

Closing Thoughts

Vishal Patel

Spring 2025

What We Covered

1. Introduction to Data Mining
2. The Data Science Process
3. Introduction to Regression
4. Linear Regression model
5. Build a model using partitioning
6. Decision trees
7. Classification Trees
8. Random Forests
9. Gradient Boosting Trees
10. Hyper-parameter optimization

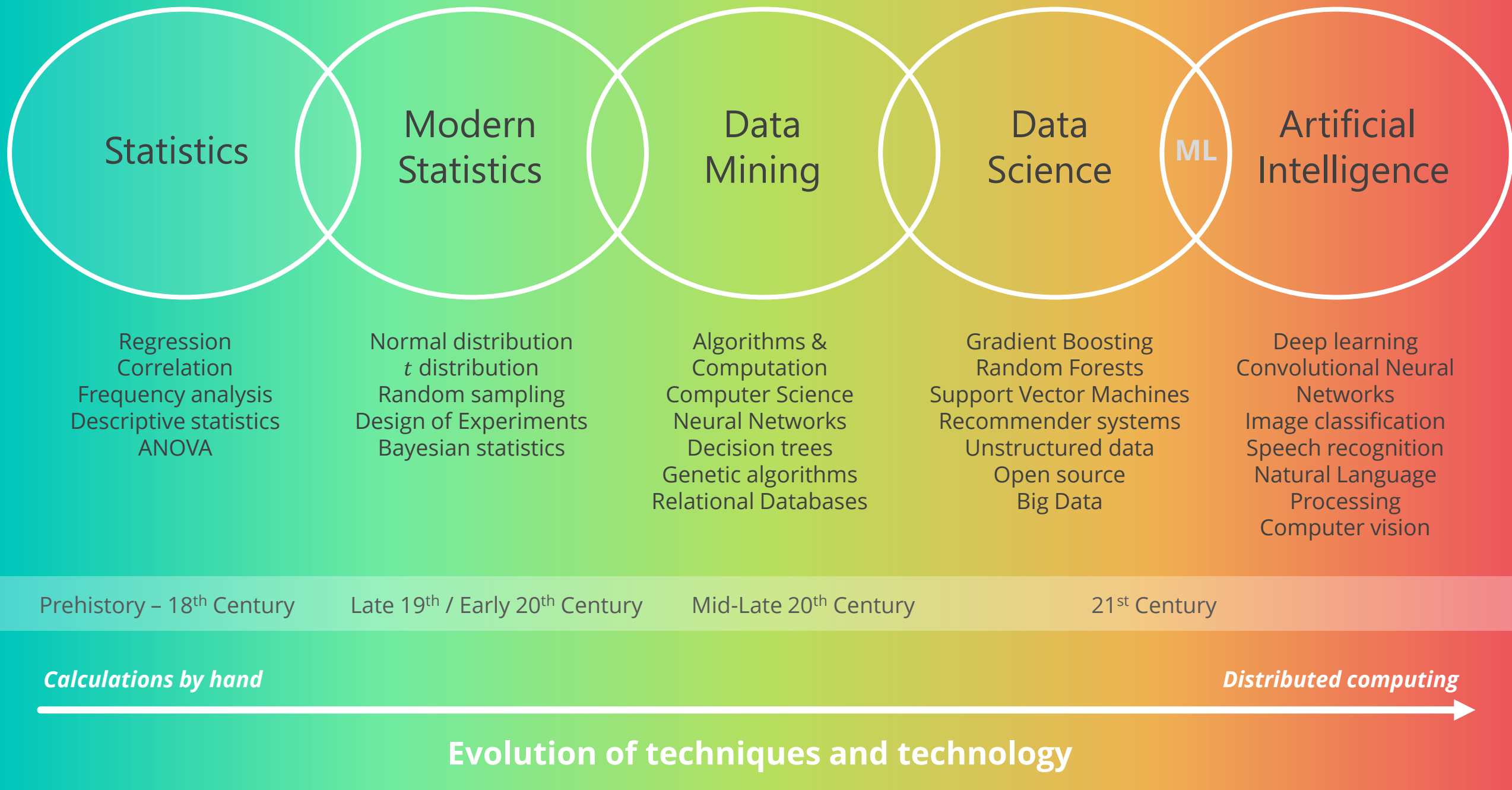
11. Introduction to Neural Networks
12. Introduction to Clustering
13. Agglomerative Clustering
14. k-means Clustering
15. DBSCAN Clustering
16. PCA
17. Association Analysis (Apriori algorithm)
18. Collaborative Filtering
19. Data Wrangling

+ 20 Jupyter Notebooks

Education: What, when and why

Training: How





"Artificial intelligence is one of the most profound things we're working on as humanity. It is more profound than fire or electricity."

