Intelligence artificielle pour l'astrophysique à l'époque du big-data

Marc Huertas-Company

l'Observatoire de Paris | LERMA
université PARIS DIDEROT
institut universitaire de France
INSTITUTO DE ASTROFISICA DE CANARIAS · IAC ·

PRACTICAL INFO

FULL ONLINE COURSE VIA ZOOM:

https://rediris.zoom.us/j/99594075110?pwd=Y0dFcTVZWFFqb0xUQm41NWpweWs5dz09

Meeting ID: 995 9407 5110
Password: 842832

MORNING ~[9h30-12h30]: THEORY / LECTURES

AFTERNOON ~[14H30-17H30]: TUTORIALS

PRACTICAL INFO

FULL ONLINE COURSE VIA ZOOM:

https://rediris.zoom.us/j/99594075110?pwd=Y0dFcTVZWFFqb0xUQm41NWpweWs5dz09

Meeting ID: 995 9407 5110
Password: 842832

SOME RULES DURING THE ZOOM MEETING:

Keep your Microphones Muted

If need to ask questions (which I recommend!) use the “raise hand” option

PRACTICAL INFO

OTHER AVAILABLE RESSOURCES:

GITHUB REPO WITH SLIDES AND TUTORIALS:

https://github.com/mhuertascompany/DL_ED127_2020

Slack Channel for interactive discussions / questions:

https://join.slack.com/t/iaed127/shared_invite/zt-f5sohajk-uP00KQtWdaPIC_Znw6H2EA

