Introduction

Elements of Data Science and Artificial Intelligence

Prof. Dr. Jens Dittrich
bigdata.uni-saarland.de

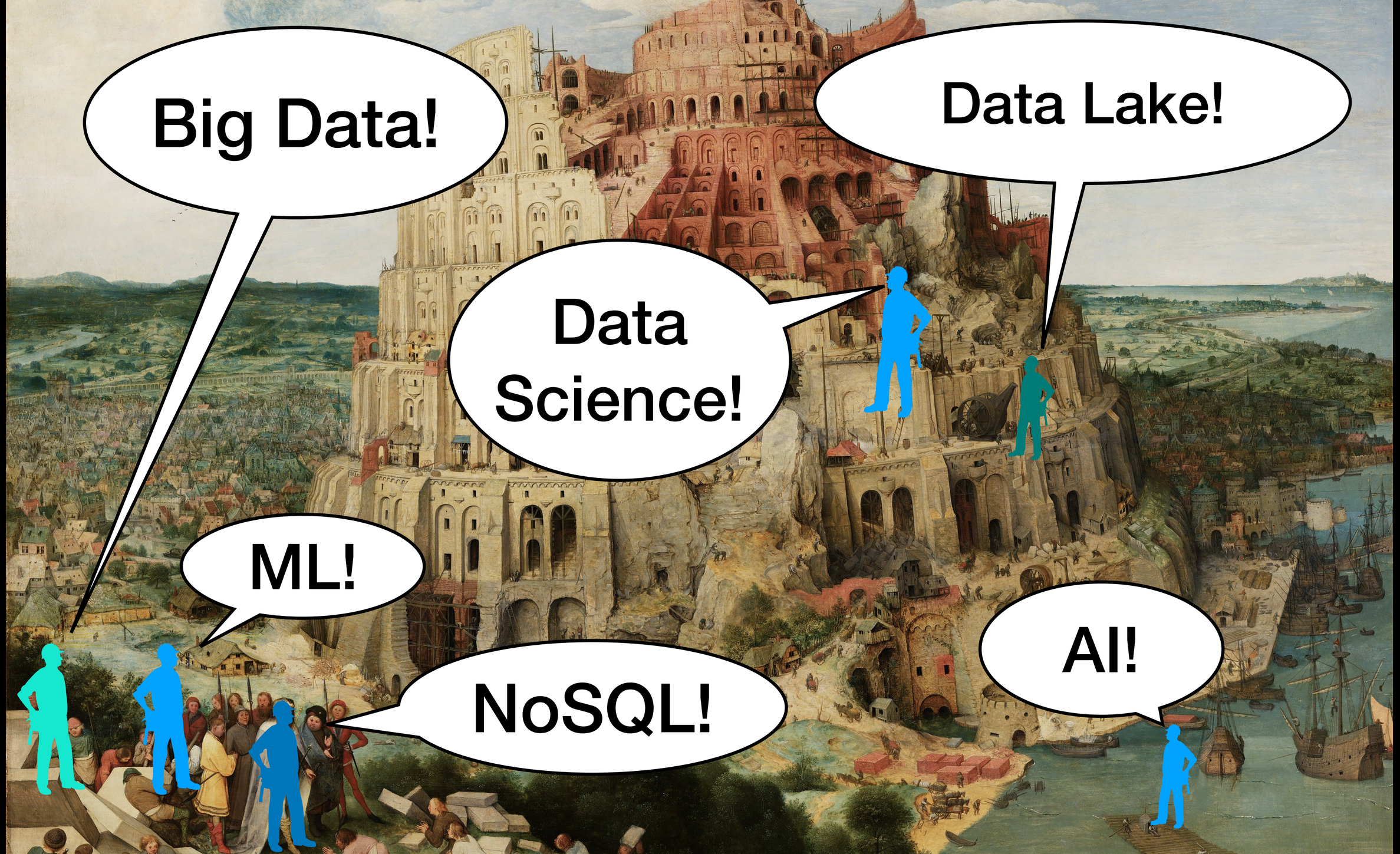
Why is this lecture&program named “Data Science and Artificial Intelligence“?*

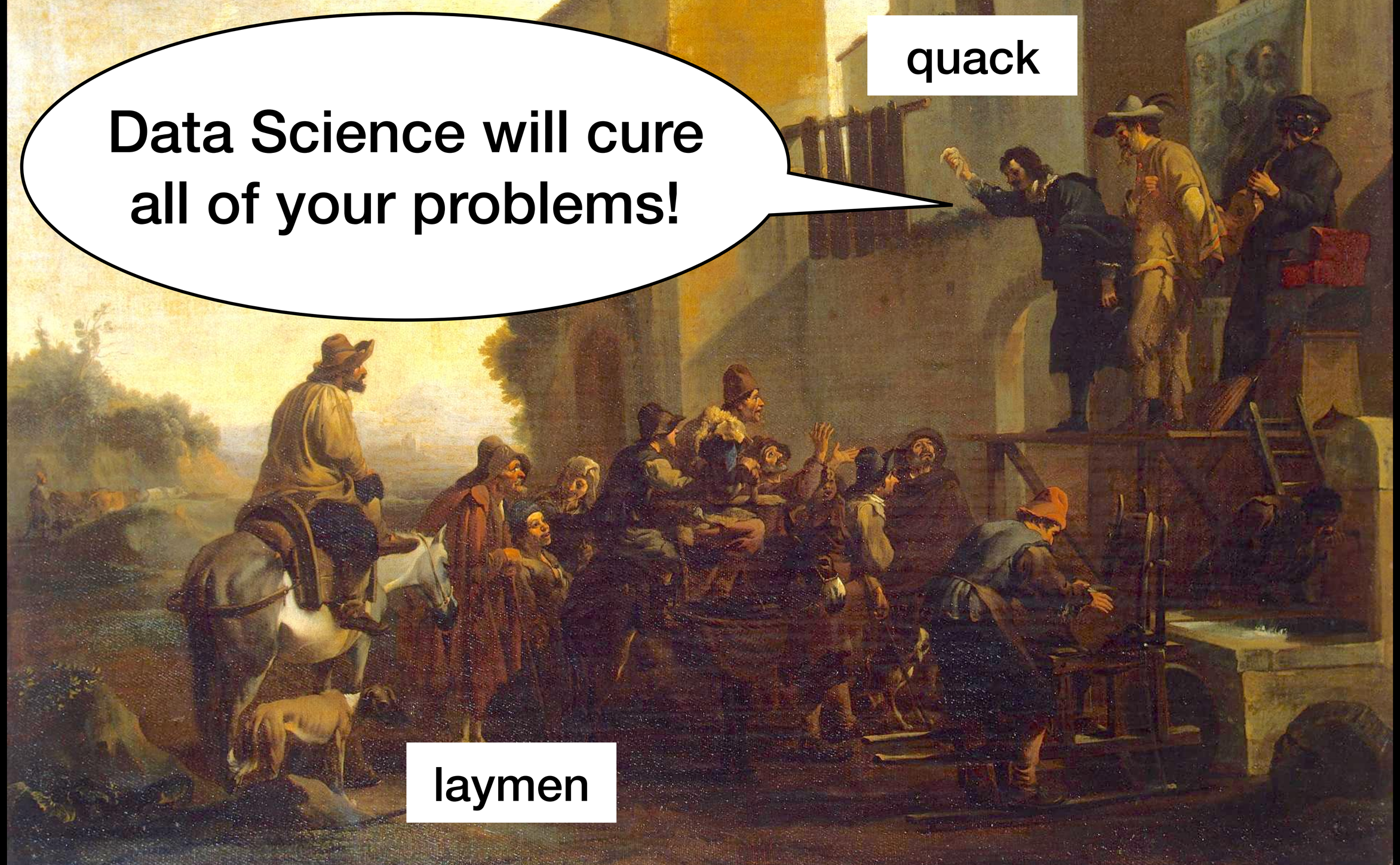
*or: how to explain your parents, friends, etc. what you are studying at SIC

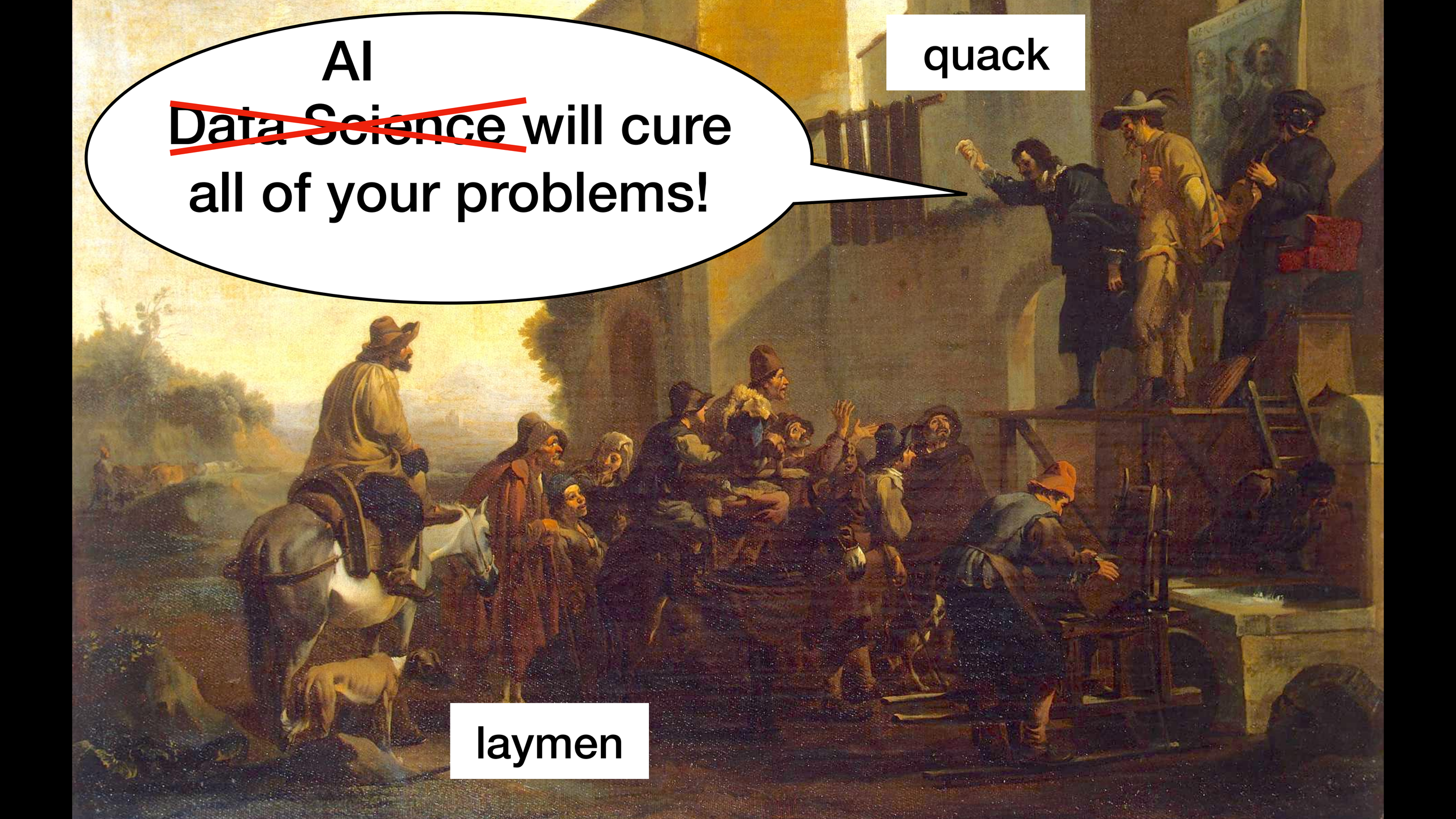
Buzzwords

KI keri KI!

© Muhammad Mahdi Karim


Big Data!
Data Lake!
Data Science!
ML!
AI!
NoSQL!


Data Science will cure all of your problems!
quack
laymen


AI
~~Data Science~~ will cure
all of your problems!
quack
laymen

The TV shop:

© istock.com JackF

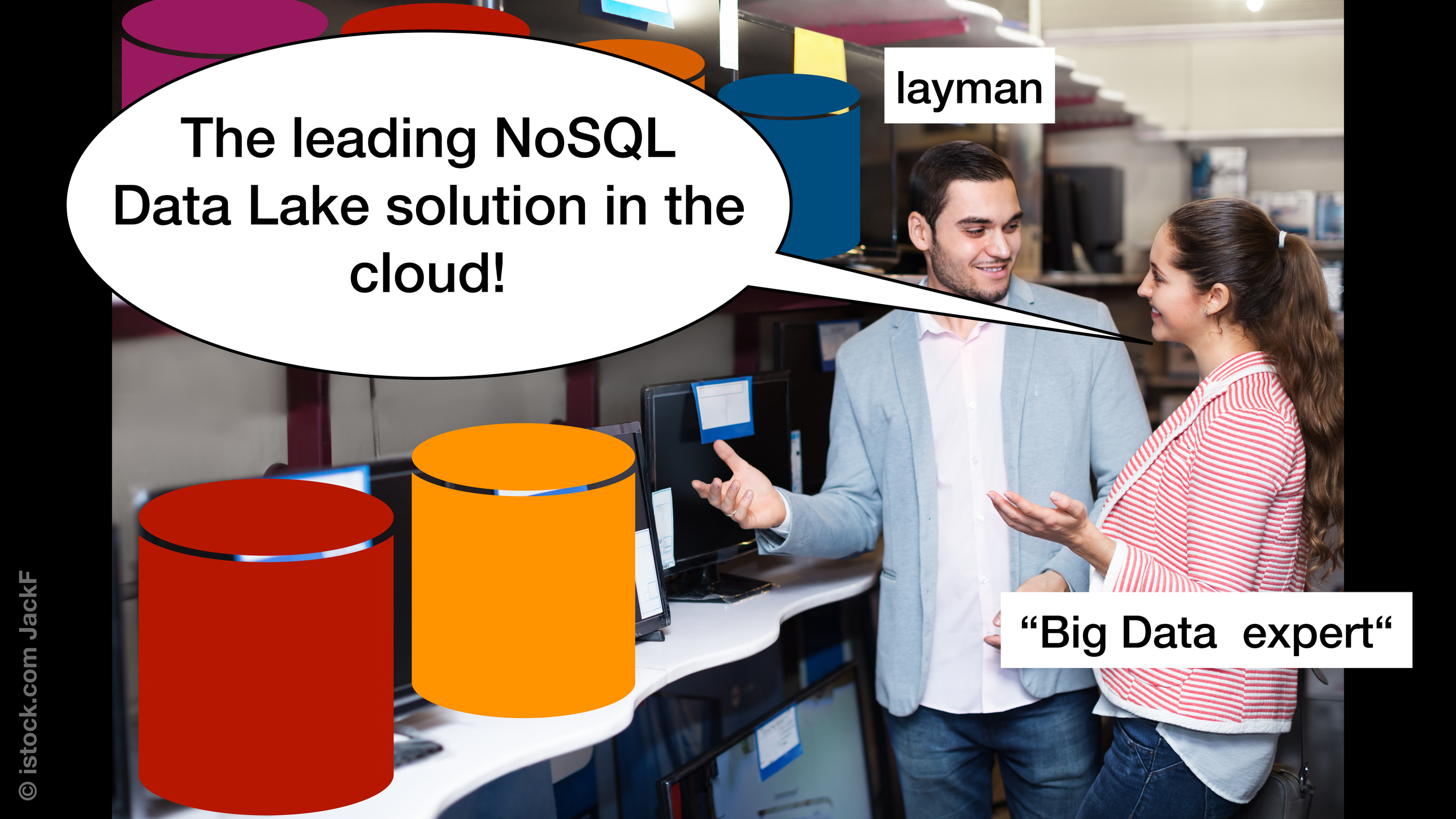
© istock.com JackF

© istock.com JackF

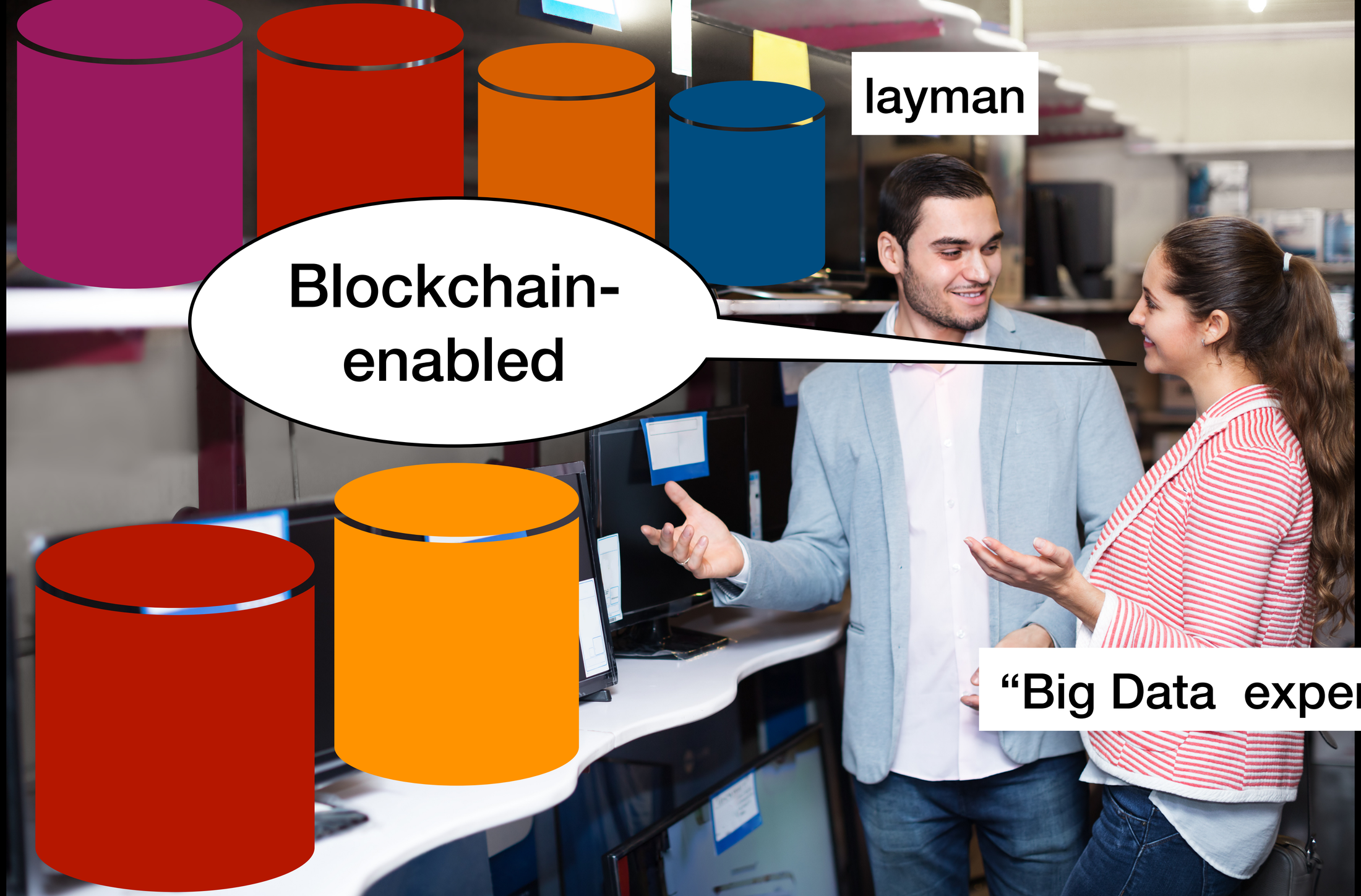
The “Big Data“ shop:

© istock.com JackF

© istock.com JackF

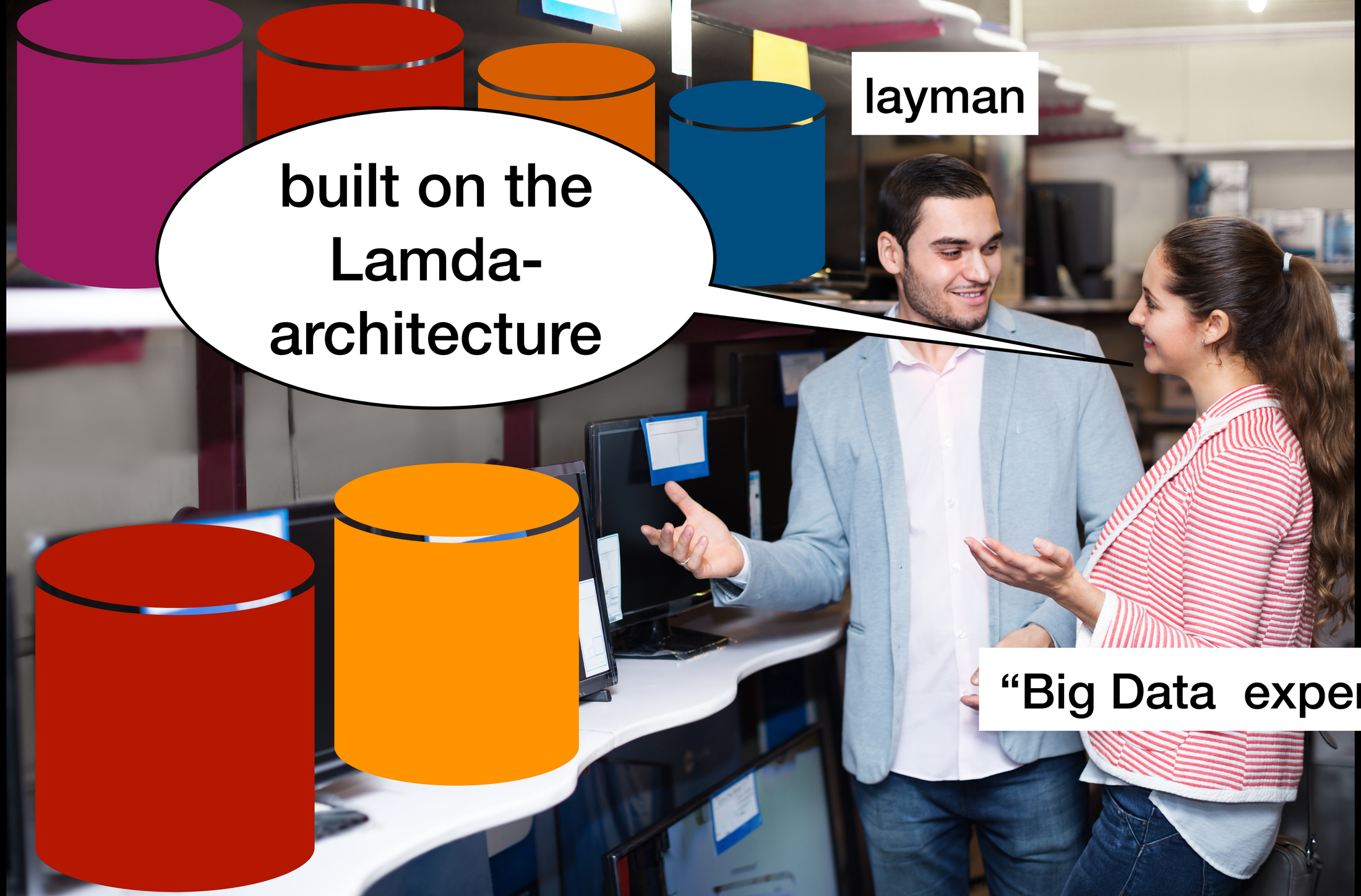


© istock.com JackF

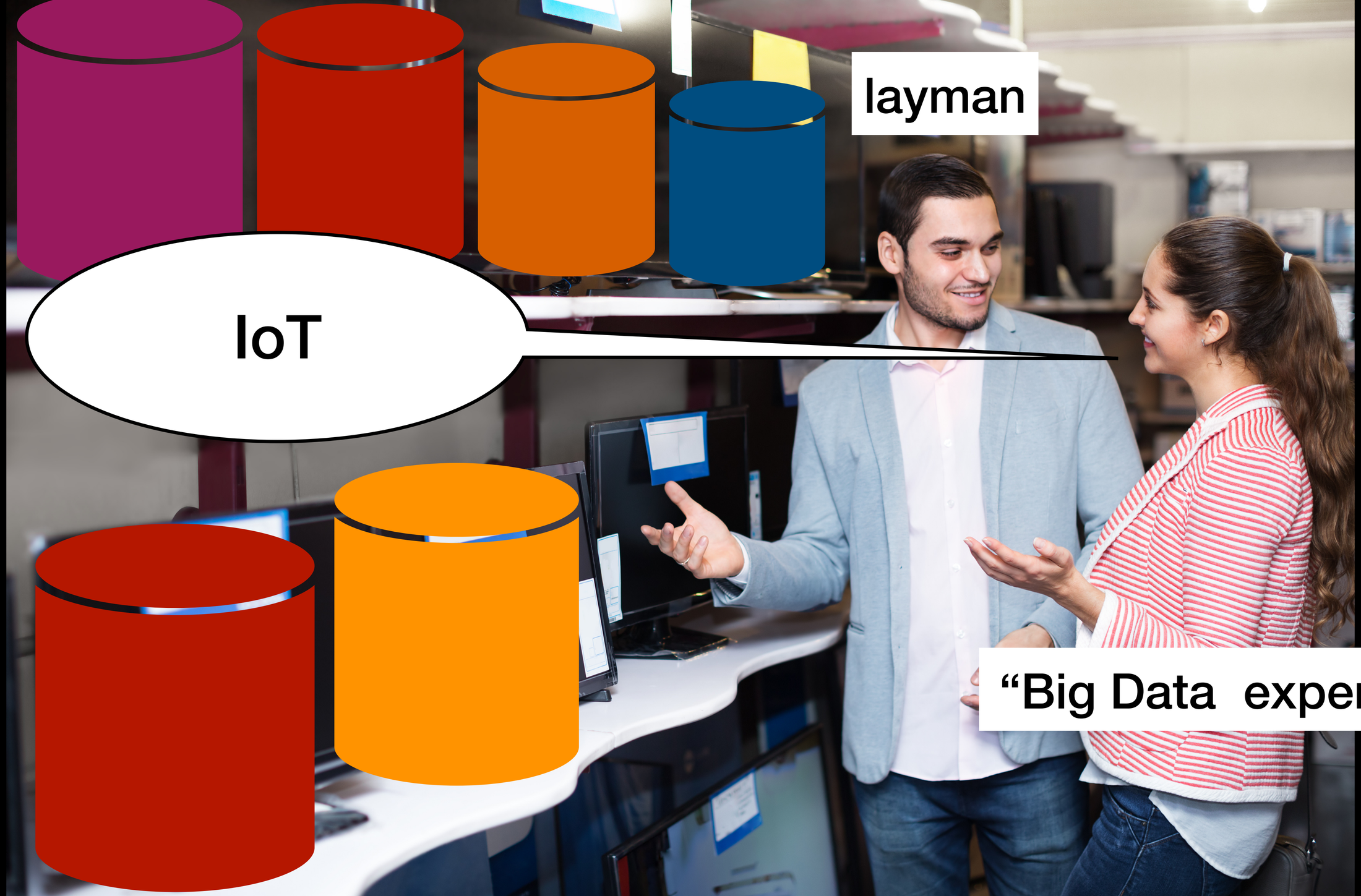


© istock.com JackF

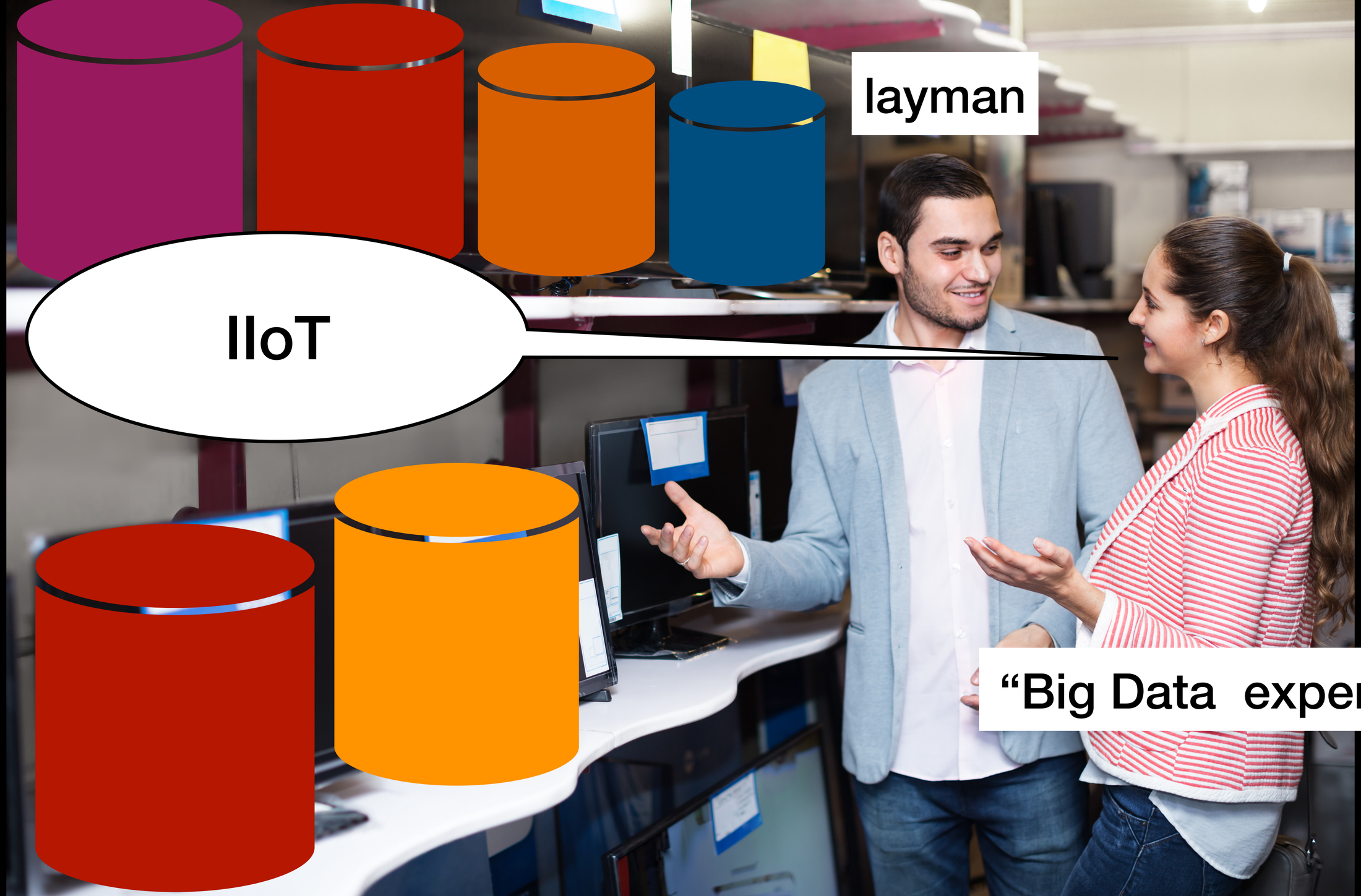
© istock.com JackF



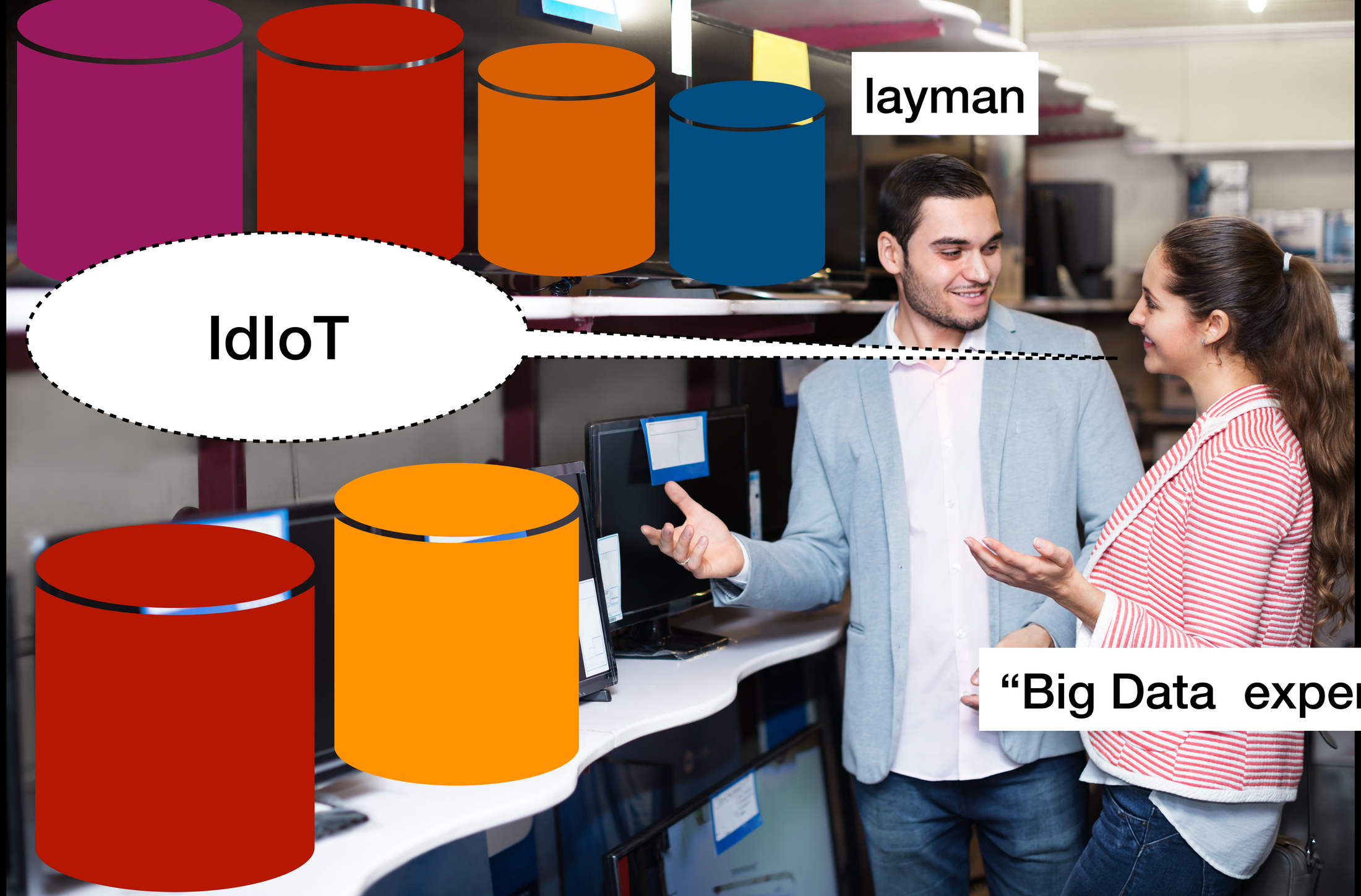
© istock.com JackF



© istock.com JackF



© istock.com JackF

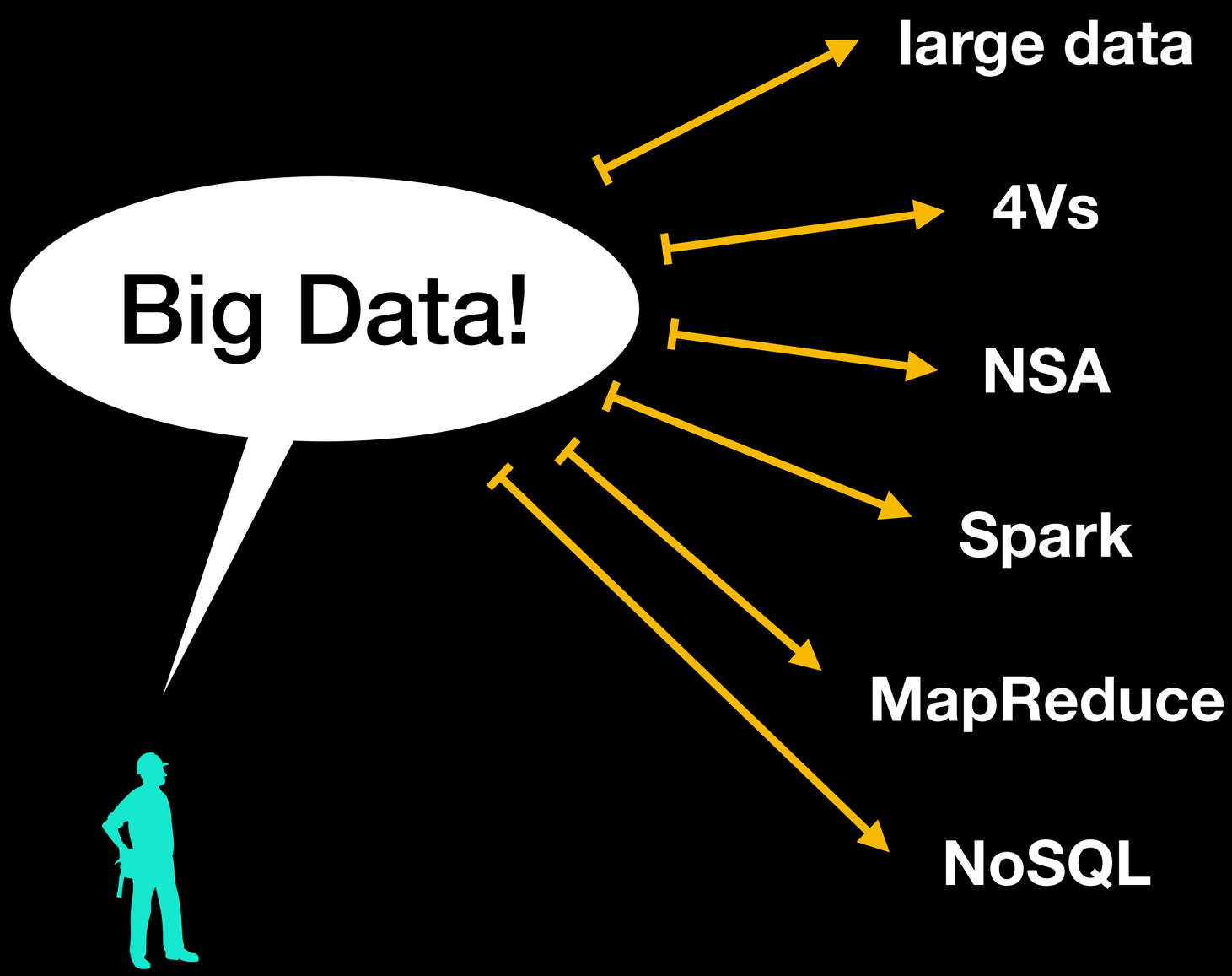


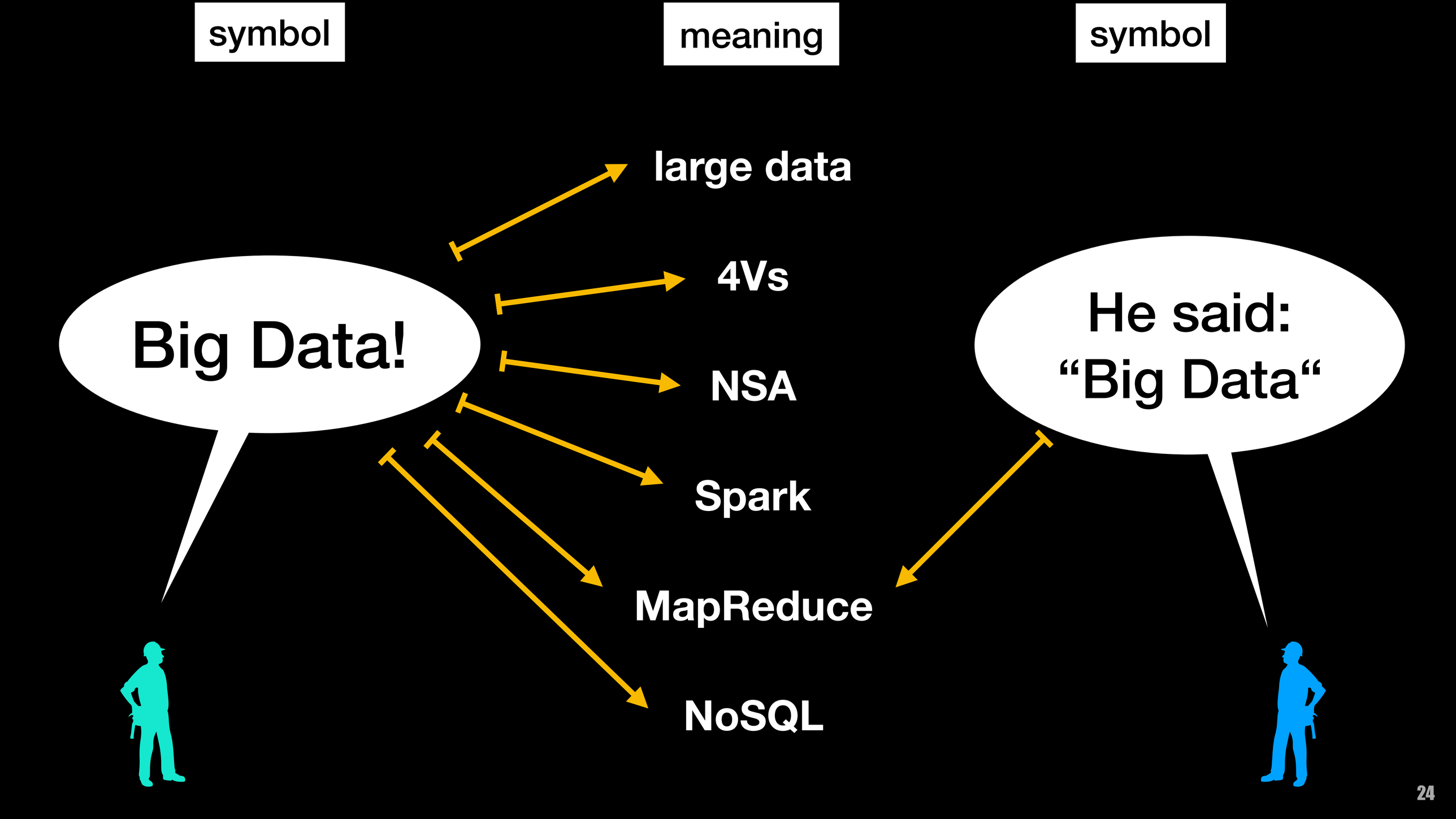
© istock.com JackF

Problem:
ambiguous communication

symbol

meaning




symbol
meaning
symbol
large data
4Vs
NSA
Spark
MapReduce
NoSQL
Big Data!
He said:
“Big Data“

symbol

meaning

symbol

large data

4Vs

NSA

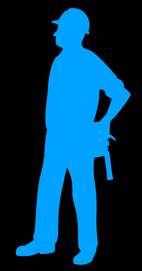
Spark

MapReduce

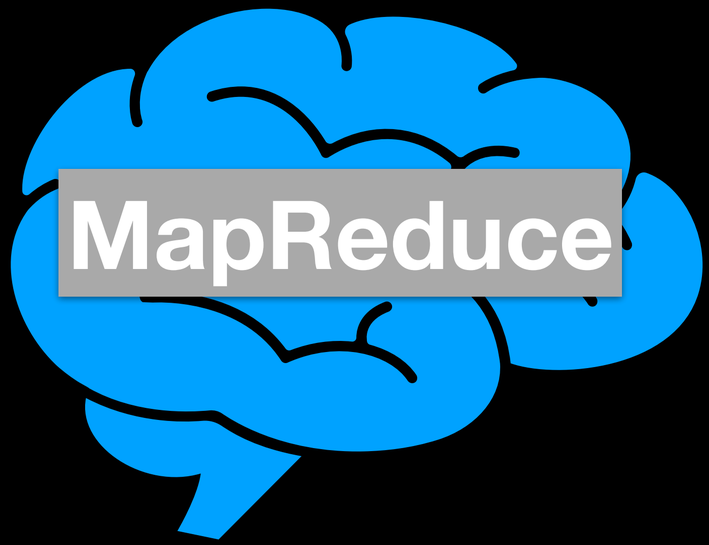
NoSQL

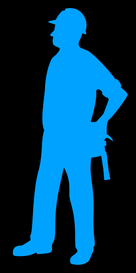
Big Data!

He said:
“Big Data“



translated to:



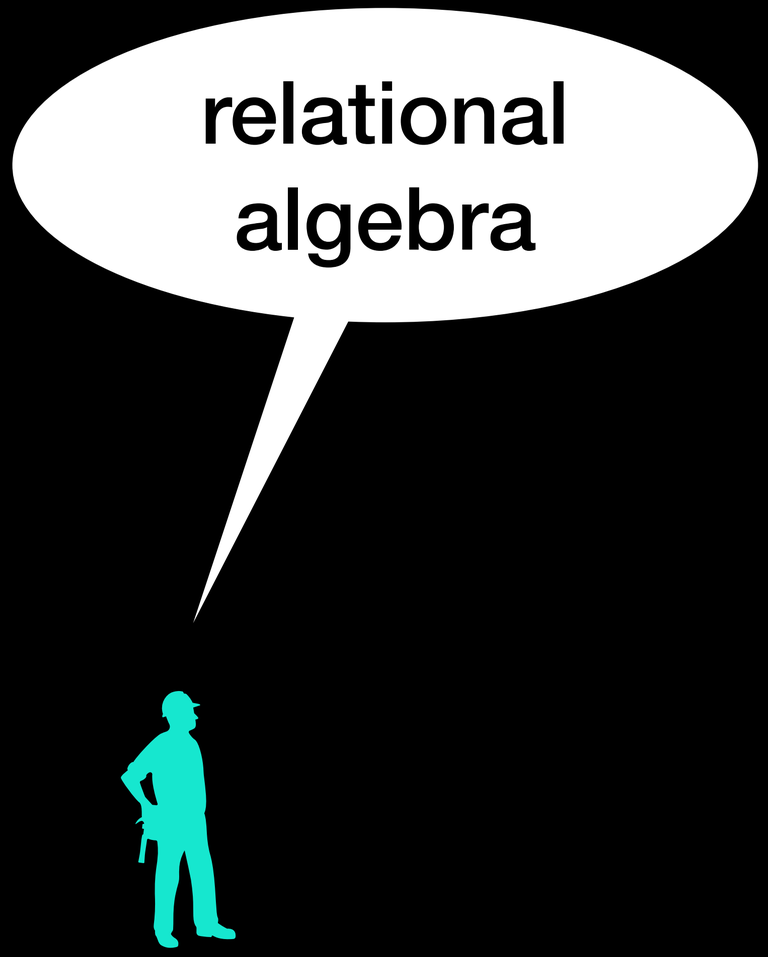


BILL MURRAY SCARLETT JOHANSSON
Lost In Translation
FOCUS FEATURES PRESENTS AN AMERICAN ZOETROPE / ELEMENTAL FILMS PRODUCTION "LOST IN TRANSLATION" BILL MURRAY SCARLETT JOHANSSON GIOVANNI RIBISI ANNA FARIS FUMIHIRO HAYASHI MUSIC PRODUCED BRIAN REITZELL COSTUME DESIGNER NANCY STEINER
R RESTRICTED UNDER 17 REQUIRES ACCOMPANYING PARENT OR ADULT GUARDIAN
Some Sexual Content
PRODUCTION DESIGNERS ANNE ROSS K.K. BARRETT EDITOR SARAH FLACK DIRECTOR OF PHOTOGRAPHY LANCE ACORD LINE PRODUCER CALLUM GREENE ASSOCIATE PRODUCER MITCH GLAZER
EXECUTIVE PRODUCERS FRANCIS FORD COPPOLA FRED ROOS PRODUCED BY ROSS KATZ SOFIA COPPOLA WRITTEN AND DIRECTED BY SOFIA COPPOLA
SOUNDTRACK AVAILABLE ON emperorNorton
FOCUS FEATURES
ARTWORK ©2003 FOCUS FEATURES, LLC. ALL RIGHTS RESERVED.
www.lost-in-translation.com
The new film written and directed by Sofia Coppola

clear communication:

symbol

meaning



relational algebra, i.e. π, σ, ⋈, …

symbol

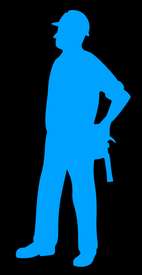
meaning

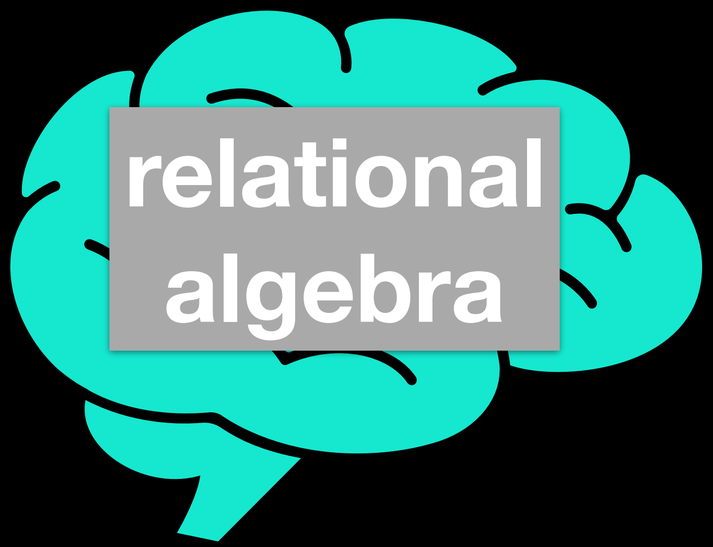
symbol

relational algebra

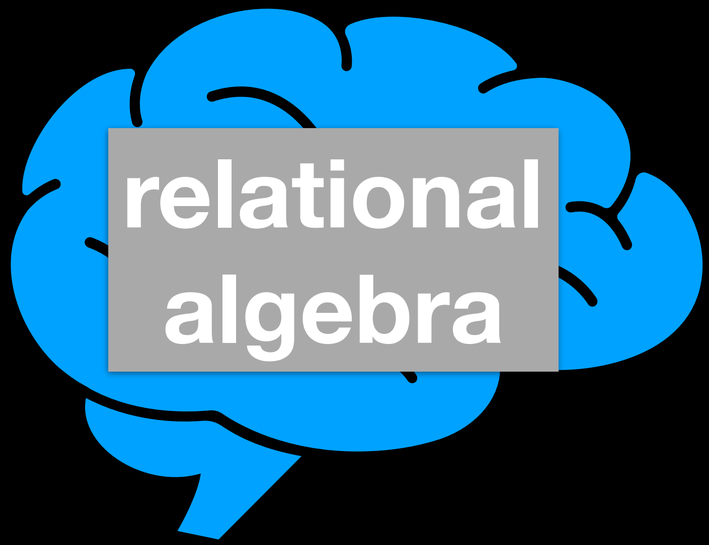
relational algebra, i.e. π, σ, ⋈, …

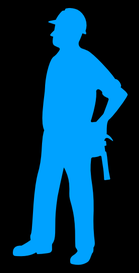
He said: “relational algebra“





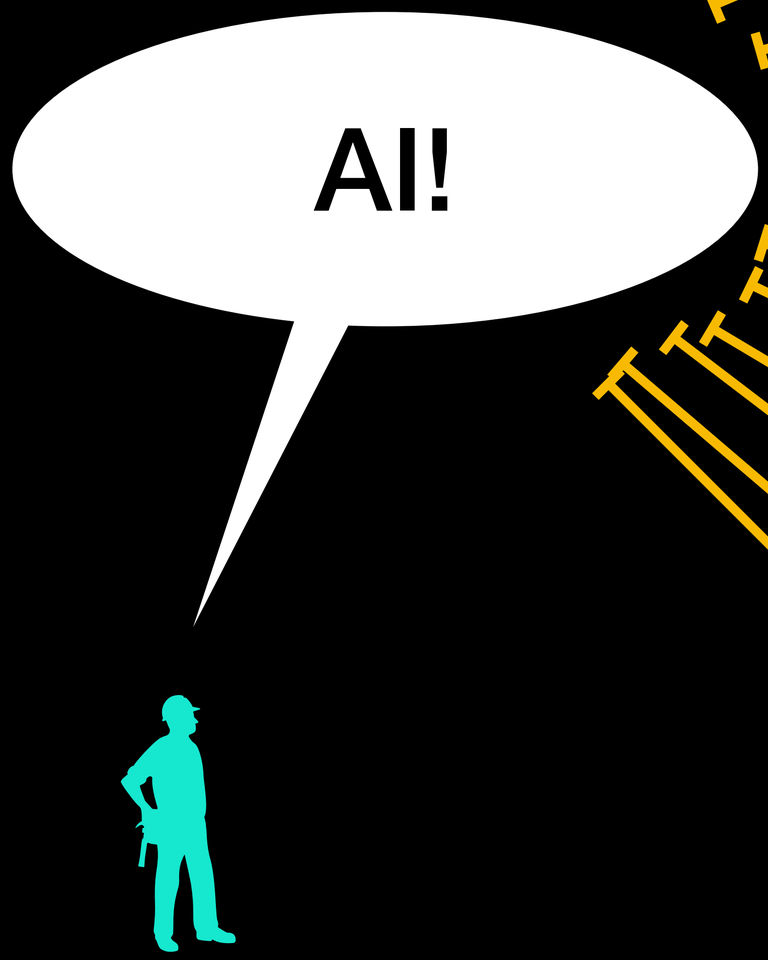
translated to:





symbol

meaning



deep learning

knowledge representation

symbolic AI

machine learning

planning

NLP

robots

statistical learning

deductive databases

computer vision

logics

symbol

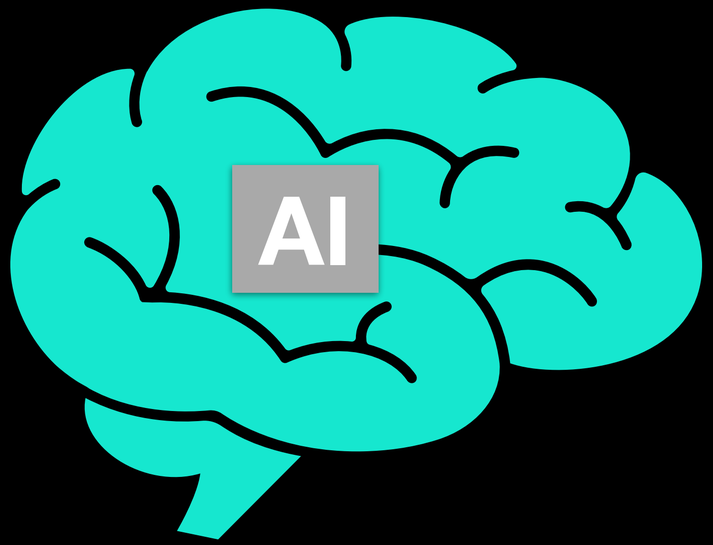
meaning

„Company X wants to solve problem Z using an AI.“

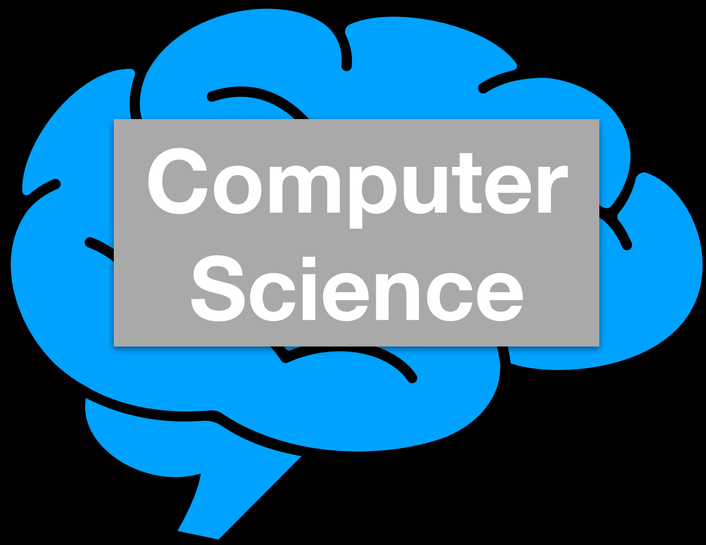
circuit

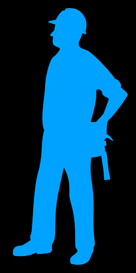
electronics

computer science

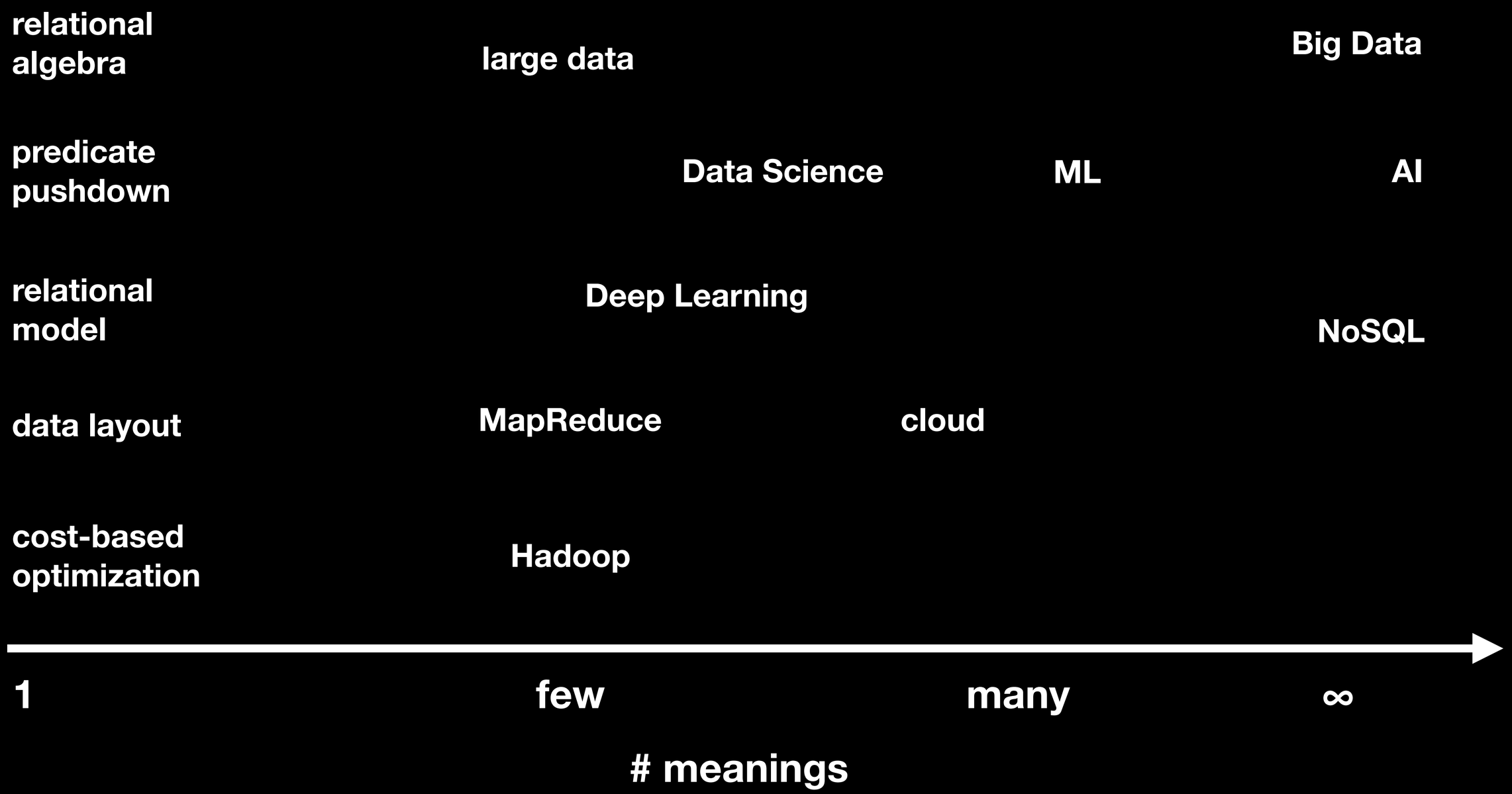


translated to:



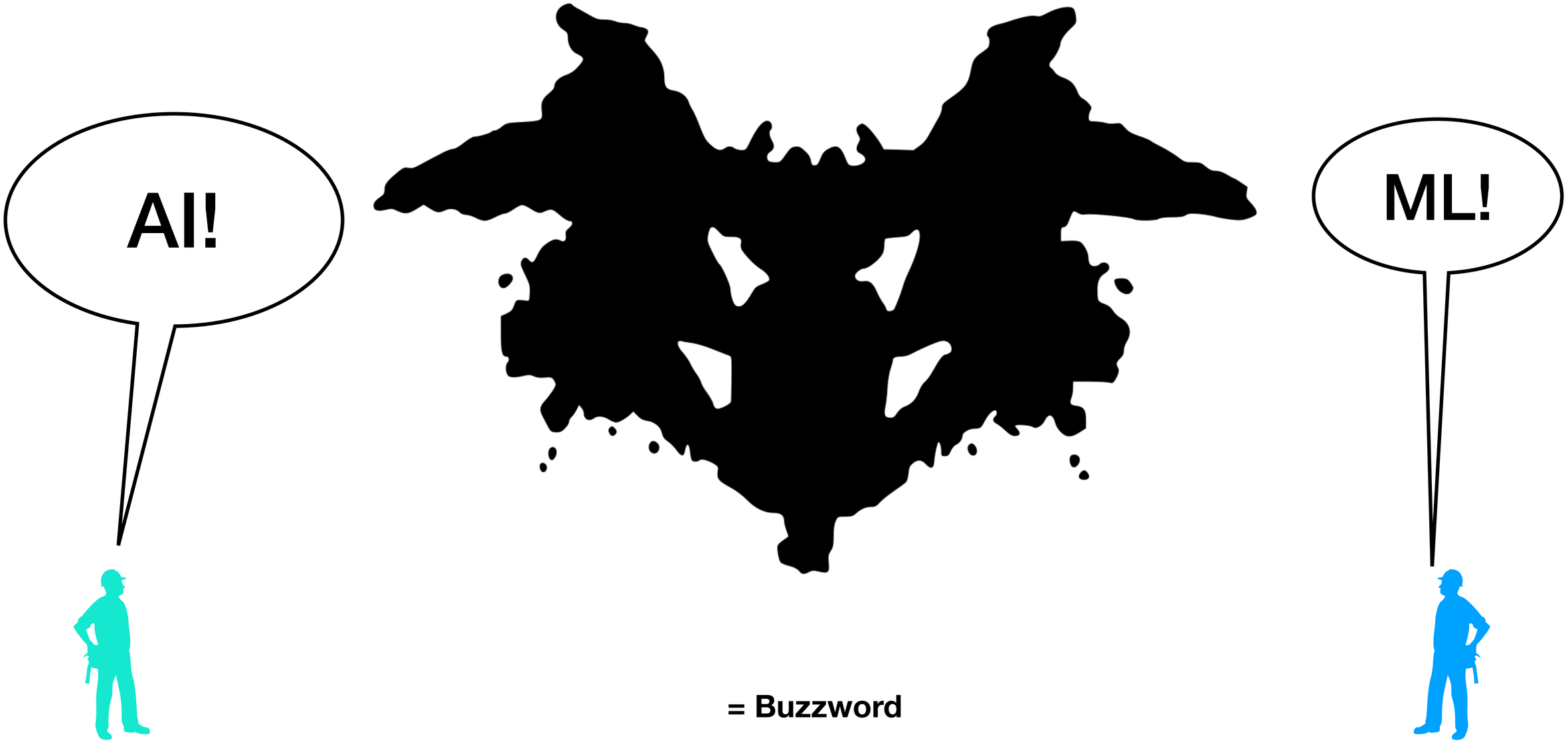


The buzzword bullshit landscape

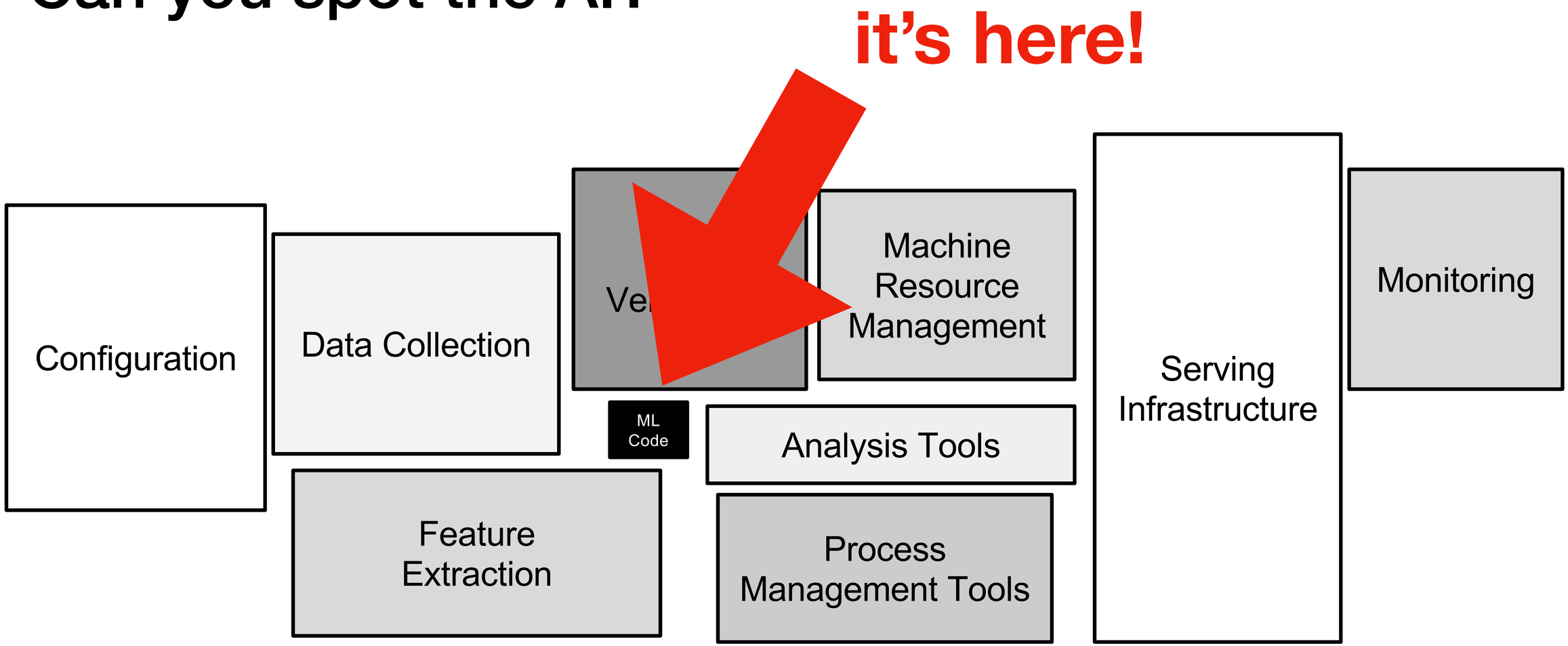


advice: only use non-ambigous terms in communication

Buzzwords as a Rorschach test



Can you spot the AI?



