

COURSEWARE

AI Literacy Courseware

ENGLISH

Rowel Gündlach & Sophie Campman

 **Van Haren**
PUBLISHING



**AI Literacy Courseware
English**

Colophon

Title: AI Literacy Courseware - English

Authors: Rowel Gündlach
Sophie Campman

Publisher: Van Haren Publishing, 's-Hertogenbosch

ISBN Hard Copy: 978 94 018 1390 7

Edition: First edition, first print, November, 2025

Design: Van Haren Publishing, 's-Hertogenbosch

Copyright: © Van Haren Publishing 2025

For further information about Van Haren Publishing please e-mail us at: info@vanharen.net or visit our website: www.vanharen.net

No part of this publication may be reproduced in any form by print, photo print, micro-film or any other means without written permission by the publisher. Although this publication has been composed with much care, neither author, nor editor, nor publisher can accept any liability for damage caused by possible errors and/or incompleteness in this publication.

Publisher about the Courseware

The Courseware was created by experts from the industry who served as the author(s) for this publication. The input for the material is based on existing publications and the experience and expertise of the author(s). The material has been revised by trainers who also have experience working with the material. Close attention was also paid to the key learning points to ensure what needs to be mastered.

The objective of the courseware is to provide maximum support to the trainer and to the student, during his or her training. The material has a modular structure and according to the author(s) has the highest success rate should the student opt for examination. The Courseware is also accredited for this reason, wherever applicable.

In order to satisfy the requirements for accreditation the material must meet certain quality standards. The structure, the use of certain terms, diagrams and references are all part of this accreditation. Additionally, the material must be made available to each student in order to obtain full accreditation. To optimally support the trainer and the participant of the training assignments, practice exams and results are provided with the material.

Direct reference to advised literature is also regularly covered in the sheets so that students can find additional information concerning a particular topic. The decision to leave out notes pages from the Courseware was to encourage students to take notes throughout the material. Although the courseware is complete, the possibility that the trainer deviates from the structure of the sheets or chooses to not refer to all the sheets or commands does exist. The student always has the possibility to cover these topics and go through them on their own time. It is recommended to follow the structure of the courseware and publications for maximum exam preparation.

The courseware and the recommended literature are the perfect combination to learn and understand the theory.

-- Van Haren Publishing

Other publications BY Van Haren Publishing

Van Haren Publishing (VHP) specializes in titles on Best Practices, methods, frameworks and standards within four domains:

- IT Management
- Architecture (Enterprise and IT)
- Business Management and
- Project Management

Van Haren Publishing is publishing on behalf of leading organizations and companies: Agile Consortium, World Commerce and Contracting, IAOP, IPMA World, KNVI, PMI-NL, NLAIC and The Open Group.

Van Haren Publishing is part of the Van Haren Group and additional to the book publishing also provides the following services: accredited training materials and e-learning through Van Haren Learning Solutions, as well as independent professional certification via examination through Van Haren Certify.

Topics are (per domain):

| | | |
|--------------------------------|---------------------------------|---|
| IT Management | IT Service Management | FitSM, ISM®, ISO/IEC20000, IT4IT®, ITIL®, VerISM®, SAF, TRIM, XLA® |
| | Data Management | Data literacy, Data visualization, DMBOK |
| | IT Asset Management | HAM, ITAM, SAM |
| | IT Security Management | BIO, ISO/IEC27001, NIS2 |
| | Test Management | CTAP |
| | Application Management | ASL |
| | Other | eCF, IT-CMF, Scrum |
| Project Management | Project Management | Half Double, ICB, ISO/IEC21500, P3.express, PM2, PMBOK Guide, Praxis, PRINCE2 |
| | Agile | Agile, Agile PM |
| | Other | PMO |
| Business Management | Operations Management | Lean, Lean Six Sigma, OBM, OMC, RASCI |
| | Contract Management | CATS CM, CATS RVM, IACCM World |
| | Business Information Management | BiSL, DID |
| | Artificial Intelligence | AI, Generative AI |
| | Outsourcing | OPBOK |
| Enterprise Architecture | Enterprise Architecture | BIAN, TOGAF |
| | Modeling | ArchiMate, BPMN |
| | Software Architecture | ISAQB |
| | Other | Open Agile Architecture |

For the latest information on VHP publications, visit our website: www.vanharen.net.

Table of Content

| | <i>-- Slide Number</i> | <i>-- Page Number</i> |
|---|------------------------|-----------------------|
| Self-Reflection | | 6 |
| Agenda | | 8 |
| Part 1 - An introduction to AI | (13) | 15 |
| What is AI? | (15) | 16 |
| Why AI literacy? | (20) | 18 |
| The origins of AI | (24) | 20 |
| | | |
| Part 2 – Data and AI: How it works and its challenges | (31) | 24 |
| Data and data quality | (33) | 25 |
| The basic workings of AI | (45) | 31 |
| | | |
| Part 3 – Opportunities and challenges of AI in your work | (68) | 42 |
| Using AI in practice – output | (70) | 43 |
| Using AI in practice – input | (77) | 47 |
| | | |
| Part 4 – Ethics and regulation | (86) | 51 |
| Ethical considerations related to using AI | (88) | 52 |
| The European AI act | (100) | 58 |
| The future of AI | (128) | 72 |
| | | |
| Take home message | (136) | 76 |

Self-Reflection of understanding Diagram

‘What you do not measure, you cannot control.’ – Tom Peters

Fill in this diagram to self-evaluate your understanding of the material. This is an evaluation of how well you know the material and how well you understand it. In order to pass the exam successfully you should be aiming to reach the higher end of Level 3. If you really want to become a pro, then you should be aiming for Level 4. Your overall level of understanding will naturally follow the learning curve. So, it’s important to keep track of where you are at each point of the training and address any areas of difficulty.