SOME PRELIMINARY NOTES

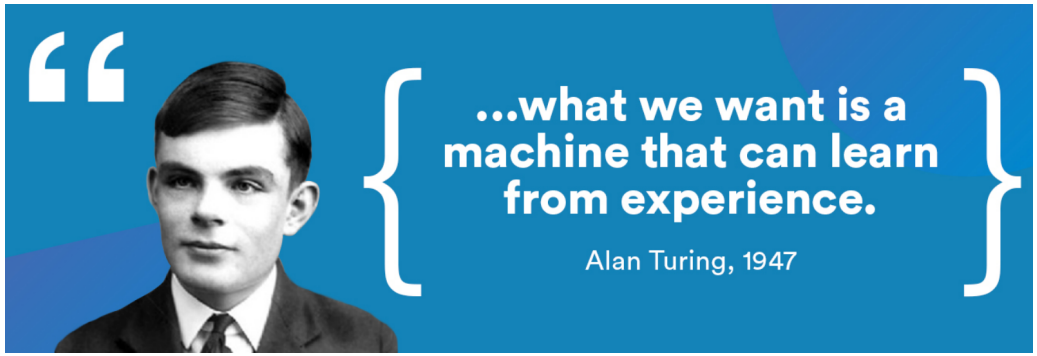
I AM NOT A MACHINE LEARNING RESEARCHER

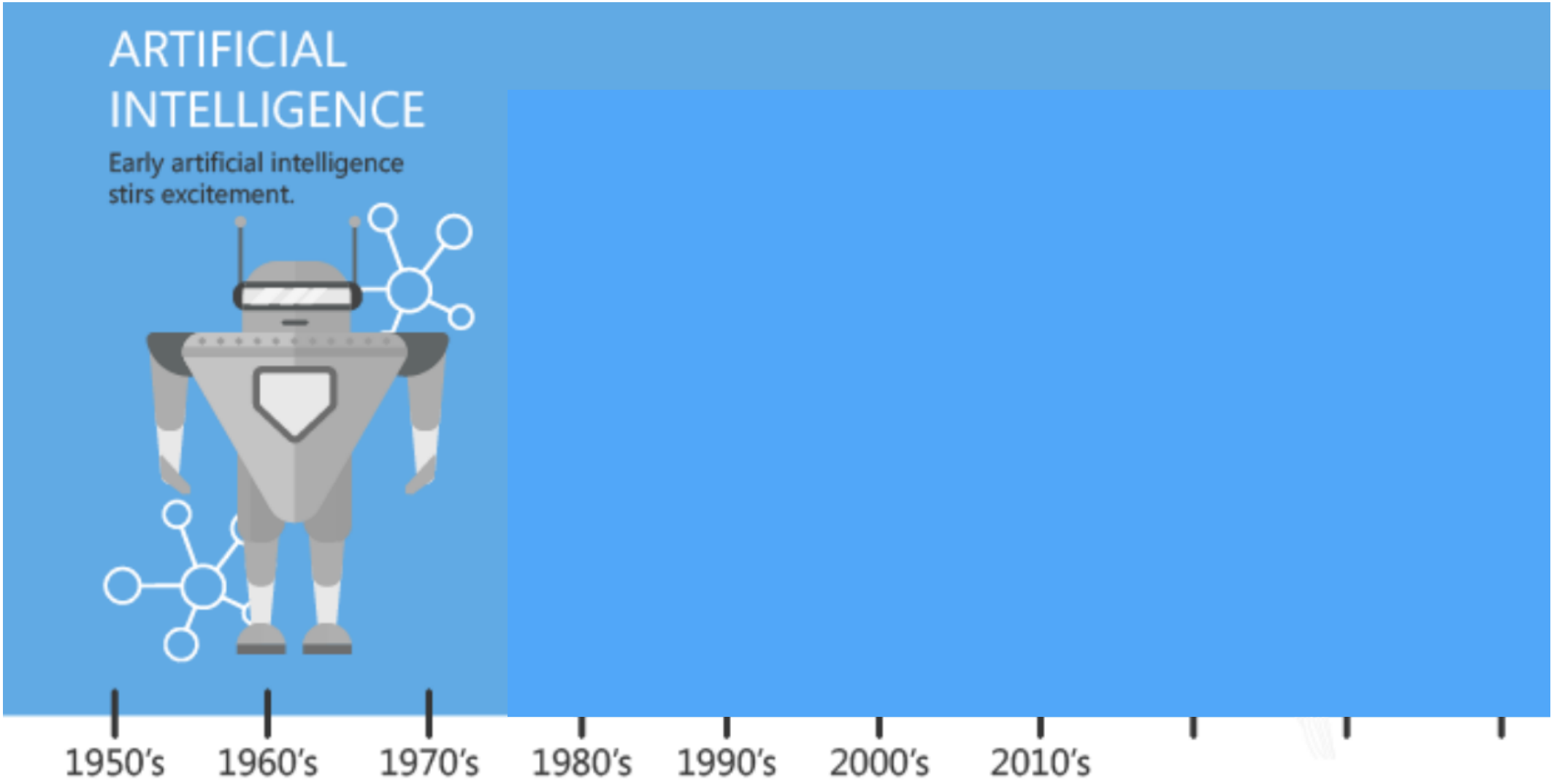
SOME PRELIMINARY NOTES

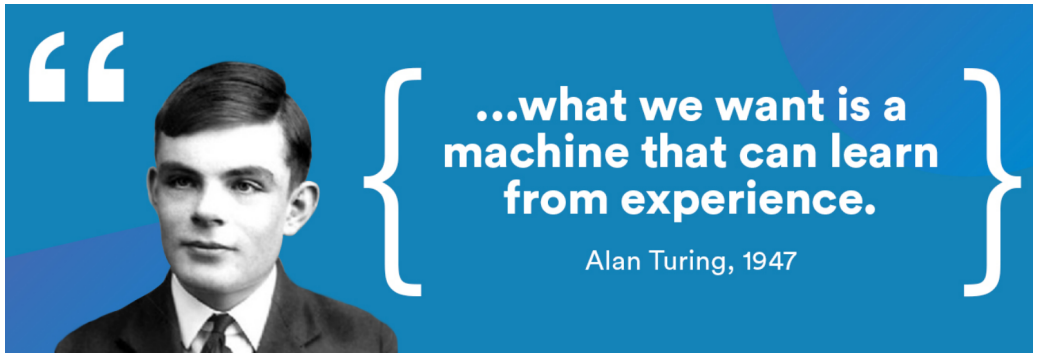
I AM NOT A MACHINE LEARNING RESEARCHER

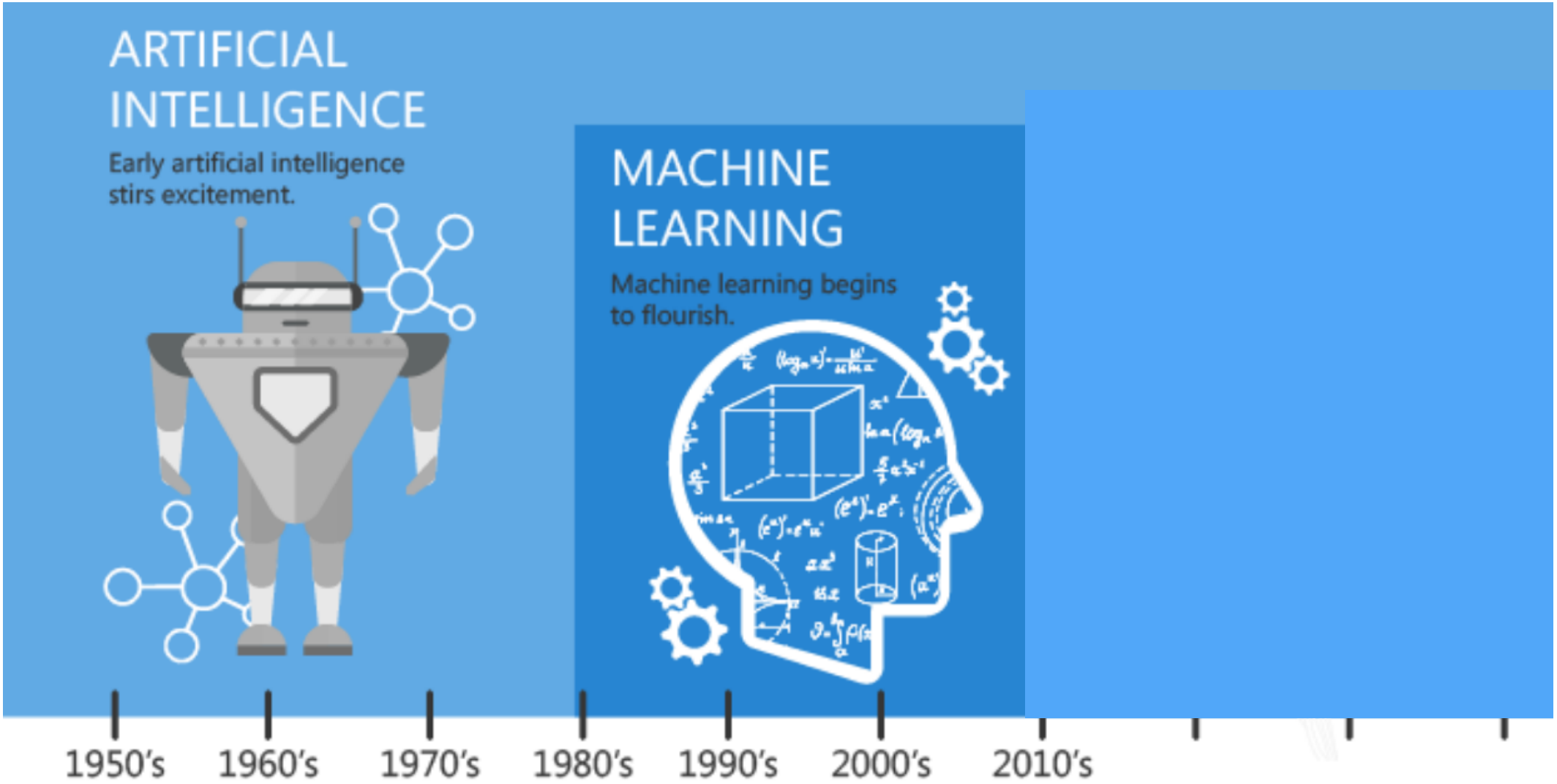
ONLY AN ASTRONOMER WHO HAS BEEN USING MACHINE LEARNING FOR THE LAST ~14 YEARS FOR MY RESEARCH

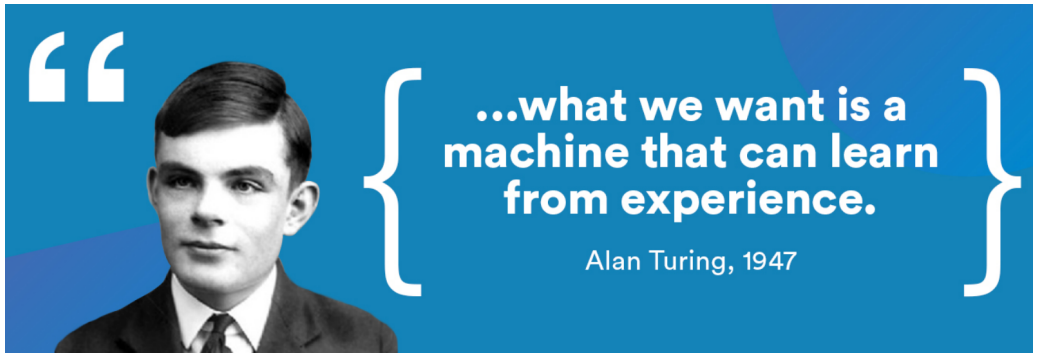
THESE LECTURES ARE INTENDED TO PROVIDE A **GLOBAL** UNDERSTANDING OF HOW AI TECHNIQUES WORK AND ESPECIALLY **HOW TO USE THEM FOR YOUR RESEARCH**