[Sculley et al. Hidden Technical Debt in Machine Learning Systems, NIPS 2015, Figure 1]

OK, if buzzwords do not help, what is DSAI about?

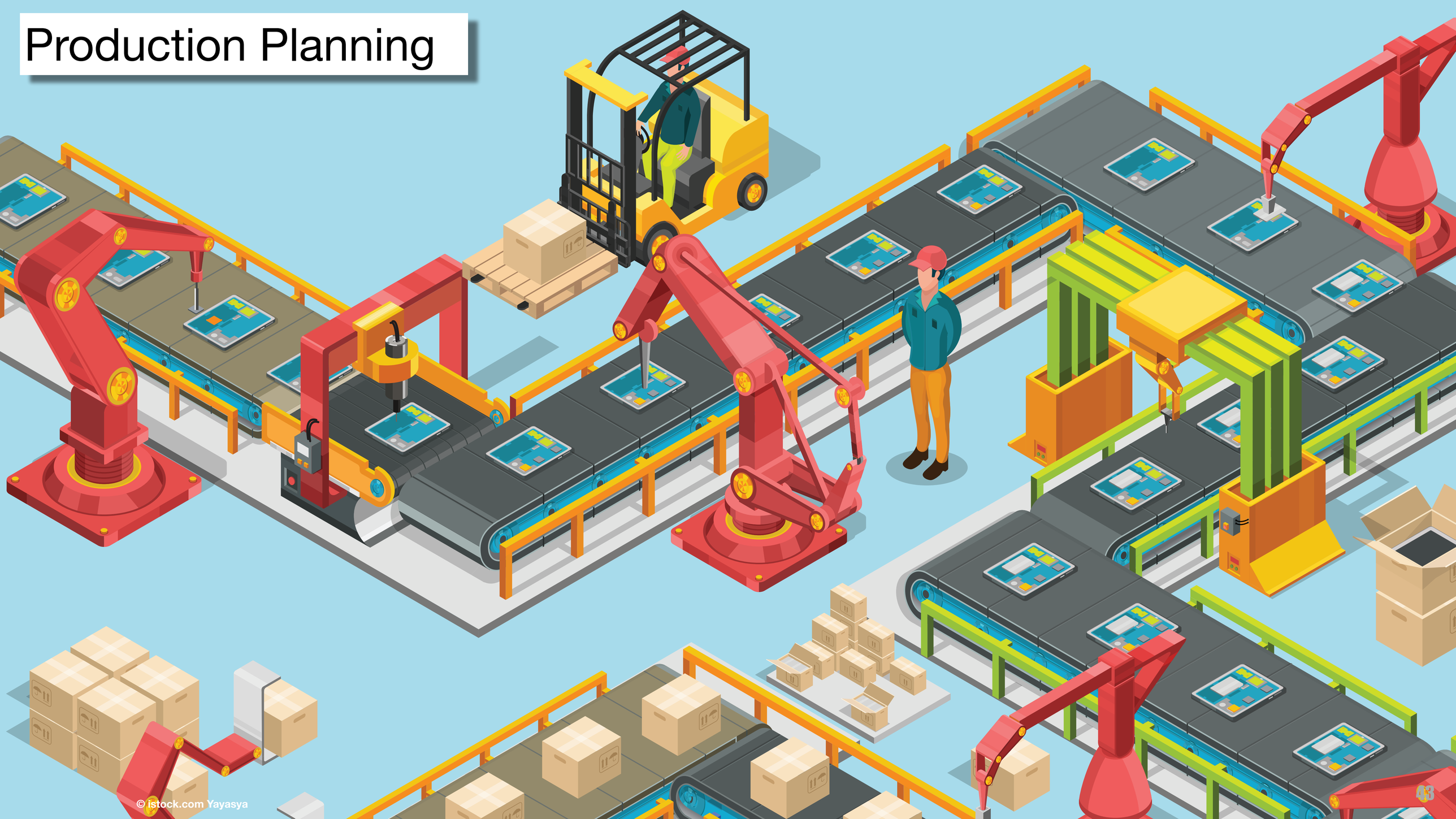
Solution:

Define DSAI by typical Applications and/or methods in that space.

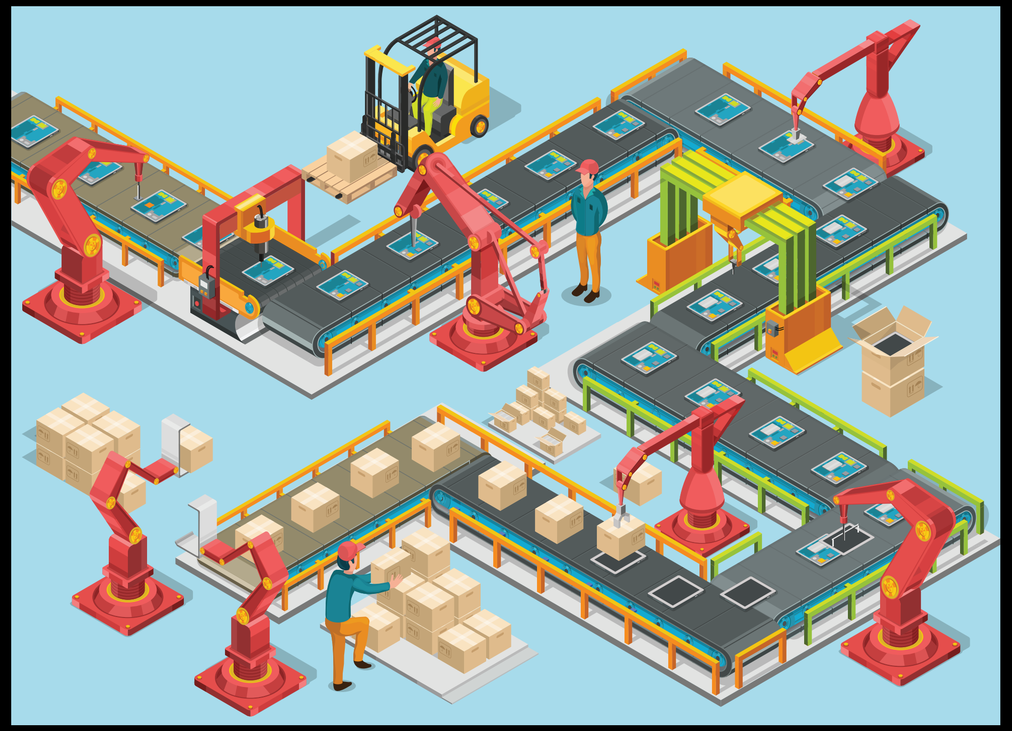
DSAI Applications:

nature
THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
At last — a computer program that can beat a champion Go player PAGE 484
ALL SYSTEMS GO

Production Planning



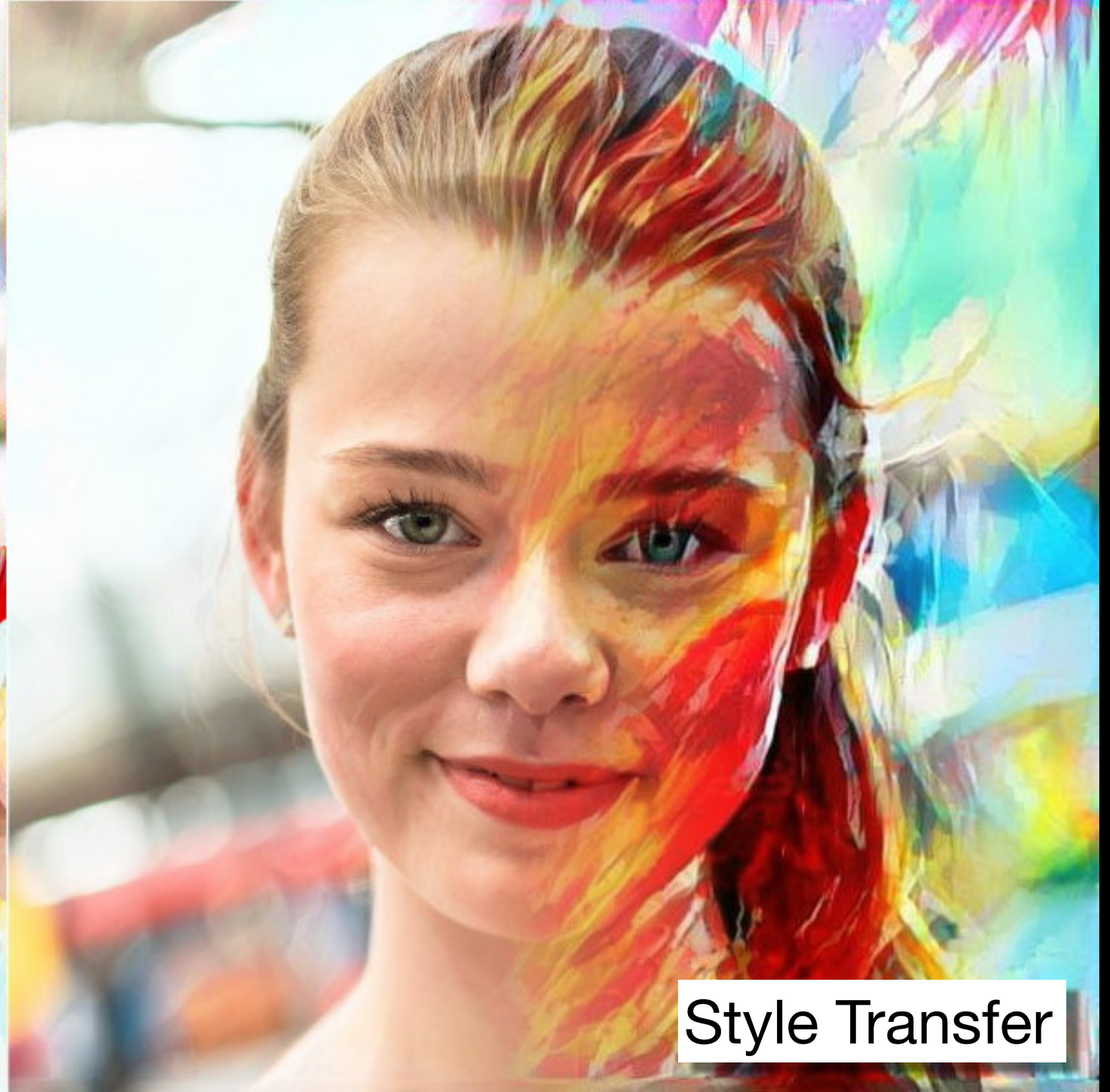
© istock.com Yayasya



+

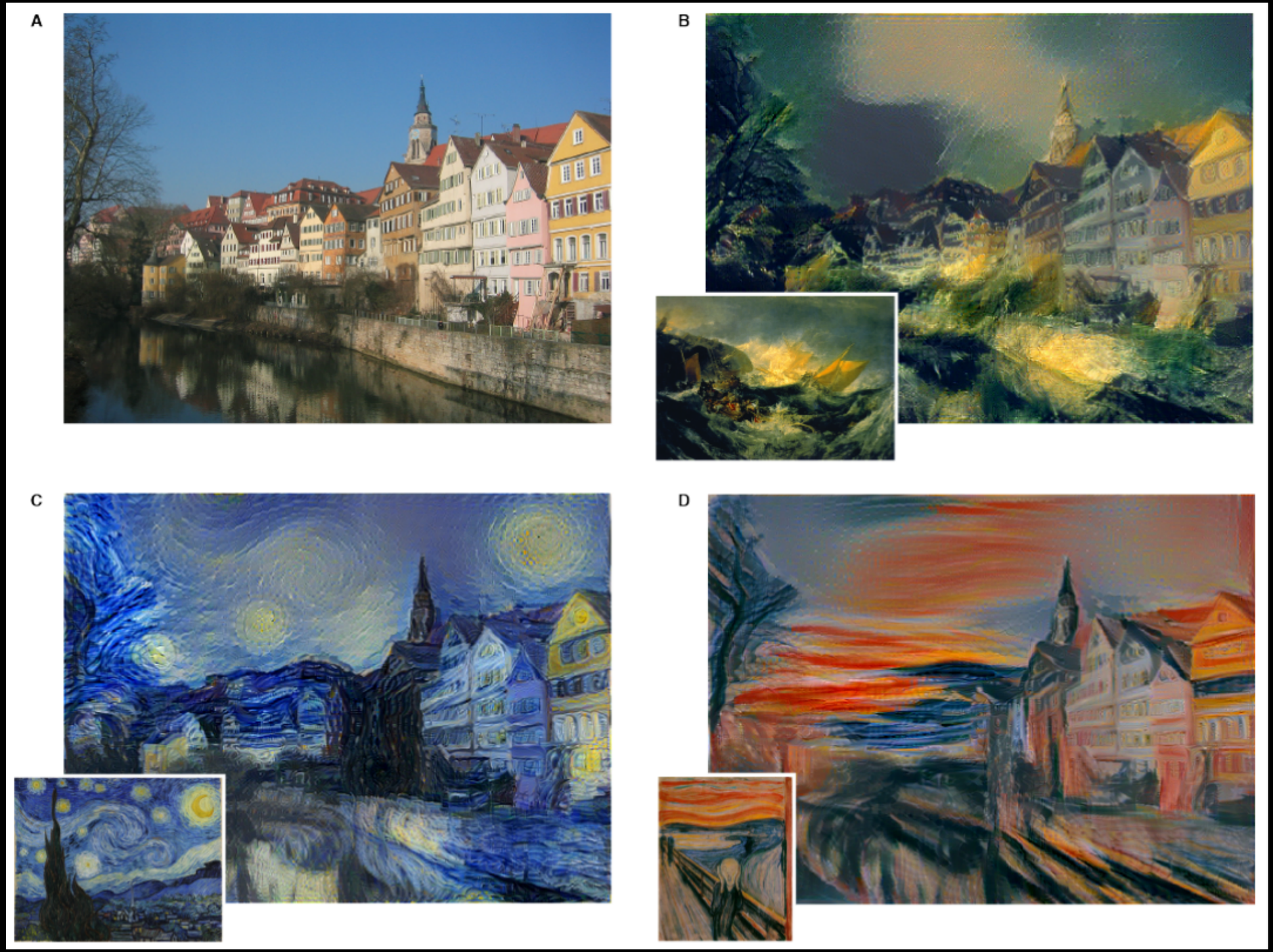
At last – a computer program that
can beat a champion Go player PAGE 484
ALL SYSTEMS GO

= d:AI:mond



Style Transfer

Style Transfer

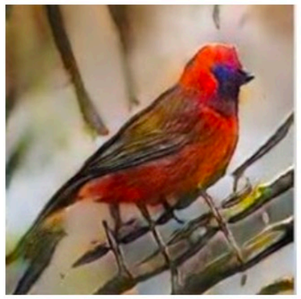
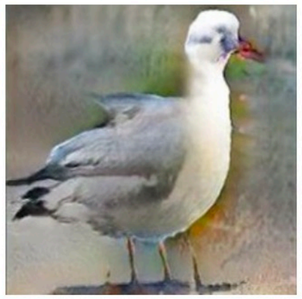
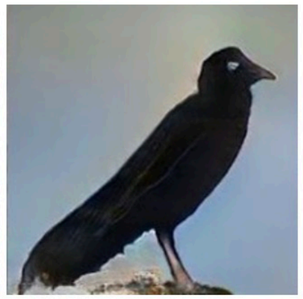
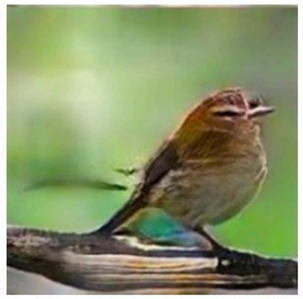
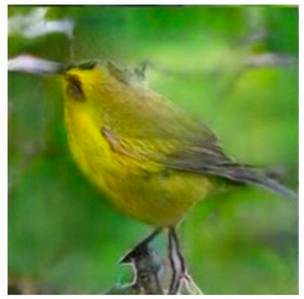


Deep Dream

Traffic Sign Recognition Benchmark

[Traffic Sign Recognition with Multi-Scale Convolutional Networks, Pierre Sermanet and Yann LeCun]

