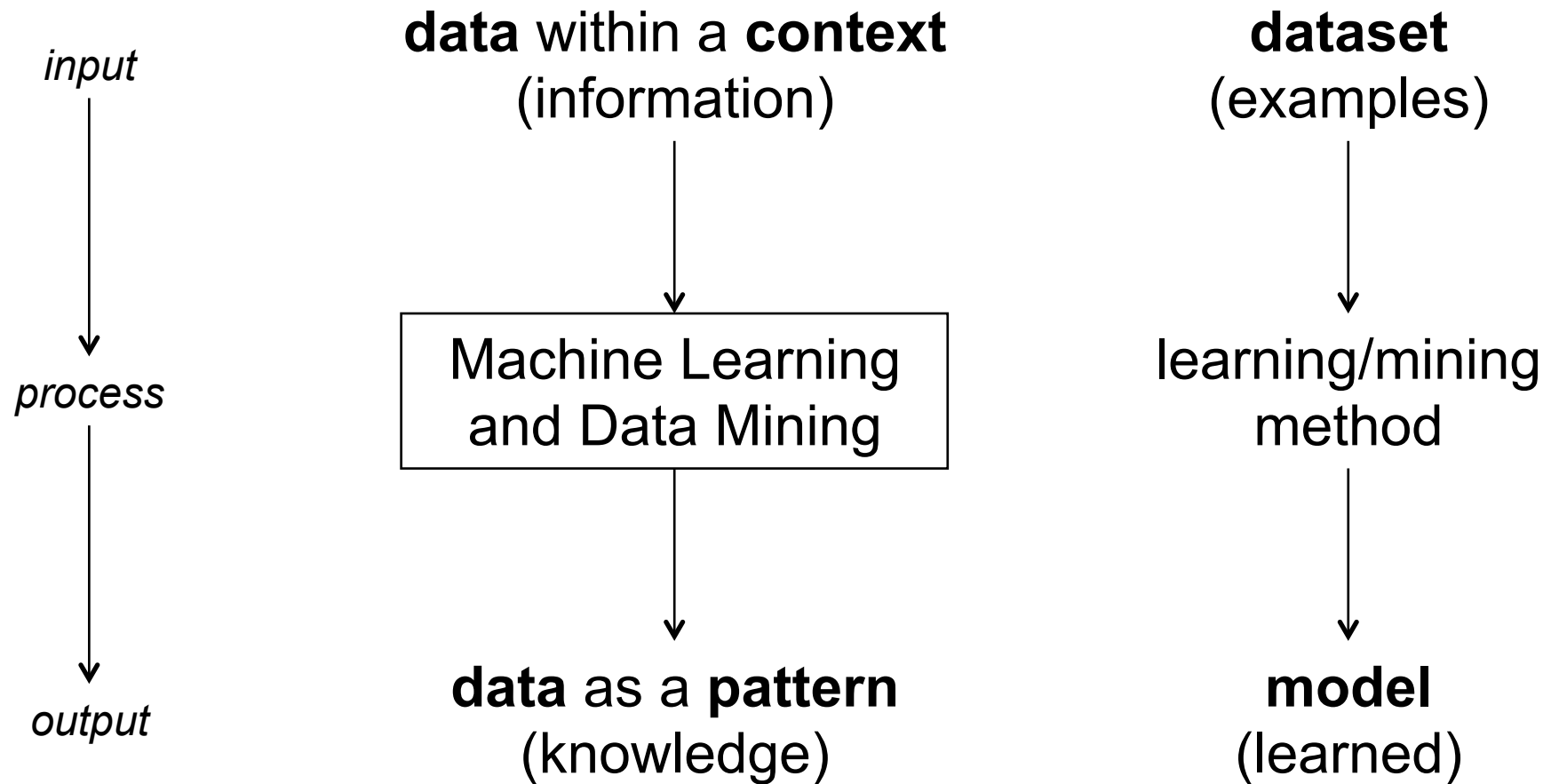


Data, Information, Knowledge

AMD – Aprendizagem e Mineração de Dados

MLDM – Machine Learning and Data Mining



Machine Learning (and Data Mining)

what is machine learning?

- A machine can learn? How?
- How do humans learn?



Machine learning is about automating automation

Machine learning is about getting computers to program themselves

A computer program is said to learn from experience E with respect to some set of tasks T and performance measure P

if

its performance at tasks in T , as measured by P , improves with experience E .

(Machine Learning and) Data Mining

what is data mining?

- *data mining is a practical topic and involves **learning** in a practical, not a theoretical, sense.*
- *we are interested in techniques for finding patterns in data, patterns that provide insight or enable fast and accurate decision making.*

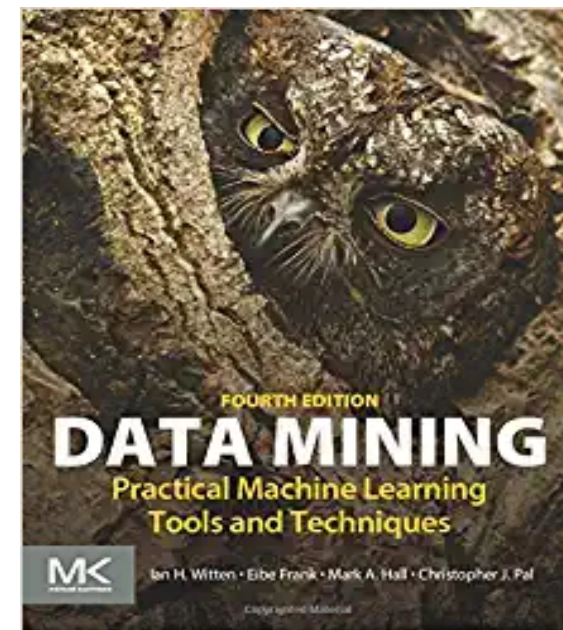
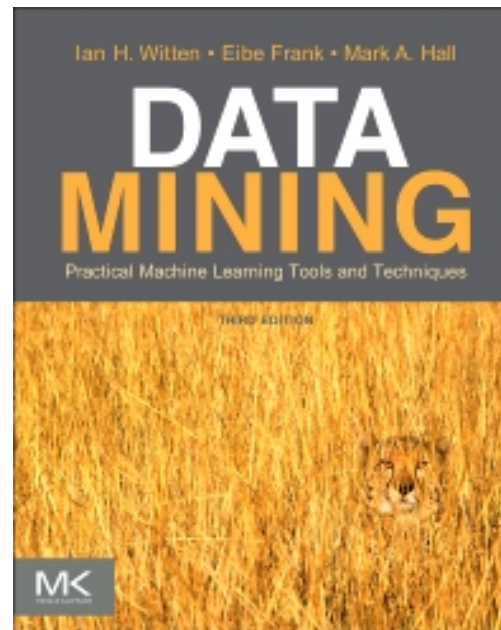