– Sundar Pichai



“With artificial intelligence
we are **summoning the demon.**”

– Elon Musk





Boston Dynamics

BASED ON YOUR
INTERNET HISTORY,
YOU MIGHT BE DUMB
ENOUGH TO ENJOY
EXTREME SPORTS.



CLICK HERE TO BUY A
TICKET TO BASE JUMP
FROM THE INTERNA-
TIONAL SPACE STATION.



I THINK
THE INTER-
NET IS
TRYING TO
KILL ME.



WE
CALL IT
"MACHINE
LEARNING."





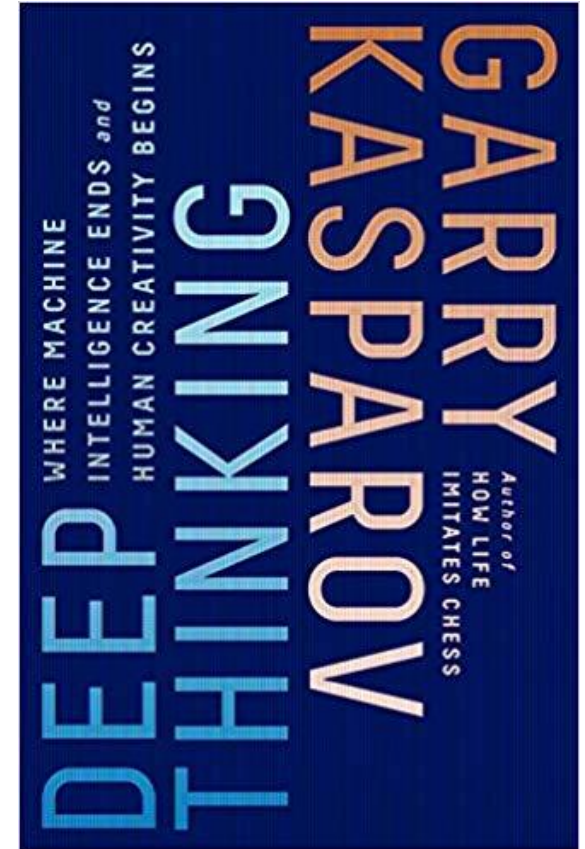
1

Jobs

2

Ethics

Good riddance, you might imagine. But the worries about operator-less elevators were quite similar to the concerns we hear today about driverless cars. In fact, I learned something surprising when I was invited to speak to the Otis Elevator Company in Connecticut in 2006. The technology for automatic elevators had existed since 1900, but people were too uncomfortable to ride in one without an operator. It took the 1945 strike and a huge industry PR push to change people's minds, a process that is already repeating with driverless cars. The cycle of automation, fear, and eventual acceptance goes on.



Everyone: AI art will make designers obsolete

AI accepting the job:



Jan 2023



Jun 2023



waze



@waze

Follow



We're working with the [@LADOTofficial](#) and our [#WazeCommunity](#) to keep the Waze Map updated with road closures (currently at 110) & open shelters (16) to help Southern Californians navigate safely around the [#wildfires](#). Please help us spread the word.

Joel Rubin  [@joelrubin](#)

"The Los Angeles Police Department asked drivers to avoid navigation apps, which are steering users onto more open routes — in this case, streets in the neighborhoods that are on fire." [fw.to/EMJxt2E](#)