Text 2 Image

Text description	This bird is red and brown in color, with a stubby beak	The bird is short and stubby with yellow on its body	A bird with a medium orange bill white body gray wings and webbed feet	This small black bird has a short, slightly curved bill and long legs	A small bird with varying shades of brown with white under the eyes	A small yellow bird with a black crown and a short black pointed beak	This small bird has a white breast, light grey head, and black wings and tail
256x256 StackGAN							

[Han Zhang et al, https://arxiv.org/abs/1612.03242]

Interpreter

Self-Driving Cars

https://en.wikipedia.org/wiki/Self-driving_car#/media/File:Waymo_Chrysler_Pacifica_in_Los_Altos,_2017.jpg

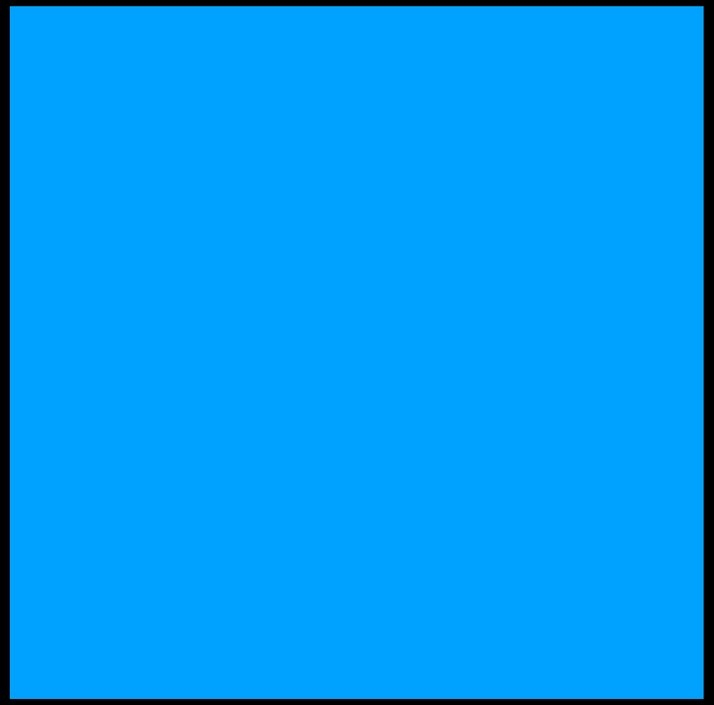
Truck drivers

Taxi drivers

Package deliverers

Cars as selective crowd-sourcers

Software 1.0 vs Software 1.2

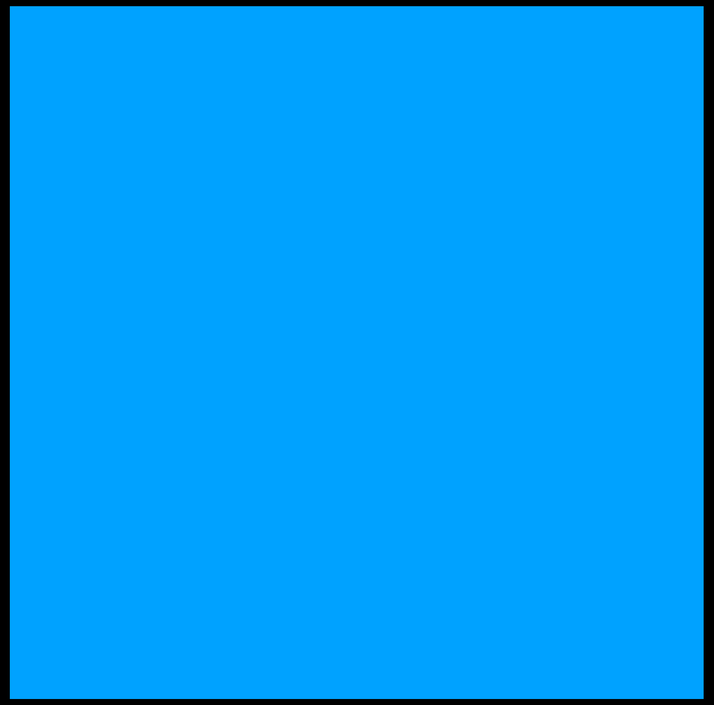


Progam code

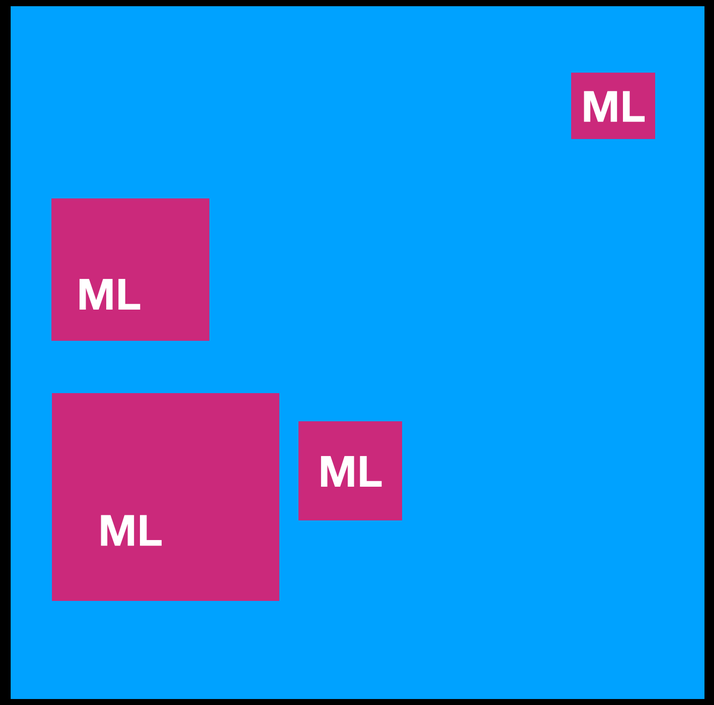
Progam code

“Software 2.0“: https://medium.com/@karpathy/software-2-0-a64152b37c35

Software 1.0 vs Software 1.4

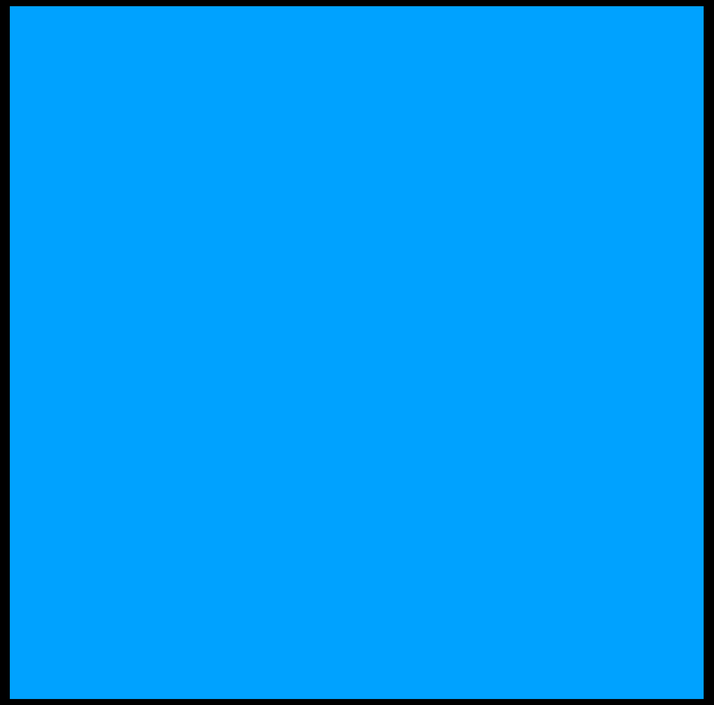


Progam code

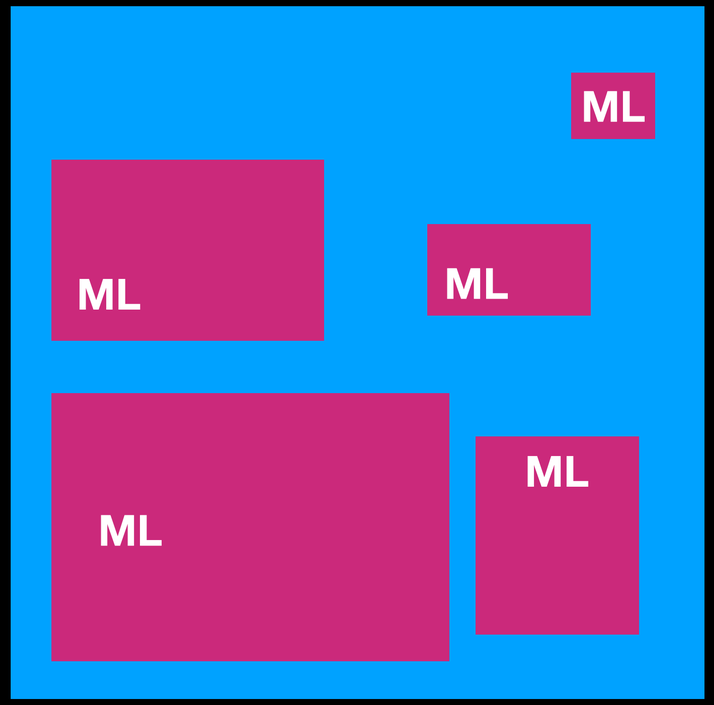


Progam code

Software 1.0 vs Software 1.7