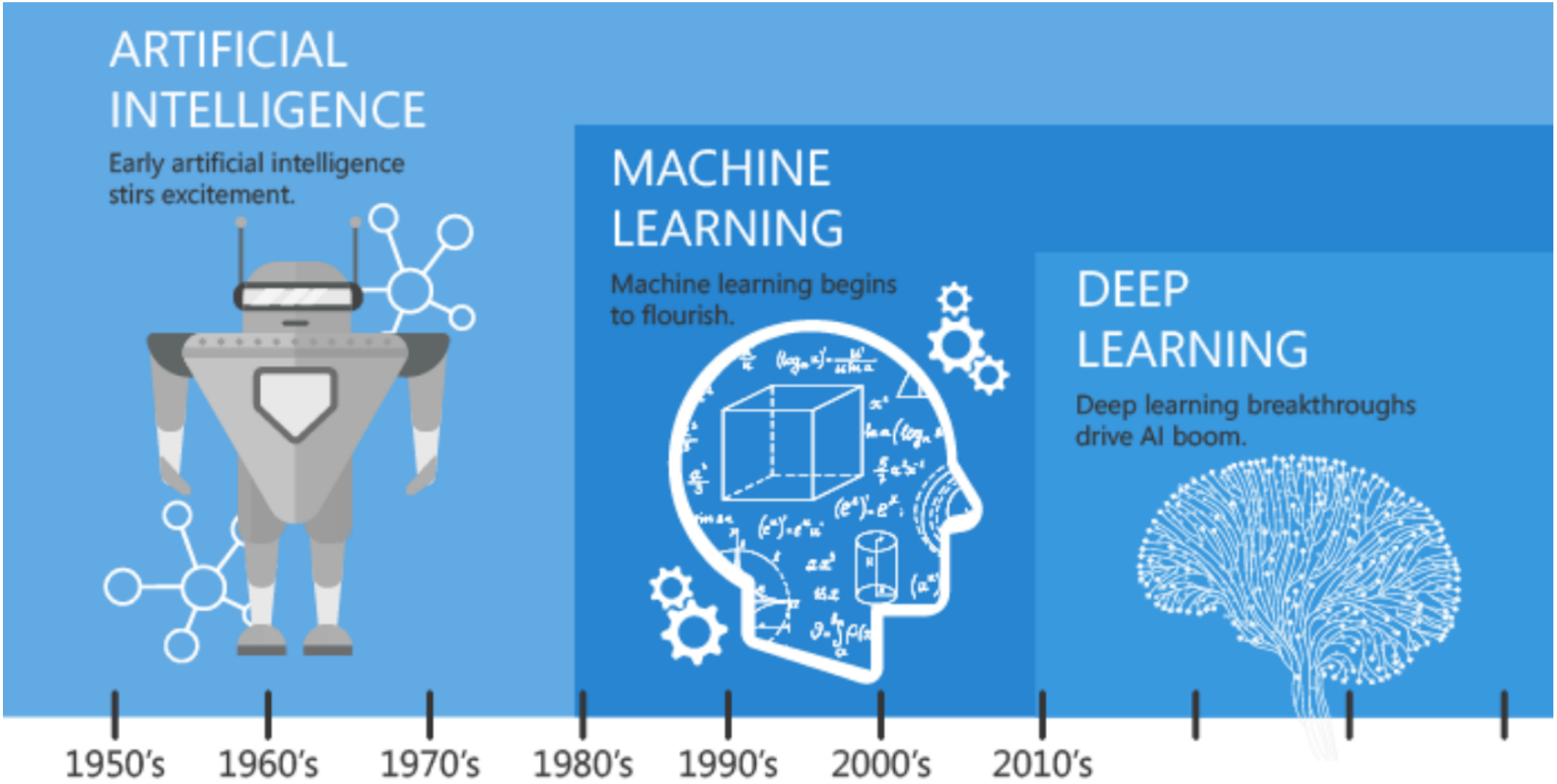

...what we want is a machine that can learn from experience.
Alan Turing, 1947


ARTIFICIAL INTELLIGENCE
Early artificial intelligence stirs excitement.
1950's
1960's
1970's
1980's
1990's
2000's
2010's


...what we want is a machine that can learn from experience.
Alan Turing, 1947


ARTIFICIAL INTELLIGENCE
Early artificial intelligence stirs excitement.
MACHINE LEARNING
Machine learning begins to flourish.
1950's
1960's
1970's
1980's
1990's
2000's
2010's


...what we want is a machine that can learn from experience.
Alan Turing, 1947


ARTIFICIAL INTELLIGENCE
Early artificial intelligence stirs excitement.
MACHINE LEARNING
Machine learning begins to flourish.
DEEP LEARNING
Deep learning breakthroughs drive AI boom.
1950's
1960's
1970's
1980's
1990's
2000's
2010's

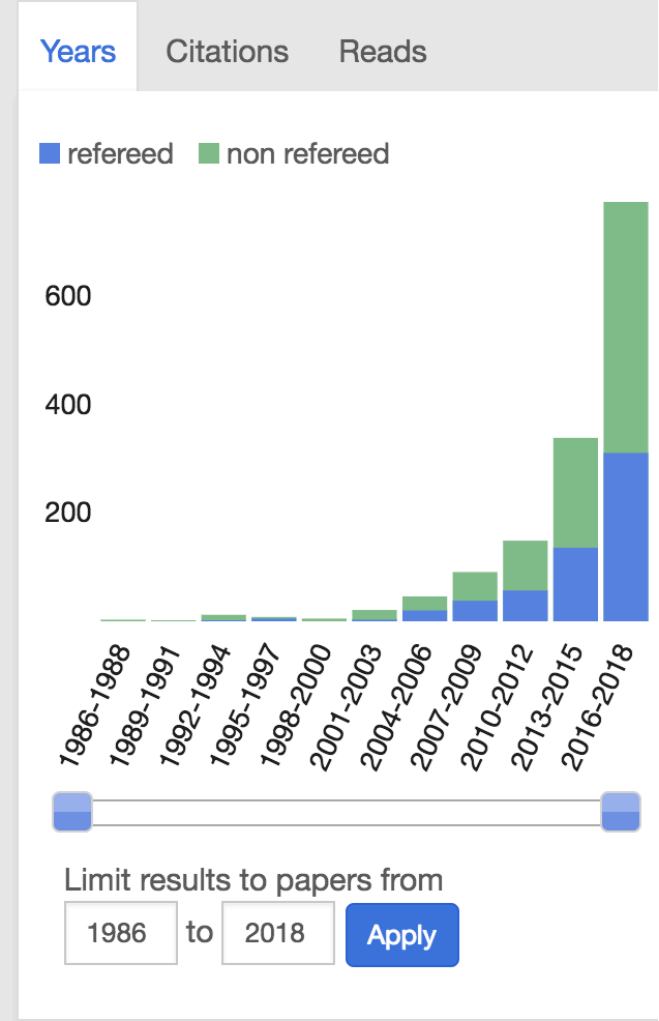
AI FEVER?

AN AMAZING MEDIA ATTENTION

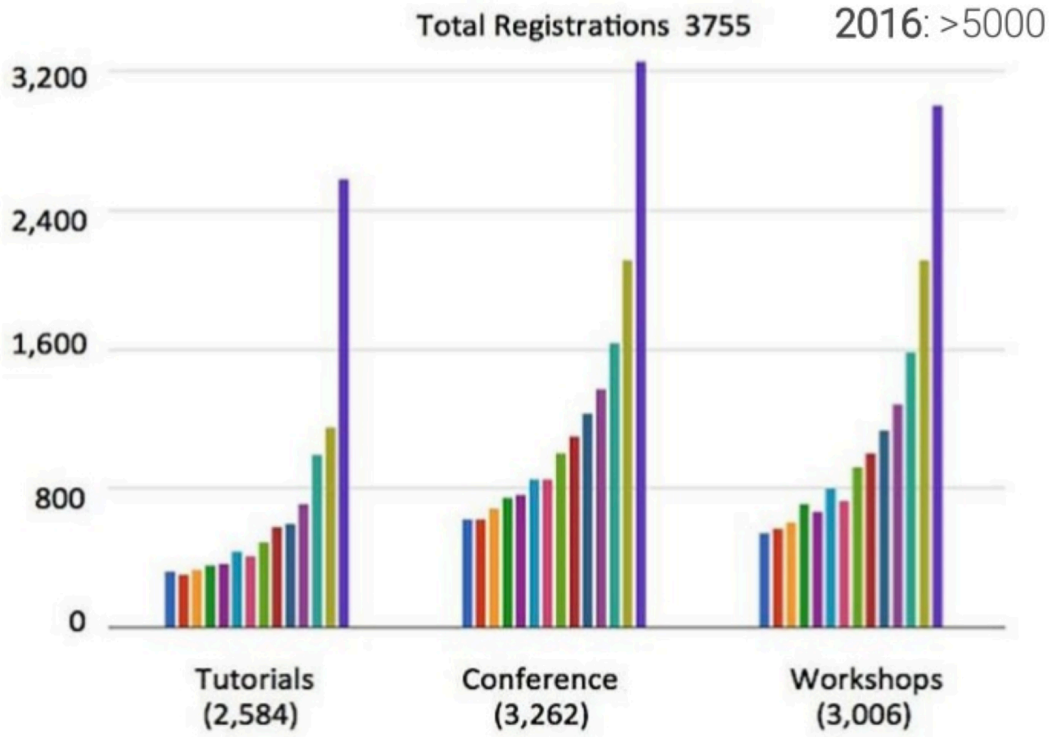
Le CNRS, Inria, l'université PSL et les entreprises Amazon, Criteo, Facebook, Faurecia, Google, Microsoft, NAVER LABS, Nokia Bell Labs, le Groupe PSA, SUEZ et Valeo font converger intérêts académiques et industriels et s'unissent pour créer, à Paris, l'Institut PRAIRIE dont l'objectif est de devenir une référence internationale de l'intelligence artificielle.

