Scientific Machine Learning

CSU/IMU Oct'2023

Mark Asch/UPJV

Plan

1. Introduction and Motivation—it's all about data.

2. What is Scientific Machine Learning?

3. How can SciML contribute to CSE?

Think out of the box!

the world has changed….

Think out of the box!

the world has changed….

Questions, questions, questions

HOW? (Analyse)
WHY? (Understand)
WHAT IF? (Predict)

1. A new approach to computational science ?

The time has come to discuss it!

The Fourth Paradigm (2009)

1. Experiment
2. Theory
3. Simulation
4. Data, data, data

From antiquity, till now.

From middle ages, till now.

20th Century, till now.

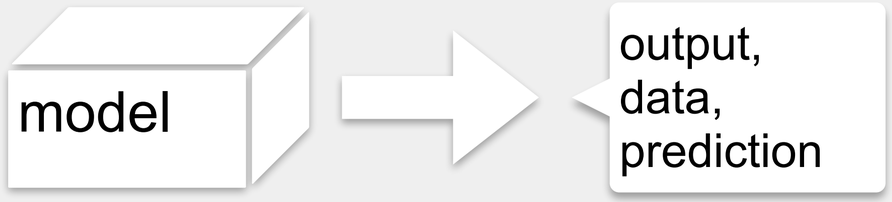
21st Century, **NOW!**

Machine Learning

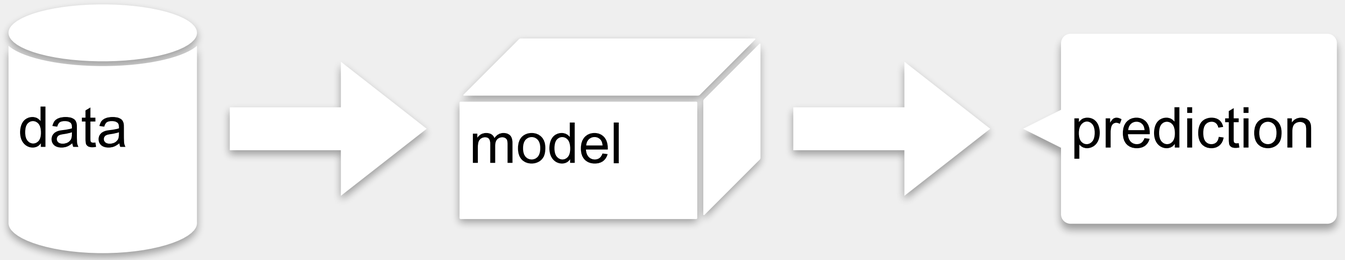
Turning our world (CSE) inside-out…

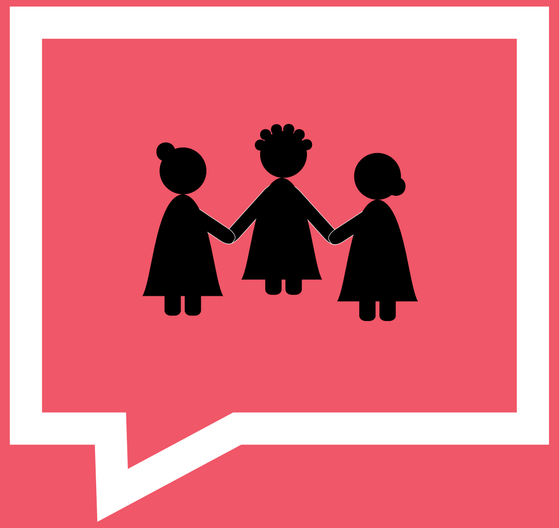
Two, diametrically opposite approaches?

- CSE: from models to data (predictions)



- ML/AI: from data to models to predictions ("Let the data talk")





African proverb

"If you want to go quickly, go alone. If you want to go far, go together."