https://en.wikipedia.org/wiki/Gold_mining

(Machine Learning and) Data Mining (cont.)

what is data mining?

- *data mining is a practical topic and involves **learning** in a practical, not a theoretical, sense.*
- *we are interested in techniques for finding patterns in data, patterns that provide insight or enable fast and accurate decision making.*



Learn from Examples (Data)

The **data** will take the form of a set of **instances/examples/individuals**

- examples of customers who have switched loyalties
- or
- situations in which certain kinds of contact lenses can be prescribed
- or
- examples of clinical medical data

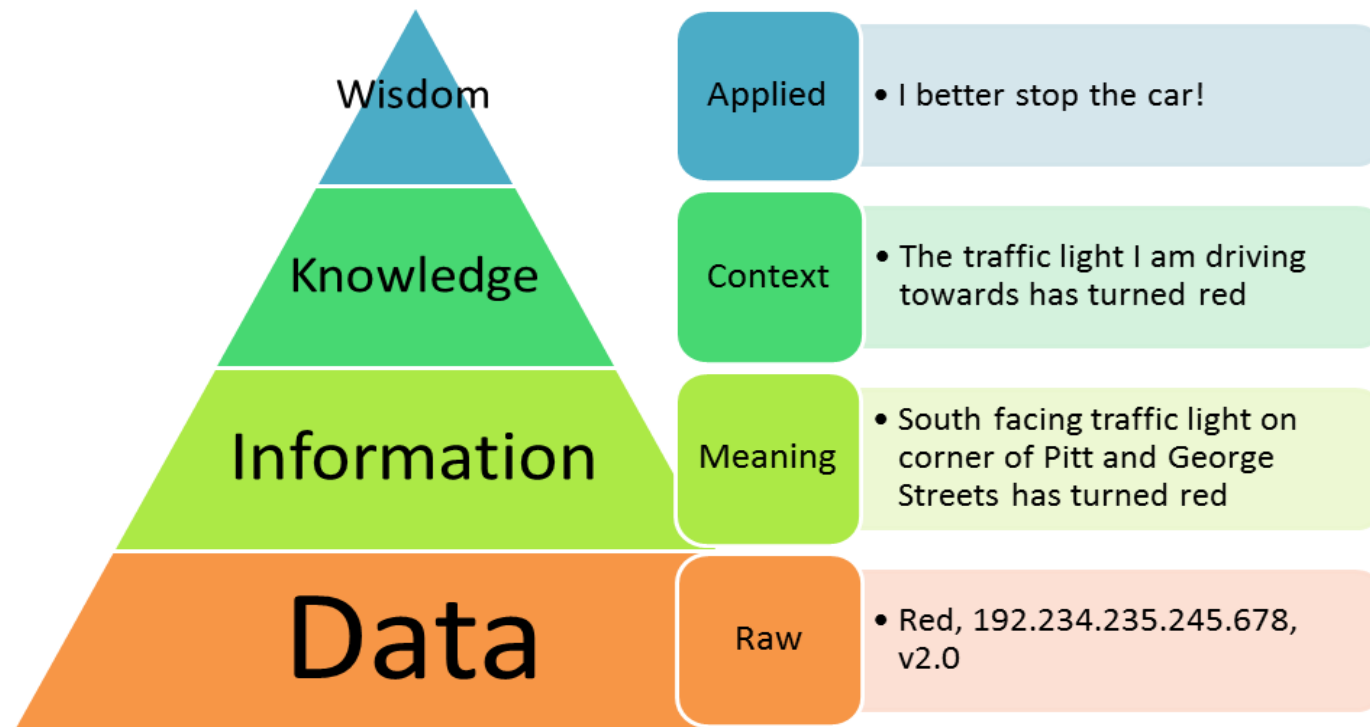
The output takes the form of **predictions** on new example

- a prediction of whether a particular customer will switch
- or
- a prediction of what kind of lens will be prescribed under given circumstances
- or
- a cancer diagnosis

The same as humans, machines learn by example

Data, Information, Knowledge, Wisdom (DIKW)

the **DIKW** hierarchy



<https://allthingy.com/data-information-knowledge-wisdom/>

© 2011 Angus McDonald

Data, Information, Knowledge, Wisdom (DIKW)

