

Introduction: The 3 Pillars

Algorithms, Data and Security
A.Y. 2025/26

Valeria Cardellini

Global Governance, 3rd year
Major in Science and Technology

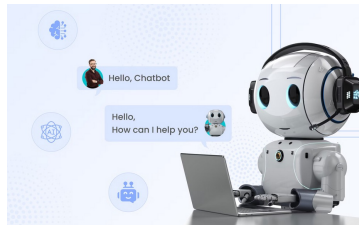
The 3 course pillars

Algorithms, Data and Security

- Why should we care?
- Digital technologies are rapidly transforming and shaping our world
 - Algorithms influence decisions
 - Data drives innovation
 - Security protects individuals, organizations, and societies

The IT revolution

- Smart living
- Smart cities
- AI and automation
- Future of work
- New experiences



Valeria Cardellini - ADS 2025/26

2

Why is the IT sector so important?

- Largest communication companies?
- Largest retailer?
- Most popular media platforms?
- Largest movie house?
- Largest software vendors?
- Largest taxi company?
- Largest accommodation provider?

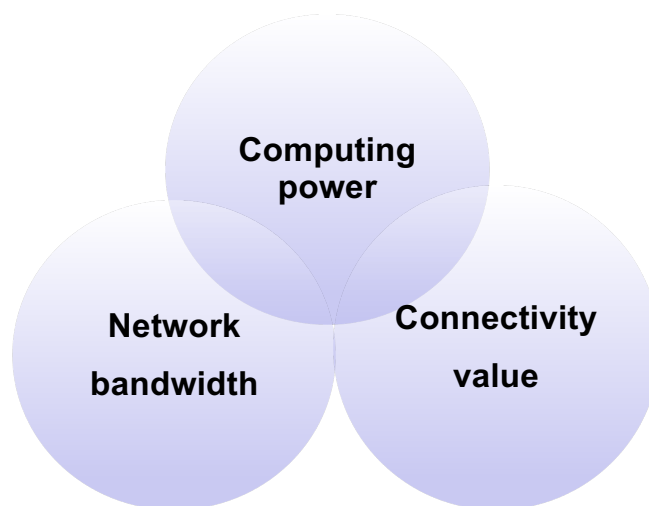
Valeria Cardellini - ADS 2025/26

3

Why is the IT sector so important?

- Largest communication companies? *No telecom infrastructure (Skype, Telegram, WeChat, WhatsApp)*
- Largest retailer? *Carries no physical inventory (Alibaba)*
- Most popular media platforms? *Create little or no original content (Instagram, TikTok, Youtube)*
- Largest movie house? *Owns no cinemas (Netflix)*
- Largest software vendors? *Rely heavily on third-party app developers (Apple, Google)*
- Largest taxi company? *Owns no cars (Uber)*
- Largest accommodation provider? *Owns no property (Airbnb)*

Why is it possible?



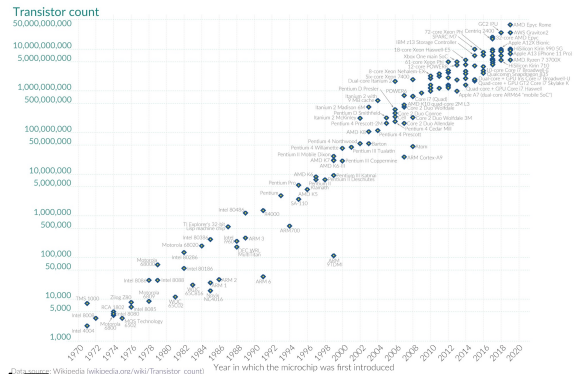
Computing power

- **Moore's law**

The number of transistors on a computer chip doubles about every 18 months

- Computers got

- Faster
- Smaller
- Cheaper
- Power efficient



- Slowed down in 2015-2017

- Physical and energy limits

- What do we do now?