AI FEVER?

PUBLICATIONS (ADS)

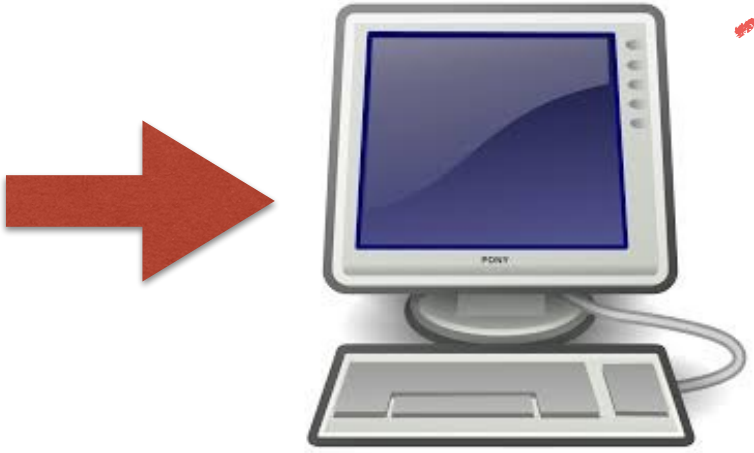


CONFERENCES



Source

BEFORE 2012....

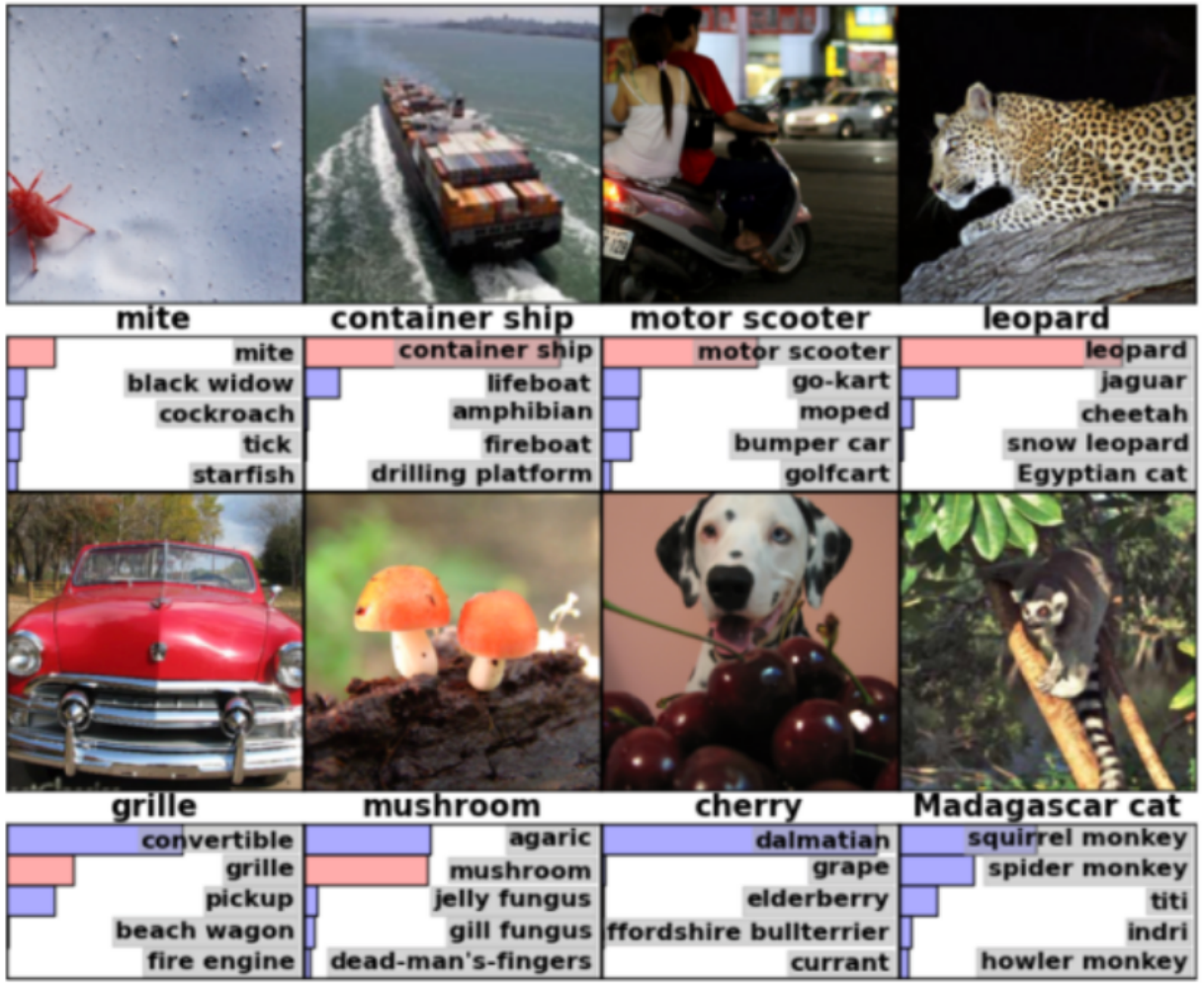


CAT?

DOG?

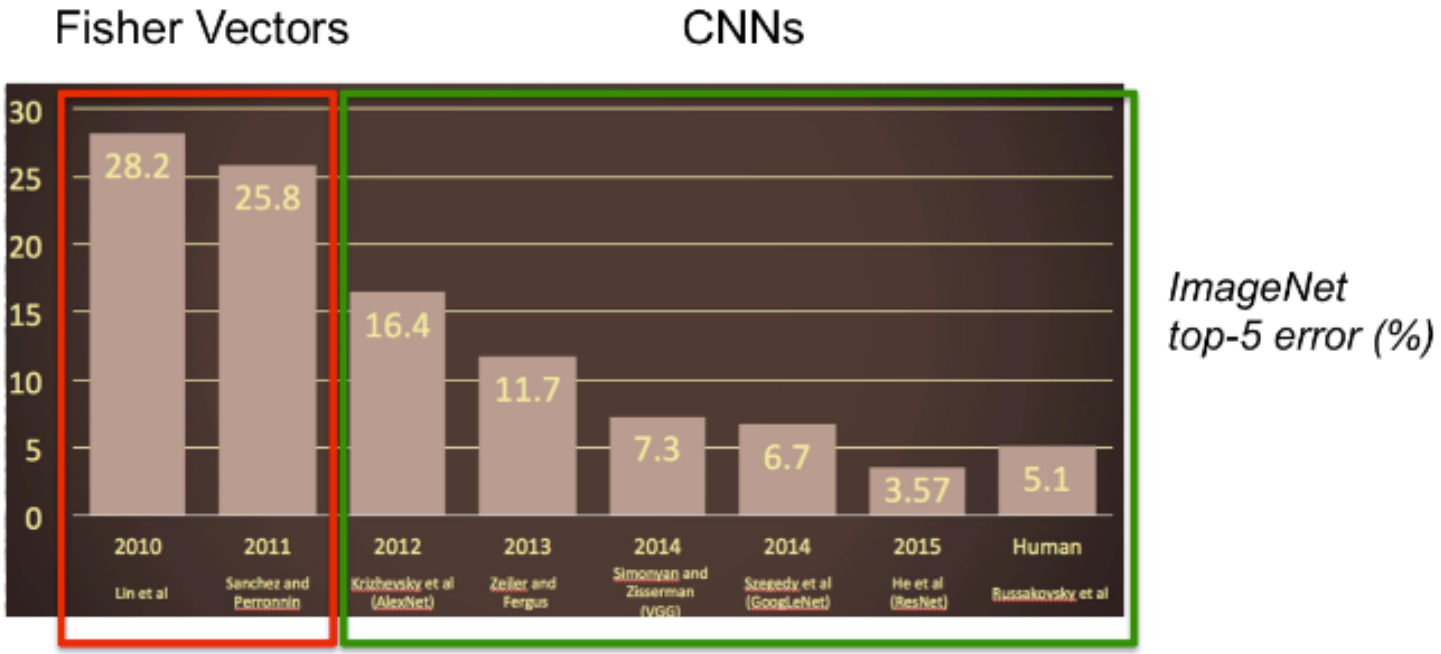
TRIVIAL HUMAN TASKS REMAINED CHALLENGING FOR COMPUTERS

