ETH-RAT

SDSC

Machine Learning Applications at SDSC

Fernando Perez-Cruz

Chief Data Scientist at Swiss Data Science Center

(ETH Zurich and EPFL)

September 24th, 2018

Swiss Data Science Center

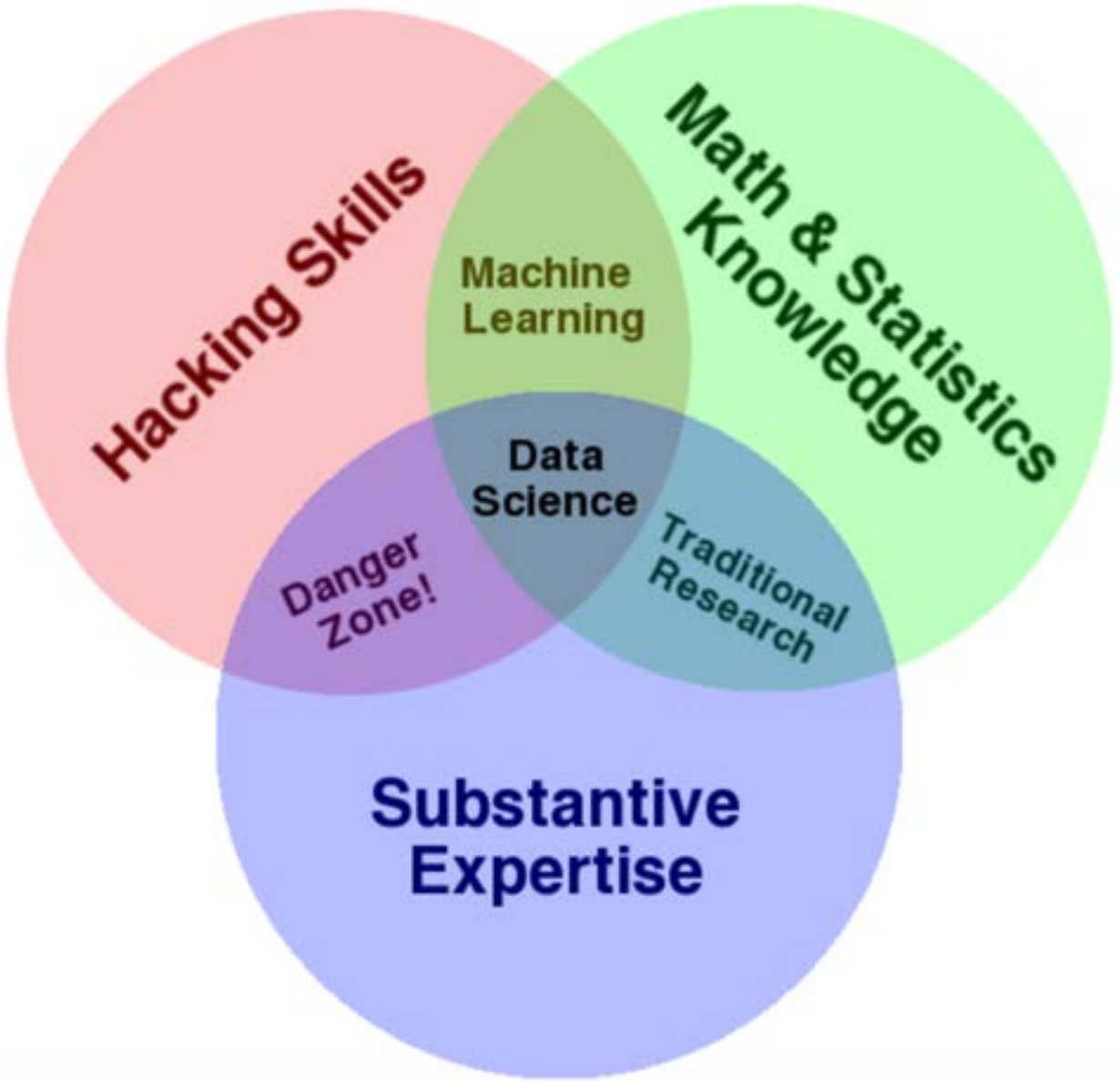
- SDSC was stablished in 2016 by the ETH board to:

  "Accelerate the adoption of Data Science in Industry and Academia"

- Why it needs accelerating?
- What are the burdens that impede widespread adoption?
- Why aren't we solving this faster?

