

INTELLIGENCE ARTIFICIELLE POUR L'ASTROPHYSIQUE À L'ÉPOQUE DU BIG-DATA

Marc Huertas-Company
IAC, OBSPM



institut
universitaire
de France



PRACTICAL INFO

5-DAY COURSE

10h-17h00?

MORNING ~[10h-12h30]: THEORY / LECTURES

AFTERNOON ~[14H-17H]: TUTORIALS

PRACTICAL INFO

5-DAY COURSE

ASSISTANCE IS MANDATORY

PLEASE LET ME KNOW IN ADVANCE IF YOU
NEED TO SKIP A SESSION

LECTURERS

- Marc HUERTAS-COMPANY [Instituto de Astrofísica de Canarias / Observatoire de Paris]: lectures on ML and CNNs + tutorials
- Alexandre BOUCAUD [APC]: Tutorials (from Wednesday mostly)
- External “SEMINARS”:
 - Etienne DECENCIÈRE [Mines-Paristech] - Wednesday
 - Emille ISHIDA [COIN, Besançon] - Thursday
 - Tom CHARNOCK [IAP] - Friday
 - Emanuel BERTIN [IAP] - Friday

PRACTICAL INFO

PLEASE WRITE DOWN THE ROOMS FOR EVERY DAY!

Location

- M: Morning: Salle J.F Denisse (Observatoire) Afternoon: Salle du Conseil (IAP)
- Tu: Morning: Salle J. F. Denisse (Observatoire) Afternoon: Salle Danjon (Observatoire)
- W: Salle Danjon (Observatoire)
- Th: Morning: Salle des Séminaires (IAP) // Afternoon: Salle J.F. Denisse (Observatoire)
- F: Morning: Salle J.F. Denisse Afternoon: Salle des Séminaires

PRACTICAL INFO

INTERNET ACCESS

PREFERRED OPTION: EDUROAM

ALTERNATIVE SOLUTION: OBSERVATOIRE-PORTAIL

[ASK ME FOR LOGIN INFO IF NEEDED]

ONLINE MATERIAL

- A GITHUB WEBPAGE TO SHARE SLIDES AND PRACTICAL INFO [<https://github.com/mhuertascompany/deeplearning4astronomy>]
- AN ADDITIONAL GITHUB WEBPAGE FOR INFORMATION REGARDING TUTORIALS

REFERENCES

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

- Deep Learning: Do-It-Yourself! [Bursuc, Krzakala, Lelarge]
- DEEPMLEARNING.AI [COURSERA, Ng, Bensouda, Katanforoosh]
- MACHINE LEARNING LECTURES [Keck]
- EPFL DEEP LEARNING COURSE [Fleuret]
- IAC XXX WINTER SCHOOL [Baron]

+ **MANY ONLINE RESOURCES!**

Thanks to all of them!

INTRODUCTION

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 ~13 YEARS FOR MY RESEARCH

THIS COURSE IS INTENDED TO PROVIDE A GLOBAL
UNDERSTANDING OF HOW AI TECHNIQUES WORK AND
ESPECIALLY HOW TO USE THEM FOR YOUR RESEARCH

SOME PRELIMINARY NOTES

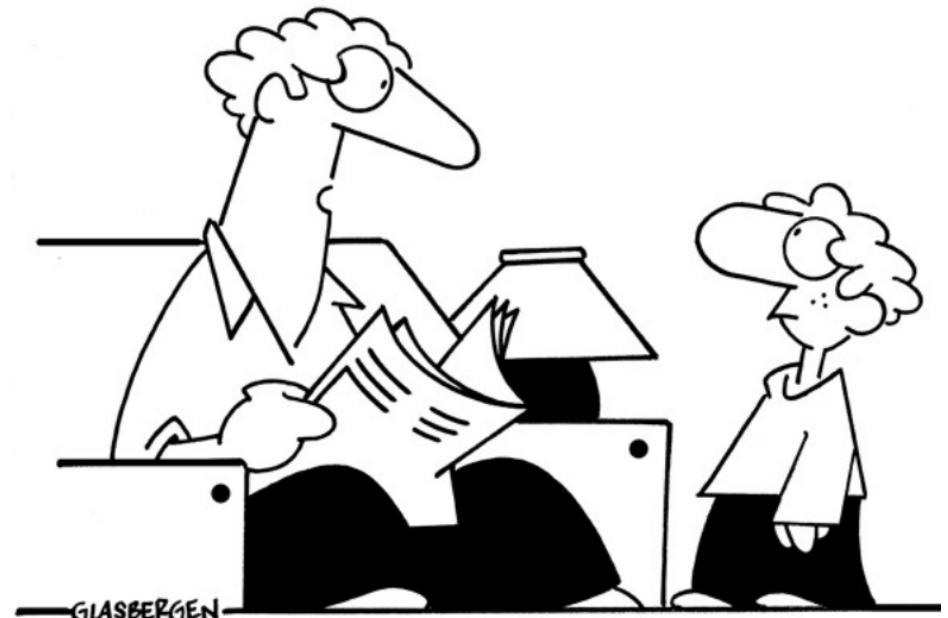
I AM NOT A MACHINE LEARNING RESEARCHER

**IMPORTANT: THIS IS NOT A MATH COURSE. THERE
WILL BE LITTLE FORMALISM. METHODS ARE
PRESENTED FROM A PRACTICAL PERSPECTIVE**

THIS COURSE IS INTENDED TO PROVIDE A GLOBAL
UNDERSTANDING OF HOW AI TECHNIQUES WORK AND
ESPECIALLY HOW TO USE THEM FOR YOUR RESEARCH

WHAT ARE WE GOING TO LEARN?

data-science
pattern-recognition
artificial-intelligence
database
data
big-data machine
data-mining
learning
clustering



*"Artificial intelligence is when you get a college degree,
but you're still stupid when you graduate."*