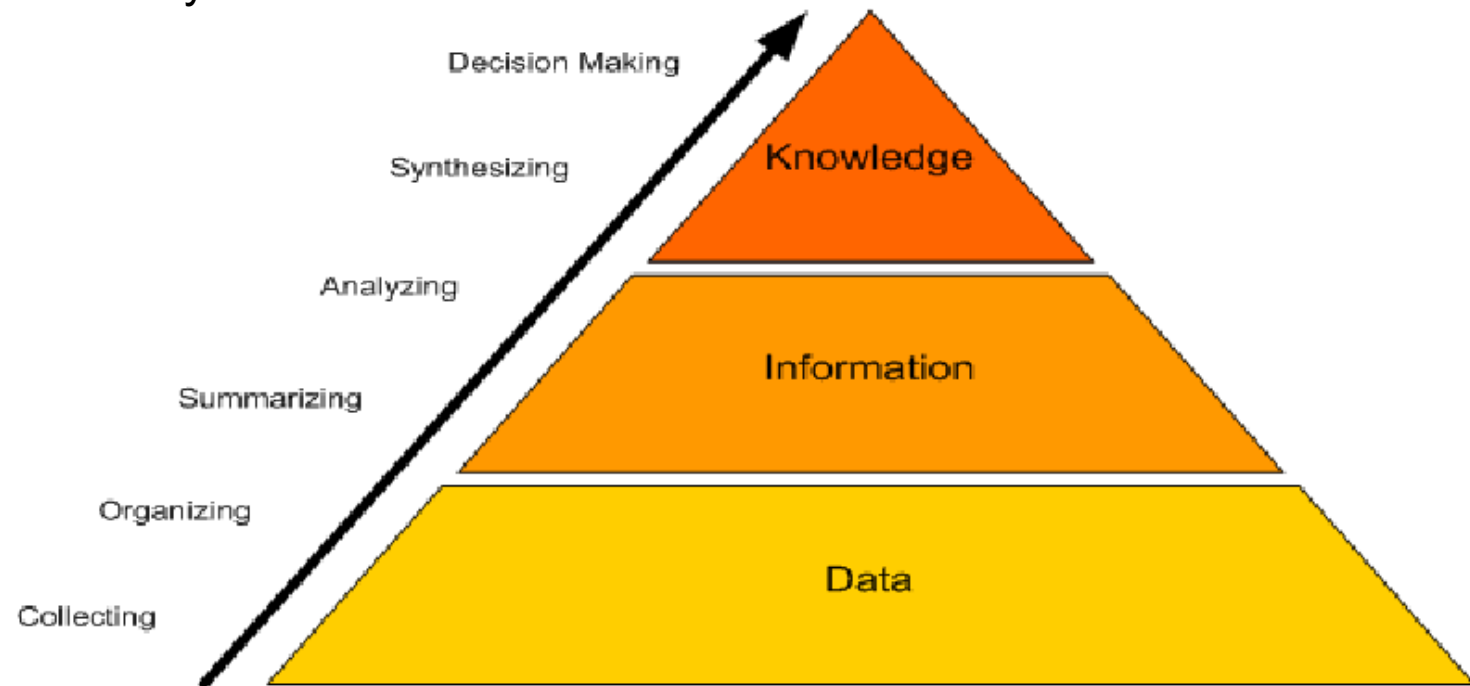
the **DIKW** hierarchy



https://www.researchgate.net/figure/The-data-information-knowledge-wisdom-DIKW-hierarchy-as-a-pyramid-to-manage-knowledge_fig6_332400827

Data, Information, Knowledge (DIK)

the **DIK** hierarchy

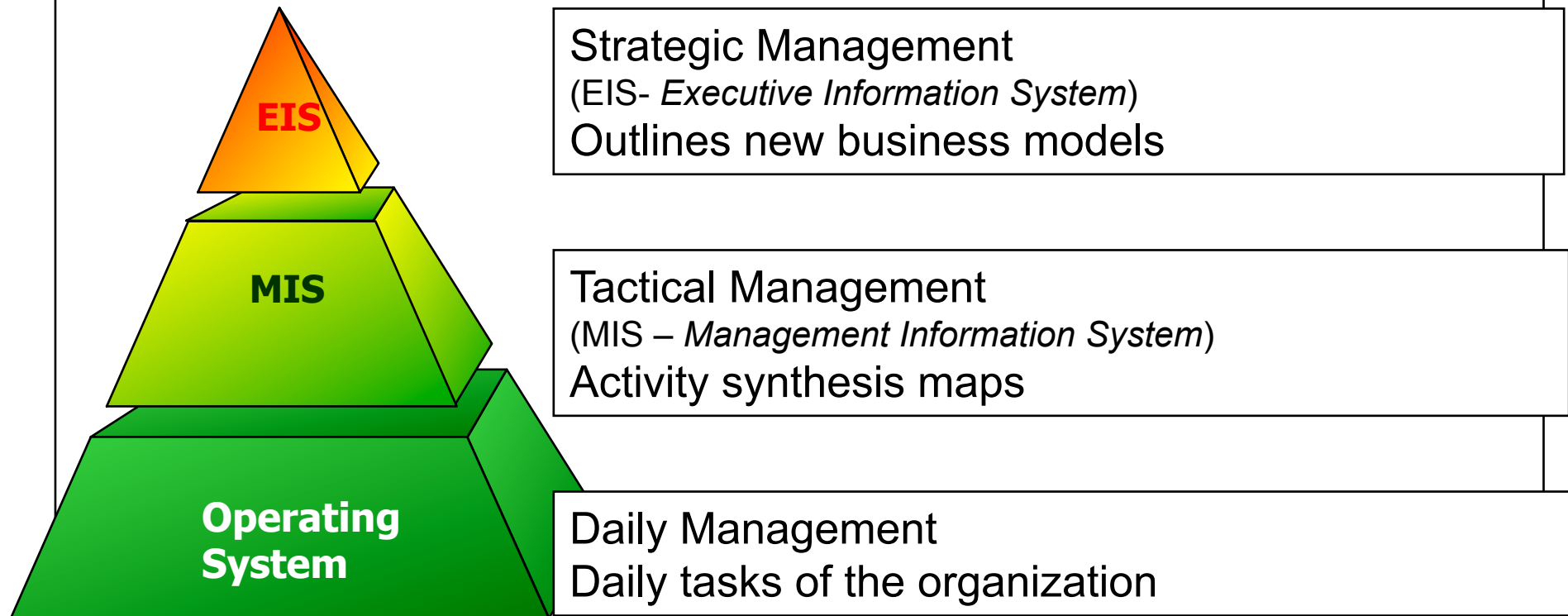


Finck, 2005)