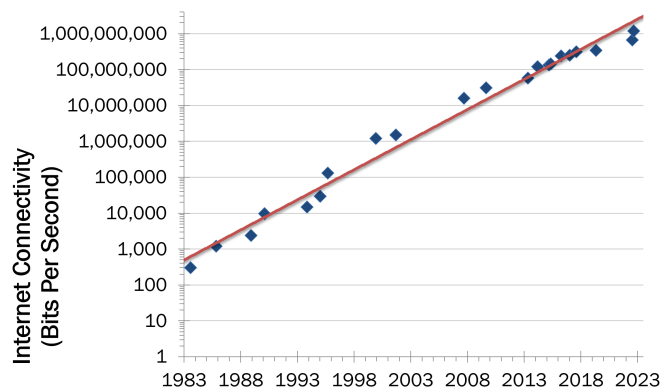
- Put multiple processors (cores) inside one chip
- Make computers work on many tasks at the same time

Network bandwidth

- **Nielsen's law**

Users' bandwidth grows by 50% per year

<https://www.nngroup.com/articles/law-of-bandwidth/>



- Why it matters

- Faster bandwidth makes new online services possible (e.g., video streaming, gaming, virtual reality)

Network bandwidth

- **5G**: super-fast low latency Internet
 - Faster than typical home broadband
 - Wireless connectivity for cities
- How fast is 5G?
 - 4G: up to 100 megabits per second (Mbps)
 - 5G: up to 10 *gigabits* per second (Gbps)
 - 100x faster than 4G (in theory)
- **Satellite Internet**
 - Internet from space (no cables or towers)
 - Covers rural and remote areas
 - High speeds (100-250 Mbps)



Connectivity value

- **Metcalfe's law**

The value of a network is proportional to the square of the number of its users

 - 2 users → 1 possible connection
 - 5 users → 10 possible connections
 - 100 users → 4,950 possible connections
- Networking is *socially* and *economically* important
 - More users → more interactions
 - More interactions → more value
 - More value → stronger incentives to join



What is an algorithm?

- Informal definition
- Sequence of **instructions**, given **step by step**, that could be executed “**easily**” and so that they produce a desired result (**output**)

What is an algorithm?

- How to solve a problem: compute (efficiently) output from input
- Problem: Wish to prepare a chocolate cake
- How to solve this problem? Recipe!

Input

200g **golden caster sugar**
200g **unsalted butter**, softened plus extra for the tins
4 **large eggs**
200g **self-raising flour**
2 tbsp cocoa powder
1 tsp **baking powder**
½ tsp **vanilla extract**
2 tbsp **milk**

Recipe
and you!

Output



A classical algorithm

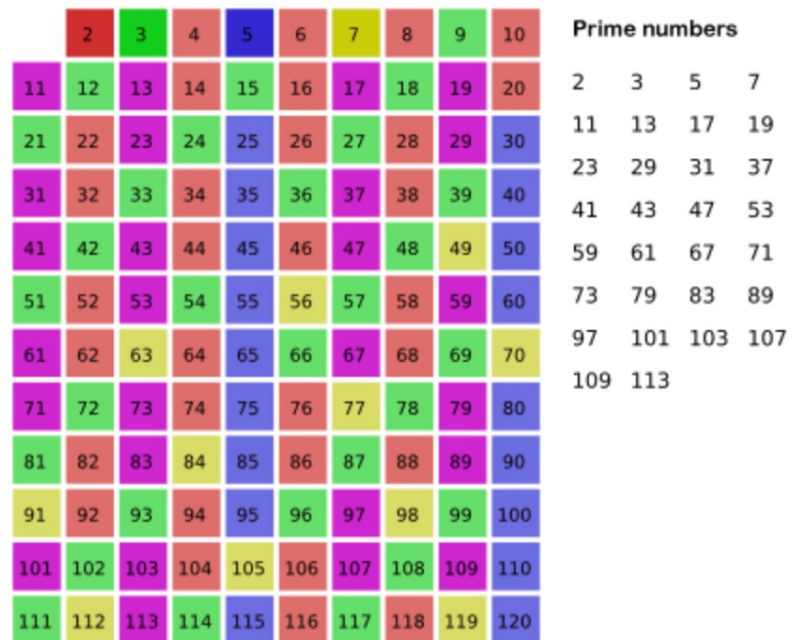
- **Eratosthenes' sieve** (275-195 BC)
 - Goal: find all prime numbers $\leq n$
1. Write a list of numbers from 2 to n
 2. Start with $k = 2$ (the first prime)
 3. Mark all multiples of k from k^2 to n (they cannot be prime)
 4. Replace k with the next smallest unmarked number $> k$
 5. If $(k^2 \leq n)$ go back to step 3
 6. Output all unmarked numbers: these are the primes