Based on where you are within the Self-Reflection of Understanding diagram you can evaluate the progress of your own training.

| Level of Understanding | Before Training (Pre-knowledge) | Training Part 1 (1st Half) | Training Part 2 (2nd Half) | After studying / reading the book | After exercises and the Practice exam |
|---|---------------------------------|----------------------------|----------------------------|-----------------------------------|---------------------------------------|
| Level 4 <i>I can explain the content and apply it .</i> | | | | | ✓ |
| Level 3 <i>I get it!</i> <i>I am right where I am supposed to be.</i> | | | | | Ready for the exam! |
| Level 2 <i>I almost have it but could use more practice.</i> | | | | | |
| Level 1 <i>I am learning but don't quite get it yet.</i> | | | | | |

(Self-Reflection of Understanding Diagram)

Write down the problem areas that you are still having difficulty with so that you can consolidate them yourself, or with your trainer. After you have had a look at these, then you should evaluate to see if you now have a better understanding of where you actually are on the learning curve.

Troubleshooting

Problem areas:

Topic:

Part 1

Part 2

**You have gone
through the book
and studied.**

**You have answered
the questions and
done the practice
exam.**

Agenda

09:00 – 10:00: An introduction to AI

09:00 – 10:00: Break

10:10 – 11:00: Data and AI: How it works and its challenges

11:00 – 11:10: Break

11:10 – 12:00: Data and AI: How it works and its challenges

12:00 – 13:00: Break

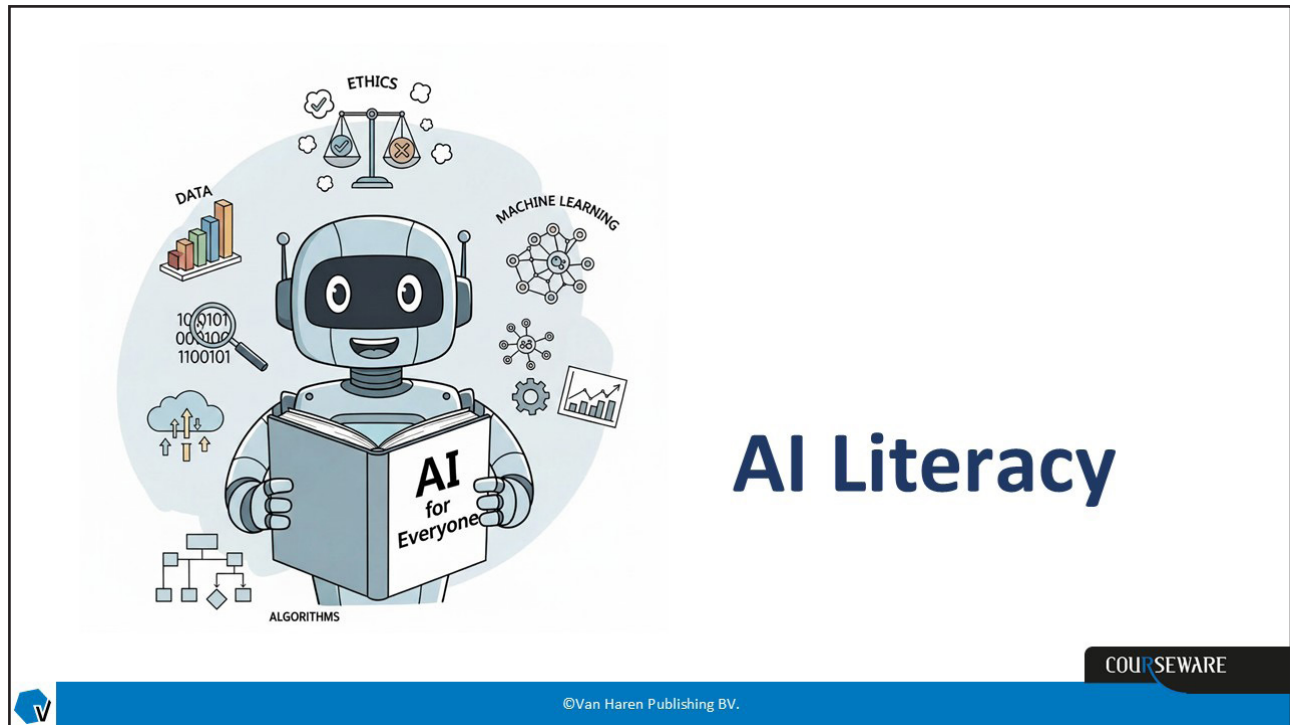
13:00 – 14:00: Opportunities and challenges of AI in your work

14:00 – 14:10: Break

14:10 – 15:00: Opportunities and challenges of AI in your work

15:00 – 15:10: Break

15:10 – 16:15: Ethics and regulation



The cover features a central illustration of a friendly robot holding a book titled "AI for Everyone". Surrounding the robot are various icons representing AI concepts: "ETHICS" (scales of justice), "DATA" (bar chart), "MACHINE LEARNING" (neural network), "ALGORITHMS" (flowchart), and binary code. The title "AI Literacy" is prominently displayed on the right side.

AI Literacy

COURSEWARE

©Van Haren Publishing BV.

Introduction

Who are you?

- Name & role
- Do you already encounter AI (at work or in daily life)
- What would you like to get out of this training?

The European AI Act, the reason we are here

- First comprehensive **law** in the world regulating Artificial Intelligence (AI)
- Classifies AI systems by **risk level**
- Sets rules for trustworthy AI in the EU
- Aimed at **developers, deployers, and users** of AI
- States that users of AI should be AI-literate



What does this mean for you?