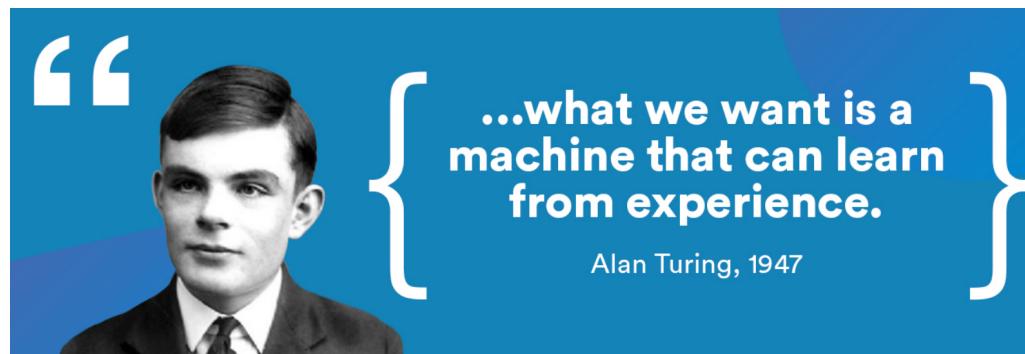
WHAT ARE WE GOING TO LEARN?

data-science
pattern-recognition
artificial-intelligence
data-mining
dc
big-data modeling
data-mining
learning
clustering

LET'S TRY TO SHED
SOME LIGHT INTO
THESE (SOMETIMES)
CONFUSING TERMS

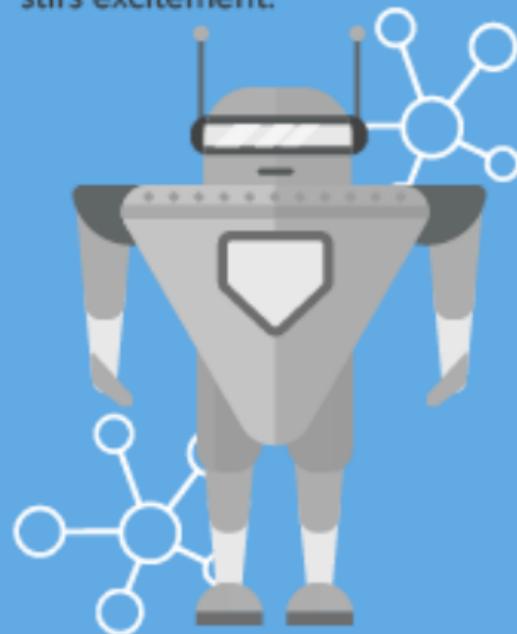


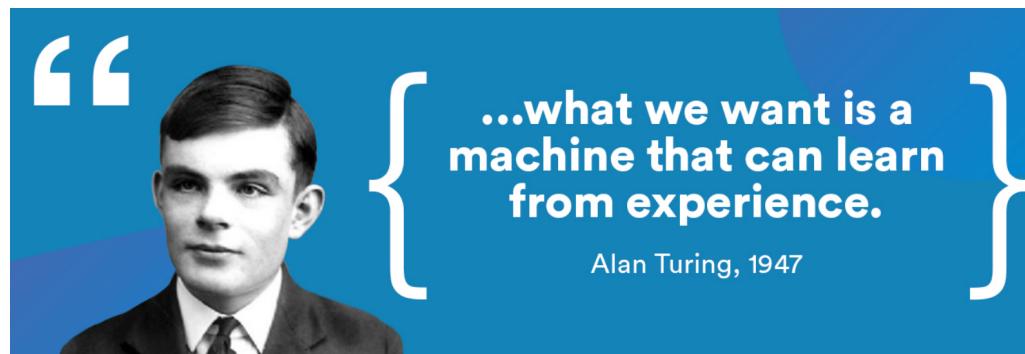
*"Artificial intelligence is when you get a college degree,
but you're still stupid when you graduate."*



ARTIFICIAL INTELLIGENCE

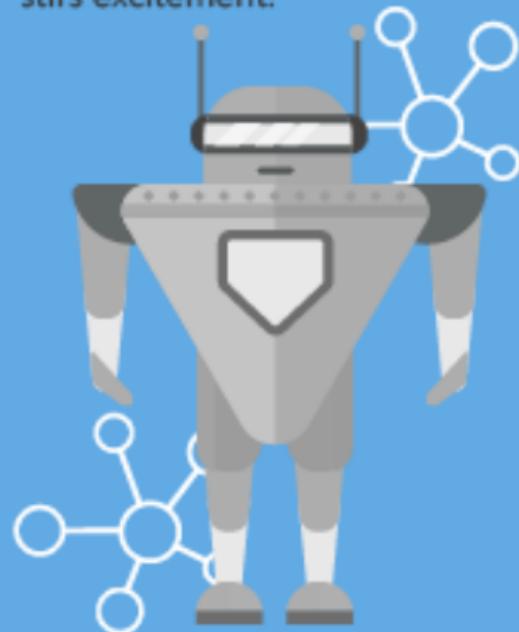
Early artificial intelligence stirs excitement.





ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

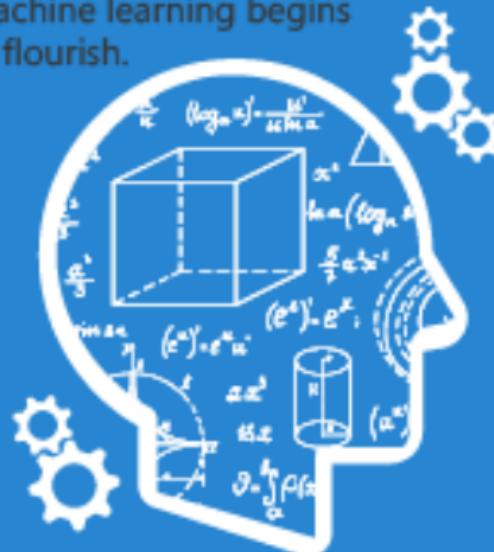
1990's

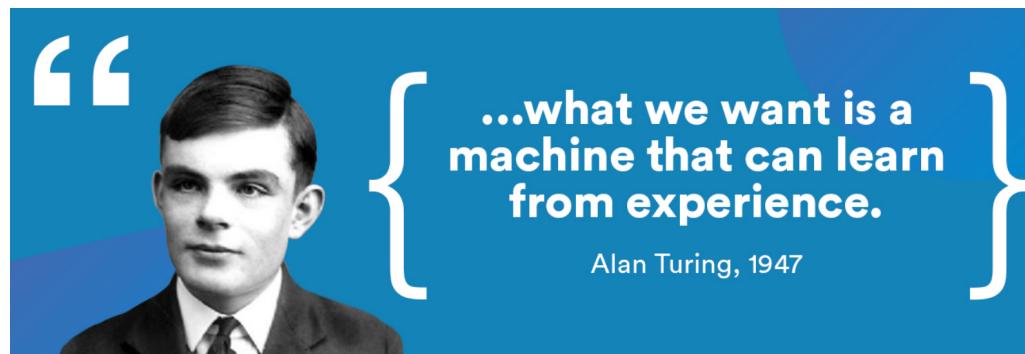
2000's

2010's

MACHINE LEARNING

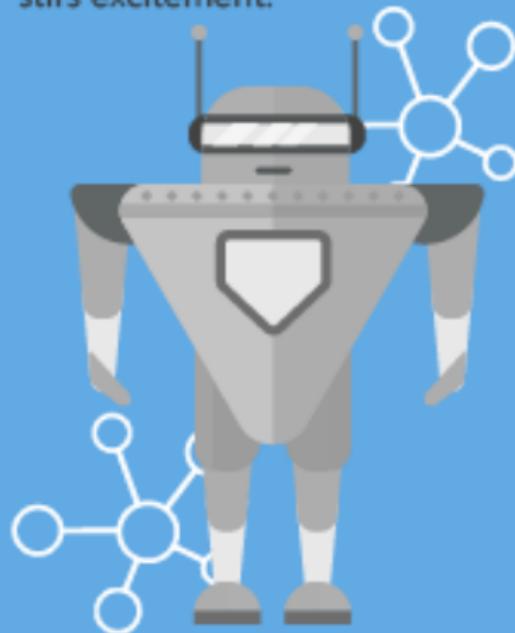
Machine learning begins to flourish.





ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

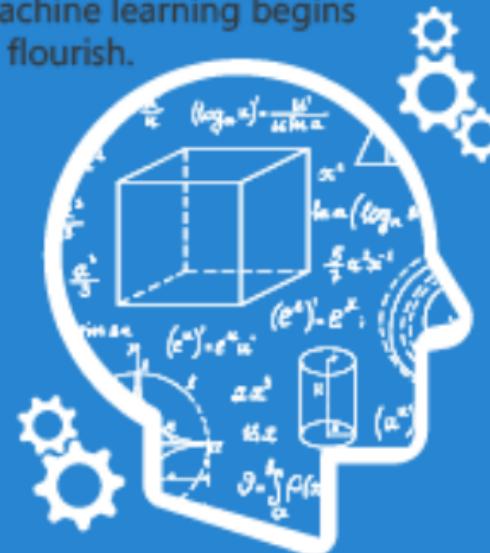
1990's

2000's

2010's

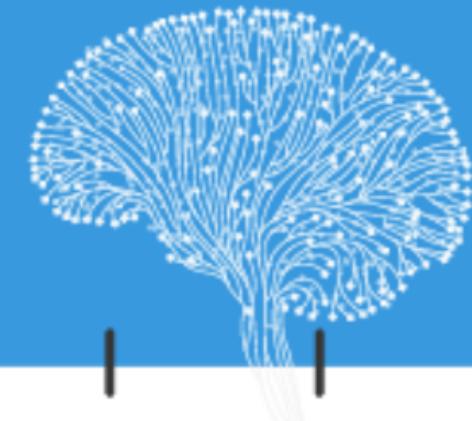
MACHINE LEARNING

Machine learning begins to flourish.

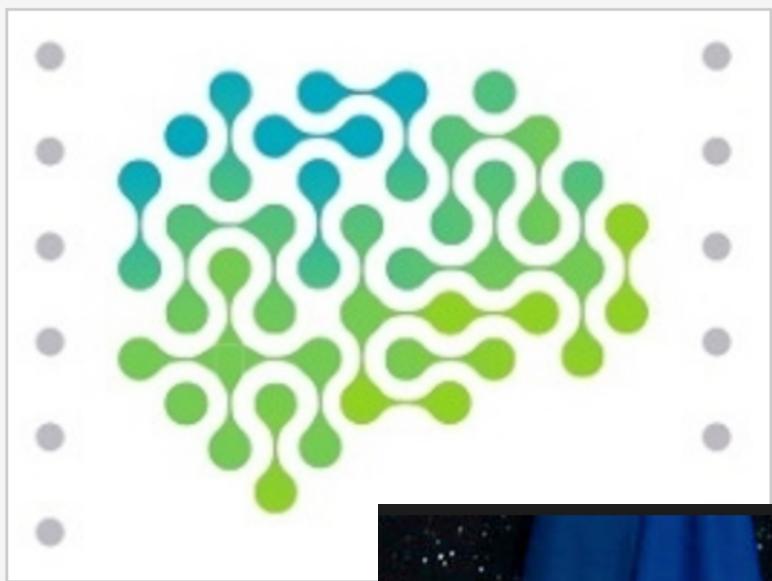


DEEP LEARNING

Deep learning breakthroughs drive AI boom.