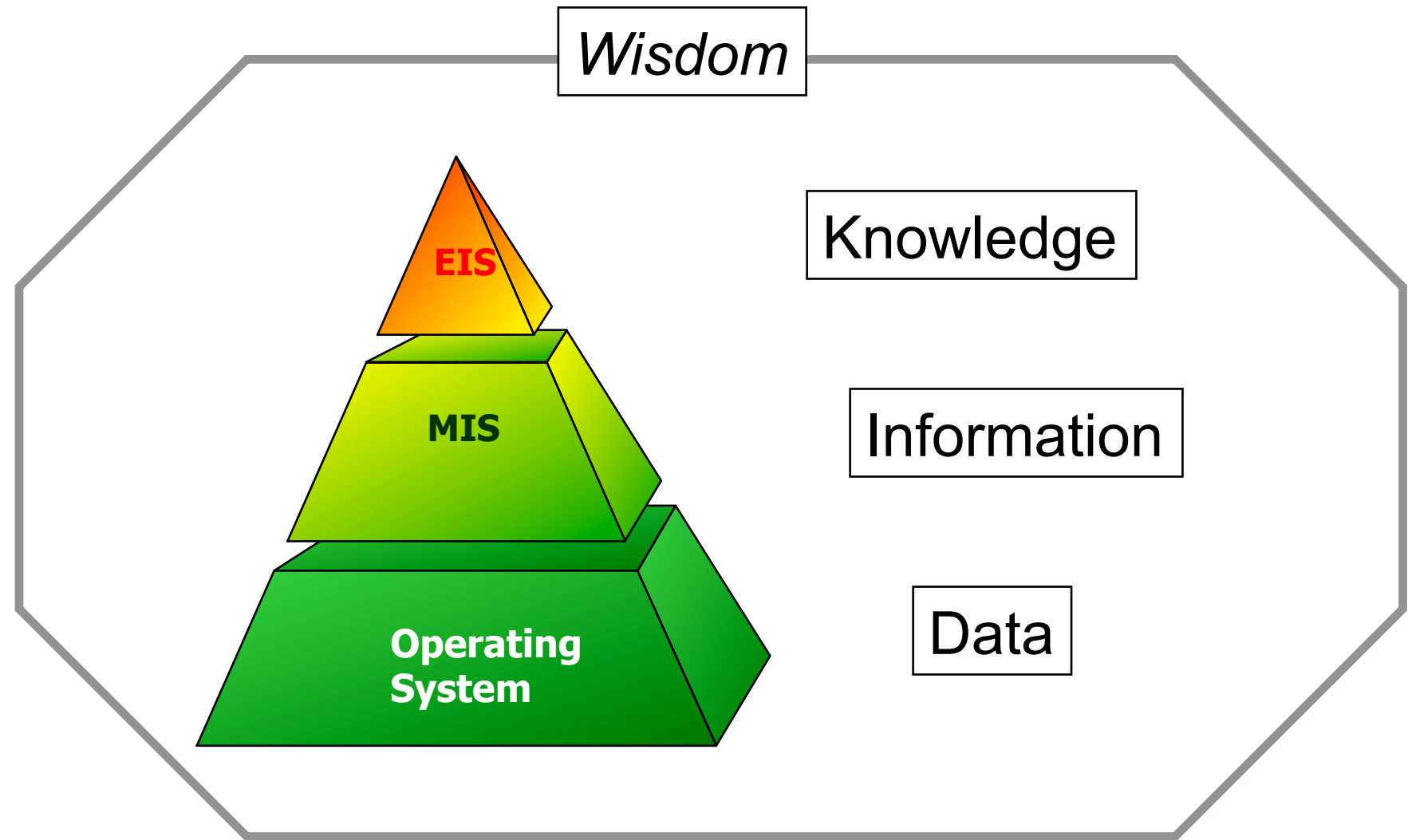
Figure 1: DIK Model

[https://www.semanticscholar.org/paper/Revisiting-Data-%2C-Information-%2C-Knowledge-and-\(-\)-Mutongi/14b904f0e28b26eaa214fe086f0df90ce502ae53](https://www.semanticscholar.org/paper/Revisiting-Data-%2C-Information-%2C-Knowledge-and-(-)-Mutongi/14b904f0e28b26eaa214fe086f0df90ce502ae53)