SDSC

Conway's Data Science Venn Diagram



SDSC

NIPS Sponsor's list 2016

SDSC

SDSC Vision

- SDSC has a team of 20 Software developers and Data Scientists
- We have three lines of actuation:
  - Renku: Open platform for collaborative Data Science.
    - Foster collaboration
    - Open Science
    - Reproducibility
  - Teaching and Mentoring.
    - CAS in Data Science in Lausanne
    - DAS in Data Science in Zurich.
  - Research with labs and industry:
    - Climate Science.
    - Personalized Medicine
    - Physic
    - Architecture
    - …
    - Predictive maintenance (Buhler and PSA).

SDSC

ETH-RAT

SDSC

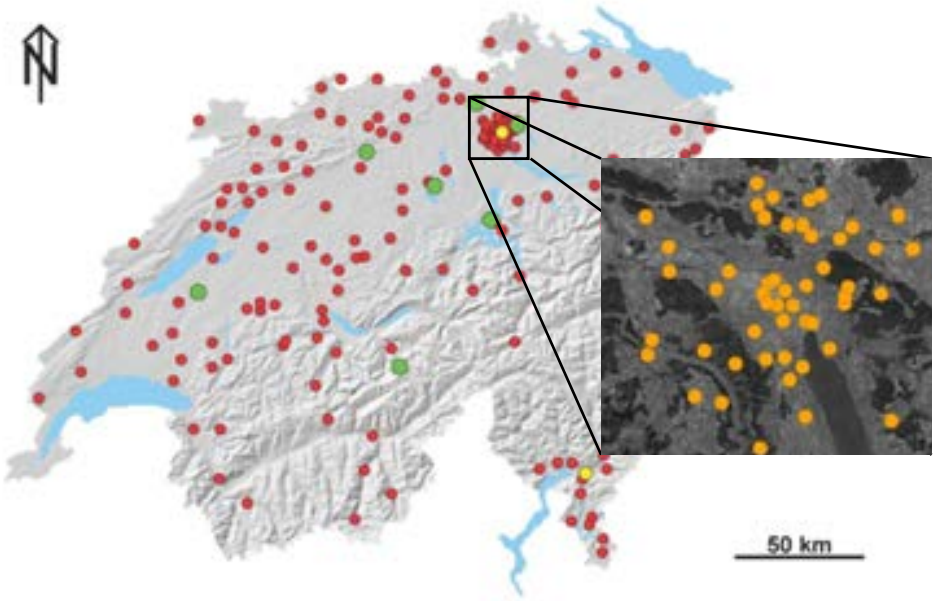
Carbosense4D

w Dominik Brunner

Empa

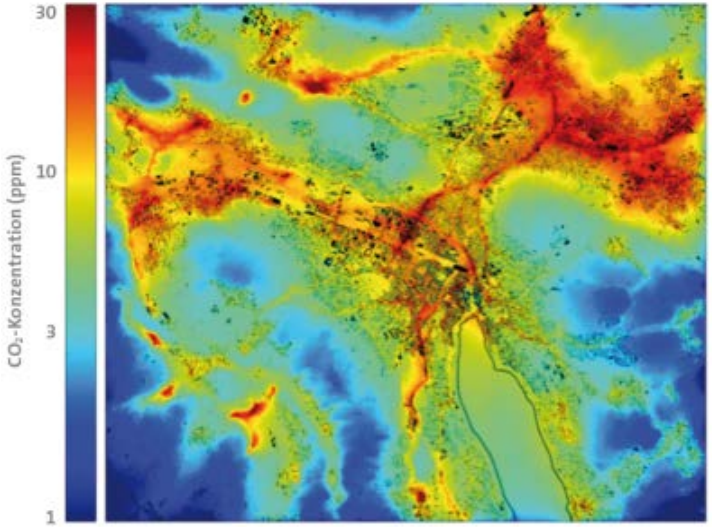
Goals

- Determine CO_2 emissions over multiple years.
- Enhance understanding of biospheric CO_2 fluxes over Switzerland.
- Describe accurately the 4-D evolution of CO_2 over Switzerland.