AN AMAZING MEDIA ATTENTION

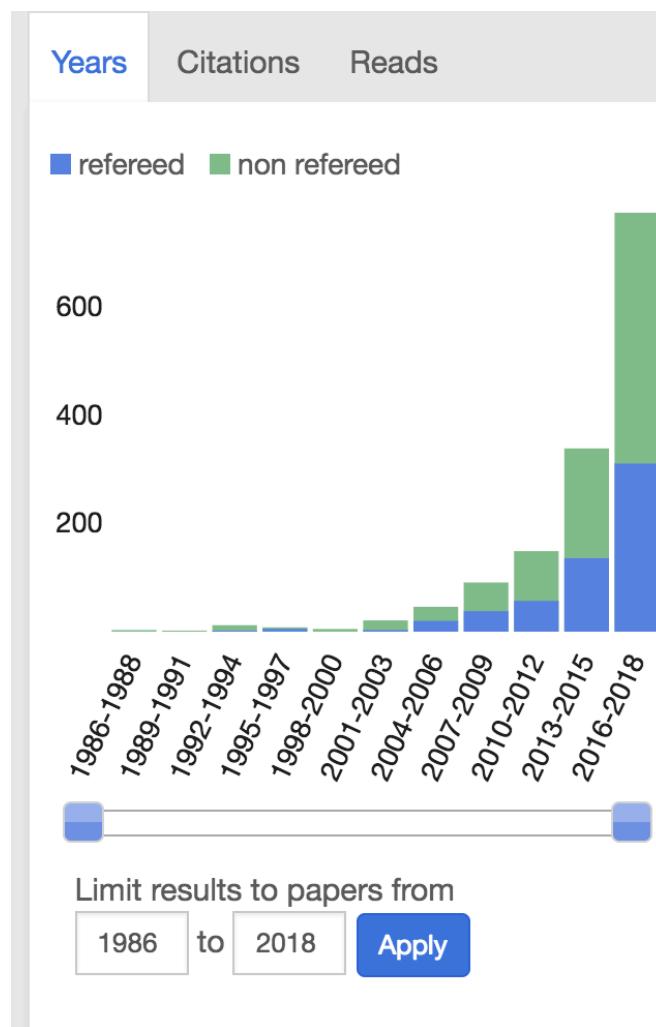


Le CNRS, Inria, l'université PSL et les entreprises Amazon, Criteo, Facebook, Faurecia, Google, Microsoft, NAVER LABS, Nokia Bell Labs, le Groupe PSA, SUEZ et Valeo font converger intérêts académiques et industriels et s'unissent pour créer, à Paris, l'Institut PRAIRIE dont l'objectif est de devenir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

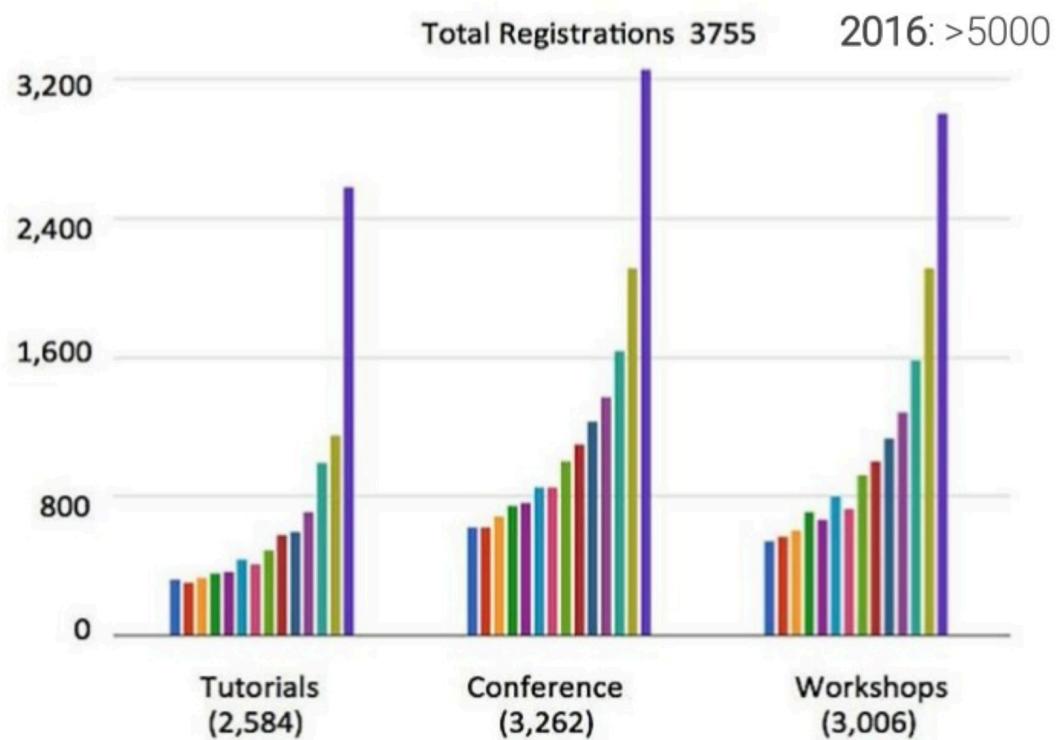


AI FEVER?

PUBLICATIONS (ADS)



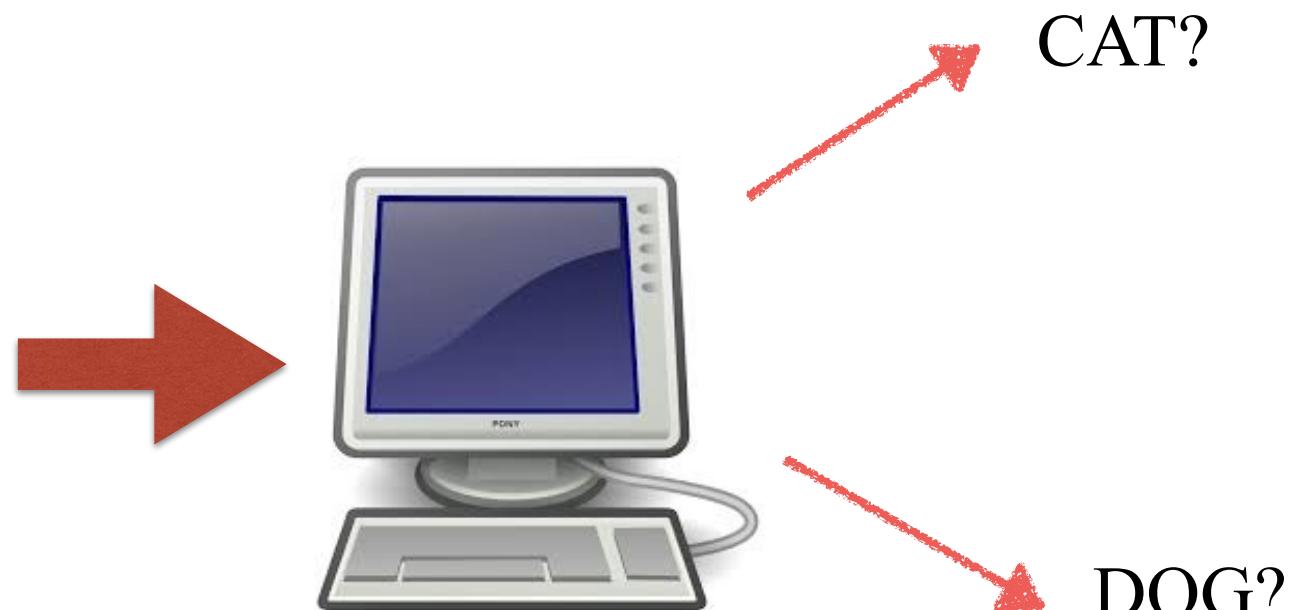
CONFERENCES



Source

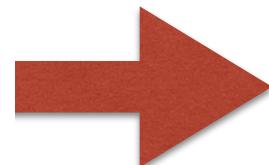
WHY?