Stratified Perspective



... from Data to Wisdom



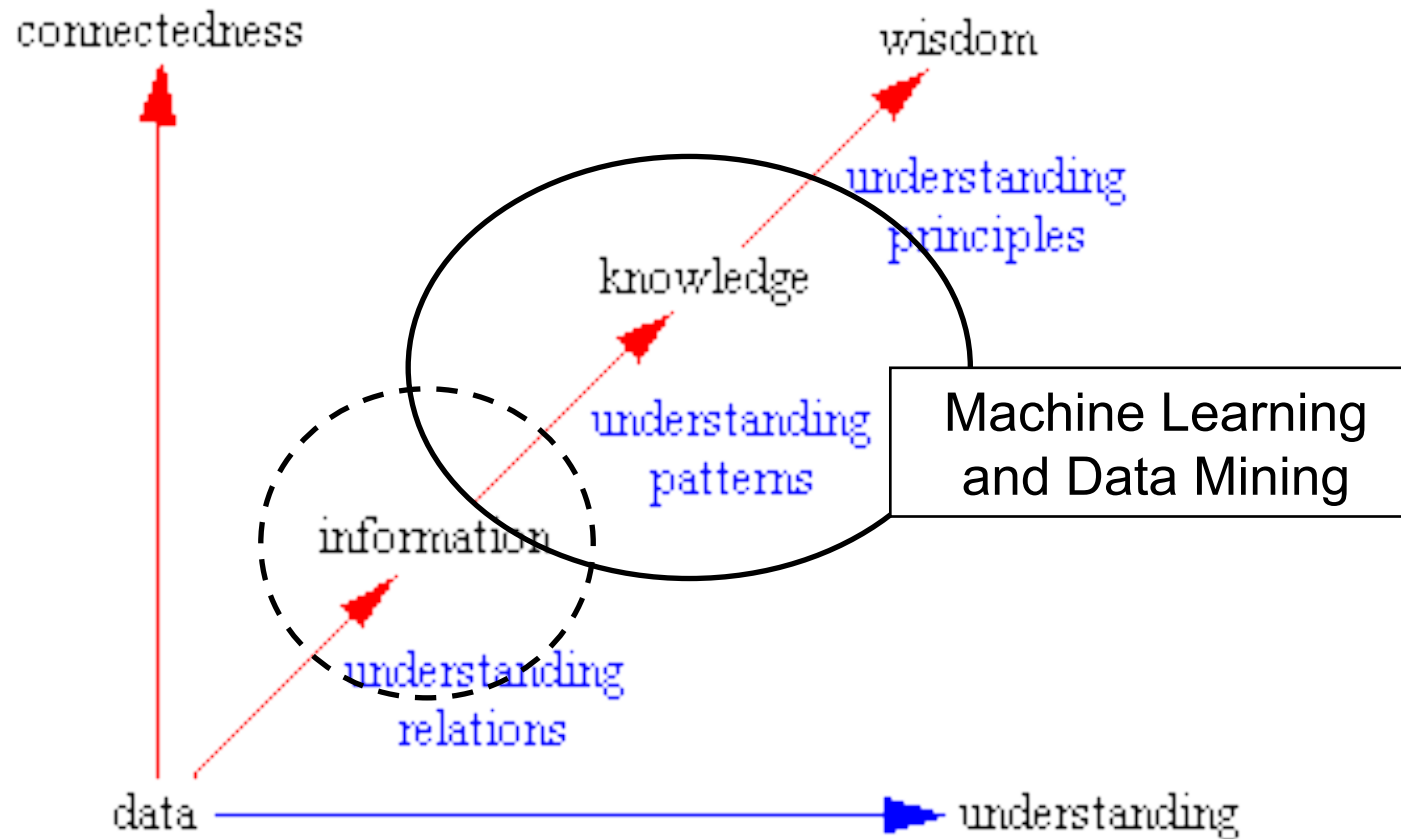
Taxonomy

- **Data**, relates with symbol perception
 - e.g., 2010.OUT.5, 155euros, Miguel, 33-MP-44
- **Information**, processing data to be useful
 - answer to the 4Ws questions
who (quem), *what* (qual), *where* (onde), *when* (quando)
- **Knowledge**, use of *data* and *information*
 - answers to questions of *how* (como)
- Comprehension
 - sensing the *how* (como) response
- **Wisdom**
 - Assessment on the comprehension made

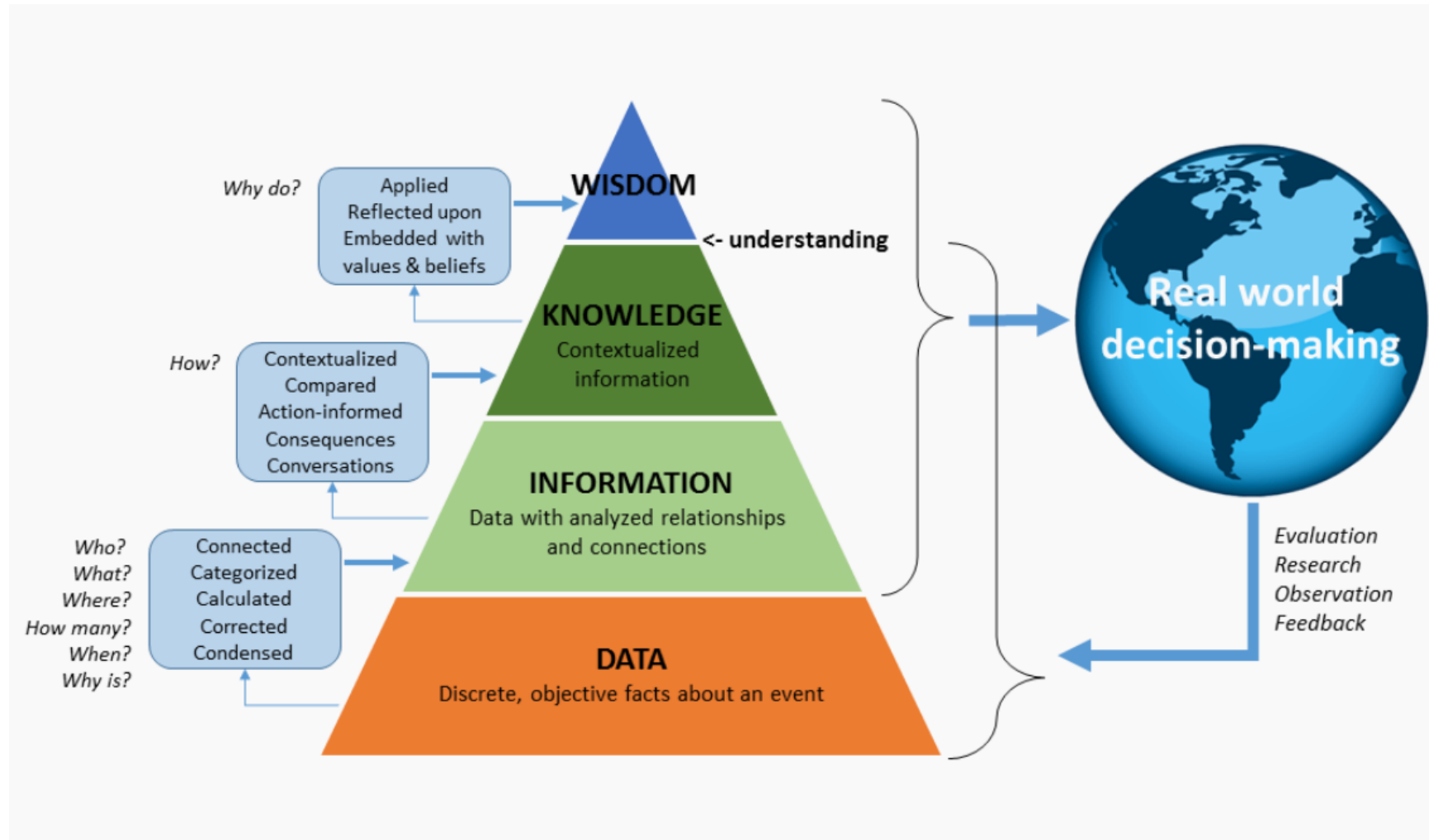
Taxonomy

- **Data** is a fact or event unrelated with “other things”
 - e.g., it is raining or the license plate is XX-YY-ZZ
- **Information** relates data on a given context
 - e.g., temperature decreases 10 degrees and it is raining
 - e.g., the license plate XX-YY-ZZ is from an electric car
- **Knowledge** incorporates a pattern that can provide prediction ability about what is being described or about the future
 - e.g., if the humidity is too high and the temperature drops sharply, then the atmosphere is likely to lose holding capacity and rain
 - e.g., if an electric car enters the highway and the battery has a limited capacity, it is likely to be refueled during the journey
- **Wisdom** includes to understand what builds the knowledge
 - e.g., the interaction between temperature gradients, evaporation, air currents and rain