7:59 AM - 7 Dec 2017

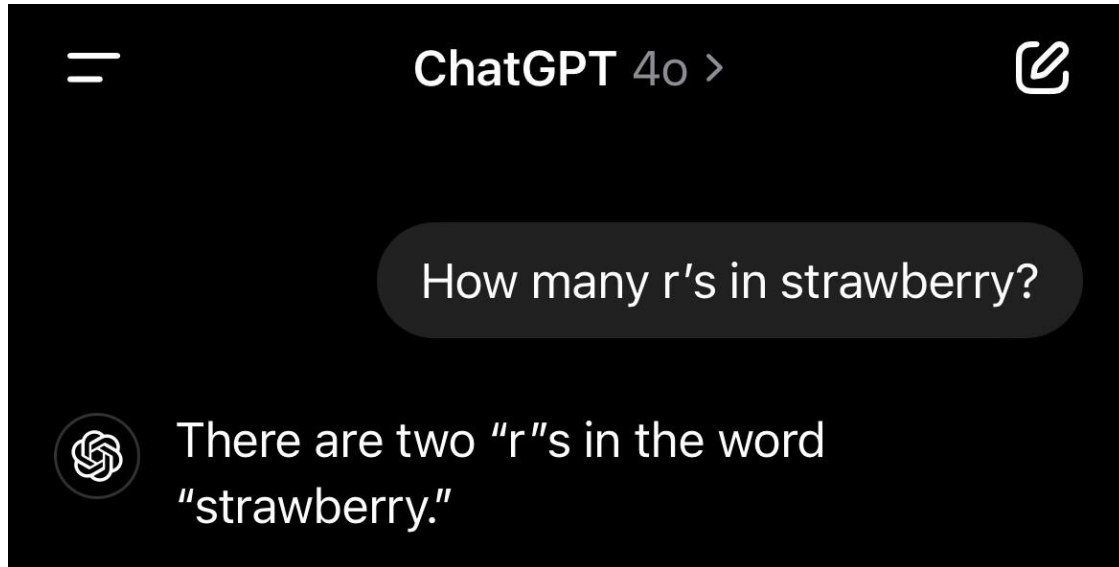


Tay was an artificial intelligence chatbot that was originally released by Microsoft Corporation via Twitter on March 23, 2016.

It caused subsequent controversy when the bot began to post inflammatory and offensive tweets through its Twitter account, forcing Microsoft to shut down the service only 16 hours after its launch.

It was soon replaced with **Zo** [2016-2019].

And now, as of 2025, **ChatGPT** has become a part of our daily life.



ChatGPT Omni prompted to "create the exact replica of this image, don't change a thing" 74 times







**Being able to
communicate
clearly in writing**



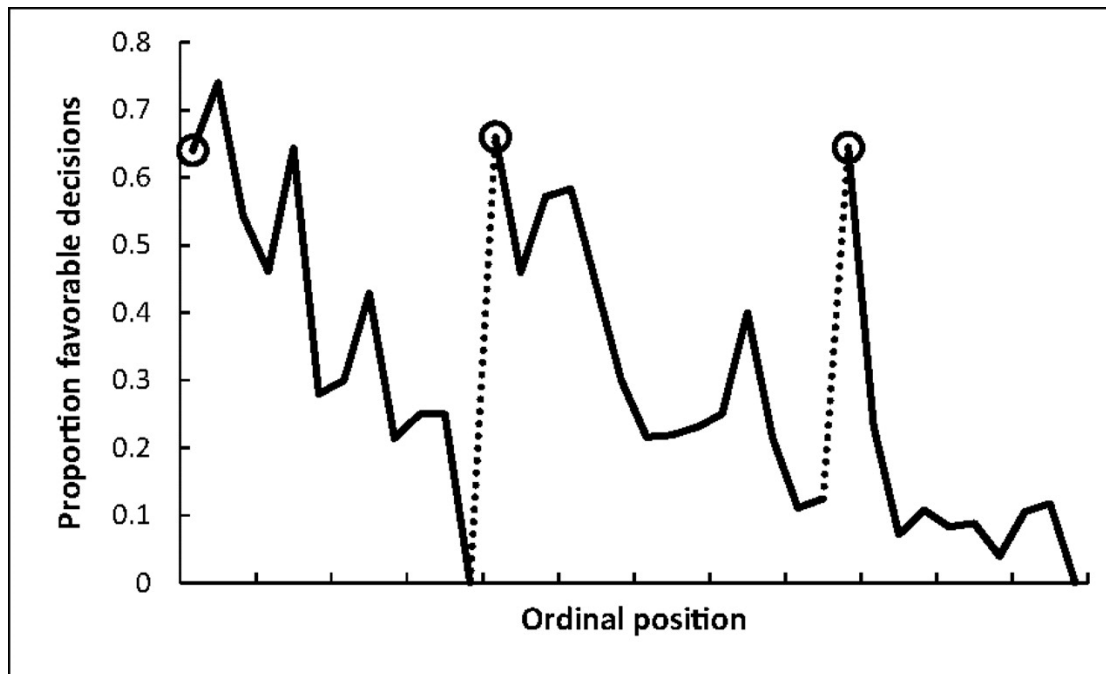
**Prompt
Engineering**

Extraneous factors in judicial decisions

Shai Danziger, Jonathan Levav , and Liora Avnaim-Pesso [Authors Info & Affiliations](#)