- Dutch companies and public organizations must **comply with EU standards**
- Citizens will **interact more often with AI systems** in work, education, and daily life
- Growing demand for **AI understanding and responsible use** across all roles
- Need to recognize **bias, misinformation, and automation risks**
- AI literacy helps individuals **use AI safely and ethically** under the new regulations

Levels of AI literacy

- There are different **levels of data and AI literacy**, as defined by the Dutch government
 - Not every employee interacts with data and AI in the same way or with the same intensity
 - The expected skills and knowledge about data and AI vary by role and responsibility
 - There is a wide range of prior knowledge, some employees may have studied analytical topics or read about data-driven work, while others have little or no exposure to the subject

Levels of AI and data literacy and their core focus

1. Data and AI aware

Basic knowledge and awareness of AI and data

2. Data and AI proficient

Critical application and evaluation of AI and analyses. Basic knowledge of data collection.

3. Data and AI fluent

In-depth understanding and knowledge of working with data & AI, the prerequisites, and the opportunities.

4. Data and AI expert

Designing, developing, and optimizing data and AI solutions.

Levels and core focus as described by the Dutch government data and AI literacy working group.

After today you will be ...

1. Data and AI aware

Basic knowledge and awareness of AI and data

2. Data and AI proficient

Critical application and evaluation of AI and analyses. Basic knowledge of data collection.

3. Data and AI fluent

In-depth understanding and knowledge of working with data & AI, the prerequisites, and the opportunities.

4. Data and AI expert

Designing, developing, and optimizing data and AI solutions.

Levels and core focus as described by the Dutch government data and AI literacy working group.

Core competences

1. Data and AI aware

1. Basic **knowledge** of common concepts related to data and AI
2. Ability to **recognize data** and **AI applications** in work processes
3. Understanding the **limitations** and **ethical considerations** of using data and AI
4. Ability to **work with** readily available **AI** and **data tools**

2. Data and AI proficient

1. Critical **evaluation** of **AI output** and data insights
2. Understanding how data and AI **support** and **influence decisions** in the organization, and the role they play in work and daily tasks
3. Ability to **identify opportunities** for using data and AI tools, formulate tasks and develop solutions
4. Basic knowledge of **AI and data risks** as well as **ethical issues** within the organization

Learning objectives

After today, you will be able to...

- Understand **common terms** related to data and AI
- Understand that AI is a **tool** and not a definitive authority
- **Critically** assess the **output** generated by AI systems
- Recognize where AI and data **applications** are present in your work and daily tasks
- Understand how AI and data **support** and influence your work and daily tasks
- Identify **opportunities** for using data and AI tools, formulate clear tasks, and facilitate the development of solutions within your organization
- Understand **limitations** and **ethical challenges** associated with using AI
- **Effectively** and **correctly** work and interact with **readily available AI** and data tools

Content overview

Part 1 – An introduction to AI

- What is AI?
- Why AI literacy?
- The origins of AI

Part 2 – Data and AI: How it works and its challenges

- Data and data quality
- The basic workings of AI

Content overview

Part 3 – Opportunities and challenges of AI in your work

- Using AI in practice – output
- Using AI in practice – input

Part 4 – Ethics and regulation

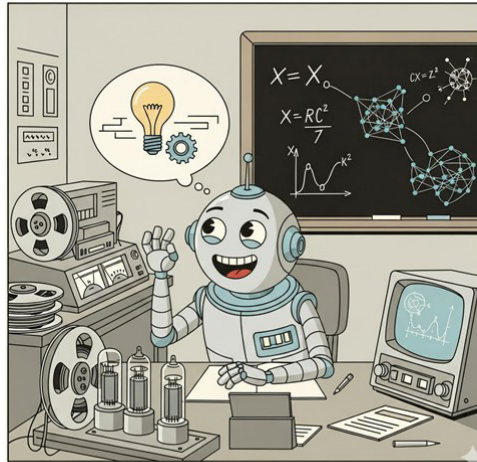
- Ethical considerations related to using AI
- The European AI act
- The future of AI

Take home message

Agenda

- 09:00 – 10:00 > An introduction to AI
- 10:00 – 10:10 > Break 01 (10 min)
- 10:10 – 11:00 > Data and AI: How it works and its challenges
- 11:00 – 11:10 > Break 02 (10 min)
- 11:10 – 12:00 > Data and AI: How it works and its challenges
- 12:00 – 13:00 > Break (60 min)
- 13:00 – 14:00 > Opportunities and challenges of AI in your work
- 14:00 – 14:10 > Break 03 (10 min)
- 14:10– 15:00 > Opportunities and challenges of AI in your work
- 15:00 – 15:10 > Break 04 (10 min)
- 15:10– 16:15 > Ethics and regulation

Part 1 – An introduction to AI



© Van Haren Publishing BV.