AFTER 2012



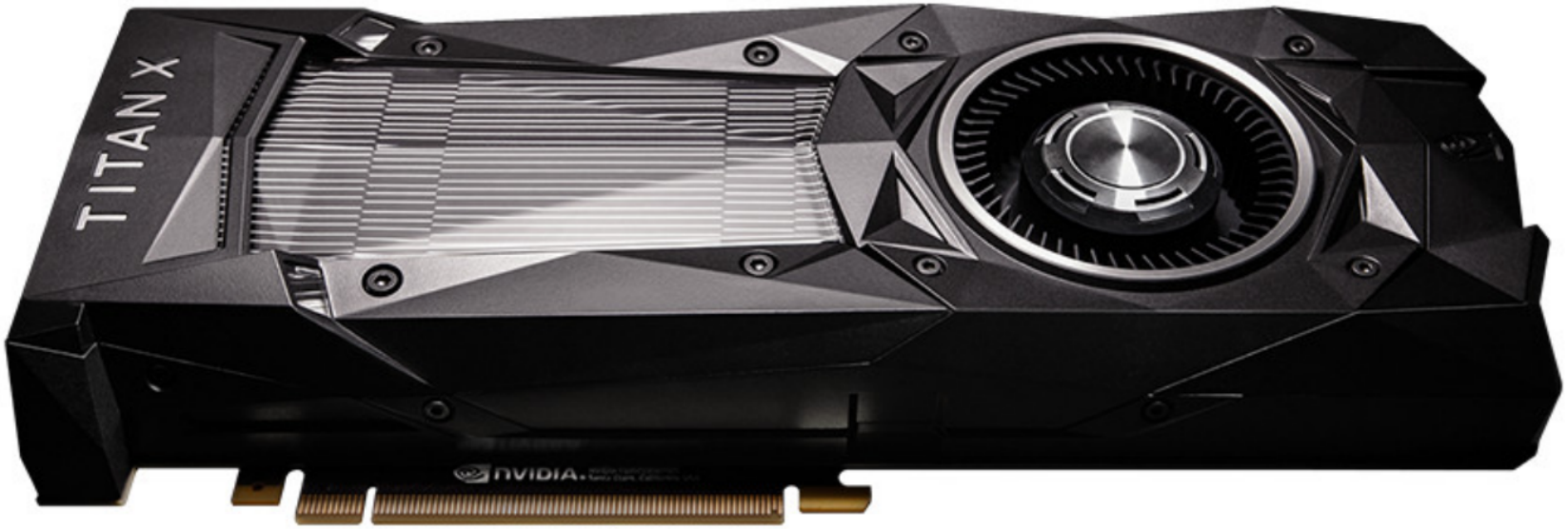
IT HAS BECOME TRIVIAL….

THIS IS A CHANGE OF PARADIGM!



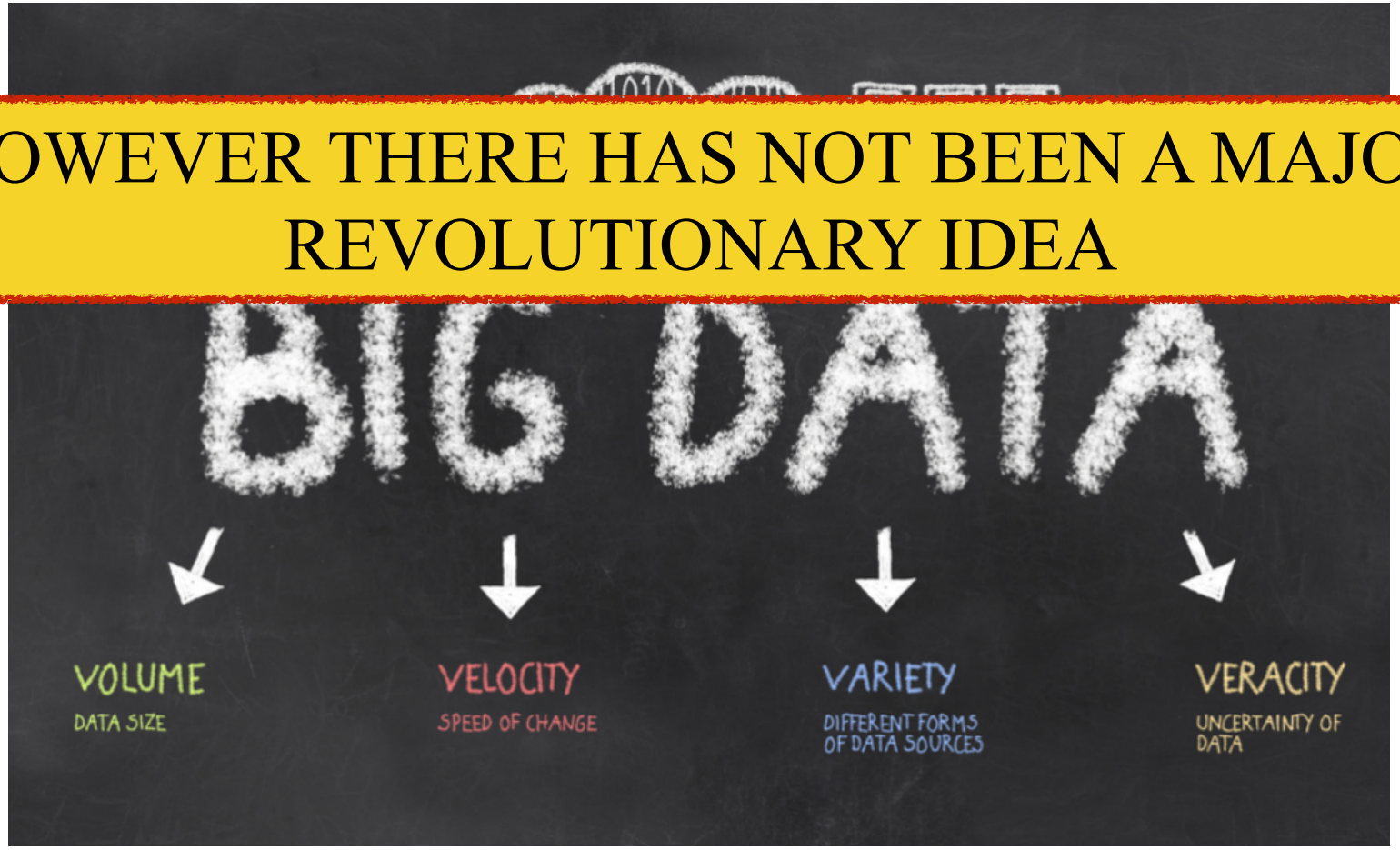
ONE OF THE MAIN REASONS OF THIS BREAKTHROUGH IS THE AVAILABILITY OF VERY LARGE DATASETS TO LEARN