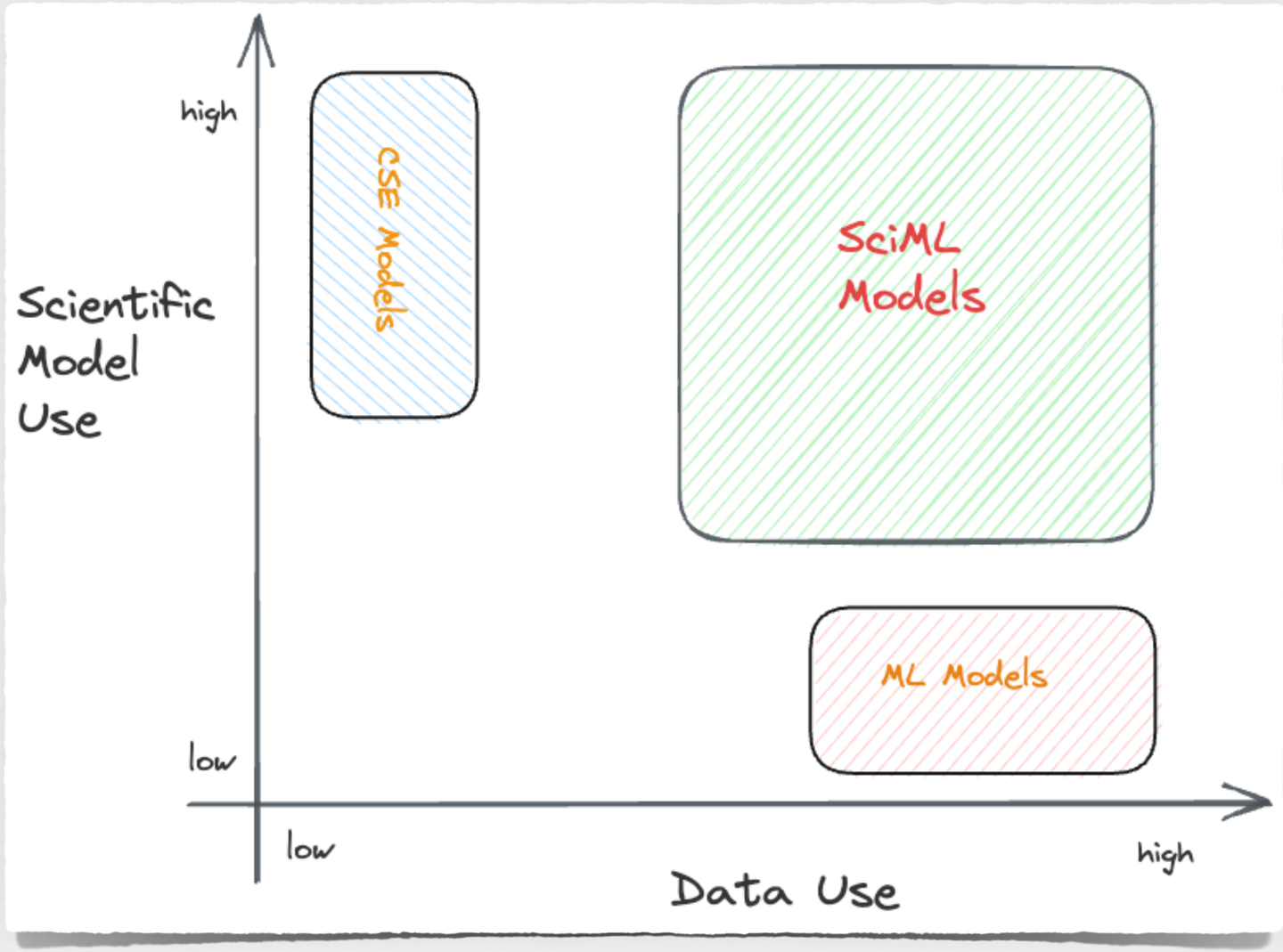
Scientific Machine Learning *(SciML) is a field of research that combines traditional scientific modeling with machine learning techniques. It aims to develop new methods and tools for solving scientific problems that are more accurate, efficient, and generalizable than traditional methods.*

SciML is still a young field, but it has the potential to make a major impact on a wide range of scientific disciplines..