13

Core competences addressed in Part 1

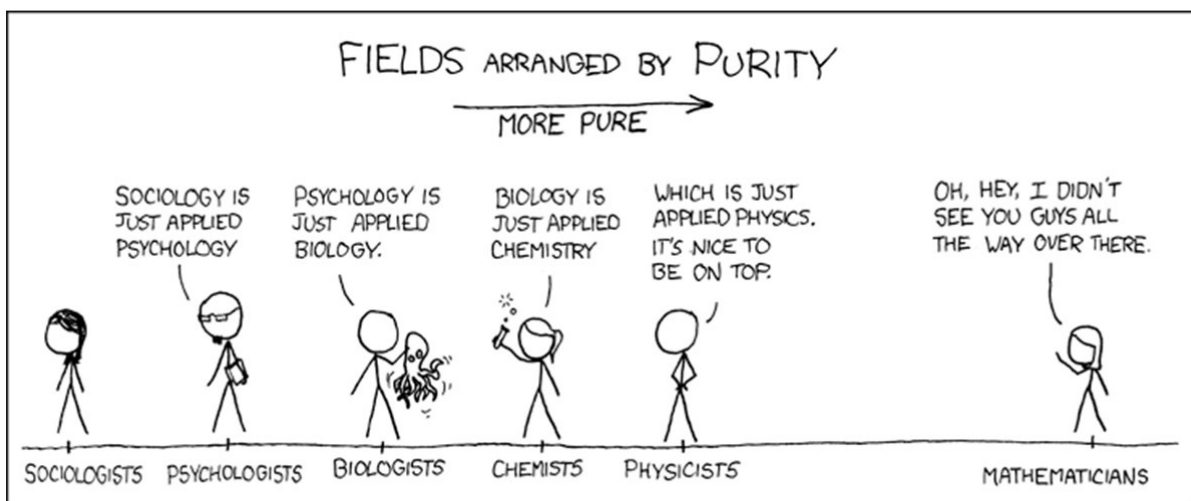
1. **Basic knowledge** of common concepts related to data and AI

© Van Haren Publishing BV.

14

What is AI?

From mind to math - the inspiration for AI

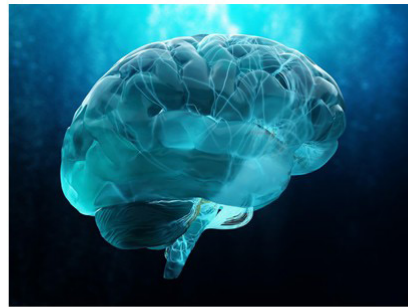


Source: Xkcd <https://xkcd.com/435/>

The definition of Intelligence

Intelligence can be defined as the ability to perceive or infer **information**, and to retain it as **knowledge** to be applied towards adaptive **behaviors** within an environment or context.

Artificial Intelligence (AI) can be defined as intelligence exhibited by **machines** with cognitive functions that are associated to humans.



The concept of AI

- AI helps machines perform tasks that typically require **human intelligence**, e.g.,
 - Understanding language
 - Recognizing images or sounds
 - Learning from data
 - Solving problems
 - Interacting with people (e.g., chatbots)
- AI learns from data
 - AI learns from examples and experiences, not from fixed rules
 - AI recognizes patterns

The concept of AI



Why AI literacy?

AI literacy

- It helps you **understand** and **effectively** use AI
- It enables you to recognize **bias**
- It helps you to think **critically** about AI and its ethical implications
- It enables you to use AI **consciously** and **responsibly**

The European AI Act

Article 4 of the EU AI act reads

“Providers and deployers of AI systems shall take measures to ensure, to their best extent, a sufficient level of AI literacy of **their staff and other persons dealing with the operation and use of AI systems** on their behalf, taking into account their technical knowledge, experience, education and training and the context the AI systems are to be used in, and considering the persons or groups of persons on whom the AI systems are to be used.”

The European AI Act

Article 4 of the EU AI act reads

“Providers and deployers of AI systems shall take measures to ensure, to their best extent, a sufficient level of AI literacy of **their staff and other persons dealing with the operation and use of AI systems** on their behalf, taking into account their technical knowledge, experience, education and training and the context the AI systems are to be used in, and considering the persons or groups of persons on whom the AI systems are to be used.”

Since 2 February 2025: **mandatory** to be AI literate  **This training!**

The origins of AI

How did we get where we are now?

The history of AI

Inspiration

- In the late 19th century, people began to understand more about human biology

The start of AI

- 1950s: The birth of the concept of AI

AI in the fifties – The Turing test

- Alan Turing published an important paper in 1950 in which he speculates about the possibility of creating machines that think
- While the basic concepts of AI are still in their early stages, he devises his famous Turing test

The Turing test is a test of a machine's ability to exhibit intelligent behavior that is indistinguishable from that of a human.



The person in the middle has to identify who is the computer and who is the human. If they fail, the Turing test is passed

The history of AI

Inspiration

- In the late 19th century, people began to understand more about human biology

The start of AI

- 1950s: The birth of the concept of AI
- 1960s: Disappointing results, little development
- 1990s: Machines start to improve at games (checkers, chess)
- 2000s: Smart machines begin to make our lives easier
- 2010s: Big data, cloud, open source

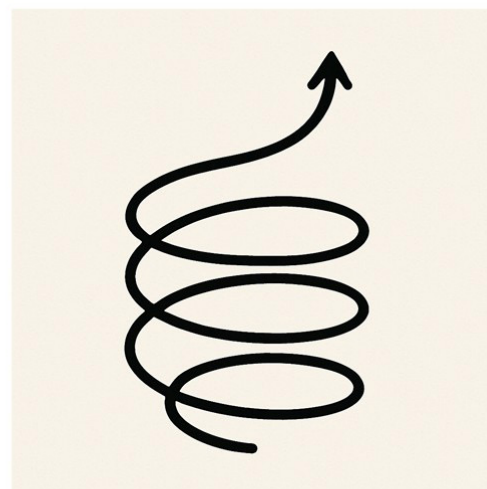
The current landscape of AI - upwards spiral

A Major Step for AI: Generative AI

- Since 2022, everyone has access to ChatGPT
- Rapid growth of generative AI

The successes of AI inspire more entities to develop and implement AI

- Companies invest in AI
- Students study and work in AI
- More technological advances in AI
- Companies invest more in AI
- ...