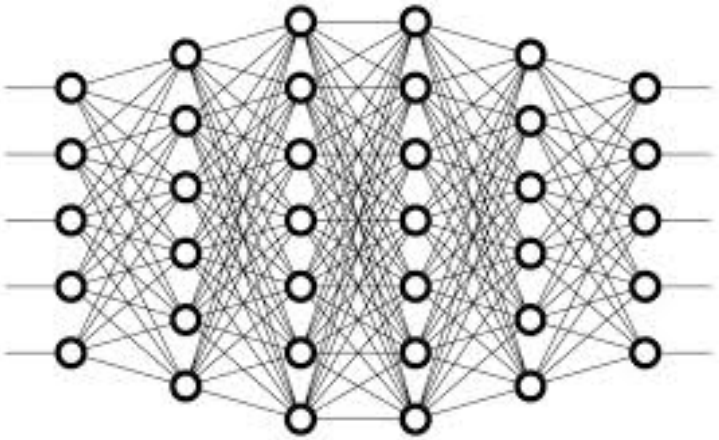


SDSC

Carbosense4D Sensors

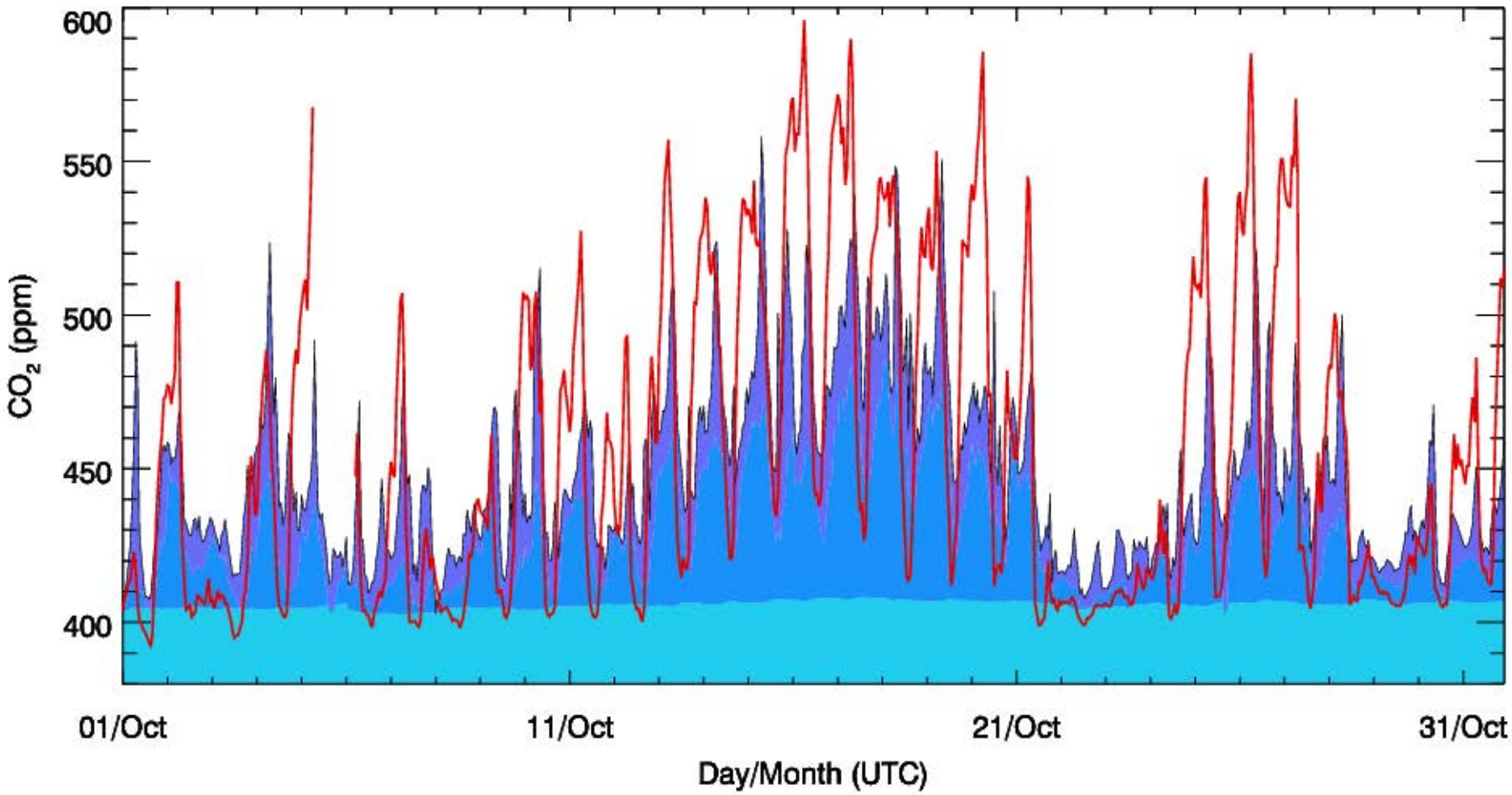


Machine Learning



SDSC

Observations versus model



SDSC

ETH-RAT

SDSC