Eratosthenes' sieve: example with $n=120$

	2	3	4	5	6	7	8	9	10	Prime numbers
11	12	13	14	15	16	17	18	19	20	
21	22	23	24	25	26	27	28	29	30	
31	32	33	34	35	36	37	38	39	40	
41	42	43	44	45	46	47	48	49	50	
51	52	53	54	55	56	57	58	59	60	
61	62	63	64	65	66	67	68	69	70	
71	72	73	74	75	76	77	78	79	80	
81	82	83	84	85	86	87	88	89	90	
91	92	93	94	95	96	97	98	99	100	
101	102	103	104	105	106	107	108	109	110	
111	112	113	114	115	116	117	118	119	120	

Eratosthenes' sieve: example with n=120

- See https://en.wikipedia.org/wiki/Sieve_of_Eratosthenes



Valeria Cardellini - ADS 2025/26

14

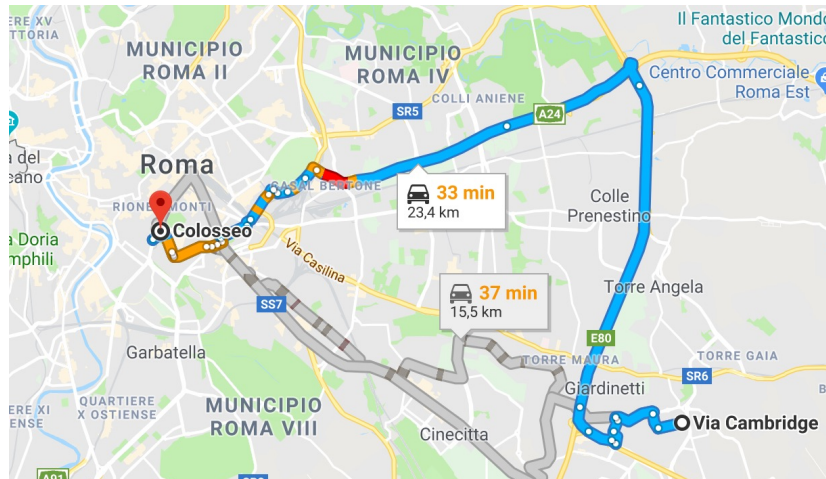
Other examples of algorithms

- Count distinct objects
 - Example list: b, a, b, c, g, d, d, g
- Sort list of names
- Search on Google
 - We will study **PageRank**, invented by Google's founders to rank web pages in search results
 - Now used in bibliometrics, social network analysis, link prediction and recommendation



Other examples of algorithms

- Find the shortest distance
- Google Maps uses Dijkstra's **shortest path algorithm** to calculate routes from A to B



Other examples of algorithms

- **AlphaGo**: AI mastering Go
 - Go: complex board game, more difficult for computers than chess
 - AlphaGo: computer program developed by DeepMind that plays Go
 - Uses machine learning (ML) to learn from vast amounts of Go game data
 - Employs artificial neural networks to evaluate board positions and select moves
 - Combines traditional search techniques with learned knowledge to make decisions
 - Breakthrough in **reinforcement learning** and **neural networks**