COMBINED WITH THE TECHNOLOGY TO PROCESS ALL THIS DATA



ONE OF THE MAIN REASONS OF THIS BREAKTHROUGH IS THE AVAILABILITY OF VERY LARGE DATASETS TO LEARN

HOWEVER THERE HAS NOT BEEN A MAJOR REVOLUTIONARY IDEA



WHAT IS THESE SERIES OF LECTURES ABOUT?

BASICS OF CLASSICAL MACHINE LEARNING

BASICS OF DEEP LEARNING
(BOTH SUPERVISED AND UNSUPERVISED)

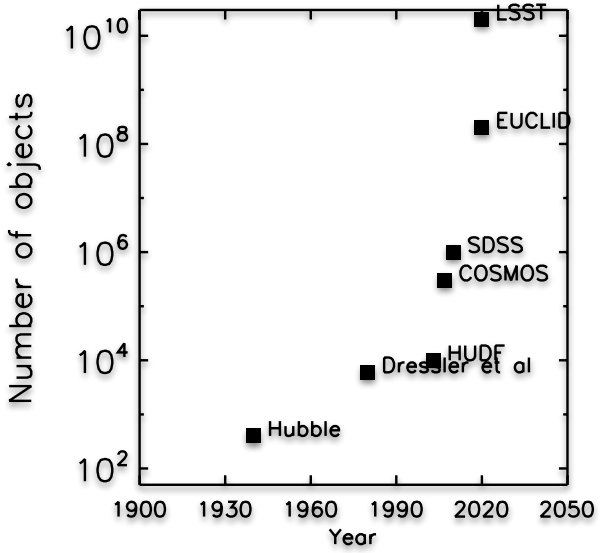
HOPING THAT THIS WOULD BE USEFUL FOR YOUR RESEARCH!

(Apologies in advance for biases on Extra-Galactic Science + imaging)

WHY DO WE NEED THESE TOOLS IN ASTRONOMY?

WHY DO WE NEED THESE TOOLS IN ASTRONOMY?