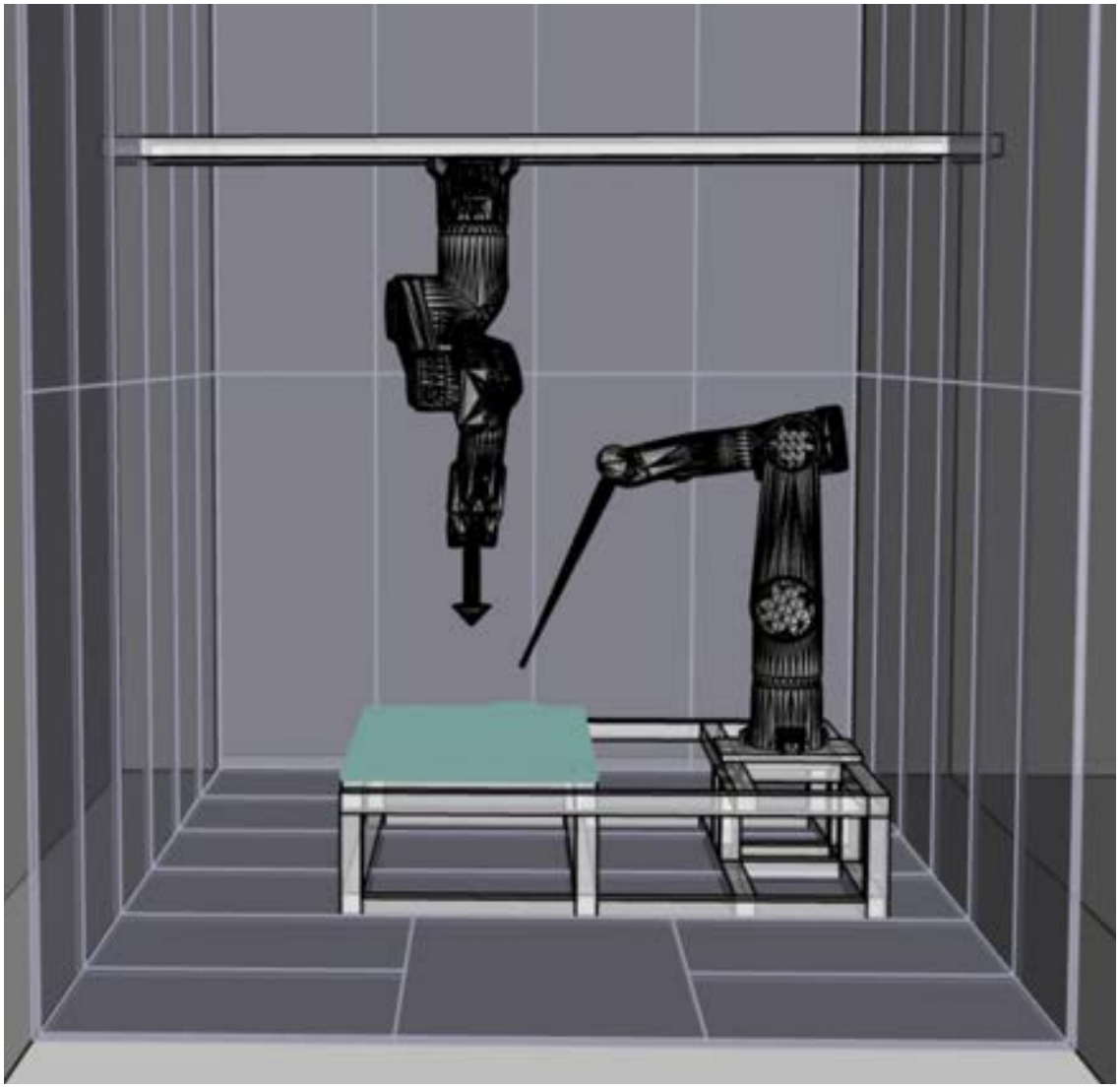
Data Driven Acoustical Design

w Matthias Kohler and Fabio Gramazio

ETHZ - Architecture

Goals

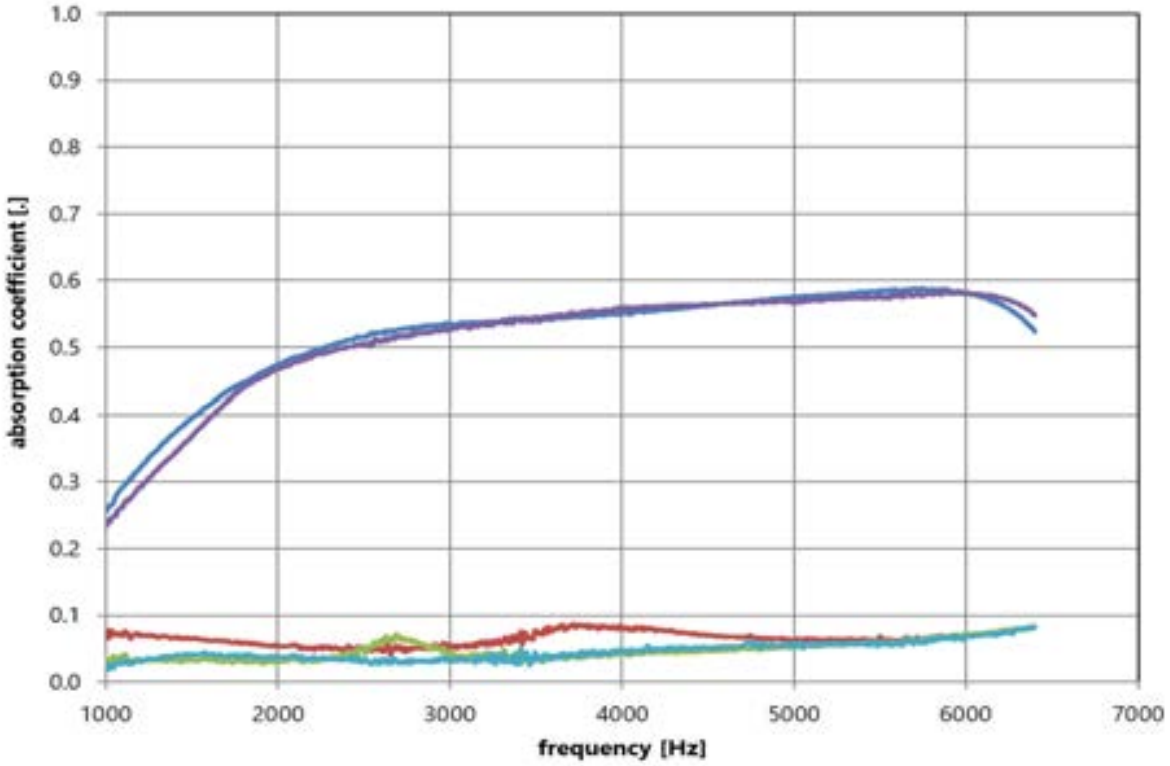
- Model sound propagation in any wall.
- Use those model to simulate realistic sounds conditions in any room design.
- Architecture for everyday use that is sound resilient.
- Bonus: Improved sound quality of VR environments.