Search for Connections to Reach Comprehension



... Reach Comprehension to Support Decision-Making



<https://universidadedatecnologia.com.br/dados-vs-informacoes-digitais/data-information-knowledge-wisdom/>

Business Analytics (for Data-Driven Decision-Making)

*Business analytics (BA) is the combination of **skills, technologies,** and **practices** used to examine an organization's **data** and performance as a way to gain insights and **make data-driven decisions** in the future using **statistical analysis**.*

*The goal of BA is to **narrow down which datasets are useful** and which can increase revenue, productivity, and efficiency.*

<https://learn.g2.com/business-analytics>

https://en.wikipedia.org/wiki/Business_analytics

Business Analytics (BA) and Business Intelligence (BI)

*Business Analytics (**BA**) makes extensive use of analytical modeling and numerical analysis, including explanatory and predictive modeling and fact-based management to drive decision making.*

***BA** may be used as input for **human decisions** or may drive fully **automated decisions**.*

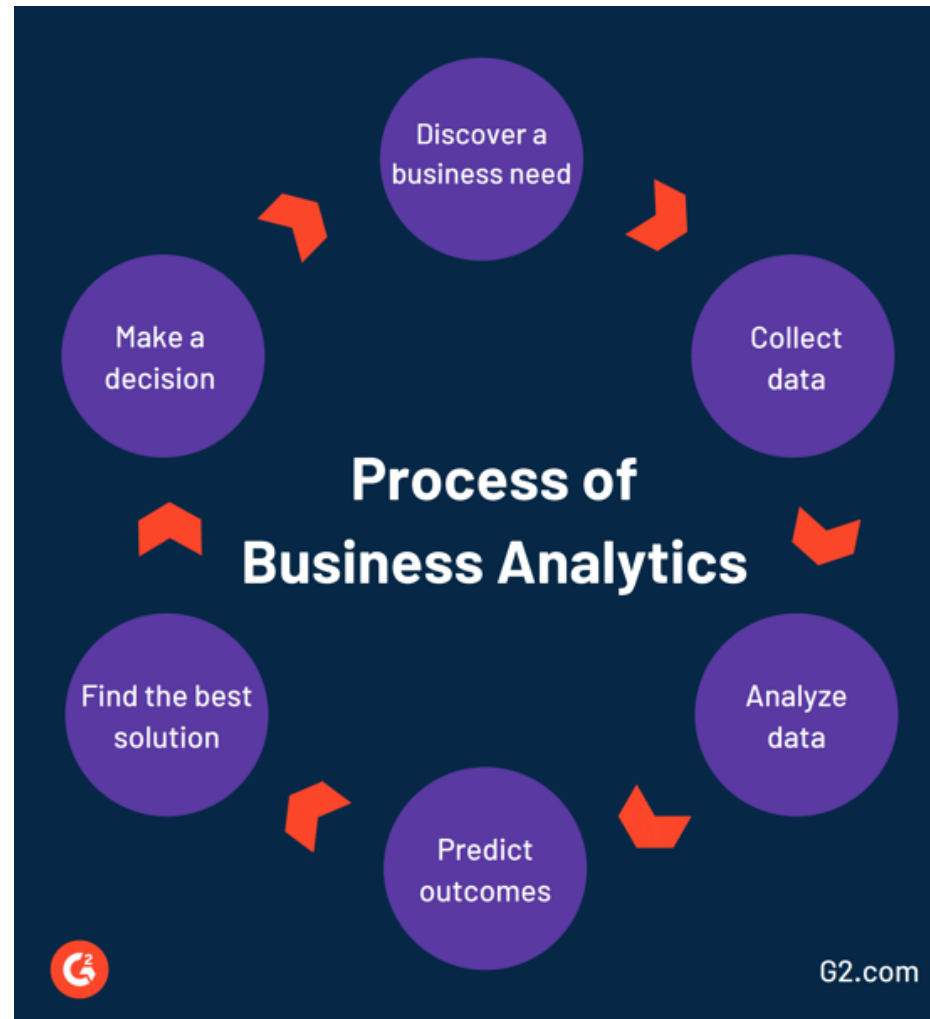
*Business Intelligence (**BI**) is querying, reporting, online analytical processing (OLAP), and "alerts".*

***BI** transforms data into actionable insights that inform an organization's **strategic** and **tactical** business decisions.*