Why now

Due to recent technological advances, we were able to make significant steps in the field of AI. The main contributing factors are

Data eruption

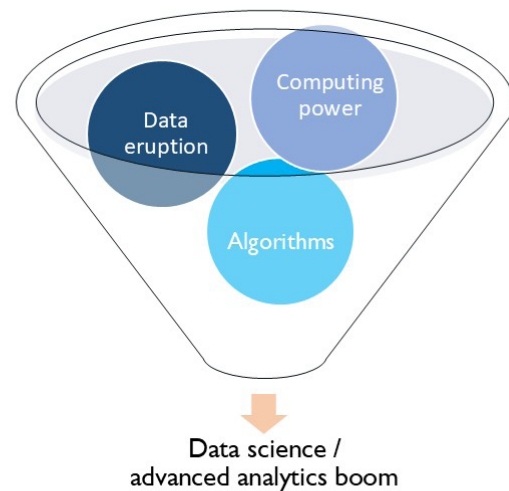
- More data available
- More storage capacity

Computing power

- Stronger CPU / GPU

Advanced algorithms

- Smarter mathematical models make AI stronger and more efficient



AI as the fourth industrial revolution



First (18th and 19th centuries)

- **Steam engine**



Second (1870 to 1914)

- **Electricity** (telephone, light bulb)



Third (1980s)

- Digital and **ICT** (information and communications technology)(personal computers, internet, automation)



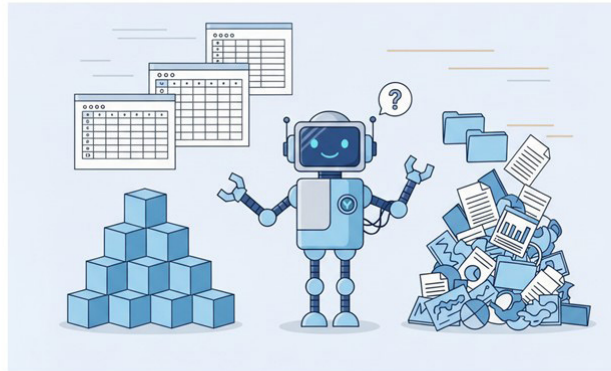
Fourth (today)

- Exploits the digital revolution and is disruptive, driven by **AI**, robotics, Internet of Things, 3D printing, nanotech, bio-engineering.

The Fourth Industrial Revolution is changing every area of our lives.

Klaus Schwab – Founder of the World Economic Forum

Part 2 – Data and AI: How it works and its challenges



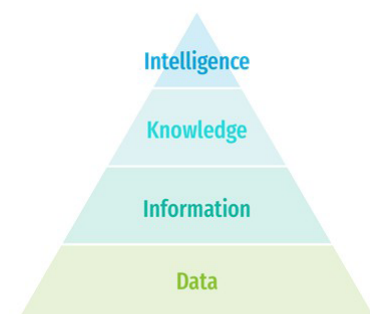
Core competences addressed in Part 2

1. **Basic knowledge** of common concepts related to data and AI
2. Ability to **recognize data** and **AI applications** in work processes

Data and data quality

Data, Information, Knowledge and Intelligence

- **Data:**
 - raw building blocks, numbers without context
- **Information:**
 - structure or order in data
- **Knowledge:**
 - understanding of the information
- **Intelligence:**
 - Ability to use the information to make decisions
 - Recognize objects and events
 - Present knowledge in a real-world model
 - Reason about the future



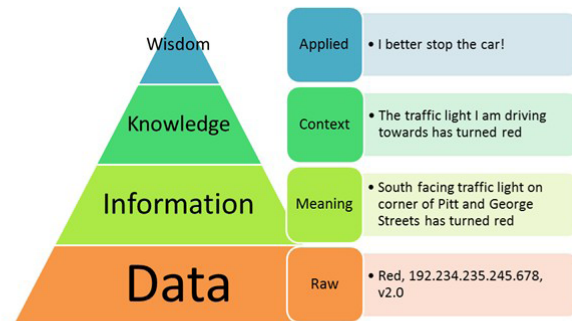
Data, Information, Knowledge and Intelligence - Example

- **Data:** 24, 25, 25, 26, 28, 28, 30
- **Information:**

| | | | | | | |
|------|-------|-------|-------|-------|-------|-------|
| 9:00 | 10:00 | 11:00 | 12:00 | 13:00 | 14:00 | 15:00 |
| 24°C | 25°C | 25°C | 26°C | 28°C | 28°C | 30°C |

- **Knowledge:** It is getting warmer
- **Intelligence:** I should turn on the AC!

- In practice, 'information' is also often called 'data' (and we will do so from now on)



Unstructured vs structured data

Unstructured

The university has 5600 students.
 John's ID is number 1, he is 18 years old and already holds a B.Sc. degree.
 David's ID is number 2, he is 31 years old and holds a Ph.D. degree. Robert's ID is number 3, he is 51 years old and also holds the same degree as David, a Ph.D. degree.

Semi-structured

```
<University>
  <Student ID="1">
    <Name>John</Name>
    <Age>18</Age>
    <Degree>B.Sc.</Degree>
  </Student>
  <Student ID="2">
    <Name>David</Name>
    <Age>31</Age>
    <Degree>Ph.D. </Degree>
  </Student>
  ....
</University>
```

Structured

| ID | Name | Age | Degree |
|----|---------|-----|--------|
| 1 | John | 18 | B.Sc. |
| 2 | David | 31 | Ph.D. |
| 3 | Robert | 51 | Ph.D. |
| 4 | Rick | 26 | M.Sc. |
| 5 | Michael | 19 | B.Sc. |

Source: https://www.researchgate.net/publication/236860222_Developing_Dynamic_Packaging_Applications_using_Semantic_Web_based_Integration

Types of Data

- **Structured data:** stored in a particular, well-defined format, fit for analysis as is:
 - Numbers, table, an Excel sheet
- **Unstructured data:** no predefined specific data model that all data points will adhere to:
 - Text, video, audio, internet content
 - Note: unstructured data is not immediately suitable for analysis
- **Semi-structured data:** no fixed data model, but metadata indicates contents:
 - JSON, XML are well known examples
 - Metadata gives semantic meaning to data:


```
{ "fruit": "banana", "color": "red" }
```

```
{ "fruit": "apple", "country": "Belgium" }
```

Big data

Big Data is the ocean of information we swim in every day – vast zettabytes of data flowing from our computers, mobile devices, and machine sensors.