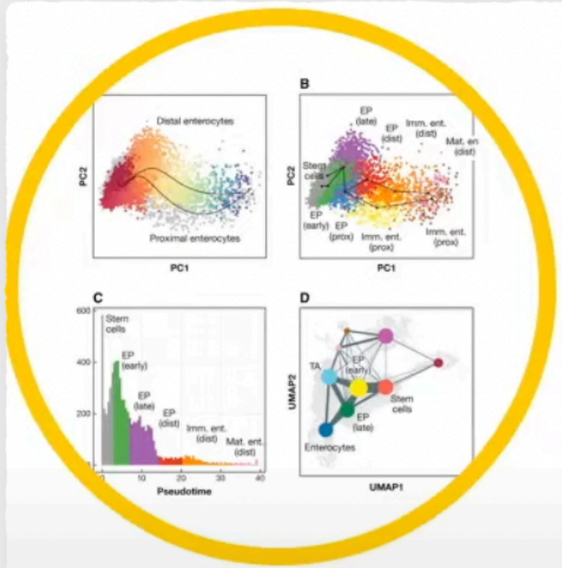
Credit: Bard…

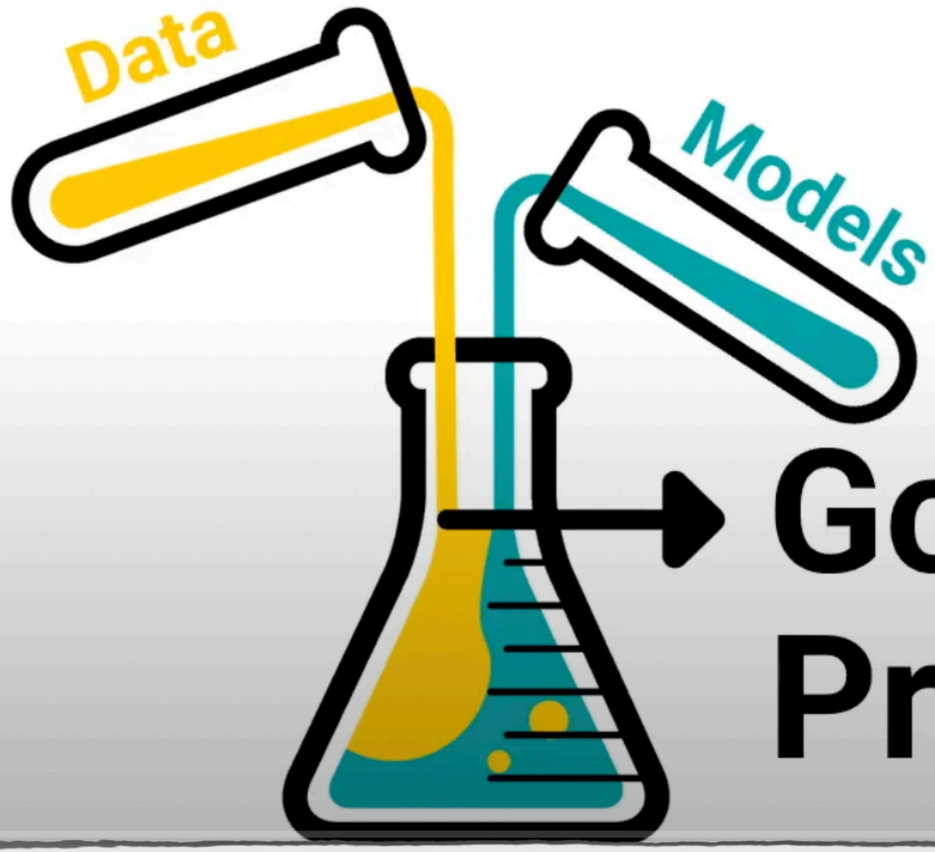
It’s a question of balance…

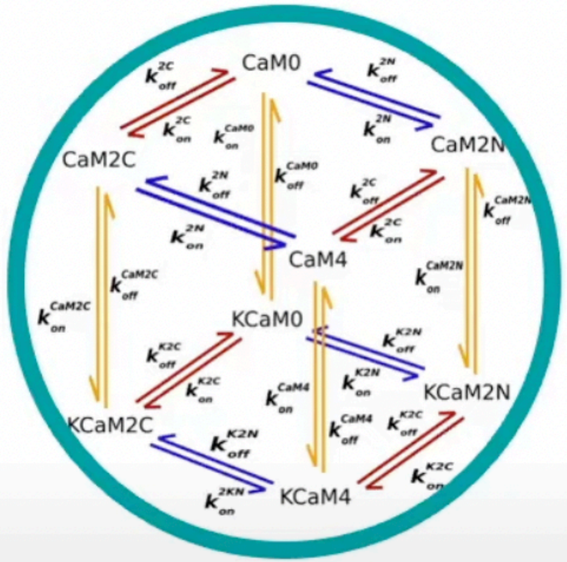
- How much is data-driven?
- How much is scientific knowledge-driven?
- Multi-objective optimization problem