<https://learn.g2.com/business-analytics>

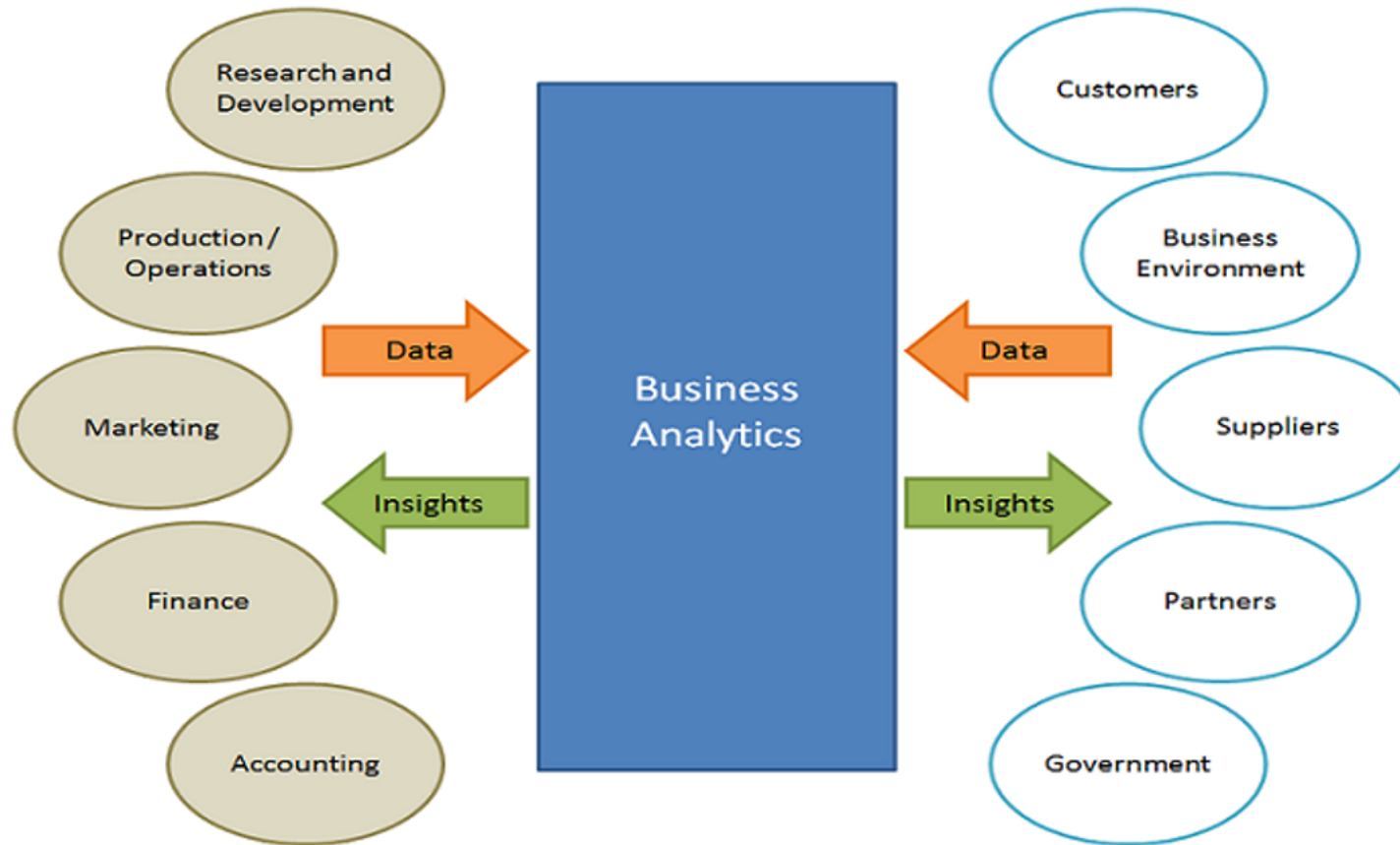
https://en.wikipedia.org/wiki/Business_analytics

Business Analytics (“a Workflow” Perspective)



<https://learn.g2.com/business-analytics>

Business Analytics (“a Service” Perspective)



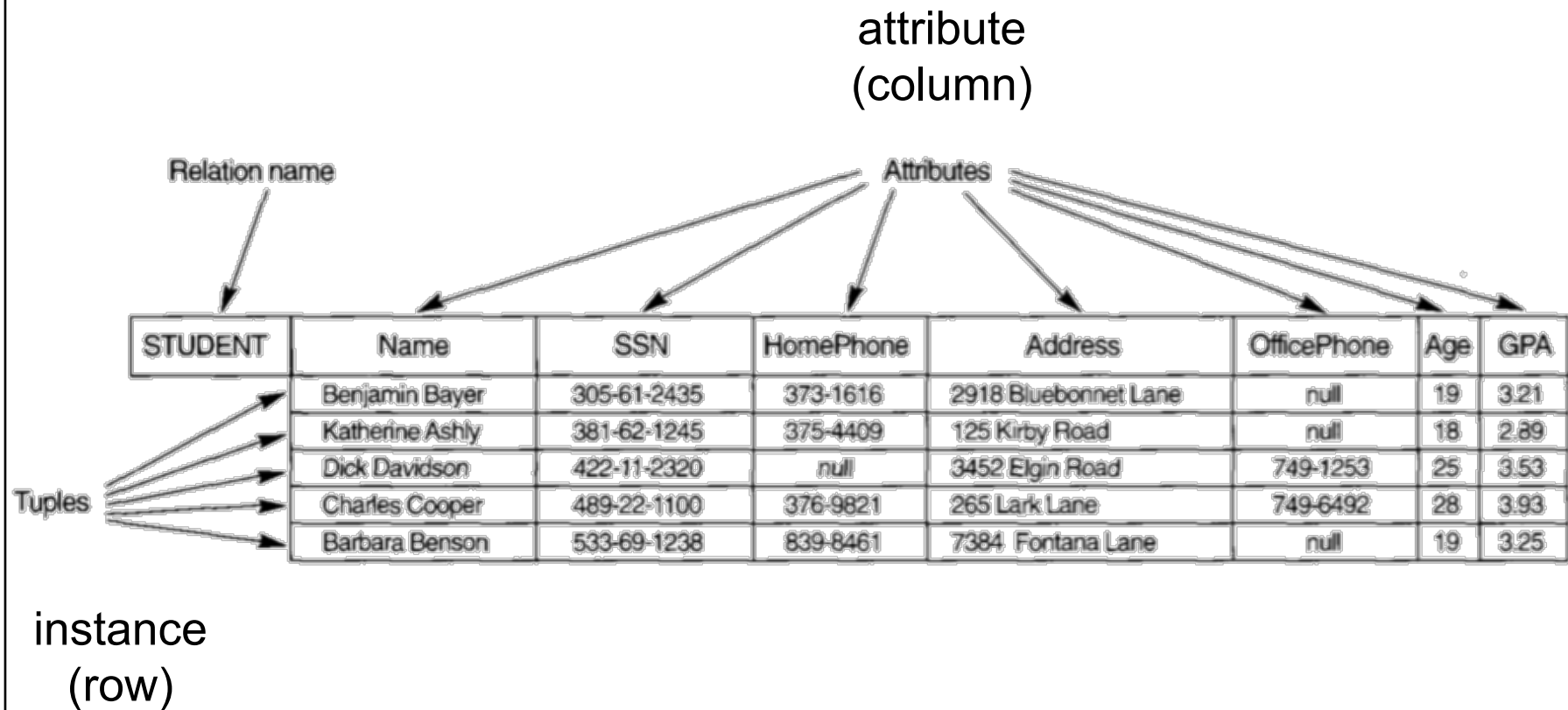
https://cio-wiki.org/wiki/Business_Analytics