AS IN MANY OTHER DISCIPLINES THE BIG-DATA REVOLUTION HAS ARRIVED TO ASTRONOMY TOO



← BIG-DATA REVOLUTION

← we are here

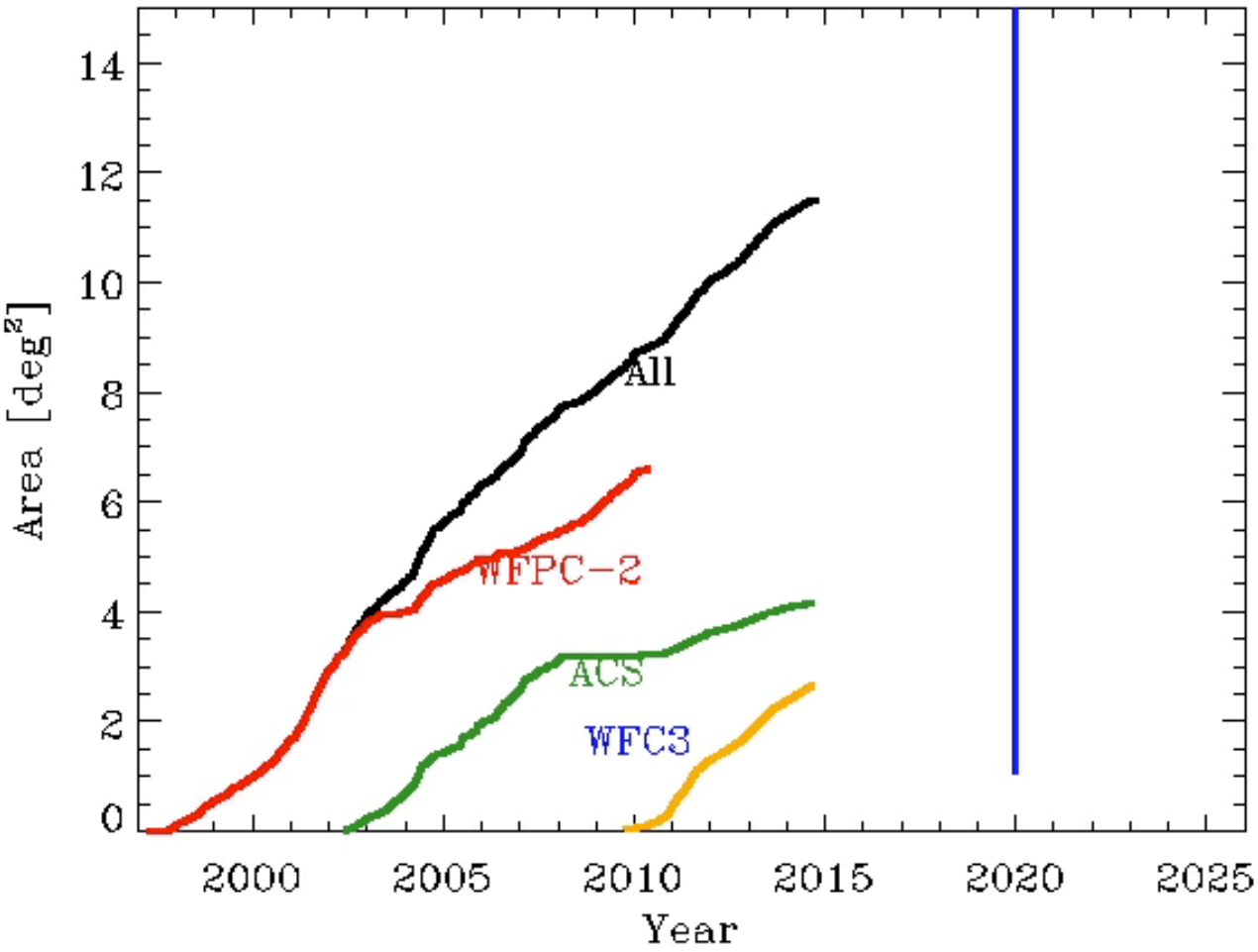
EXTREMELY LARGE IMAGING SURVEYS DELIVERING BILLIONS OF OBJECTS IN 2-5 YEARS

Large Synoptic Survey Telescope

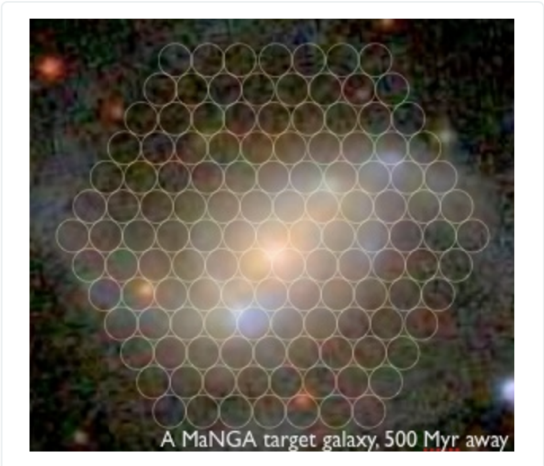
euclid consortium

WIDE-FIELD INFRARED SURVEY TELESCOPE
WFIRST
ASTROPHYSICS • DARK ENERGY • EXOPLANETS

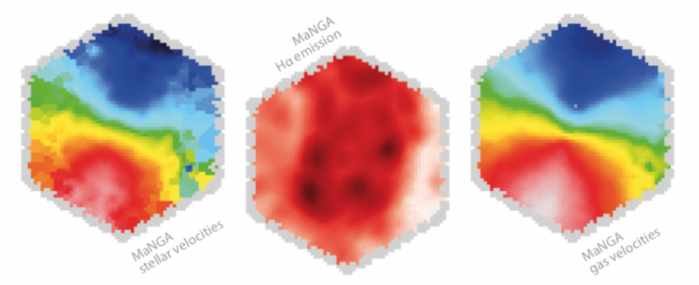
LSST simulation



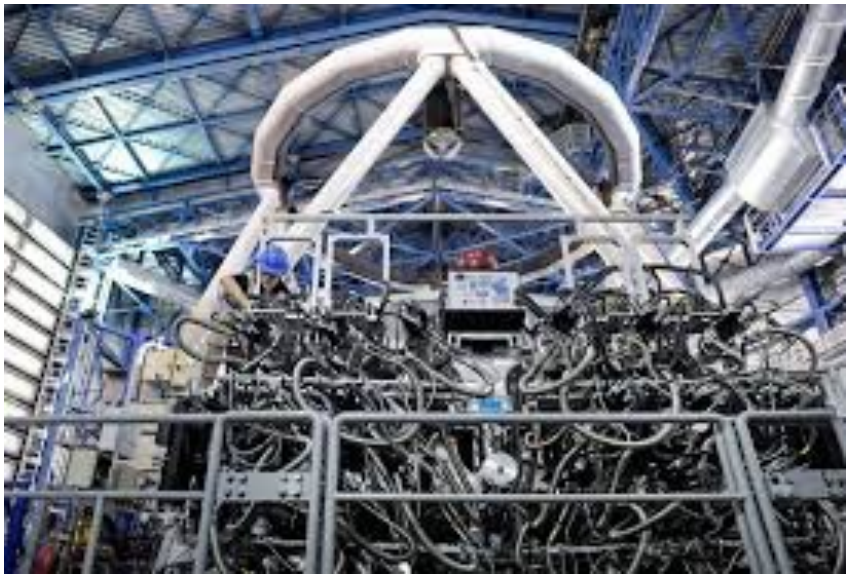
(Thanks to J. Brinchmann)