<https://www.nytimes.com/2017/05/23/business/google-deepmind-alphago-go-champion-defeat.html>



Artificial neural network

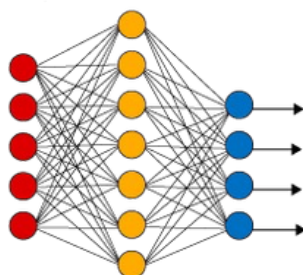
- **Artificial neural network**: computing system inspired by biological neural networks that constitute animal brains
 - It “learns” to perform tasks by studying examples, without being explicitly programmed with specific rules
 - Example: image classification
 - The network studies many labeled images, like “cat” or “dog” (**training**)
 - It learns patterns that help it recognize these animals
 - Later, it can identify cats or dogs in new, unseen images (**inference**)



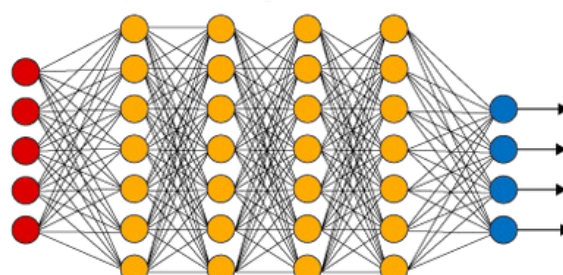
Artificial neural network

- Neural networks consist of many connected nodes, organized in layers
 - **Input layer**: receives the raw data
 - **Output layer**: produces the final result or prediction
 - **Hidden layers**: each hidden layer transforms the data using **nonlinear functions**, enabling the network to learn complex patterns

Simple Neural Network



Deep Learning Neural Network



● Input Layer

● Hidden Layer

● Output Layer

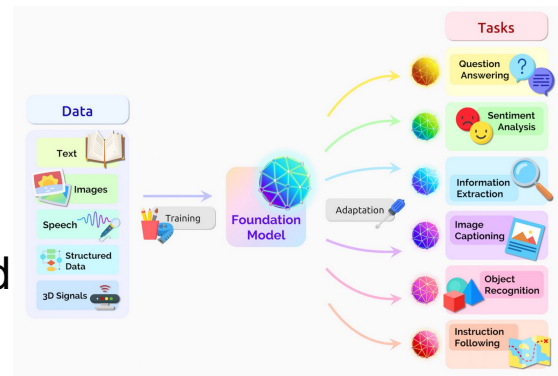
Transformers and foundation models

- **Transformer**

- A type of neural network that learns context and meaning by tracking relationships in sequential data (like words in this sentence)

- **Foundation models** (aka large pretrained models, e.g., GPT-4)

- Large-scale AI models trained on massive unlabeled data that can be adapted for many tasks
- Learn general knowledge first, which can then be fine-tuned for specific applications



"The sheer scale and scope of foundation models over the last few years have stretched our imagination of what is possible" <https://crfm.stanford.edu/report.html> 20

Algorithms and Big Data

- Why algorithms matter more today: Big Data is bigger than ever

- Early recognition

- The Economist, 2010: [The Data Deluge](#)
- NYT in 2012: [The Age of Big Data](#)



- Today

- Humans can't sift through massive data manually
- AI and ML extract patterns at scale
- Business and science depend on data-driven decisions

Why Big Data?

How much data is created every single minute of the day?

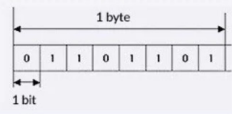
Global Internet population in Jan. 2026: 6.04 billion (73% of world population)
1 billion in 2005



Source: <https://www.domo.com/data-never-sleeps>

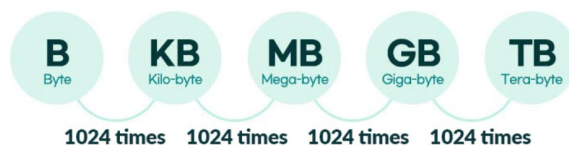
Demystifying data units

- Bit and byte

Bit (binary digit, bit)	Byte
Measurement unit that can only have two values, 0 and 1	Unit that indicates the amount of data, consisting of 8 bits
 0 OFF FALSE  1 ON TRUE	