Big data refers to **large and complex datasets** that cannot be easily processed using traditional data processing techniques. These datasets often include data from a variety of sources, such as social media, sensor data, and transactional data.

This data is used by organizations to drive decisions, improve processes and policies, and create customer-centric products, services, and experiences.



Data quality

- Data is the key driving force for AI
- Based on existing data, an **algorithm** finds certain patterns that is used to make future predictions

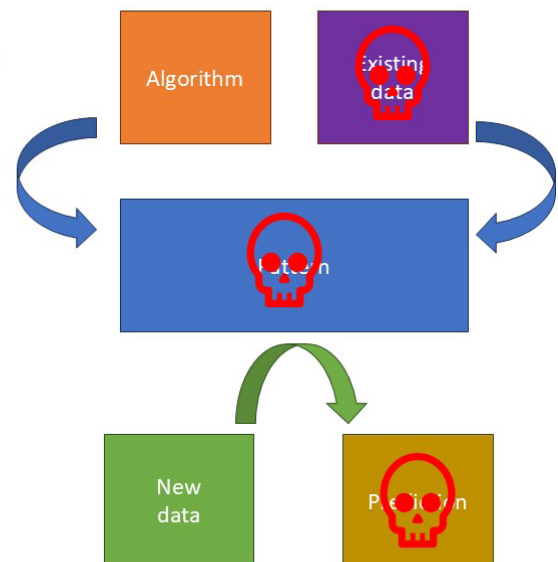
We make two key observations

- To make accurate predictions, an algorithm needs to find complex patterns in the data, which it only can when there is a large amount of data available
- The quality of the predictions depends on the quality of the data
- An algorithm can get stuck in a certain pattern

Data quality rule of thumb

The result of a machine learning algorithm is, at most,
as good as the data that it uses

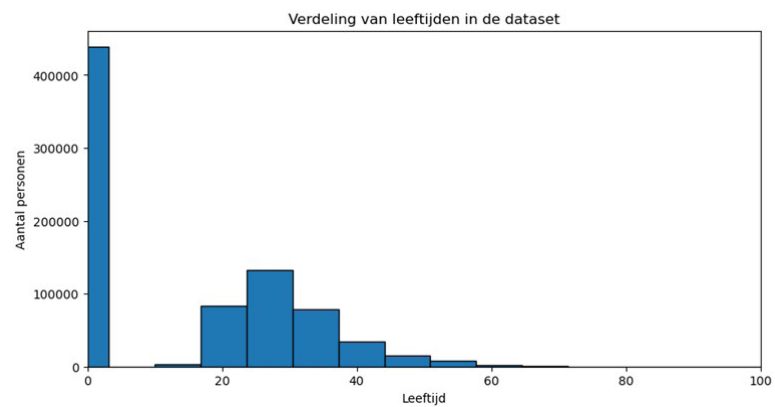
- Bad data = Bad results
- However, getting “good” data can be challenging



Challenge - Accuracy

Is it correct?

- Distribution of age



Sometimes data fields are not filled in and get a 0 by default. If this is not noticed, one may make predictions based on people with an age of 0!

Challenge - Human error

Is it correct?

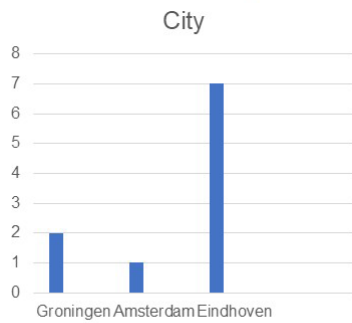
- Distribution of age
- Addresses

| ID | Owner | City | Zip code |
|----|---------------|-----------|----------|
| 1 | D. de Jong | Groningen | 9718 AC |
| 2 | M. Bas | Amsterdam | 1011 NK |
| 3 | D. de Waard | Groningen | 9718 AC |
| 4 | S. van Houten | Eindhoven | 1556 PE |

When it comes to manually inserting data (for example zip codes) it is very easy to make a mistake. Eindhoven has a region code of 56, so there was a typo, but this is not easy to notice.

Challenge - Uniqueness

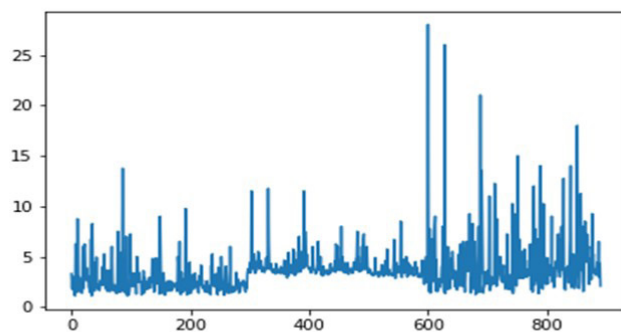
- Is it free from duplication?