UNTIL 2012....



TRIVIAL HUMAN TASKS REMAINED
CHALLENGING FOR COMPUTERS

SINCE 2012



IT HAS BECOME TRIVIAL....

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

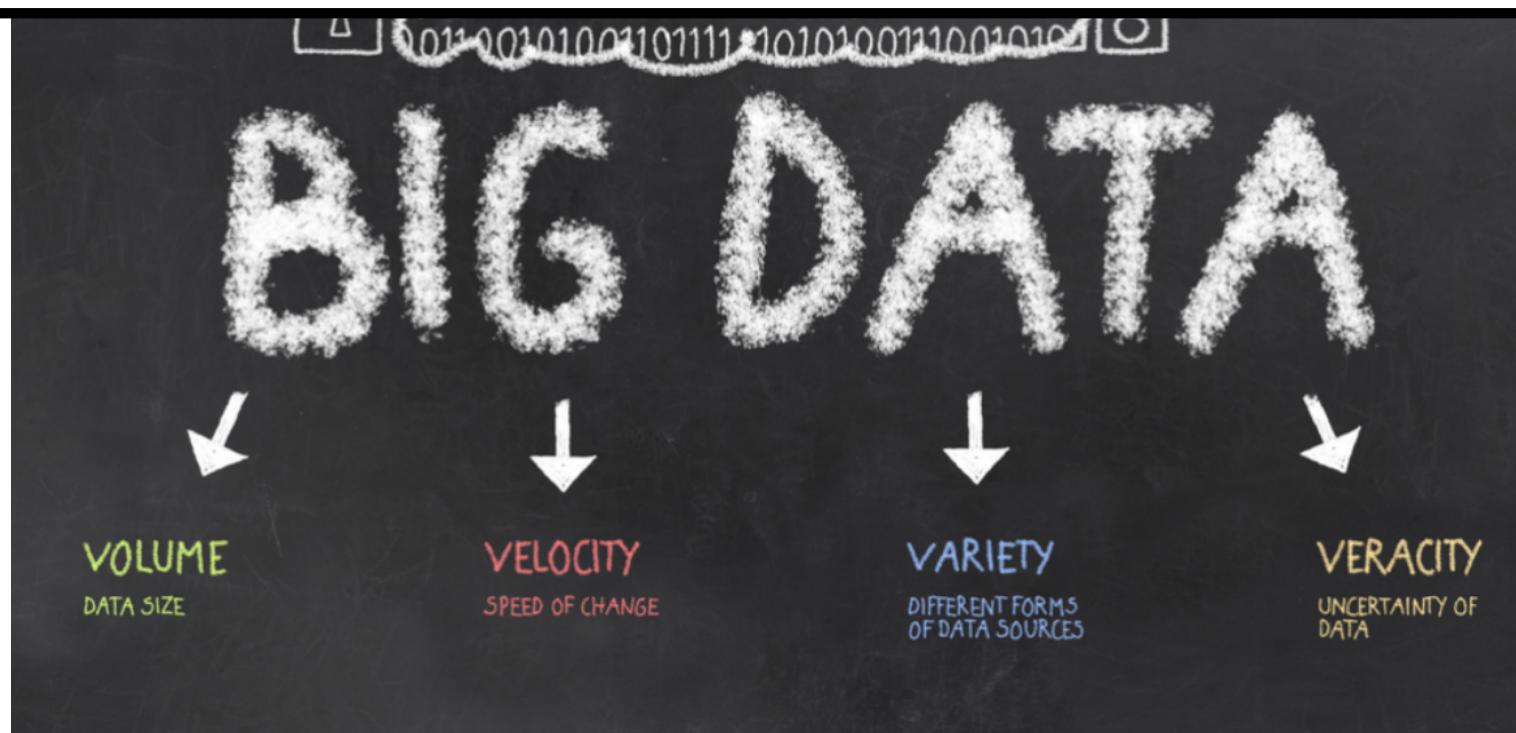


COMBINED WITH THE TECHNOLOGY TO
PROCESS ALL THIS DATA



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 ARE WE GOING TO LEARN?

BASICS OF CLASSICAL MACHINE LEARNING

BASICS OF DEEP LEARNING
(BOTH SUPERVISED AND UNSUPERVISED)

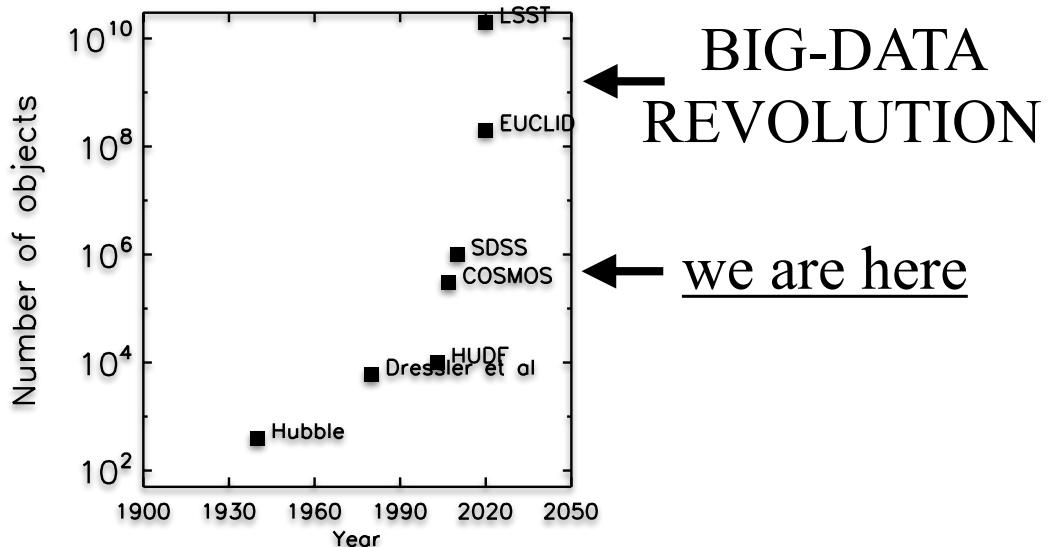
**HOPING THAT THIS WOULD BE USEFUL FOR YOUR
RESEARCH!**