<https://cio-wiki.org/images/4/41/BusinessaAnalytics.png>

Information and Information Systems (Core Concepts)

- Information Systems are an industry
 - with a strong foundation in the Database Management System
- The Database Management System
 - has on the relational model its formal support
- In the Relational Model, information is structured as
 - a collection of relations (schemes); a relation resembles a tables of values
- A relation scheme is
 - a sorted list of attributes, which one with its type
 - ... you can see the attribute as the role of a type in a context
- The relation schemes
 - are populated by instances (rows)
 - are “tied together” via (referential) integrity constraints among them

... an example of a relation scheme (table)



from: "Fundamentals of Database Systems"; Ramez Elmasri, Shamkant B. Navathe

Database and Database Management System

- A **database** is a highly structured collection of data stored and accessed electronically from a computer system.
- ... as databases become more complex, they are developed using formal modeling and design techniques.
- A **database management system (DBMS)** is a software that interacts with end users, applications, and the database itself to manipulate the data.
- ... the DBMS provides various functions that support CRUD (create, read, update, delete) operations; i.e., storage and retrieval of large quantities of data and ways to manage how that data is organized and managed.

extracted and adapted from:

<https://en.wikipedia.org/wiki/Database>

Database System and Database (“usual terminology”)

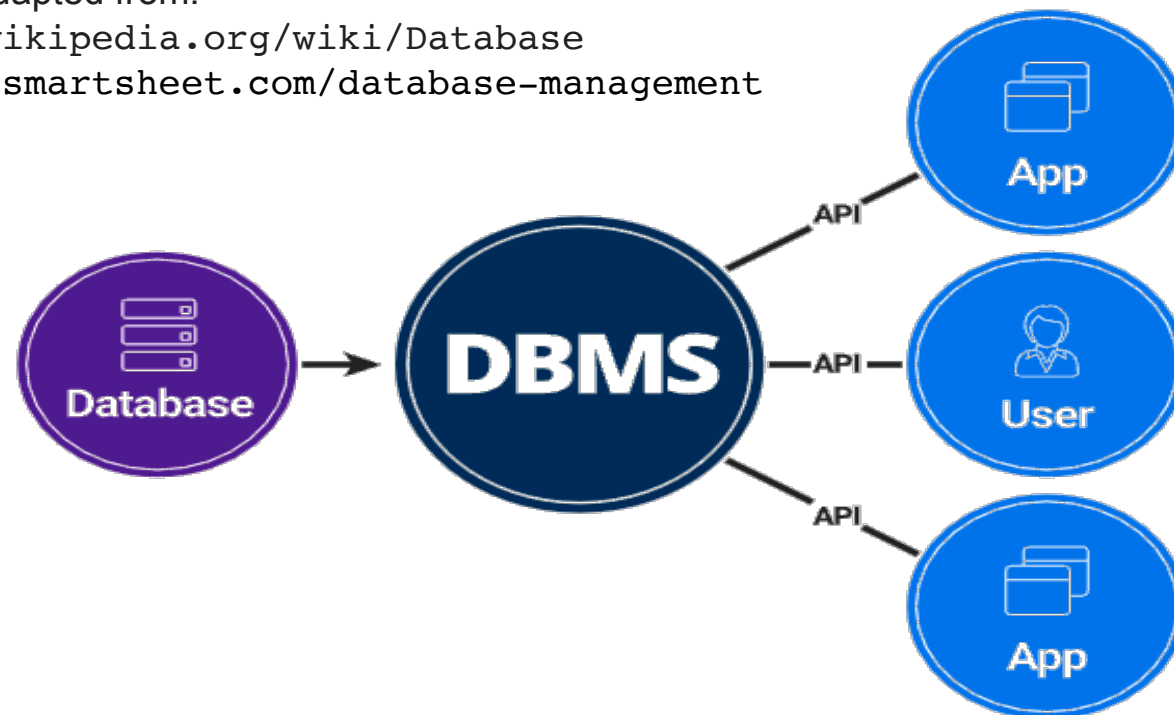
Often, the sum of the database, the DBMS and the associated applications can be referred to as a "database system".

Often, the term "database" is also used loosely to refer to any of the DBMS, the database system or an application associated with the database.

extracted and adapted from:

<https://en.wikipedia.org/wiki/Database>

<https://www.smartsheet.com/database-management>



Bibliography; Relational Model

Logical Database Design Principles;
John Garmany, Jeff Walker, Terry Clark
Auerbach Publications, 2005
(capítulos aconselhados: 1, 2, 3)

Beginning Databases with PostgreSQL; From Novice to Professional
Richard Stones, Neil Matthew
Apress, 2nd Edition, 2005
(capítulos aconselhados: 1, 2, 4, 6, 7)

Fundamentals of Database Systems
Ramez Elmasri, Shamkant B. Navathe
Addison-Wesley, 2007
(capítulos aconselhados: 1, 2, 3, 8, 9)

Bibliography; Knowledge Extraction

Data Mining - Practical Machine Learning Tools and Techniques

Ian H Witten, Eibe Frank, Mark Hall, Chris Pal

4th edition

Morgan Kaufmann, 2016

Data Mining Methods and Models

Daniel T. Larose

Wiley-IEEE Press, 2006