Edited* by Daniel Kahneman, Princeton University, Princeton, NJ, and approved February 25, 2011 (received for review December 8, 2010)

April 11, 2011 | 108 (17) 6889-6892 | <https://doi.org/10.1073/pnas.1018033108>

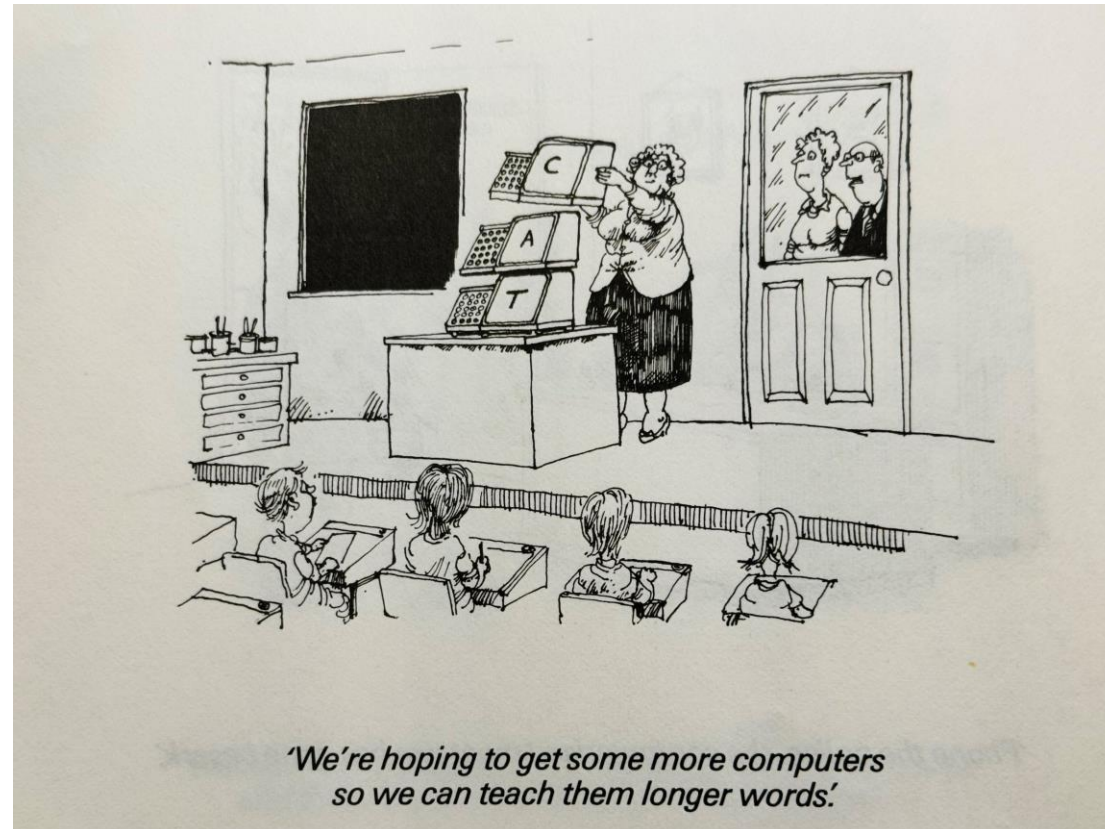


The authors of the peer-reviewed paper looked at more than 1,000 rulings made in 2009 by eight judges. They found that the **likelihood of a favorable ruling peaked at the beginning of the day**, steadily declining over time from a probability of about 65% to nearly zero, before **spiking back up to about 65% after a break for a meal or snack**.

Social and Economical Bias

1. Errors (and biases) propagate quickly
2. Far reaching consequences
3. 'Bias' is not just a technical issues, it's a **socio-technical** issue
4. There's no silver bullet-solution.
 1. What do we mean by "bias"?
 2. Who gets to decide what's fair?

- Generative AI is a type of artificial intelligence technology that uses deep learning algorithms to **generate** content: **text, images, speech, and video**.
- Okay, but what are we supposed to do with this thing...?



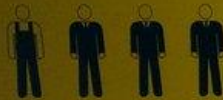
Ein Bauer ernährt zusätzlich ...