It’s a question of balance…





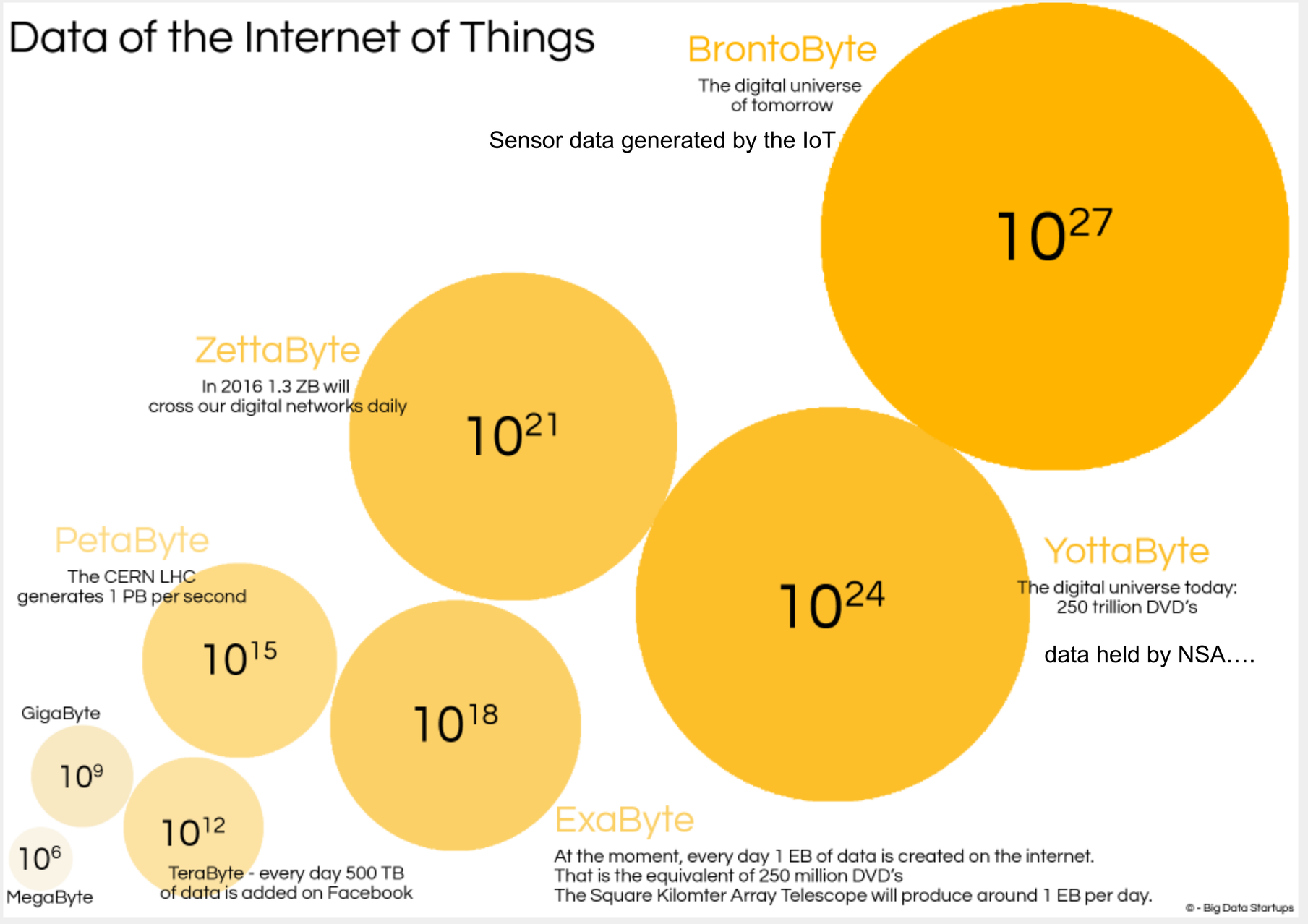


Good Predictions

Credit: C. Rackauckas

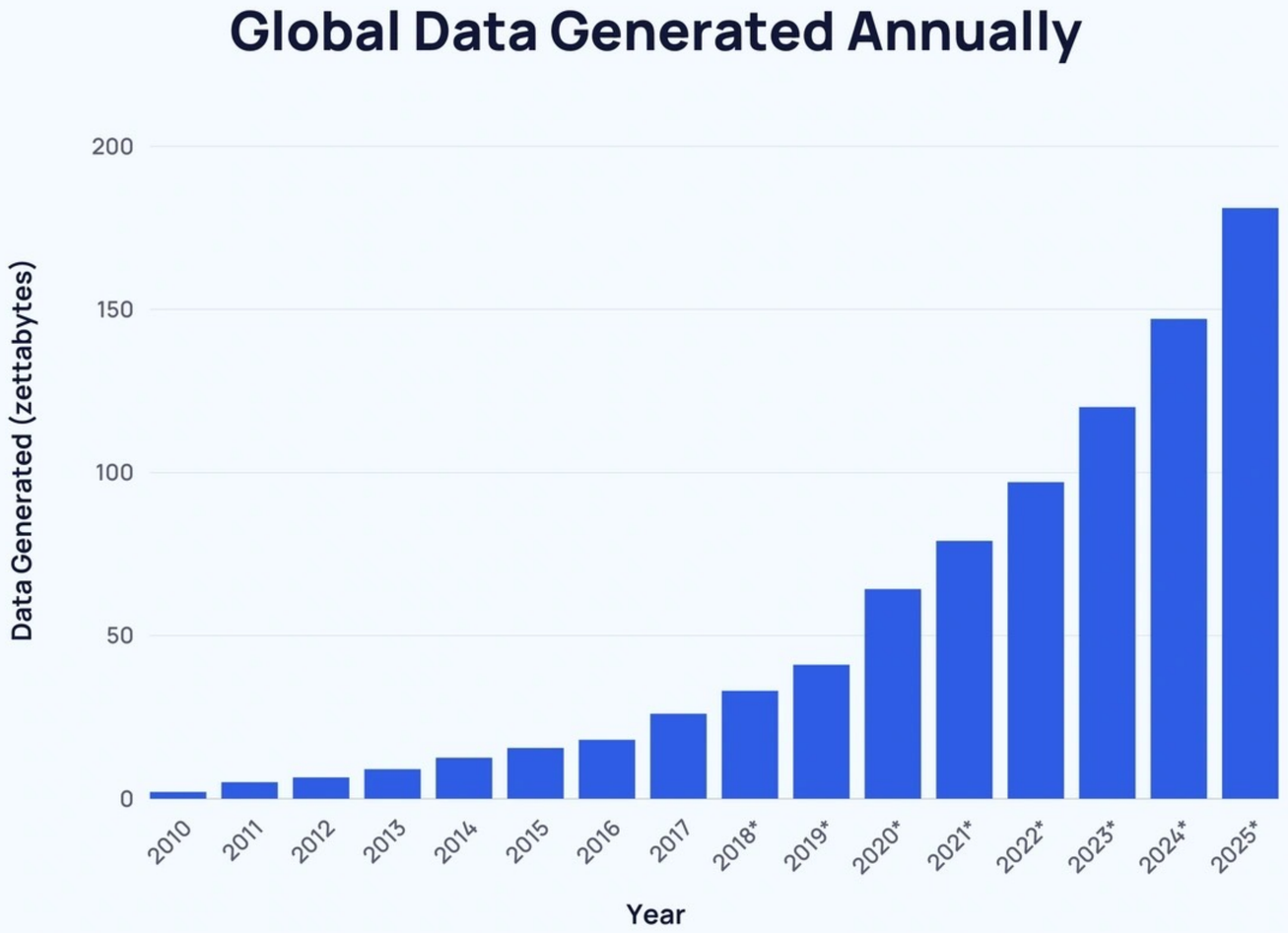
2. Data, data, data

THE dimensioning factor…

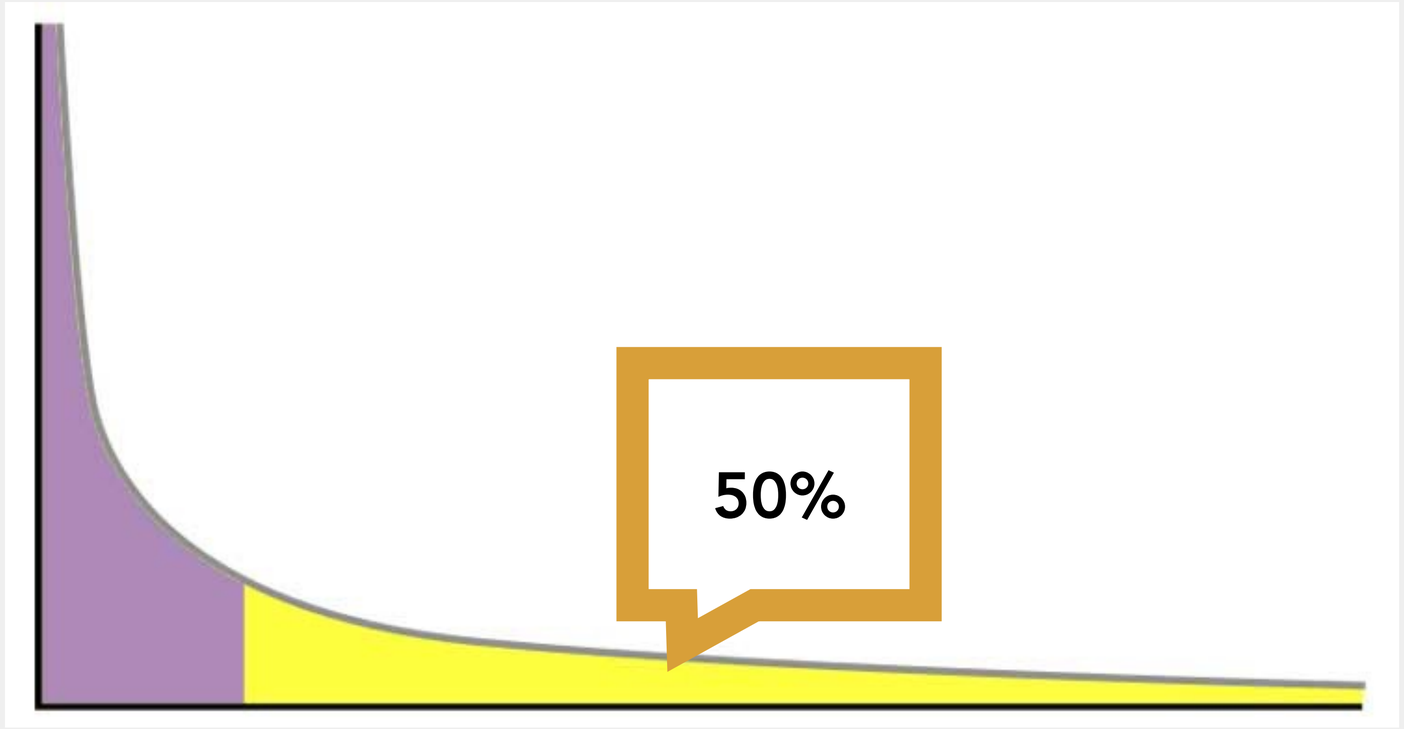

Data of the Internet of Things
BrontoByte
The digital universe of tomorrow
Sensor data generated by the IoT
10^{27}
ZettaByte
In 2016 1.3 ZB will cross our digital networks daily