- Each byte unit is 1,024 times the previous

= Data Unit =



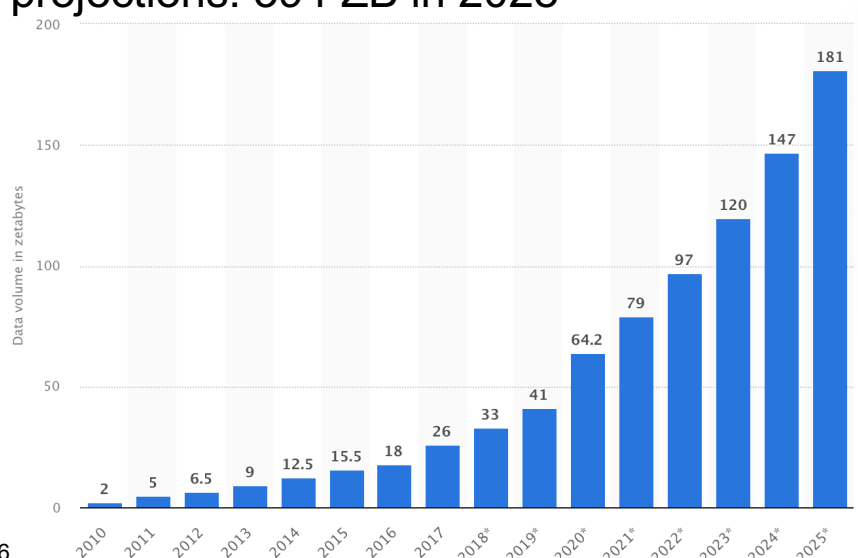
Unit	Definition	Storage space size
Bit	0 or 1	Yes/No
1 Byte	8 bit	Alphabets and one number
1 kilobyte (KB)	1,024 Byte	A few paragraphs
1 megabyte (MB)	1,024 KB	One minute-long MP3 song
1 gigabyte (GB)	1,024 MB	30 minute-long HD movie
1 terabyte (TB)	1,024 GB	About 200 FHD movies

How much data?

- Big data volume: from Terabytes to Zettabytes
 - How big is a Zettabyte?
 - **1 ZB** = 2^{70} B = $(2^{10})^7$ B $\approx$ $(10^3)^7$ B = **10^{21} B**
 - Recall that $2^{10} = 1024 \approx 10^3$
 - Also, **1 ZB** = 2^{70} B = $2^{40} 2^{30}$ B = **2^{40} GB** (being 2^{30} B = 1 GB)
- 181 Zettabytes of data generated by 2025
 - 181 Zettabytes ($181 \times 2^{70} \approx 181 \times 10^{21}$) ...
 - $\approx$ 181,000 Exabytes ($181,000 \times 10^{18}$) ...
 - $\approx$ 181,000,000 Petabytes ($181,000,000 \times 10^{15}$) ...
 - $\approx$ 181,000,000,000 Terabytes ($181,000,000,000 \times 10^{12}$) ...
 - $\approx$ 181,000,000,000,000 Gigabytes ($181,000,000,000,000 \times 10^9$) ...
 - $\approx$ 181,000,000,000,000,000,000,000 bytes!