SDSC

Set up

- 3D print a wall per week 1m x 0.6 m x 0.5m (1:10).
- Measure sound response in over 8 thousand microphone-speaker pairs.
- Every measurement takes over 11s (all just over a day).



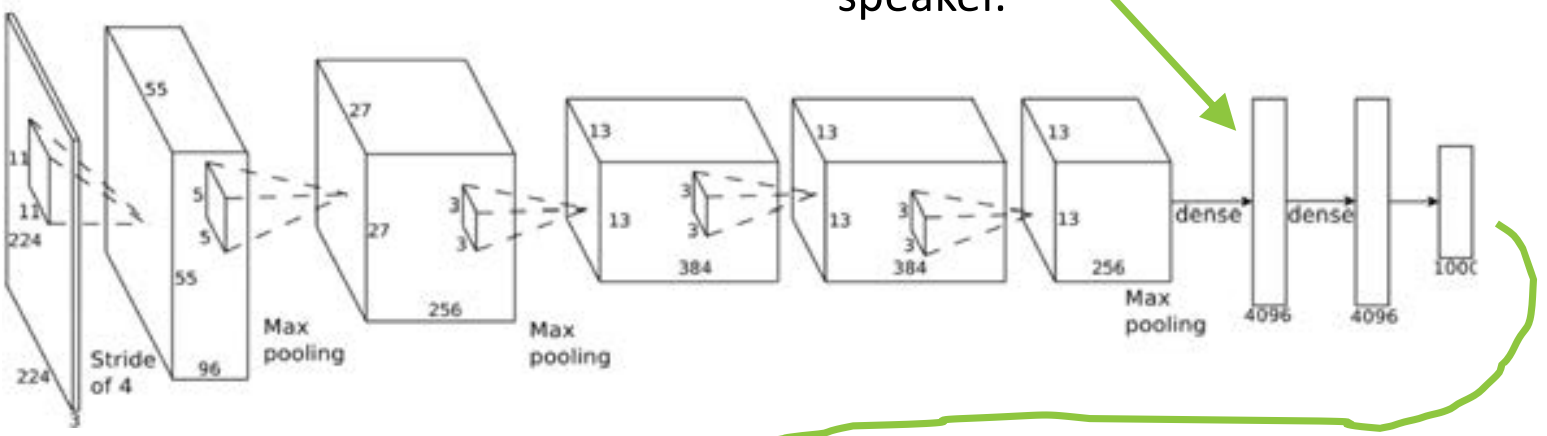
SDSC

Convolutional Neural Network

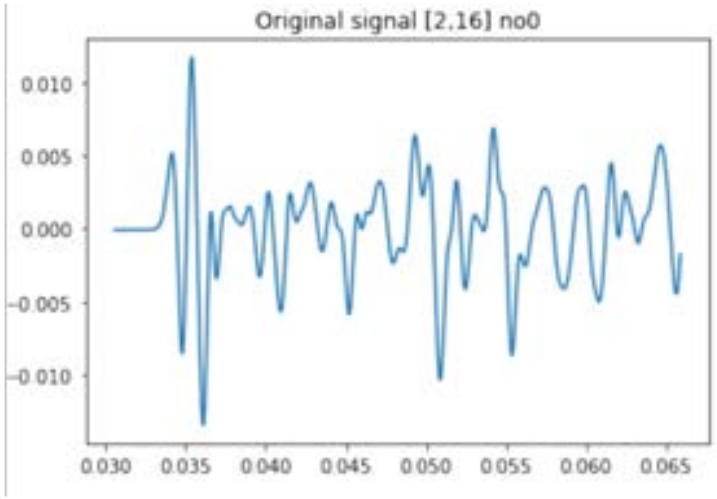
Wall tensor

$100 \times 60 \times 10$

Position of the microphone and speaker.



