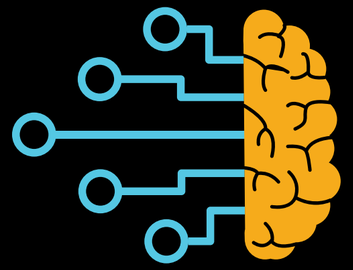
APRENDIZAJE AUTOMÁTICO.

Fundamentos y Aplicaciones en Meteorología del Espacio

DCAO, UBA
08-12 AGOSTO 2022



Dra María Graciela Molina
FACET-UNT / CONICET
Tucumán Space Weather Center - TSWC

https://spaceweather.facet.unt.edu.ar/
IG -> @spaceweatherargentina

gmolina@herrera.unt.edu.ar

Info

RESCD-2022-974-E-UBA-DCT#FCEN

- 35 horas - 1.5 puntos carrera doctorado
- Formato híbrido: aula 1105 del pabellón cero-inifito (presencial).
- Hands-on (teo + lab)
- Plataforma

Dra. María Graciela Molina
gmolina@herrera.unt.edu.ar

Lic. Jorge H. Namour
jnamour@herrera.unt.edu.ar

Dr. Sergio Dasso (coordinador)
sergio.dasso@gmail.com

Info

	LU	MA	MI	JU	VI
10 - 13 hs	X	X	X	X	X
14.30 - 16 hs	X			X	

Temas:

I. Intro . Clasificación y Regresión. Hiperparámetros. Ajuste (underfitting, overfitting, Regularización). Herramientas: Python, Sklearn.

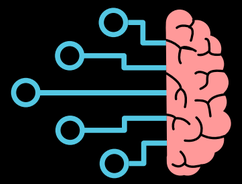
II. Redes neuronales: definiciones. topologías, parámetros e hiperparámetros (funciones de costo, métricas, algoritmos de optimización). Métodos de entrenamiento: backpropagation. Herramientas: Keras, ejemplos.

III. Deep Learning: Conceptos básicos sobre redes recurrentes (LSTM y GRU), redes convolucionales (CNN).

IV. Estado del arte: Machine Learning aplicado a Space Weather en el mundo y en Argentina. Principales preguntas y desarrollos actuales. Preparación de datos. Pronóstico. Casos de estudio.

Info

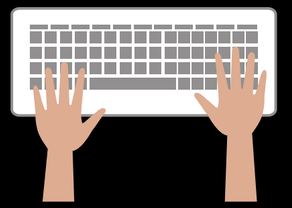
Que SI y que NO vemos en este curso



- AI - ML - DL
- Concepts & recommendations

- SWx concepts
- How to see SWx with data scientist glasses
- Open discussions

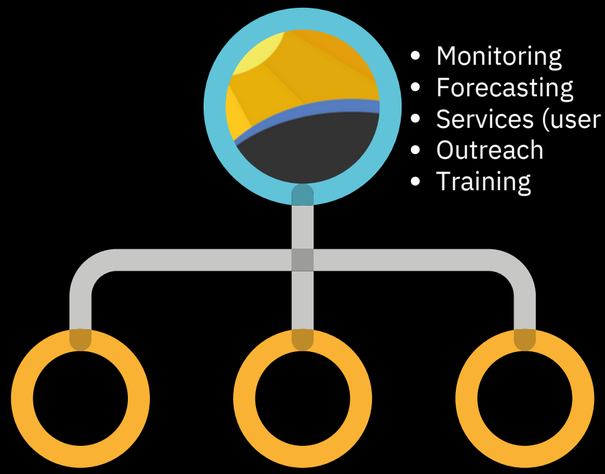
- Tools & best practices



- Hands-on

To pass the course: present TPs

About us ...



- Monitoring
- Forecasting
- Services (user tailored)
- Outreach
- Training

Instrumentation

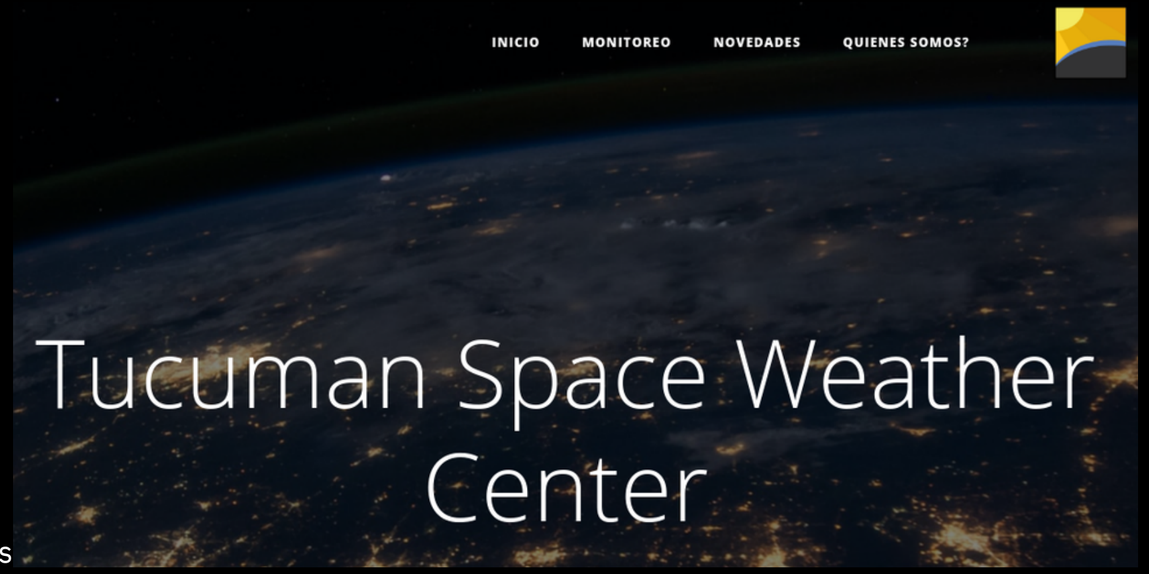
- Deployment
- Design and Development
- Networks
- Radar systems development