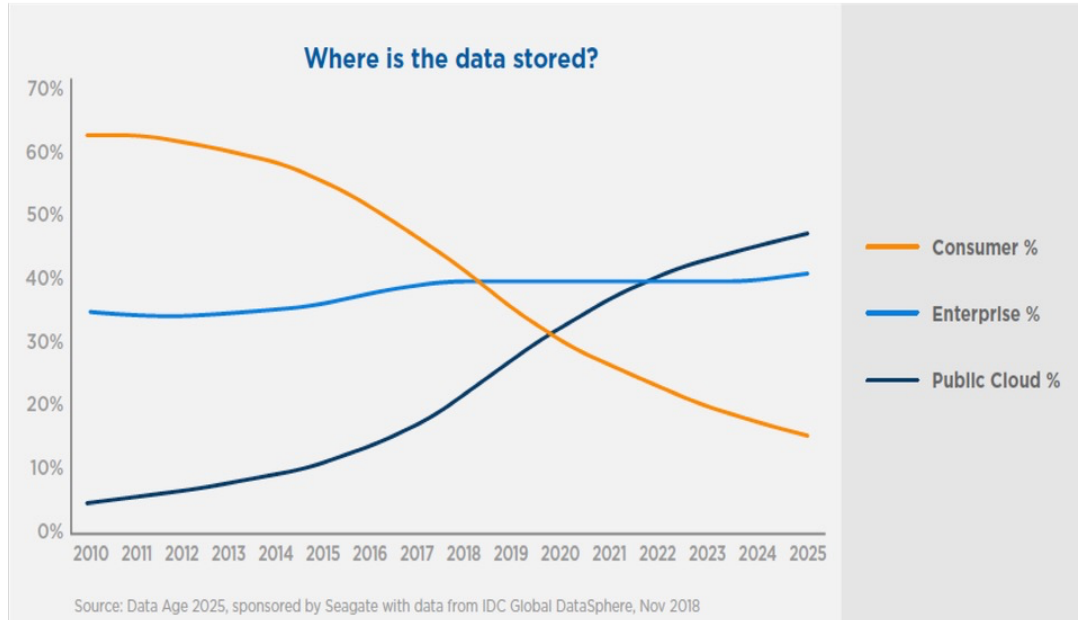
How much data?

- The explosion of data
 - 2013: 90% of all data in the world generated in the previous 2 years
 - From 2010 to 2025: 90x growth in data volume
 - Future projections: 394 ZB in 2028



Where is data stored?

- The majority of data is now stored in the Cloud
- Where are Cloud servers located? In large data centers around the world (not spread evenly)

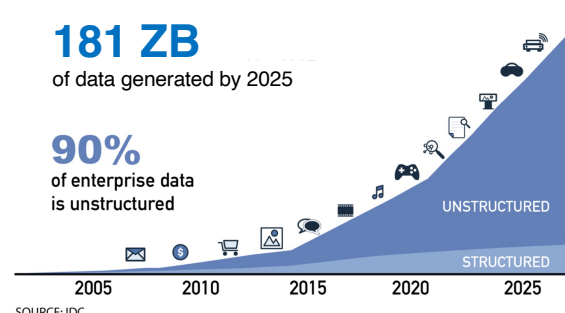


Valeria Cardellini - ADS 2025/26

26

Big data driving factors

- Why Big Data is growing fast
 - **Smartphones**: 7.5 billion smartphone user, each user generates ~25 GB of data per month
 - **Social networks**
 - **Internet of Things (IoT)**
 - **AI**: the new “engine”, from analyzing data to generating data
- Exponential growth in **unstructured** data
 - Unstructured data: information that does not have a fixed, organized structure
 - E.g., social media post, photo



Valeria Cardellini - ADS 2025/26

27

IoT and data explosion

- IoT is everywhere and largely contributes to increase Big Data challenges
 - Proliferation of data sources: by 2025 over 21 billion IoT devices
- Example: self-driving cars
 - A single car can generate 1 TB of data per hour
 - Where to process and store?
 - Collect everything is unfeasible
 - Process and discard
 - 99% of data is processed at the edge (in the car)
 - Only 1% of data is sent to the cloud (high-value events, corner cases)