From the wall, we predict the signal that would be received on each microphone



SDSC

ETH-RAT

SDSC

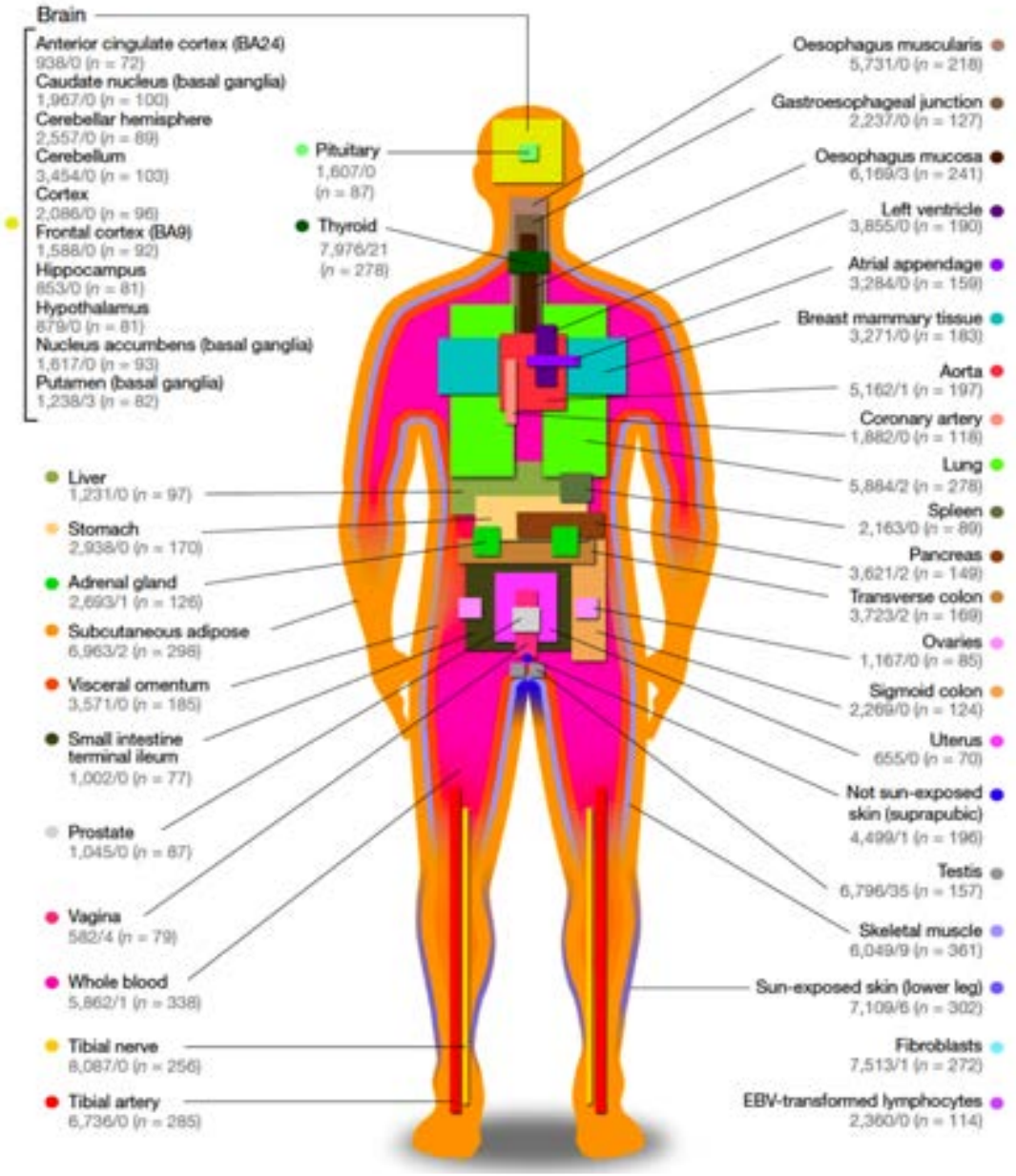
T4Med

w Didier Trono

EPFL – Life Sciences

Goals

- Can we predict from a tissue gene and TE expression data where it is coming from?
- If so, we might be able to detect metastasis.
- Can we also detect sub-cancer types?
- Data:
  - 9000+ samples from 549 cadavers.
  - 54 body sites.
  - Genes and TE expression data.