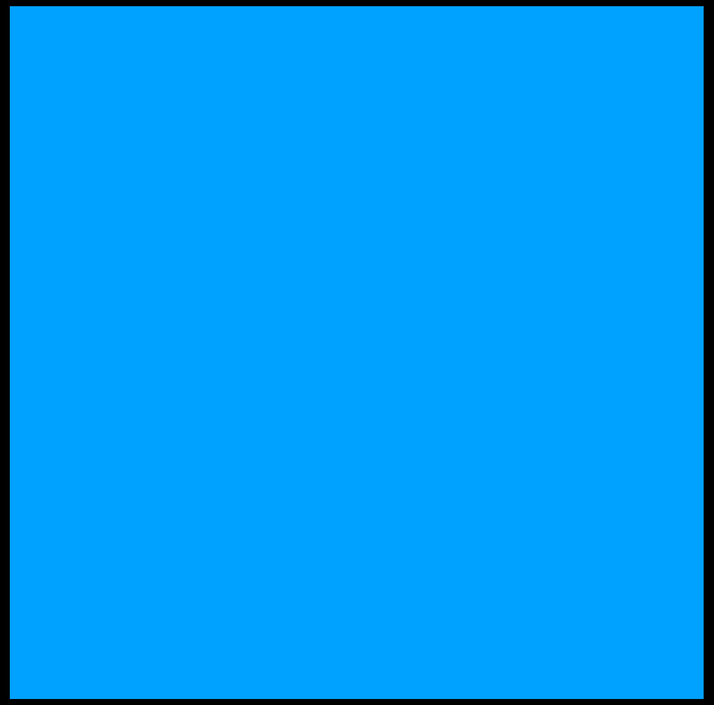


Progam code

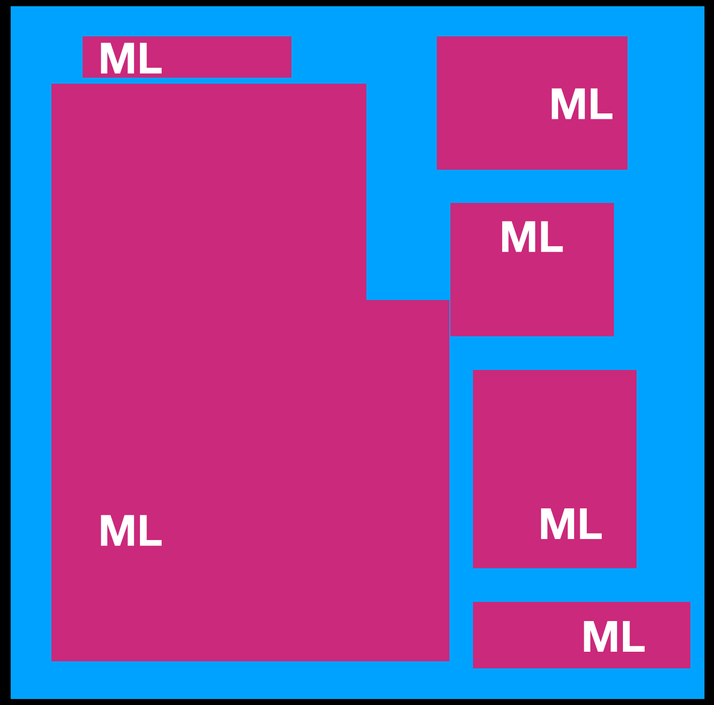


Progam code

Software 1.0 vs Software 2.0



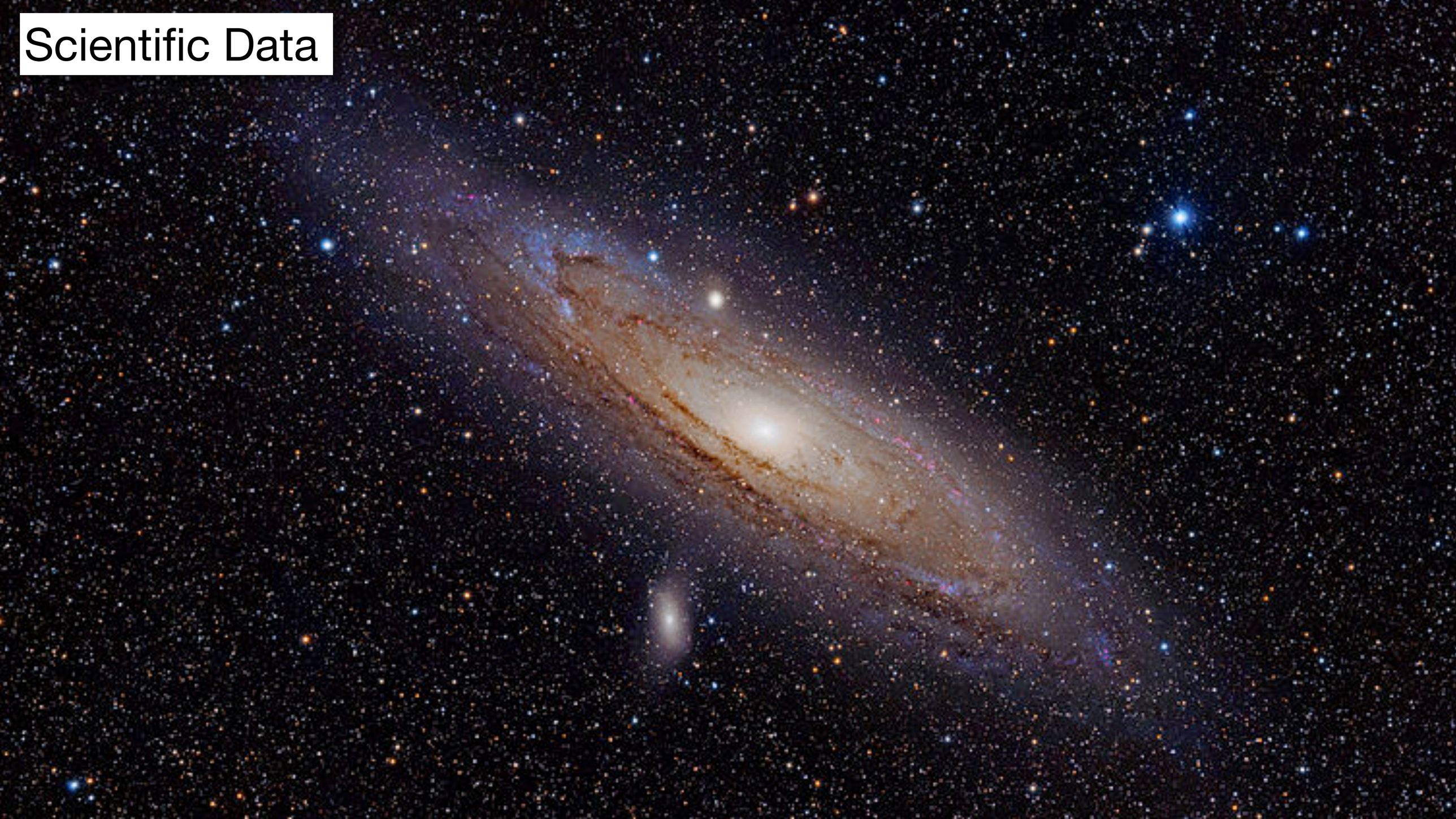
Progam code



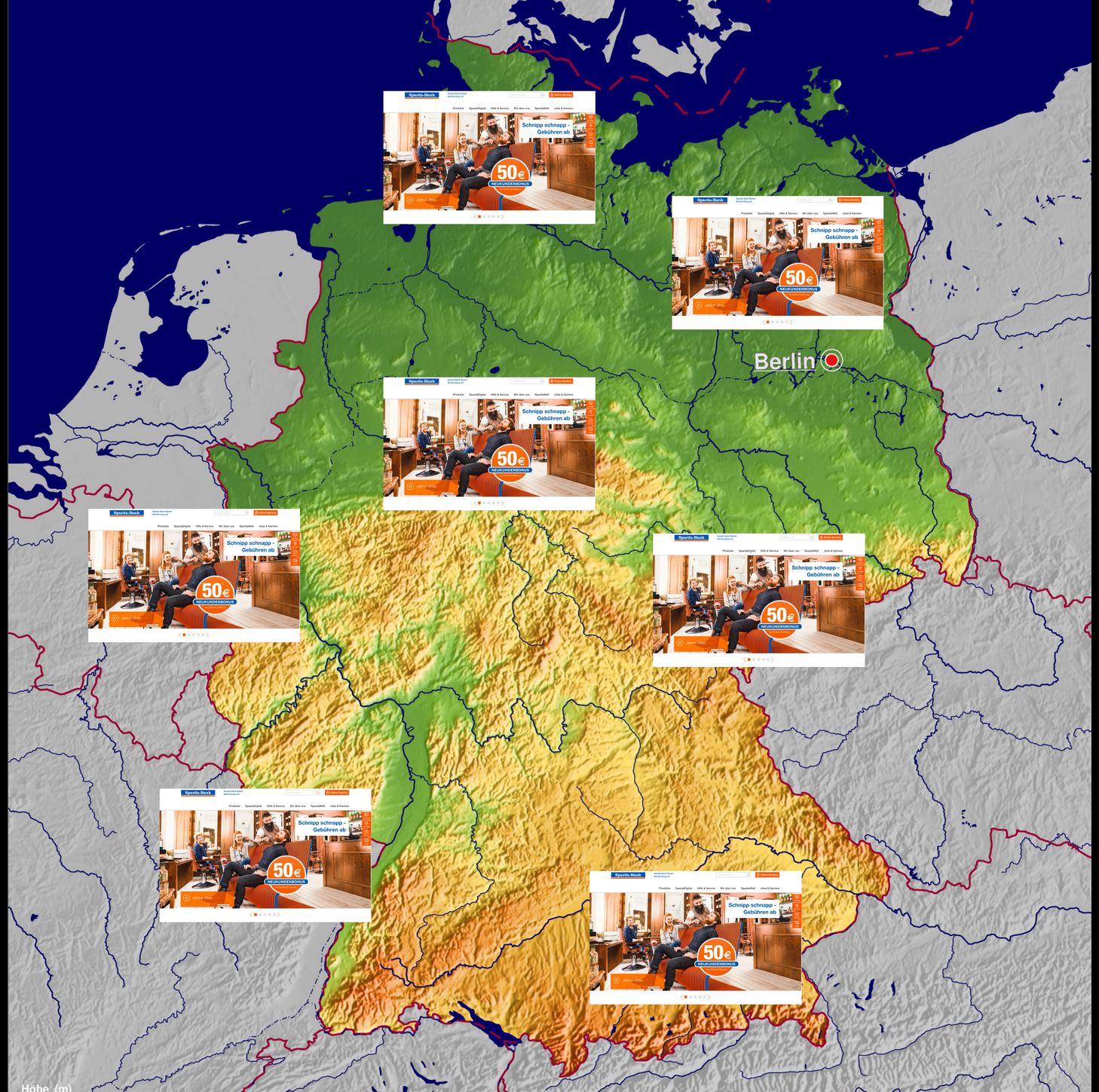
Progam code

Scientific Data

Scientific Data

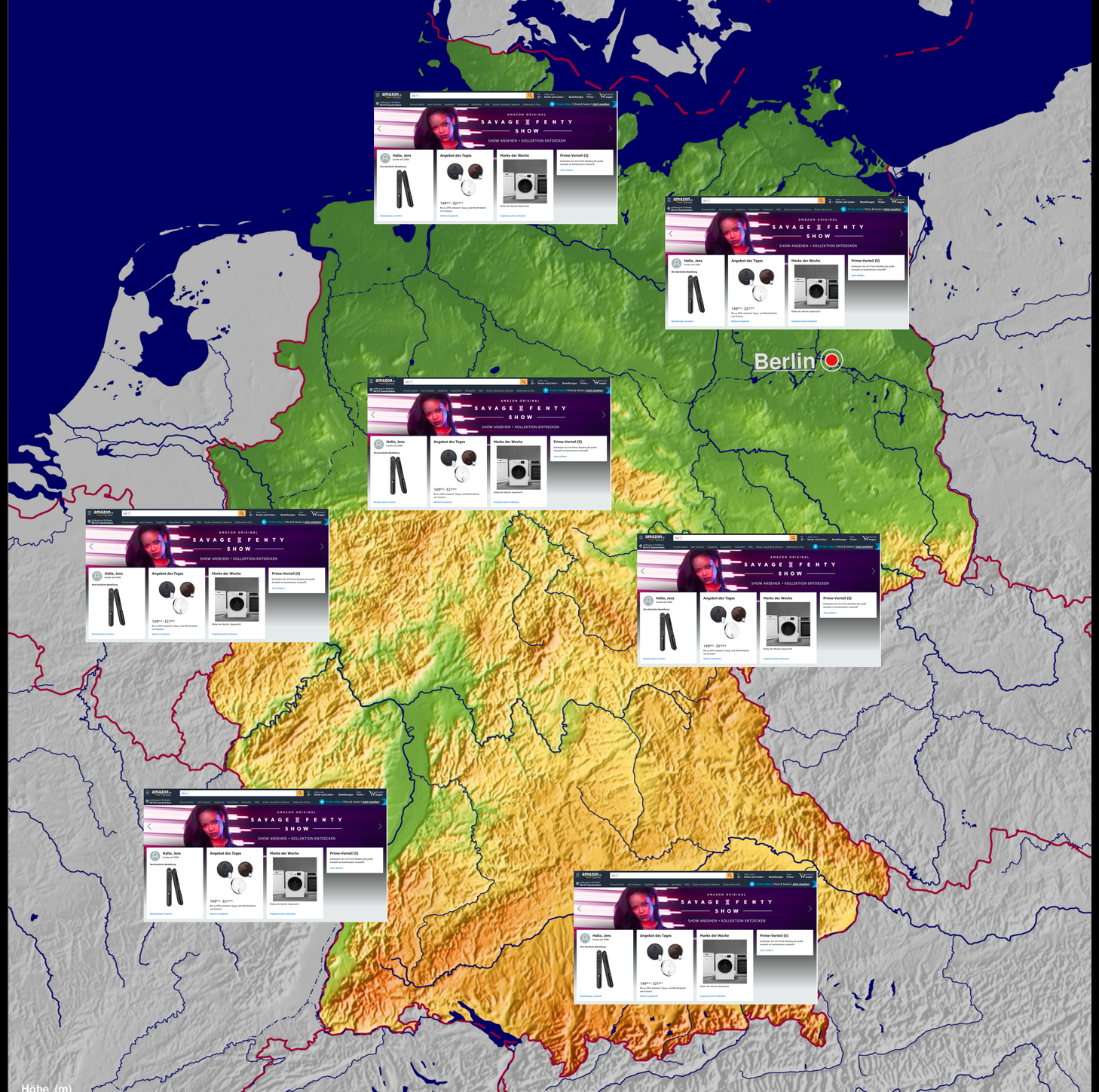


Banking



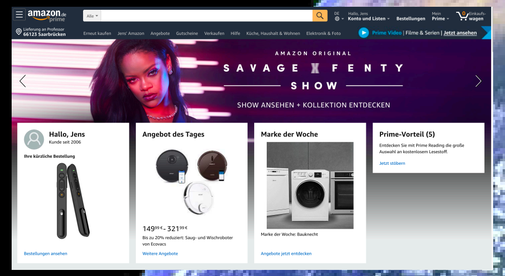
Shopping

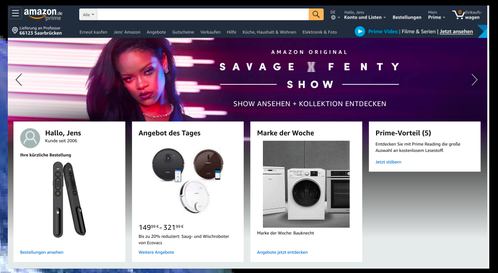
country-scale

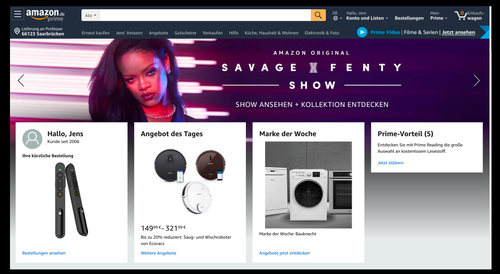


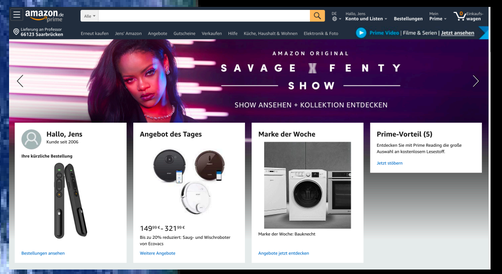
Shopping

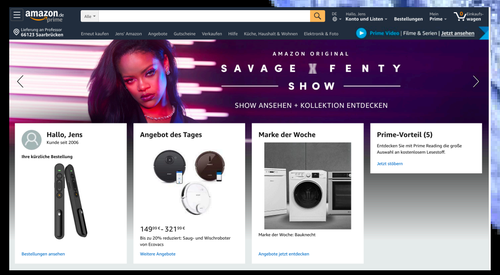
planet-scale

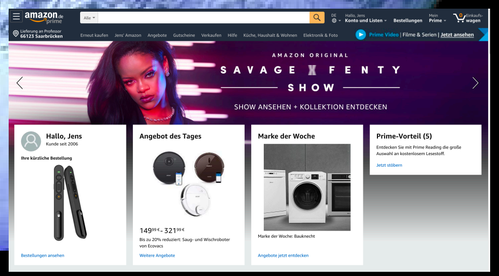






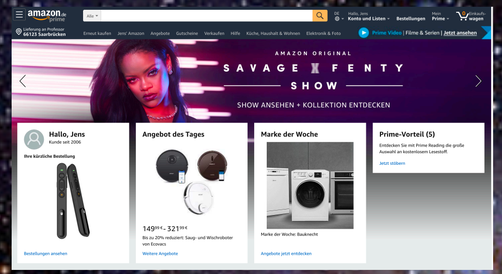


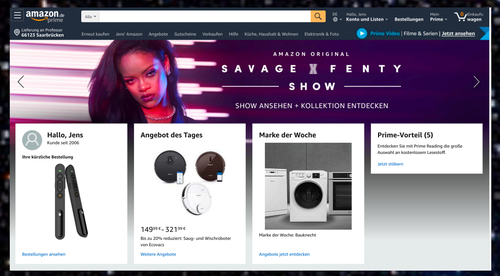


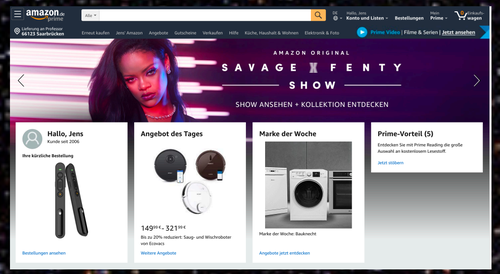


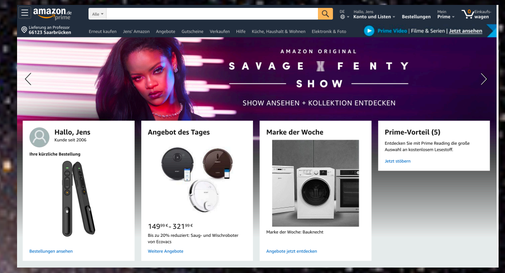
Shopping

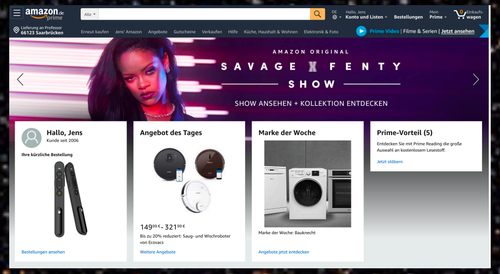
galaxy-scale

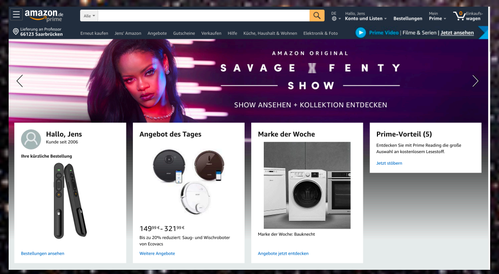




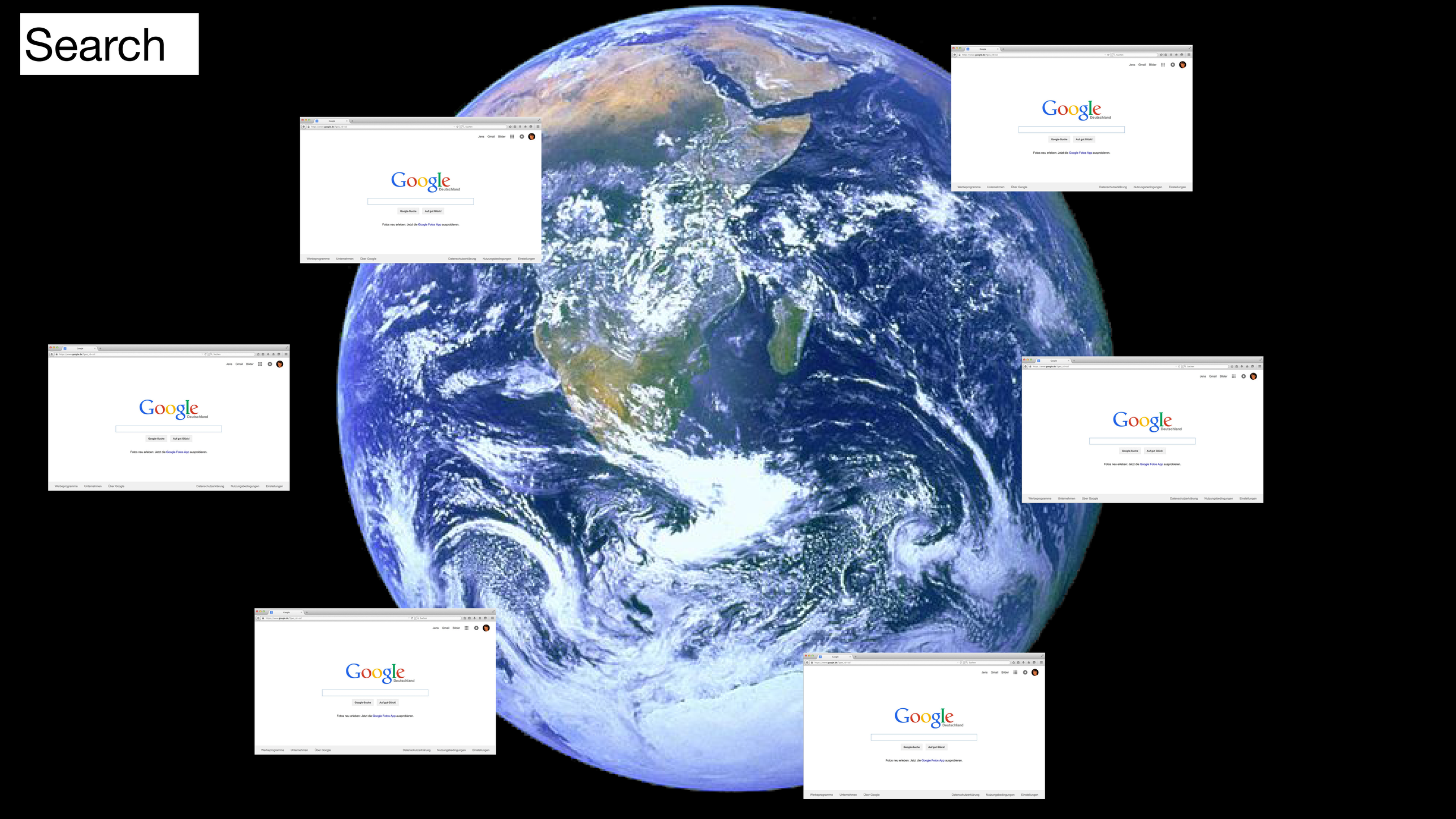




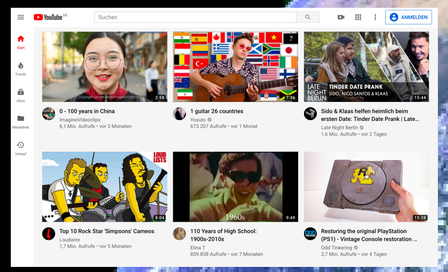


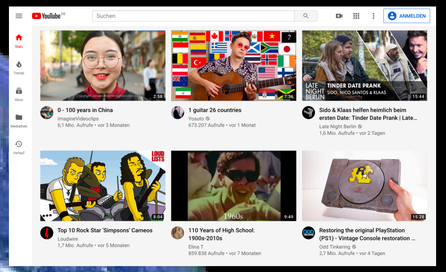


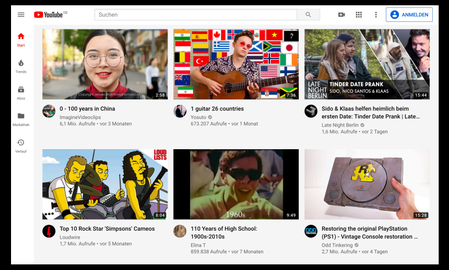
Search

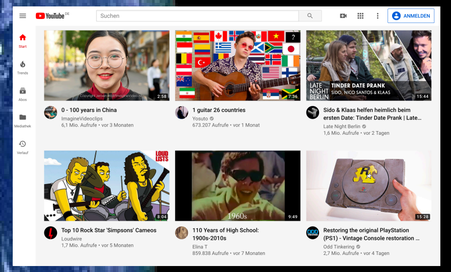

Google
Deutschland
Google-Suche
Auf gut Glück!
Fotos neu erleben: Jetzt die Google Fotos App ausprobieren.