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

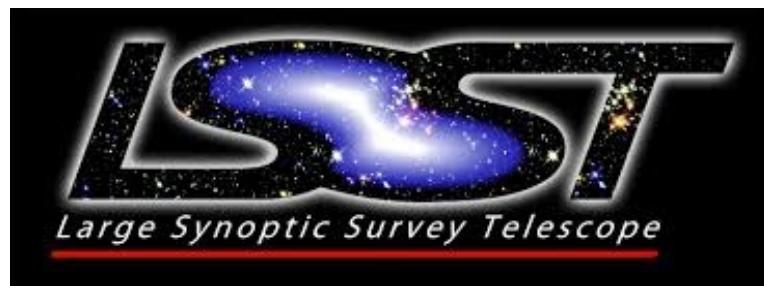
WHY DO WE NEED THESE TOOLS IN ASTRONOMY?

WHY DO WE NEED THESE TOOLS IN ASTRONOMY?

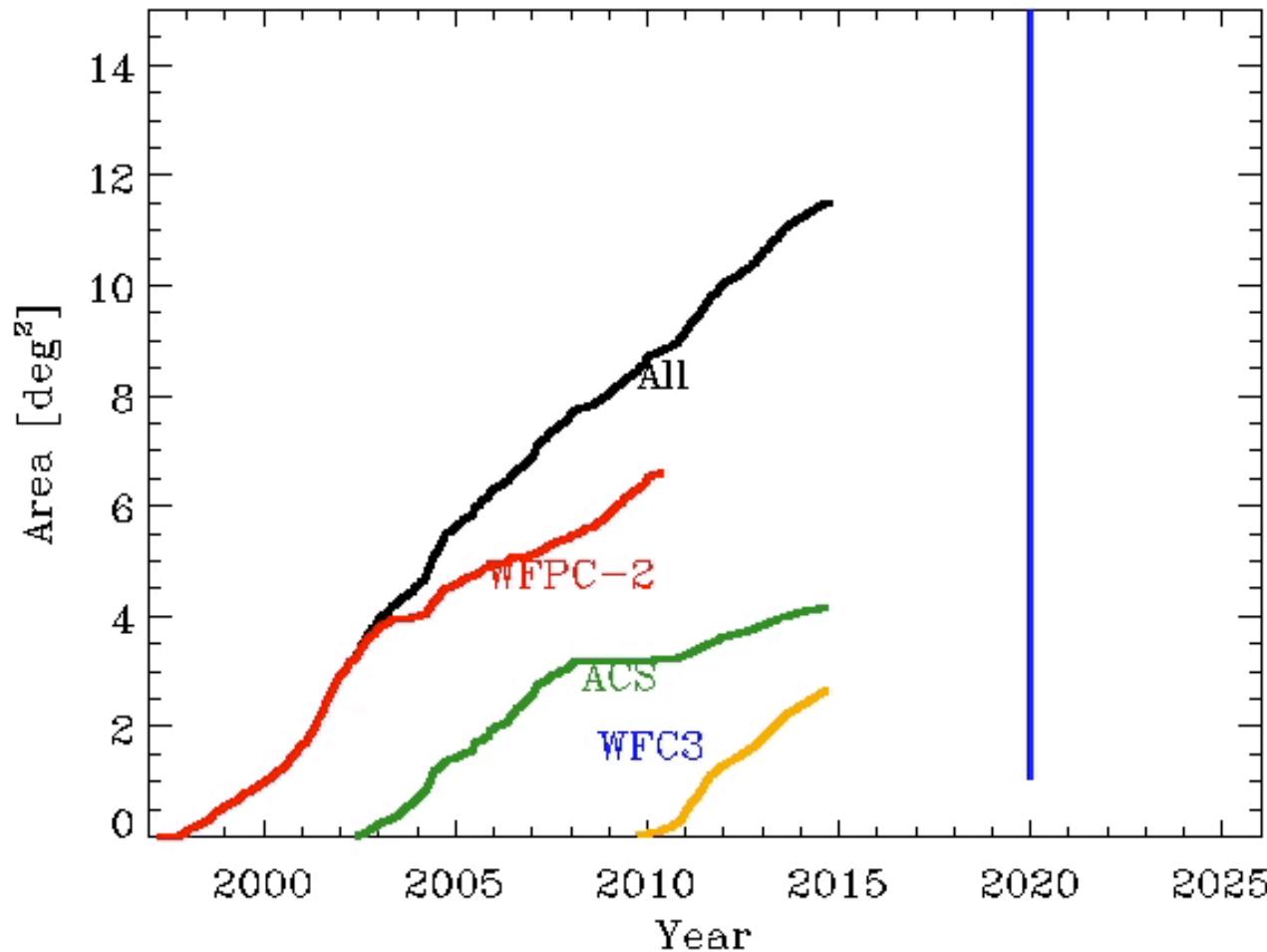
AS IN MANY OTHER DISCIPLINES THE BIG-DATA REVOLUTION HAS ARRIVED TO ASTRONOMY TOO