10^{21}
PetaByte
The CERN LHC generates 1 PB per second
10^{15}
10^{18}
10^{24}
YottaByte
The digital universe today: 250 trillion DVD's
data held by NSA….
GigaByte
10^{9}
10^{12}
10^{6}
MegaByte
TeraByte - every day 500 TB of data is added on Facebook
ExaByte
At the moment, every day 1 EB of data is created on the internet.
That is the equivalent of 250 million DVD's
The Square Kilomter Array Telescope will produce around 1 EB per day.
© - Big Data Startups

How much data is generated (daily, annually)?

Daily volume:
328 Exabytes



The Long Tail of Research Data



3. SciML

bringing it all together…

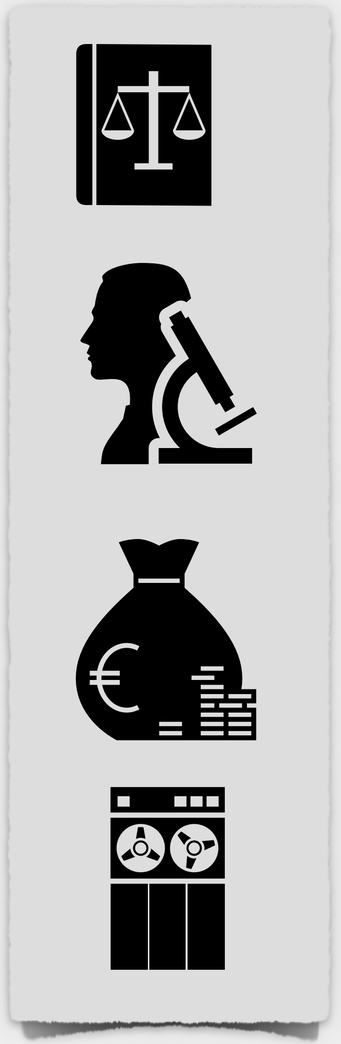
What's it all about?

1. Balancing CSE and ML.

2. Advancing science for society.

3. Budgets: cost vs. speed.

4.Needs computing capacity.

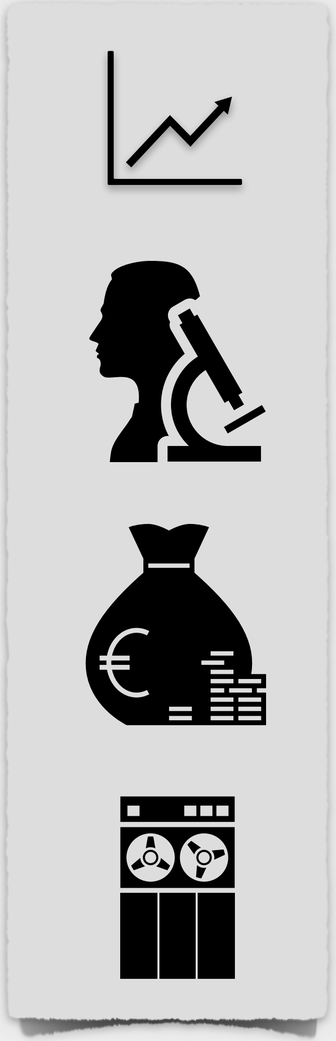


4. SciML Course Outline

What’s coming up next?

Outline of SciML course

1. Gradients, adjoints and optimization.
2. Differentiable programming.
3. ML methodologies
4. Direct and Inverse Problems.
5. Physics Constrained approaches.
6. Other approaches.
7. Strengths and weaknesses.



Thank you!

Questions/Discussion ?

You can find me at:

mark.asch@u-picardie.fr

Credits

- Presentation template by SlidesCarnival
- Photographs by Unsplash