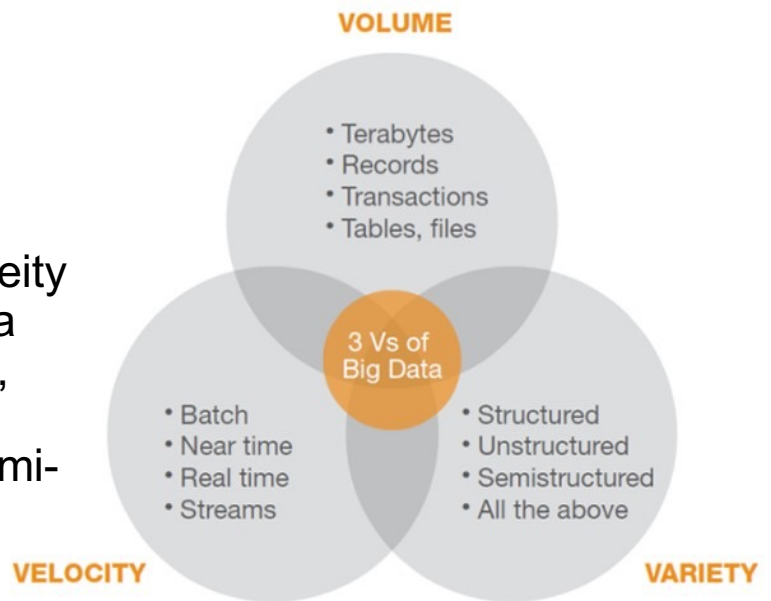
Big Data definitions

Several definitions

- “Big data primarily refers to data sets that are **too large or complex** to be dealt with by traditional data-processing software.” *Wikipedia, 2026*
- “Big data is mostly about taking numbers and using those numbers to **make predictions about the future**. The bigger the data set you have, the more accurate the predictions about the future will be.” *Anthony Goldbloom, Kaggle’s founder*

3V model for Big Data

1. **Volume**: data size
challenging to store and
process (how to index,
retrieve)
 2. **Variety**: data heterogeneity
because of different data
types (text, audio, video,
record) and degree of
structure (structured, semi-
structured, unstructured
data)
 3. **Velocity**: data generation
rate and analysis rate
- Defined in 2001 by D. Laney



The extended (3+n)V model

4. **Value**: Big data can generate huge competitive
advantages
 - “The bigger the data set you have, the more accurate the
predictions about the future will be” (A. Goldbloom)
5. **Veracity**: the uncertainty of accuracy and authenticity
of data
 - E.g., deep fakes and AI hallucinations
6. **Visualization**

Visualization

- Presentation of data in a pictorial and graphical format
- Why? Our brain processes images 60,000x faster than text
- Example



Valeria Cardellini - ADS 2025/26

32

Environmental costs of digital economy

- Rapidly rising energy consumption
 - 2022: data centers used ~1-1.3% of global electricity, cryptocurrency mining 0.4%
 - 2026 estimate: data centers will consume ~2% of global electricity
 - Similar to Japan
- The hidden cost of AI
 - A single ChatGPT query (~0.3 Wh) uses 10x more power than a Google search
- Is this sustainable?
 - The carbon gap: AI demand is growing faster than green grids can be built
 - Water impact for server cooling

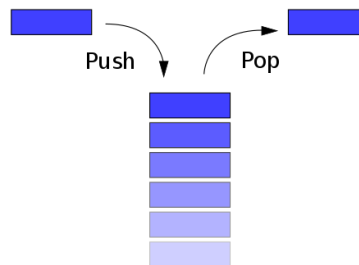
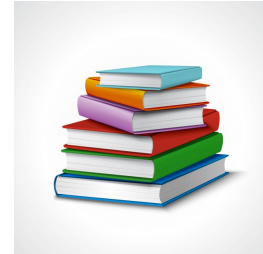
<https://www.nytimes.com/2023/10/10/climate/ai-could-soon-need-as-much-electricity-as-an-entire-country.html> , NYT, Oct. 2023

Valeria Cardellini - ADS 2025/26

33

Data structure

- Organizes and stores data
- Enables efficient access and modification
- Supports specific operations
- Example: **stack**
 - 3 basic operations:
 - Push: insert item on top
 - Pop: remove top item
 - Peek: view top item (without removing)



What about security?