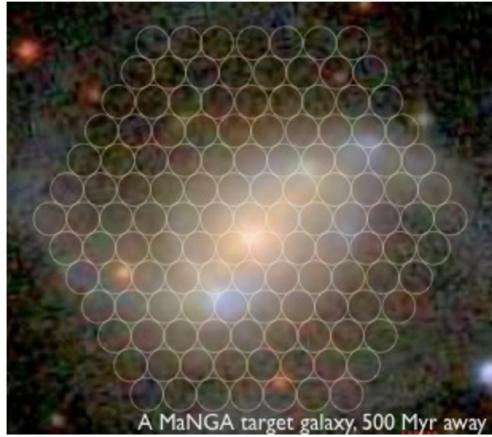
EXTREMELY LARGE
IMAGING SURVEYS
DELIVERING BILLIONS
OF OBJECTS IN 2-5 YEARS



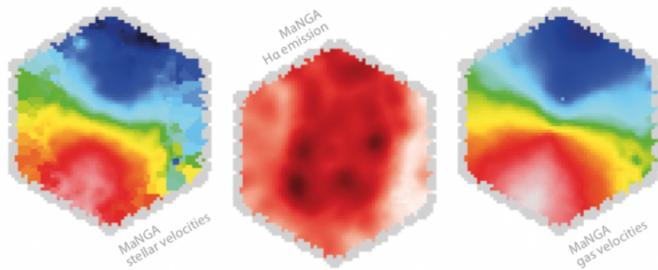
LSST simulation



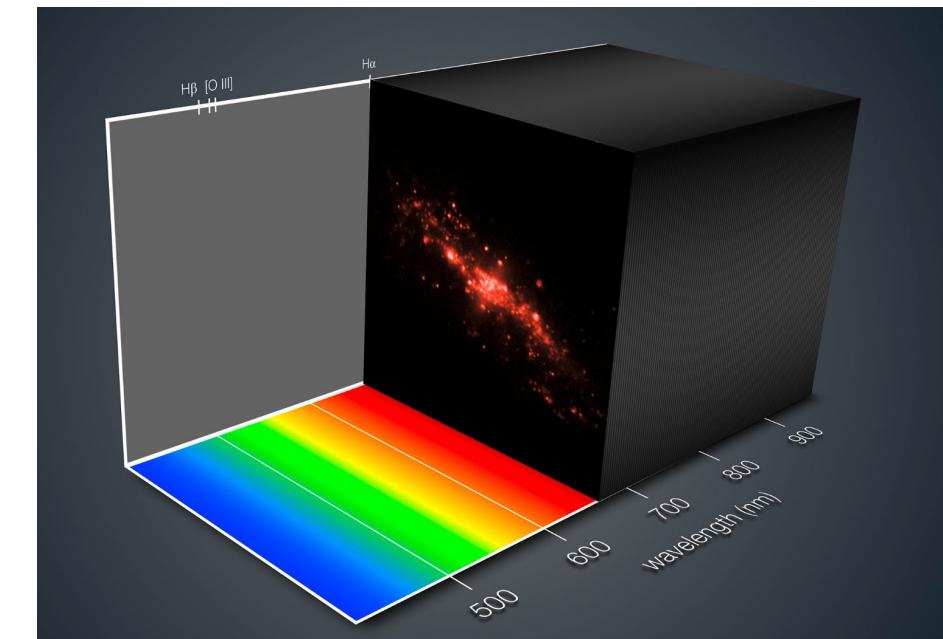
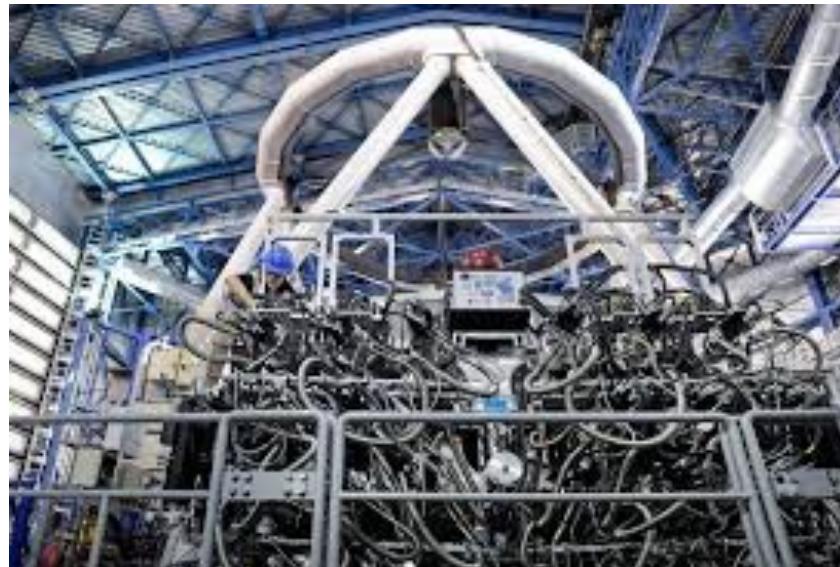
(Thanks to J. Brinchmann)



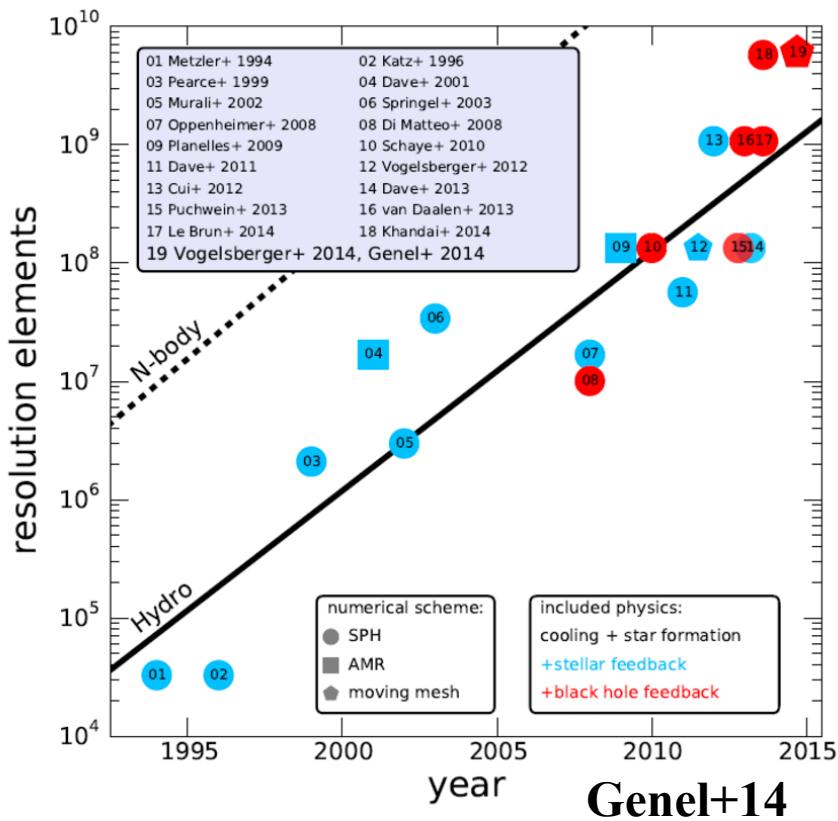
NOT ONLY VOLUME: AN INCREASING COMPLEXITY OF DATA