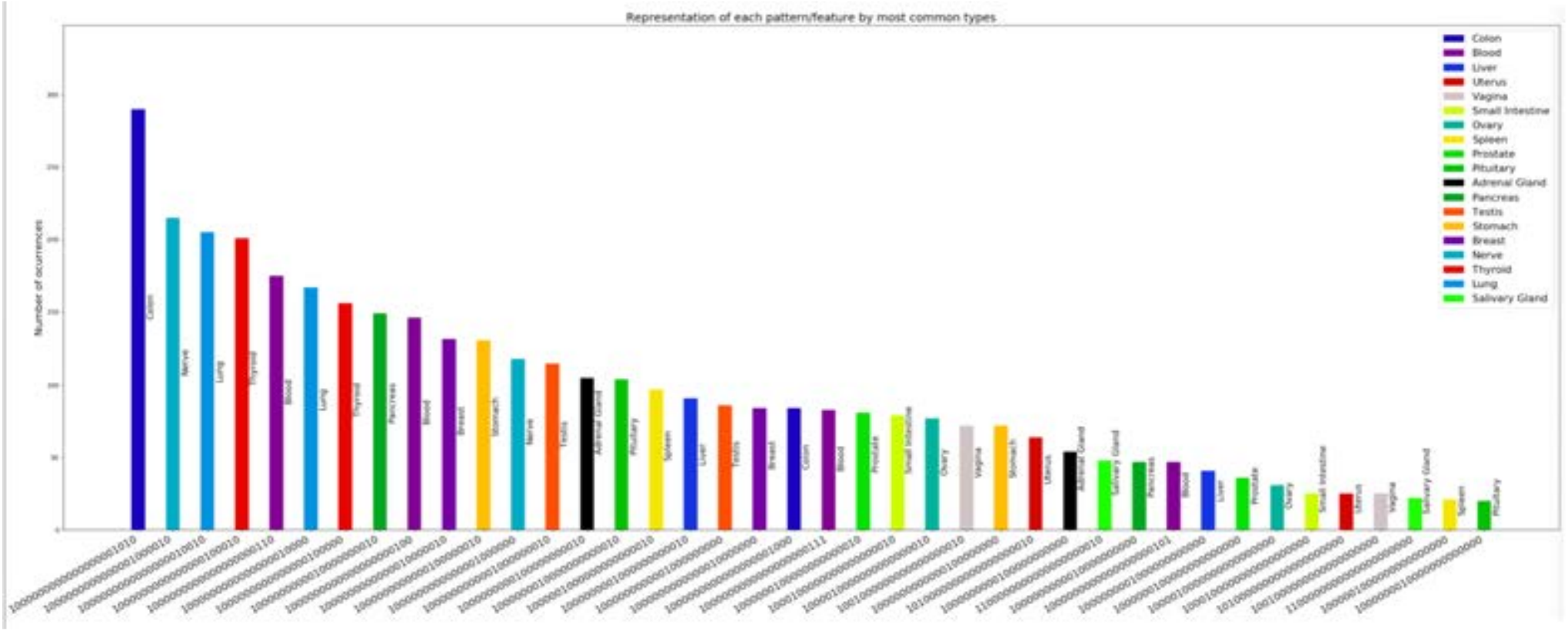


SDSC

Transposable Elements

- TEs are typically referred as junk DNA.
- TEs represent 80% of the human genome.
- They are repetitive and it is thought to be left-over material from the last common ancestor or viruses.
- TEs have regulatory and active in the expression of each cell tissue.
- We are using TEs:
  - Detect tissue type.
  - Detect cancer and sub-cancer type.
  - Detect metastasis.

Promising results



SDSC

ETH-RAT

SDSC

Questions?