| ID | Owner | City | Zip code |
|----|---------------|-----------|----------|
| 1 | D. de Jong | Groningen | 9718 AC |
| 2 | M. Bas | Amsterdam | 1011 NK |
| 3 | D. de Waard | Groningen | 9718 AC |
| 4 | S v houten | Eindhoven | 1556 PE |
| 5 | S v houten | Eindhoven | 5615 PE |
| 6 | S van houten | Eindhoven | 5615 PE |
| 7 | S van Houten | Eindhoven | 5615 PE |
| 8 | S. van Houten | Eindhoven | 5615 PE |

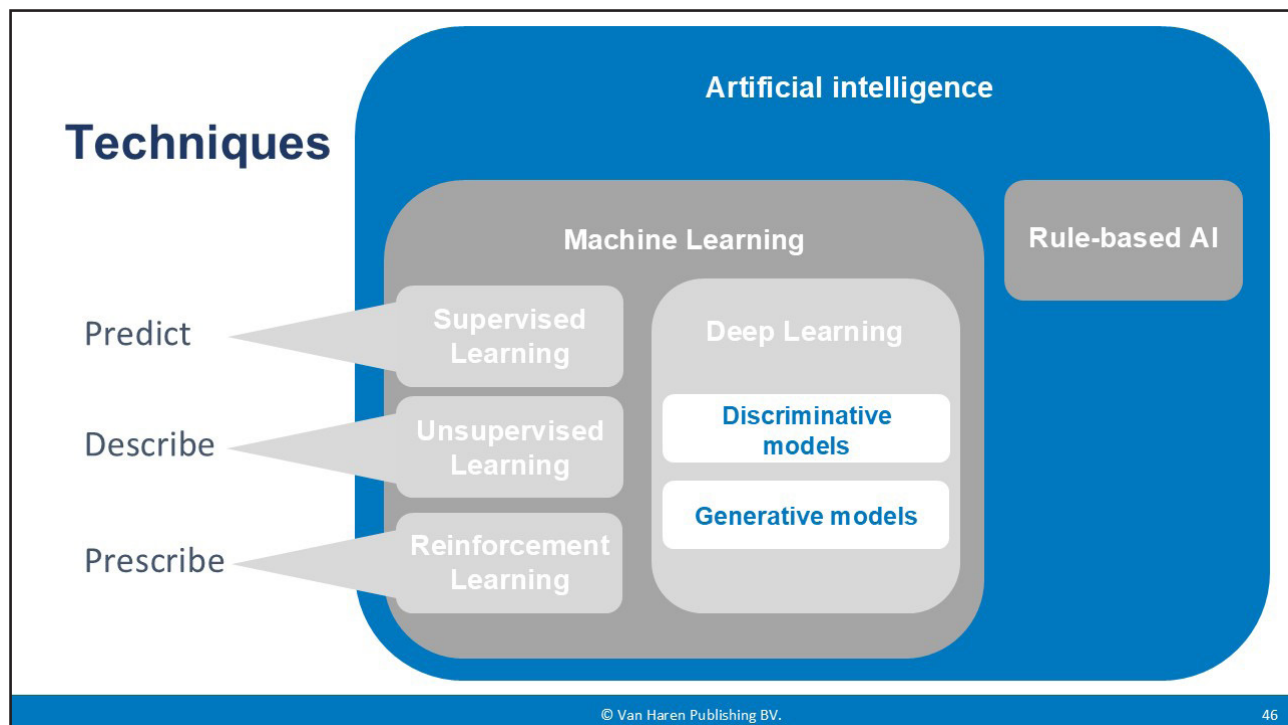
Duplication can introduce bias. In this example one may think the majority of the owners are from Eindhoven. However, identifying duplicates can sometimes be challenging (for example two of the exact same transaction that are made 2 seconds after each other)

Challenge - Automated data collection



Sometimes unforeseen factors can break a stable trend in a process. For example, when an airco breaks for an hour, the temperature can systematically rise for a certain period of time. However, this behavior is typically not what we want to include in our analysis

The basic workings of AI



Machine learning definition

Tom Mitchell's definition is widely quoted:

*"the field of machine learning is concerned with the question of how to construct computer programs that automatically **improve with experience.**"*

"a computer program is said to learn from experience, E, with respect to some class of tasks, T, and performance measure, P, if its performance at tasks, T, as measured by P, improves with experience, E."

Machine learning models

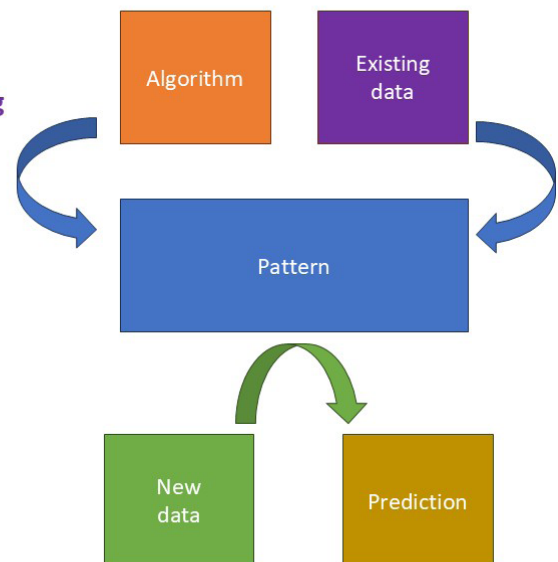
Machine learning models combine **algorithms** (mathematical techniques) in combination with **existing data** to make predictions, such as:

- The answer to a question (ChatGPT)
- The likelihood of someone committing fraud
- If this object is a dog or a cat

Existing data contains hidden patterns

Algorithms uncovers hidden patterns in an efficient manner

Idea: Uncover hidden patterns based on existing data, then use this pattern to **predict** based on **new data**