Durch die zunehmende Mechanisierung ist folgende Entwicklung möglich geworden:



Bauer Städter

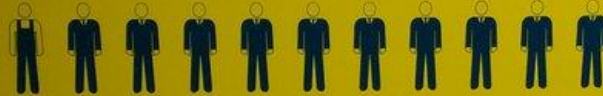
1900



+3

Ein Bauer ernährt sich und zusätzlich drei Städter.

1950



+10

Ein Bauer ernährt sich und zusätzlich zehn Städter.

2005



+142

Ein Bauer ernährt sich und zusätzlich 142 Städter.

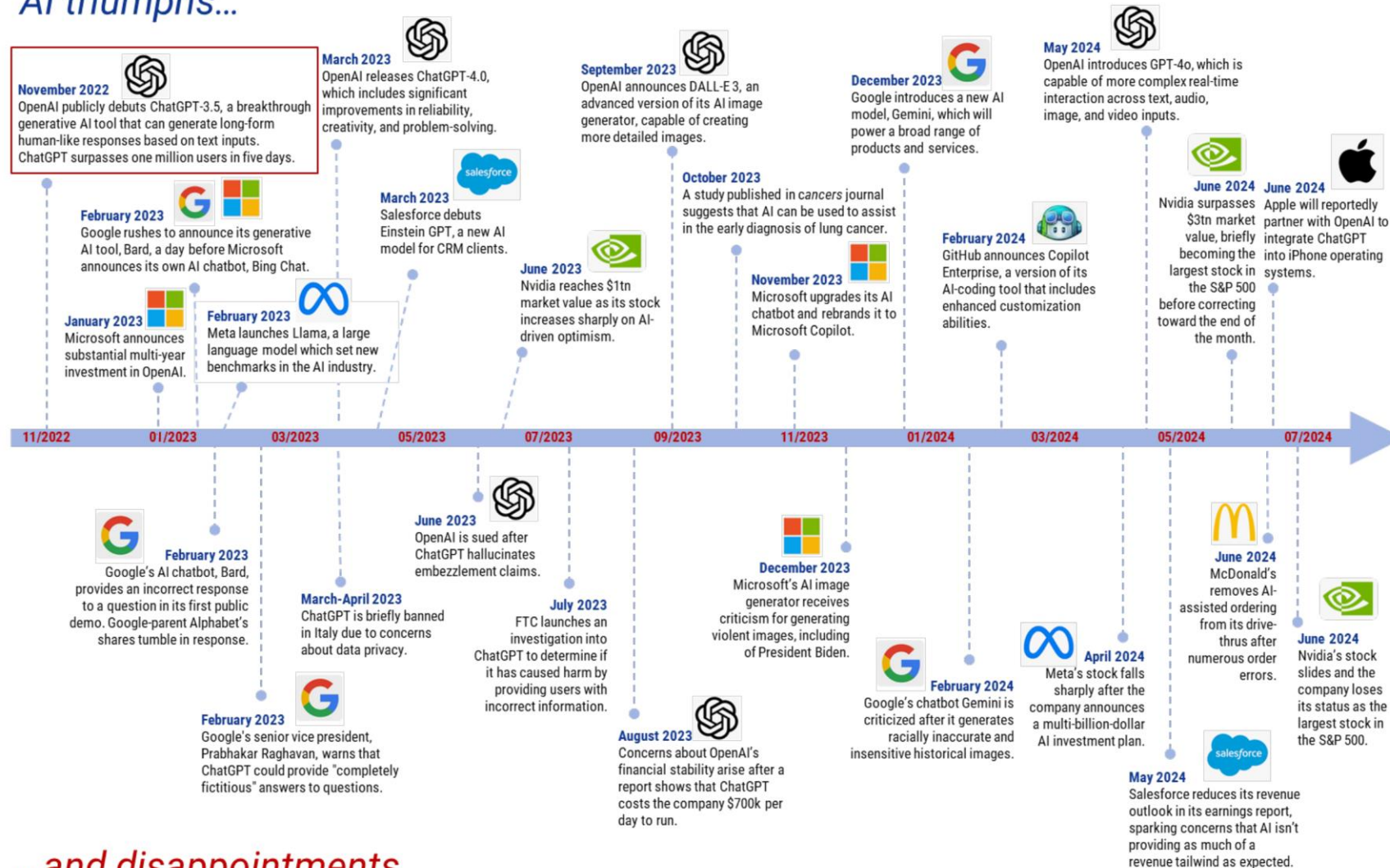
The number of people one farmer could feed