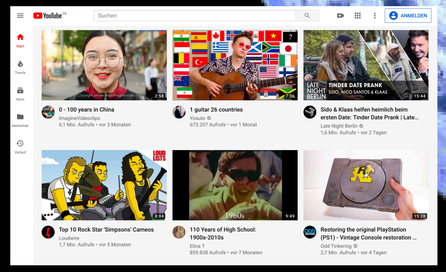
Content Delivery

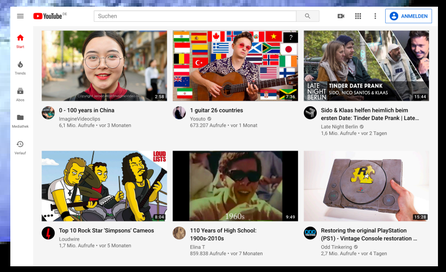


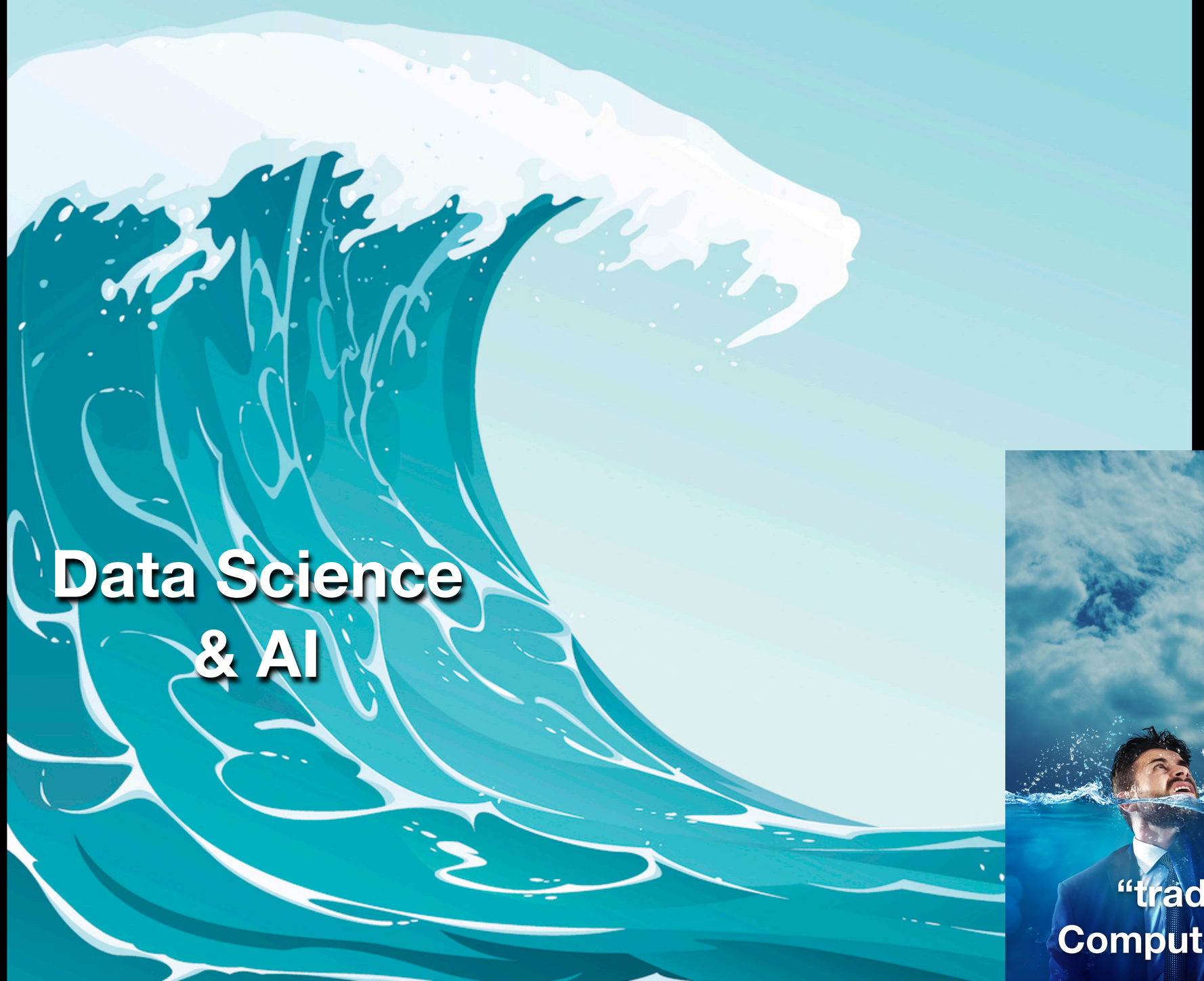











Data Science
& AI

“traditional“
Computer Science

Radiology

radiology

RADIOLOGY BUSINESS
FOR LEADERS NAVIGATING VALUE-BASED CARE

NEWS MAGAZIN

AI can differentiate between tuberculous, pyogenic spondylitis as well as radiologists

September 10, 2018 | *Michael Walter* | *Artificial Intelligence*

DIAGNOSTIC IMAGING

Enter your keywords

Topics | Image IQ | Case Studies | Blog | Conferences

Machine Outperforms Radiologists for Staging Liver Fibrosis Using CT

By **Diagnostic Imaging Staff**

Sep 19, 2018

AI algorithm spots abnormal chest X-rays with 90 percent accuracy

September 19, 2018
by Thomas Dworetzky , Contributing Reporter

Artificial intelligence can determine lung cancer type

Tool also IDs genetic changes in each patient's tumor

Spending for machine learning in radiology expected to soar

Dermatologist
Cardiologist
Pathologist
Laboratory physician
Surgeon

medicine

AI predicts risk of death from heart disease more accurately than experts

The machine learning model uses 600 variables with patient's data whereas human-constructed models made predictions based on 27, researchers say.

PUBLIC RELEASE: 21-SEP-2018

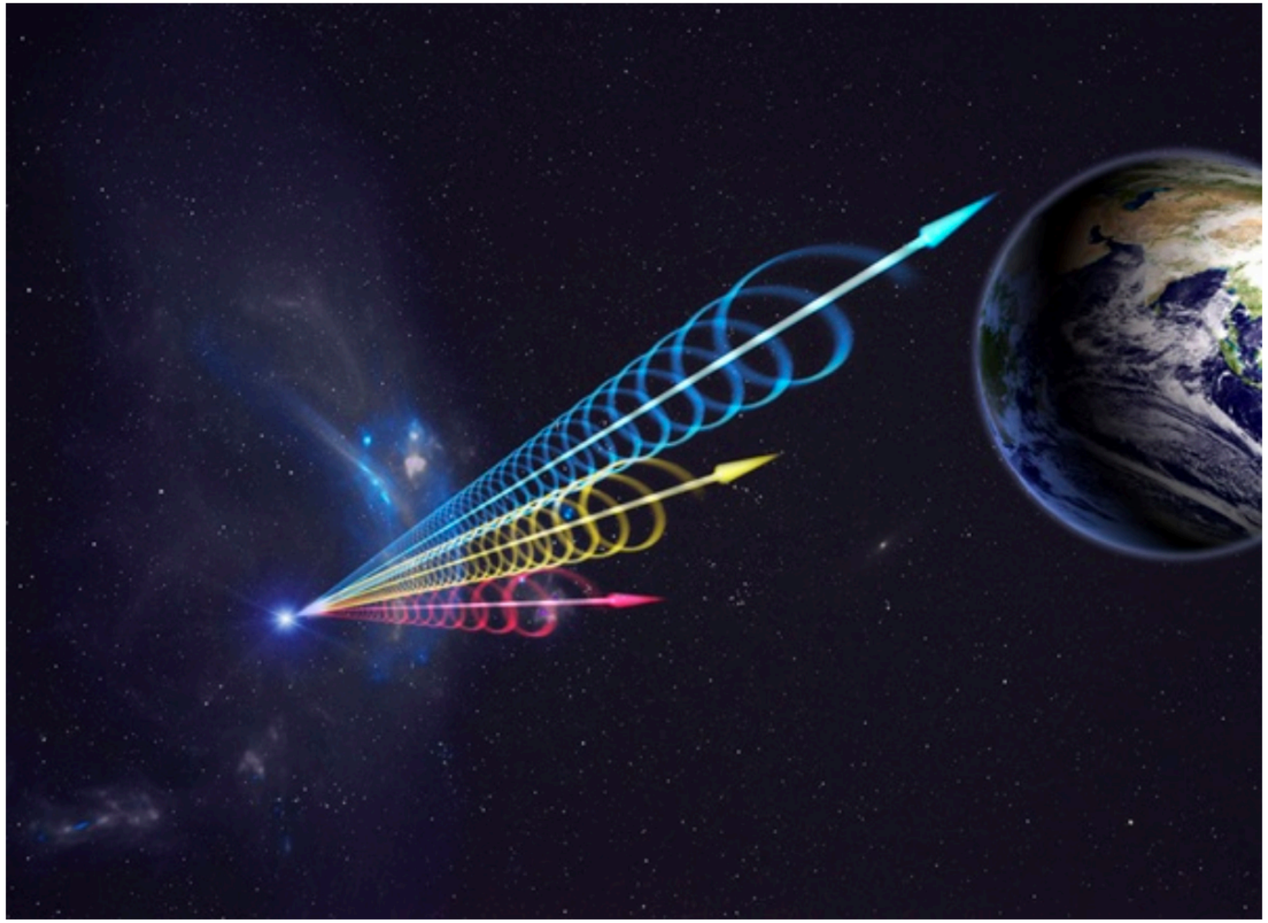
AI used to detect fetal heart problems

Sounding the Alarm: How Deep Learning Helps Doctors Detect Pediatric Sepsis

AI detects 72 fast radio bursts from a distant, unknown source

Machine learning technology can help astronomers identify signals amidst the noise of the universe.

By Amber Jorgenson | Published: Tuesday, September 11, 2018



An artist's illustration shows the signal from a fast radio burst zooming its way toward Earth.

Jingchuan Yu, Beijing Planetarium

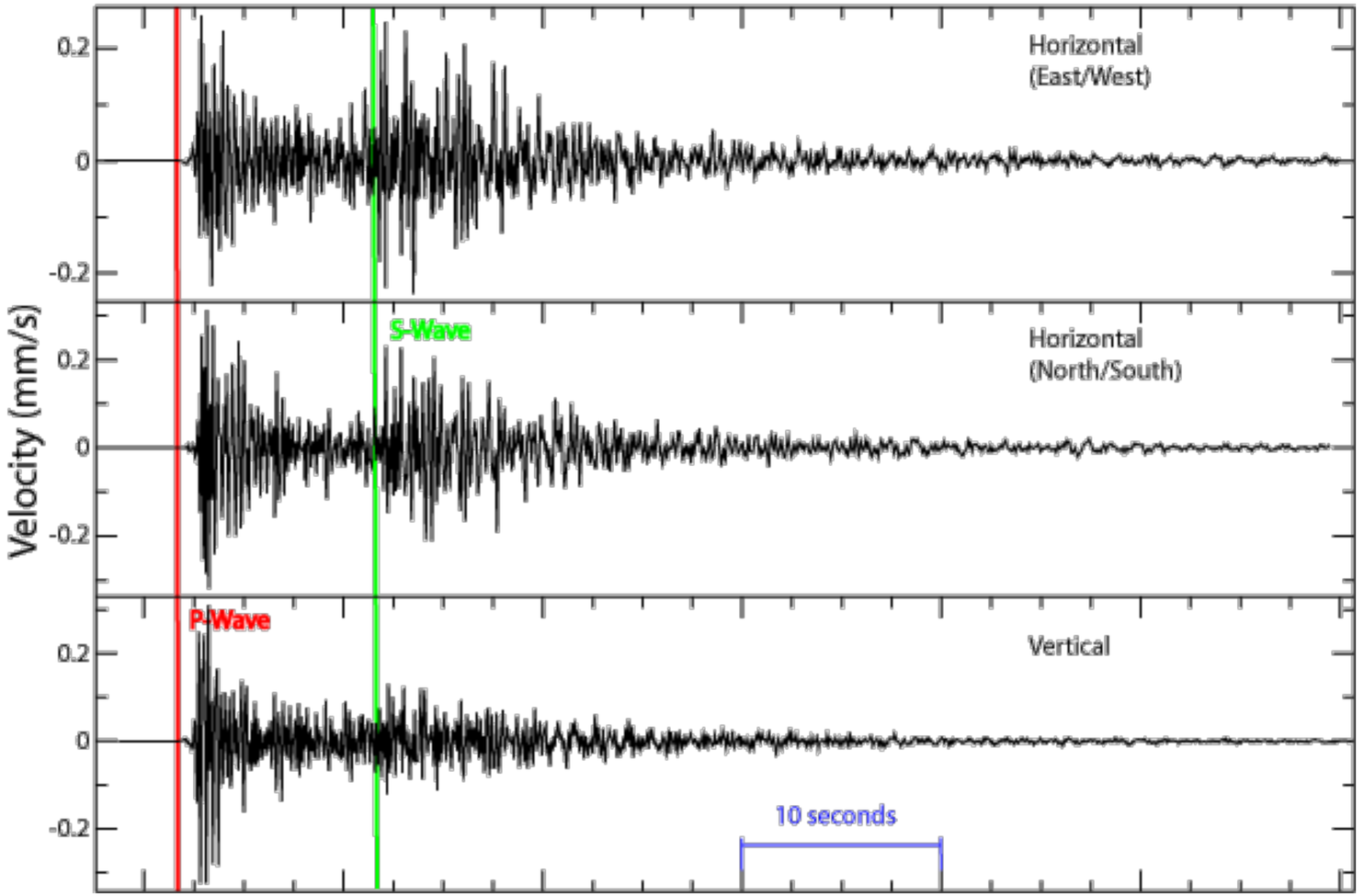
physics

Lightning Prediction

[Schön Christian, Dittrich Jens, Müller Richard. The Error is the Feature:
how to Forecast Lightning using a Model Prediction Error. SIGKDD 2019.]

© istock.com guvendemir

Seismology


Velocity (mm/s)
0.2
0
-0.2
Horizontal
(East/West)
S-Wave
Horizontal
(North/South)
P-Wave
Vertical
10 seconds

Seismology

Forecasting earthquake aftershock locations with AI-assisted science

Detecting earthquakes over a seismic network using single-station similarity measures

Karianne J Bergen ✉, Gregory C Beroza

Breaking Cascadia's Silence: Machine Learning Reveals the Constant Chatter of the Megathrust

Authors: Bertrand Rouet-Leduc[1*], Claudia Hulbert[1*], Paul A. Johnson[1]
[1]Affiliation: Los Alamos National Laboratory, Geophysics Group, Los Alamos, New Mexico, USA

Locality-Sensitive Hashing for Earthquake Detection: A Case Study of Scaling Data-Driven Science

Kexin Rong*, Clara E. Yoon†, Karianne J. Bergen‡, Hashem Elezabi*,
Peter Bailis*, Philip Levis*, Gregory C. Beroza†
Stanford University

The dark side:

Automatic warfare

Social Scoring

Global Mass-Surveillance

Data Science
& AI

“traditional“
Computer Science

“traditional“
Sciences