- IT and data are everywhere



IT and data everywhere

- New cybersecurity threats



Self-driving cars

IoT wearables



Valeria Cardellini - ADS 2025/26

36

Attacking Machine Learning algorithms

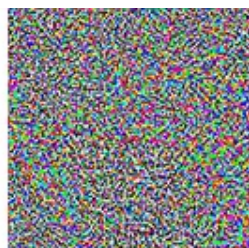
- Adversarial examples are special inputs to fool a ML model (e.g., neural network), intentionally designed by an attacker to cause a model to make a mistake
 - Like an optical illusion
 - Example: by changing only 0.04% of pixel values in an image (undetected by human eyes), the neural network changes its classification prediction from “panda” to “gibbon”
 - How to protect against these attacks?



“panda”

57.7% confidence

+ ϵ



=



“gibbon”

99.3% confidence

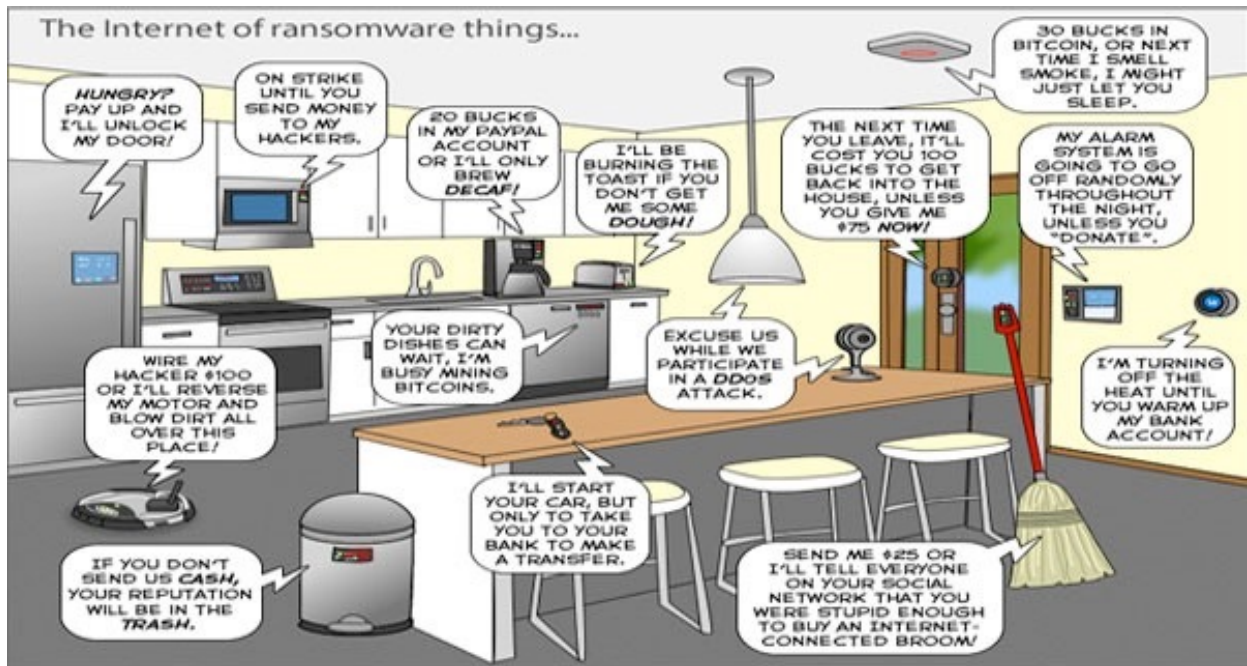


This is a gibbon!

Valeria Cardellini - ADS 2025/26

37

Internet of Things and ransomware



Ransomware: type of malware that blocks access to a victim's digital assets (e.g., files, systems, data) and demands payment ("ransom") to restore access