Museum of Bread Culture, Germany



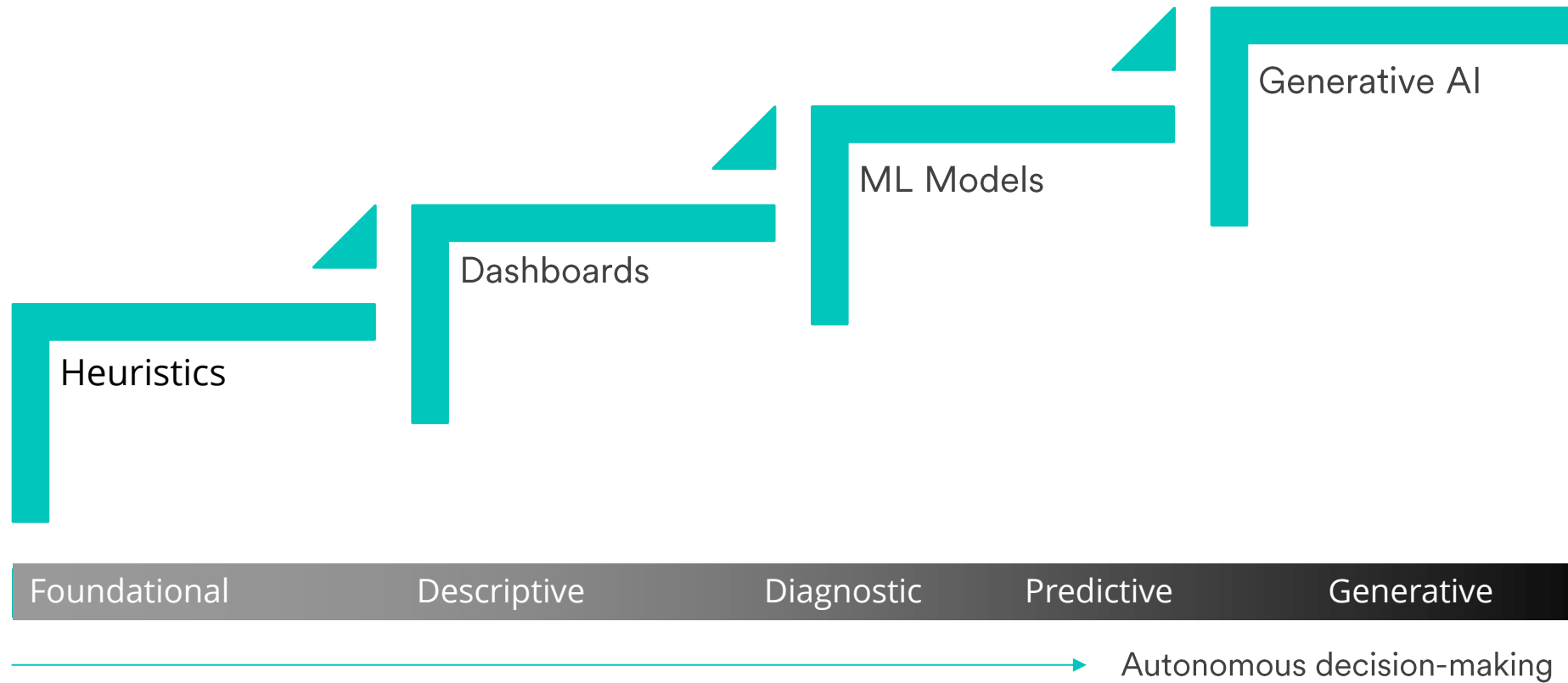
A Short History of AI Developments

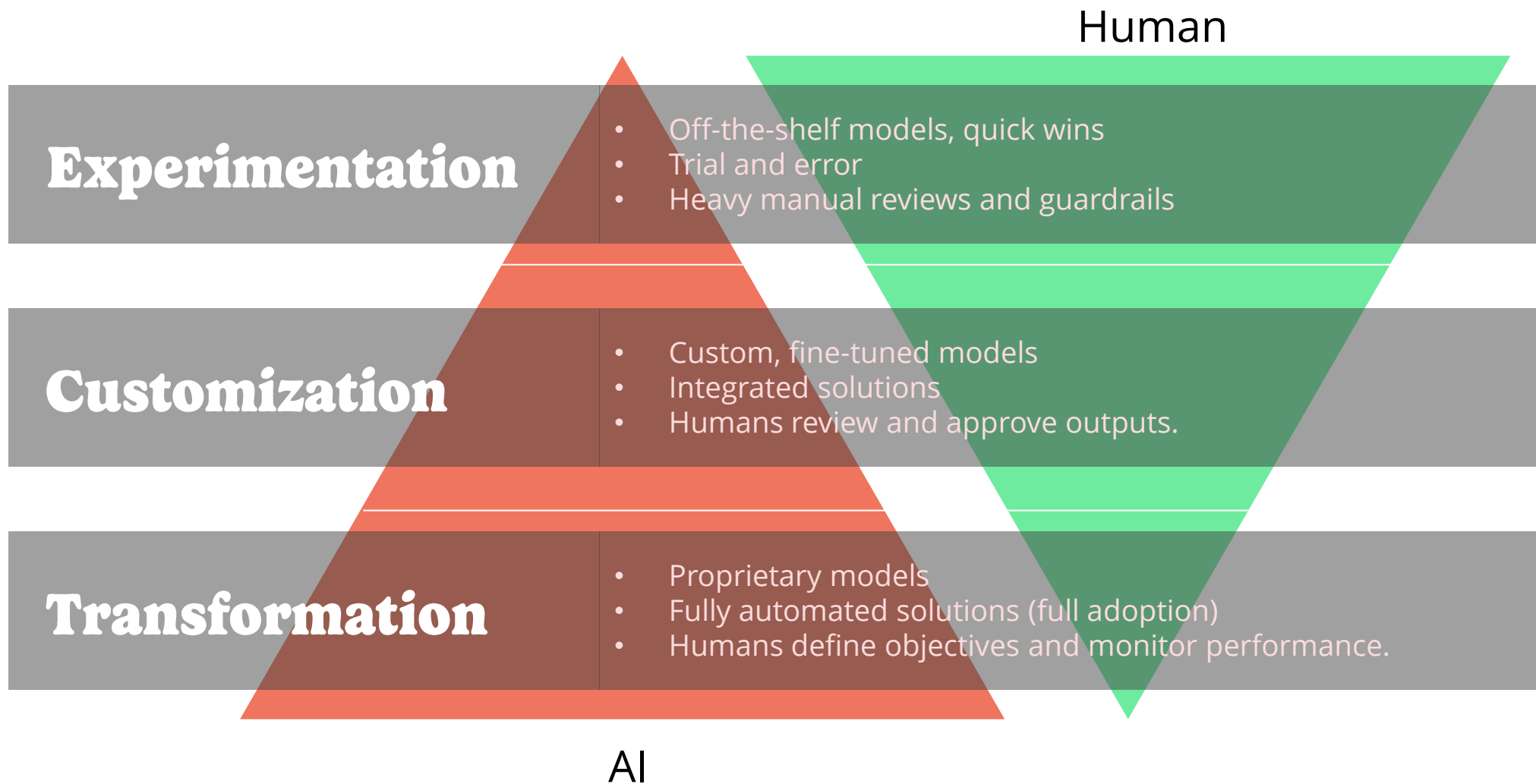
AI triumphs...



...and disappointments

Ladder of Sophistication





Qualities of an Ideal [Data Scientist]

1. Curious, but skeptical
2. Critical thinking skills
3. Logical reasoning ability
4. Creativity
5. Strategic thinking
6. Ability to read the facts and understand assumptions
7. Understands necessary information and how to act on it
8. Engages with people to understand root of the problem
9. Ability to apply appropriate [data science] tools



Mulla Nasrudin

Someone saw Nasrudin searching for something on the ground.

"What have you lost?, he asked.

"My key", said the Mulla.

So they both went down on their knees and looked for it.

After some time, the man asked: "Where exactly did you drop it?"

"In my own house".

"Then why are you looking here?"

"There is more light here than inside my own house".