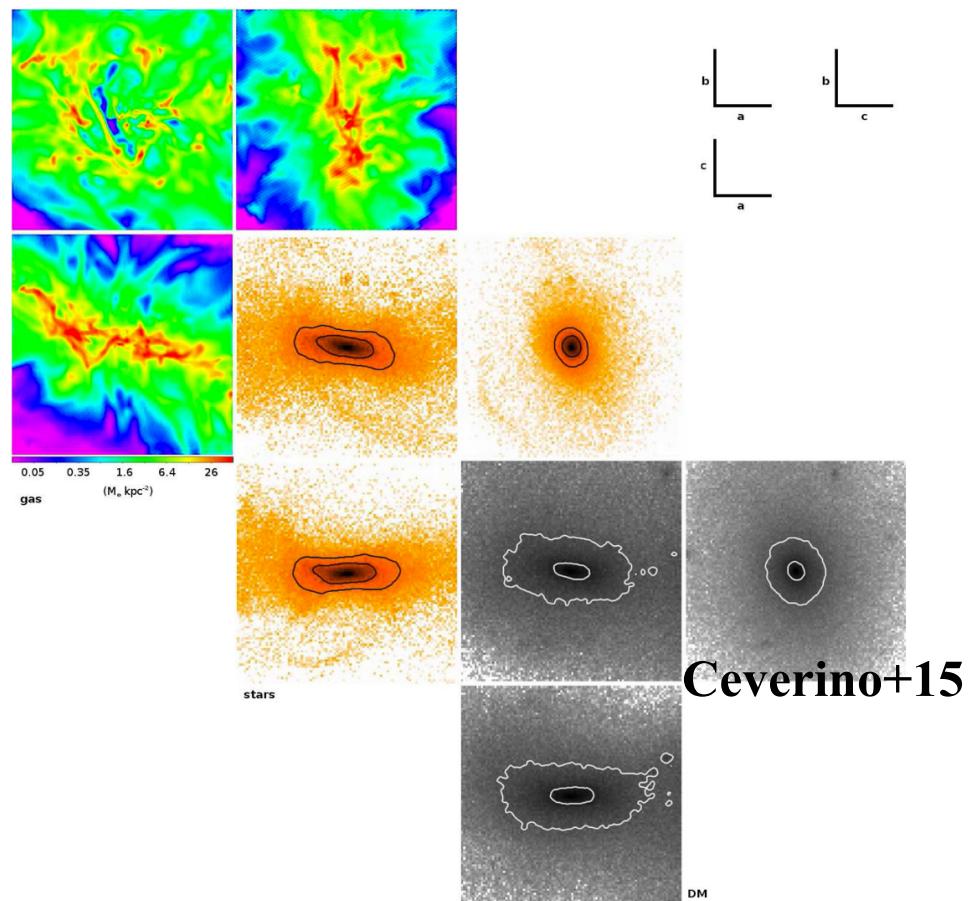
MANGA Survey



MUSE@VLT



AND ALSO
SIMULATIONS!



PROGRAM FOR THE WEEK

- PART I: QUICK INTRODUCTION TO
'CLASSICAL' MACHINE LEARNING
 - UNSUPERVISED / SUPERVISED
 - GENERAL STEPS TO “TEACH A MACHINE”
 - “CLASSICAL” CLASSIFIERS

PROGRAM FOR THE WEEK

- PART II: FOCUS ON ‘SHALLOW’ NEURAL NETWORKS
 - PERCEPTRON, NEURON DEFINITION
 - LAYER OF NEURONS, HIDDEN LAYERS
 - ACTIVATION FUNCTIONS
 - OPTIMIZATION [GRADIENT DESCENT, LEARNING RATES]
 - BACKPROPAGATION

PROGRAM FOR THE WEEK

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

PROGRAM FOR THE WEEK

- **PART IV: IMAGE TO IMAGE NETOWRKS + INTRODUCTION TO UNSUPERVISED DEEP LEARNING**
 - NETWORKS FOR IMAGE SEGMENTATION
 - UNSUPERVISED MACHINE LEARNING
 - AUTO-ENCODERS
 - GENERATIVE ADVERSARIAL NETWORKS
 - ANOMALY DETECTION

PROGRAM FOR THE WEEK

- **PART V: SOME PRACTICAL CONSIDERATIONS**
 - HOW DO I SETUP MY CNN?
 - HOW LARGE DO TRAINING SETS NEED TO BE?
 - OPTIMIZING YOUR NET: HYPER PARAMETER SEARCH
 - VISUALIZING CNNs [DECONVNETS, INCEPTIONISM, INTEGRATED GRADIENTS]

HANDS-ON SESSIONS

- P1 [Monday]: MORPHOLOGY OF GALAXIES WITH CLASSICAL MACHINE LEARNING
- P2 [Tuesday]: MORPHOLOGY OF GALAXIES WITH DEEP LEARNING

HANDS-ON SESSIONS

- T1 [Monday]: MORPHOLOGY OF GALAXIES WITH CLASSICAL MACHINE LEARNING
 - T2 [Tuesday]: MORPHOLOGY OF GALAXIES WITH DEEP LEARNING
-
- T3 [Wednesday-Friday]: SOURCE DEBLENDING WITH DEEP LEARNING

HANDS-ON SESSIONS

- P3 [Wednesday-Friday]: SOURCE DEBLENDING WITH DEEP LEARNING



THIS WILL BE ORGANIZED AS A “CHALLENGE”

YOU WILL SUBMIT JOBS TO SEVERAL GPU SERVERS

A DEBRIEFING WILL BE DONE ON FIRDAY AFTERNOON

PROGRAM FOR THE WEEK

LET'S TRY TO DISCUSS AS MUCH AS POSSIBLE!

AT THE END OF EVERY TUTORIAL YOU WILL BE ASKED TO
DO SHORT PRESENTATIONS EXPLAINING THE SOLUTIONS
FOUND

SOFTWARE REQUIREMENTS

- PYTHON 3 OR GREATER
- TENSORFLOW FOR DEEP LEARNING
- KERAS - HIGH LEVEL LIBRARY WHICH MAKES GPU CODING TRANSPARENT - SIMPLIFIES THINGS A LOT AND MOST OF THE TIME ENOUGH FOR OUR APPLICATIONS
- PLEASE FOLLOW INSTRUCTIONS ON GITHUB

QUESTIONS?