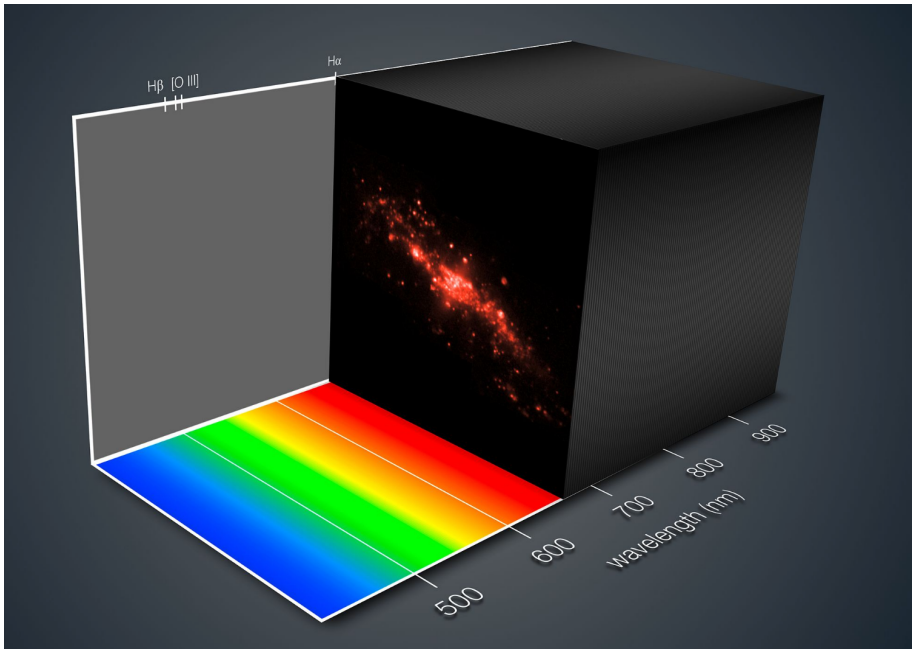


NOT ONLY VOLUME: AN INCREASING COMPLEXITY OF DATA

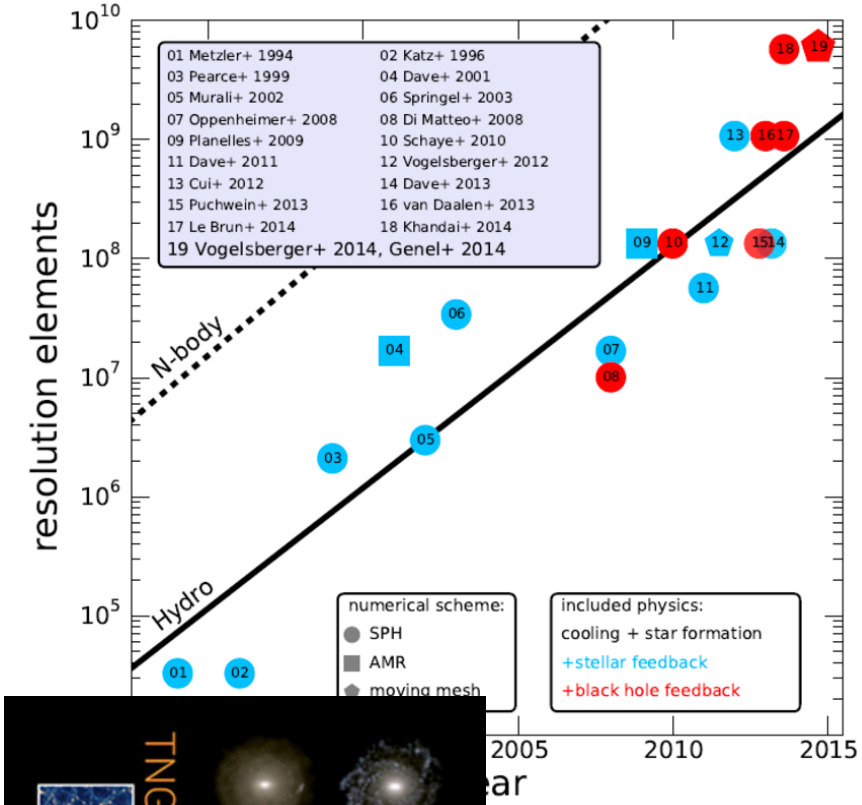


MANGA Survey



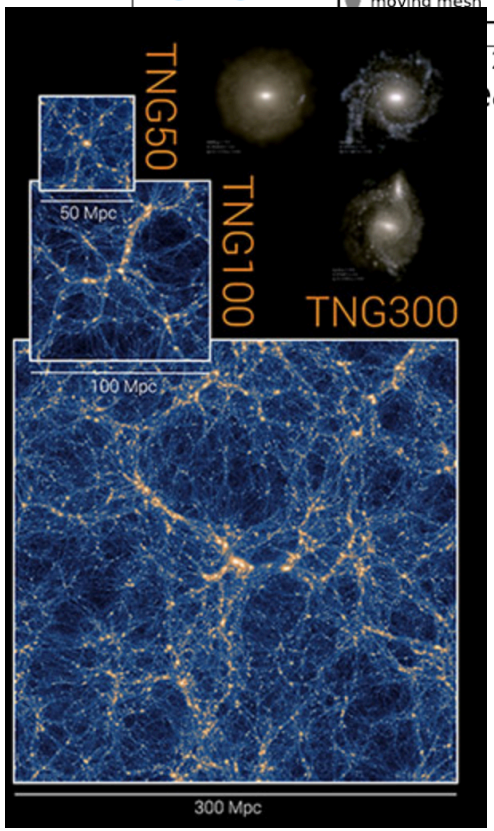


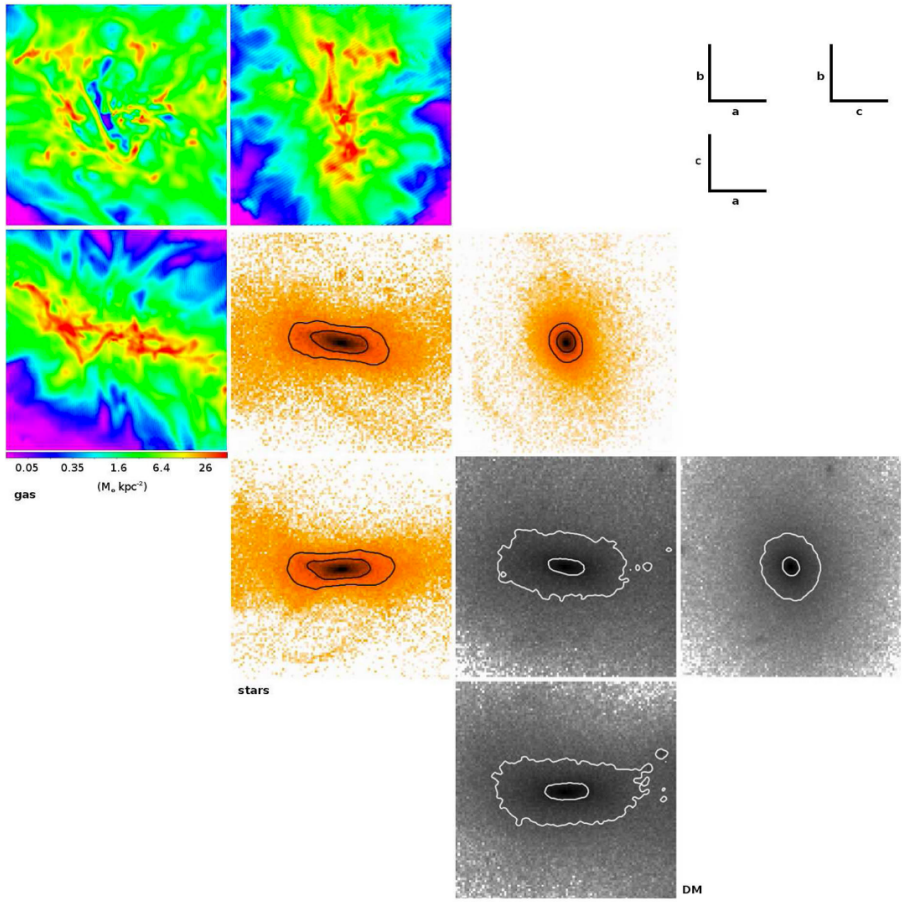
MUSE@VLT



Genel+14

AND ALSO SIMULATIONS!





Ceverino+15

PROGRAM FOR THE WEEK

- **PART I: A VERY QUICK INTRODUCTION TO 'CLASSICAL' SUPERVISED MACHINE LEARNING**
  - UNSUPERVISED / SUPERVISED
  - GENERAL STEPS TO "TEACH A MACHINE"
  - "CLASSICAL" CLASSIFIERS: RFs, KERNEL MACHINES

PROGRAM FOR THE WEEK

- **PART II: FOUNDATIONS OF 'SHALLOW' NEURAL NETWORKS**
  - PERECPTRON, NEURON DEFINITION
  - LAYER OF NEURONS, HIDDEN LAYERS
  - ACTIVATION FUNCTIONS
  - OPTIMIZATION [GRADIENT DESCENT, LEARNING RATES]
  - BACKPROPAGATION
  - LOSS FUNCTIONS

PROGRAM FOR THE WEEK

- **PART III: CONVOLUTIONAL NEURAL NETWORKS**
  - CONVOLUTIONS AS NEURONS
  - CNNs [POOLING, DROPOUT]
  - VANISHING GRADIENT / BATCH NORMALIZATION
  - CNN VISUALIZATION

PROGRAM FOR THE WEEK

- **PART IV: BEYOND CLASSIFICATION WITH CNNs**
  - FCNNs (IMAGE2IMAGE NETWORKS)
  - OBJECT DETECTION
  - INSTANCE AND SEMANTIC SEGMENTATION

PROGRAM FOR THE WEEK

- **PART V: INTRODUCTION TO UNSUPERVISED MACHINE (DEEP) LEARNING**
  - CLUSTERING ALGORITHMS
  - DEEP GENERATIVE MODELS
  - DEEP PROBABILISTIC MODELS

PROGRAM FOR THE WEEK

- **PART VI: SOME PRACTICAL CONSIDERATIONS**
  - HOW DO I SETUP MY CNN?
  - HOW LARGE DO TRAINING SETS NEED TO BE?
  - OPTIMIZING YOUR NET: HYPER PARAMETER SEARCH

HANDS-ON SESSIONS

WE WILL TRY TO PRACTICALLY IMPLEMENT SOME OF THE NOTIONS LEARNED

TUTORIALS WILL USE **GOOGLE COLAB**. YOU WILL NEED A GOOGLE ACCOUNT AND SOME SPACE IN YOUR GDRIVE. WORKS ALSO BETTER ON CHROME.

NOTEBOOKS ARE GENERALLY SELF EXPLANATORY. YOU WILL BE ASKED TO FILL PIECES OF CODE AT DIFFERENT STAGES. SOLUTIONS ARE ALSO PROVIDED.

LET'S TRY TO DISCUSS AS MUCH AS POSSIBLE!

HANDS-ON SESSIONS

LET'S TRY TO DISCUSS AS MUCH AS POSSIBLE EVEN IF VIRTUALLY!

THERE WILL BE VIRTUAL ROOMS IN ZOOM WITH A REDUCED NUMBER OF STUDENTS (3-4). I WILL GO THROUGH THE ROOMS DURING THE TUTORIAL SESSIONS SO THAT COMMUNICAITON IS MORE FOCUSSED AND PERSONALIZED

AT THE END OF EVERY SESSION - IF TIME ALLOWS - ONE GROUP WILL BE ASKED TO SUMMARIZE THEIR FINDINGS, IMPRESSIONS ETC **VERY INFORMAL!**

REFERENCES

SEVERAL SLIDES / INFOS SHOWN HERE ARE INSPIRED/ TAKEN FROM OTHER WORKS / COURSES FOUND ONLINE

- Deep Learning: Do-It-Yourself! [Bursuc, Krzakala, Lelarge]
- DEEPLEARNING.AI [COURSERA, Ng, Bensouda, Katanforoosh]
- MACHINE LEARNING LECTURES [Keck]
- EPFL DEEP LEARNING COURSE [Fleuret]
- DEEPMIND / UCL Lecture series [DeepMind]

+ many others!

Thanks to all of them!