– Sufi parable [Idris Shah, citing Mulla Nasrudin]

WORKING WITH AVAILABLE LIGHT

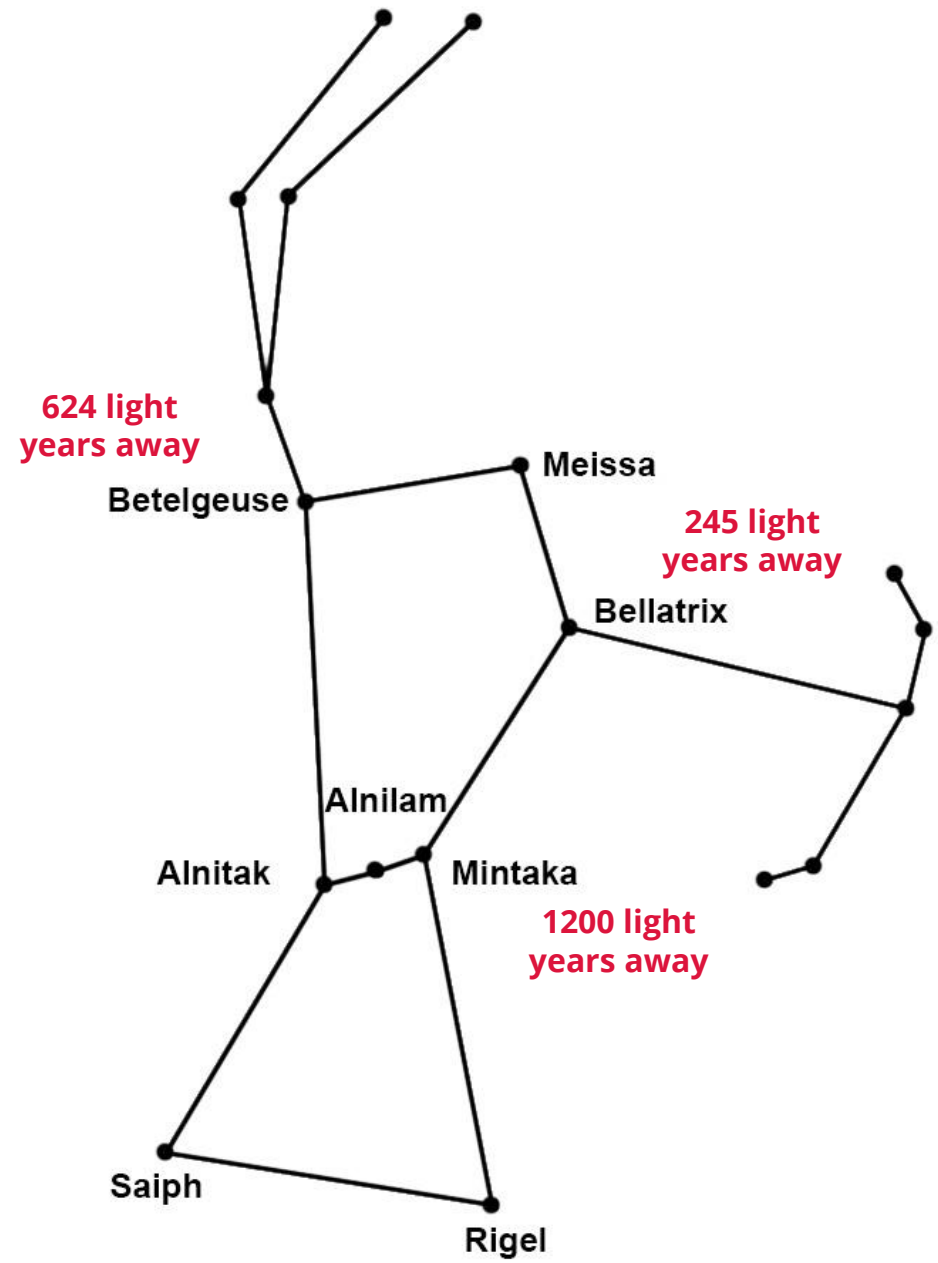
Numbers don't lie!!

If you torture the data long enough,
it will confess to anything.

From Darrell Huff's book *How to Lie With Statistics* (1954)

Truth is much too complicated
to allow anything but approximations.

John von Neumann



There is no substitute for
getting to **know your data.**

– Witten and Frank

The Role of Data Science in AI

1

EDA

Domain Expertise
Data Understanding
GIGO

2

EXPERIMENTATION

Framing Hypotheses
Testing
Iterating

3

EVALUATION

What to Measure
How to Measure
How to Interpret
Statistical Rigor

4

EXPLAINABILITY

Communicating the 'why'
Block box → Glass box



“Before starting any project, I ask: What’s the geology, what’s the geomorphology, what’s the history, where does the wind come from, where does the sun come from, what are the shadow patterns, what’s the drainage system, what’s the flora?

...I try to be a **contextual** architect.” – **Glenn** Murcutt

Age of
CONTENT



Age of
CONTEXT

Contextual Data Science

THANK YOU!!

vishal@derive.io