Machine learning models - Example

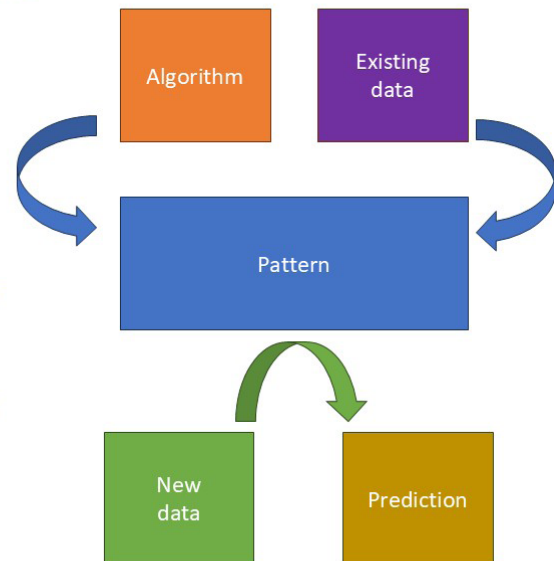
Goal: Predict if a customer is going to pay back a loan

Existing data: Data of past customer, indicating who has and who has not paid back their loans

Algorithm: I look at 5 people in my data set that most closely resemble the current customer, if the majority has paid back their loan, I approve the loan

New data: A new customer comes in, I note all properties of this customer

Prediction: Based on the previously defined algorithm, I approve the loan or not



© Van Haren Publishing BV.

49

Machine learning models - Exercise

For the following applications think of how AI can be used to analyze the following

- Use AI to increase customer satisfaction
- Use AI to when to do car maintenance
- Use AI to determine which places in the city need cleaning

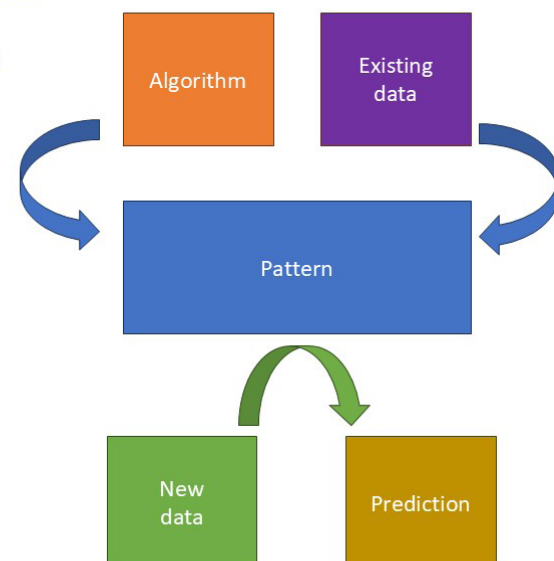
Find what is the:

Existing data:

Algorithm: [Done by the trainer]

New data:

Prediction:



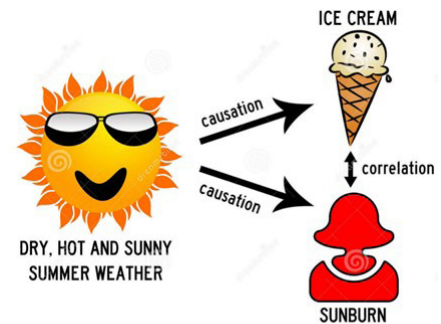
© Van Haren Publishing BV.

50

Correlation and causation

Be careful how to interpret such a pattern!

- Correlation means that **a relation between two variables exists**
 - Positive correlation: if X is bigger, then Y is also bigger
 - Negative correlation: if X is bigger, then Y is smaller
- Causation is when **one event leads to another**
 - If X happens, then Y is also more likely because X can cause Y to happen



Correlation **does not mean** causation: cause and effect are often influenced by other variables.

Machine learning will never tell you a causal relationship, only correlations!

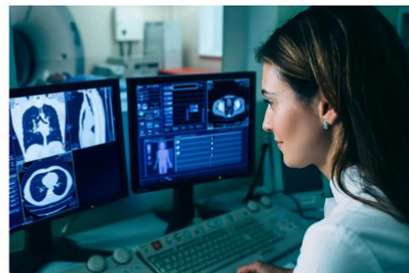
Prediction (supervised learning)

Based on past data AI can recognize certain **patterns** in the data to **predict** a certain outcome

- Predicting sales
- Spam filtering
- Fraud detection
- Predicting potential crime
- Predicting optimal candidates for a position
- Predicting health based on x-ray images

Artificial intelligence better than humans at spotting lung cancer

Researchers have used a deep-learning algorithm to detect lung cancer accurately from computed tomography scans. The results of the study indicate that artificial intelligence can outperform human evaluation of these scans.



By Aca Sandolu on May 20, 2019 — Fact checked by Isabel Godfrey

Latest news

- Does a vegetarian diet increase the risk of bone fractures?
- Sundowning in Alzheimer's: Light sensitivity may worsen late-day symptoms
- Can the MIND diet or cutting daily calories improve cognition?
- Pecans may have protective effects against obesity, diabetes
- Could a bodybuilding supplement help stave off dementia?

Prediction (supervised learning)

A typical data set used to make predictions, contain

- Labels: The outcome, **the property you want to predict**, for example:
 - How many sales are we going to make?
 - Is this a fraudulent transaction?
 - Is this a healthy lung or not?
- Features: The information, **the properties that you know**, for example
 - What is the weather, which day of the week is it ...
 - What is the transaction amount, to which country is the transaction...
 - Which parts of the X-ray picture are which color ...

Description (unsupervised learning)

Based on data, AI can find **clusters** in data of objects that typically occur together

- Customer segmentation
- Anomaly detection
- Market basket analysis