Scientific goals

- Atmospheric physics
- Space Weather chain of events understanding. Sun-Earth connections
- Geomagnetism
- Radiopropagation
- ML and stats methods for SWx
- XAI

Data Managment and Operative tools

- Data infraestructure
- Tailored software development
- DS approach: AI modeling/ forecasting, data bases. Operative AI
- ML assisted forecasters tools



https://spaceweather.facet.unt.edu.ar/

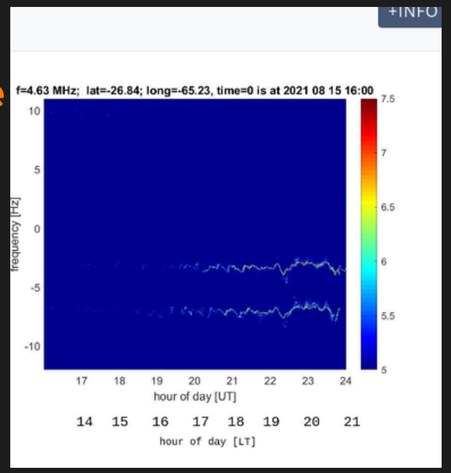
Instagram: /spaceweatherargentina

- Instrumentation development
- Data-driven modeling
- Physics
- Tailored software development
- Operative for ICAO
- Training & Capacity building

About us ...

- Ionogram scaling correction by experts (tool for ionospheric interpreters+ DB)
- Cosmic Ray detector: LAGO collaboration
- Ionospheric conditions forcasting (f0F2 and TEC, single station, regional and global) - AI modelling (status: concept testing)
- NeQuick ingested -> f0F2
- Automatic Flare Analysis Tool (UCR)
- TIDs automatic detection
- TEC calibration service
- New GNSS receiver - New 3D Doppler radar
- ...

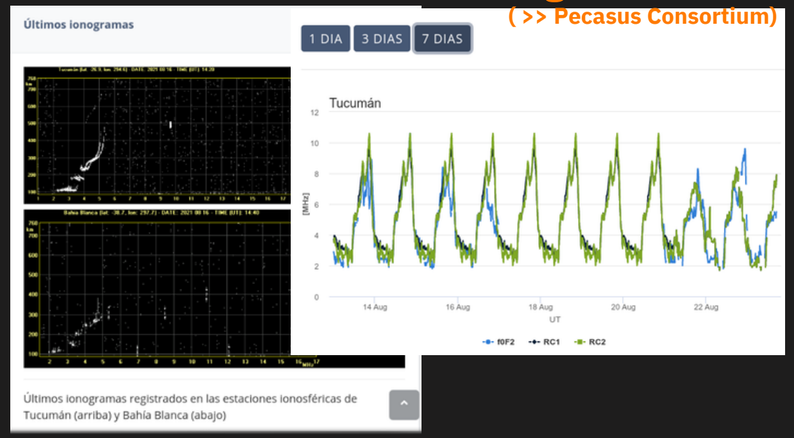
AGWs automatic detection



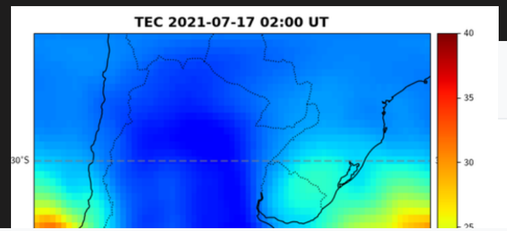
Solar wind - magnetosphere coupling monitoring

Ionosonde Argentine Net

(>> Pecasus Consortium)

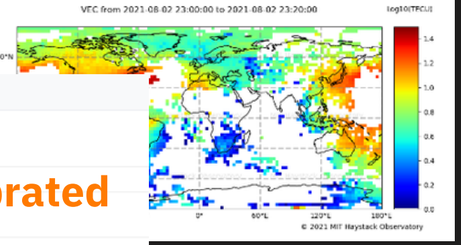


Global and Regional TEC maps

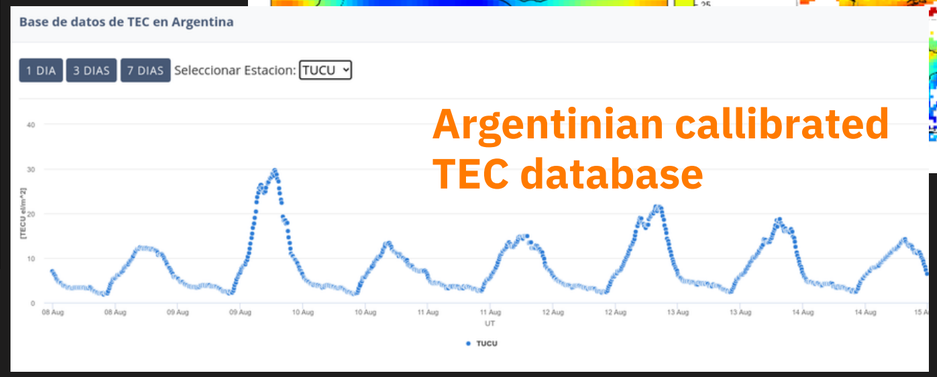


TEC Global

- Madrigal DB

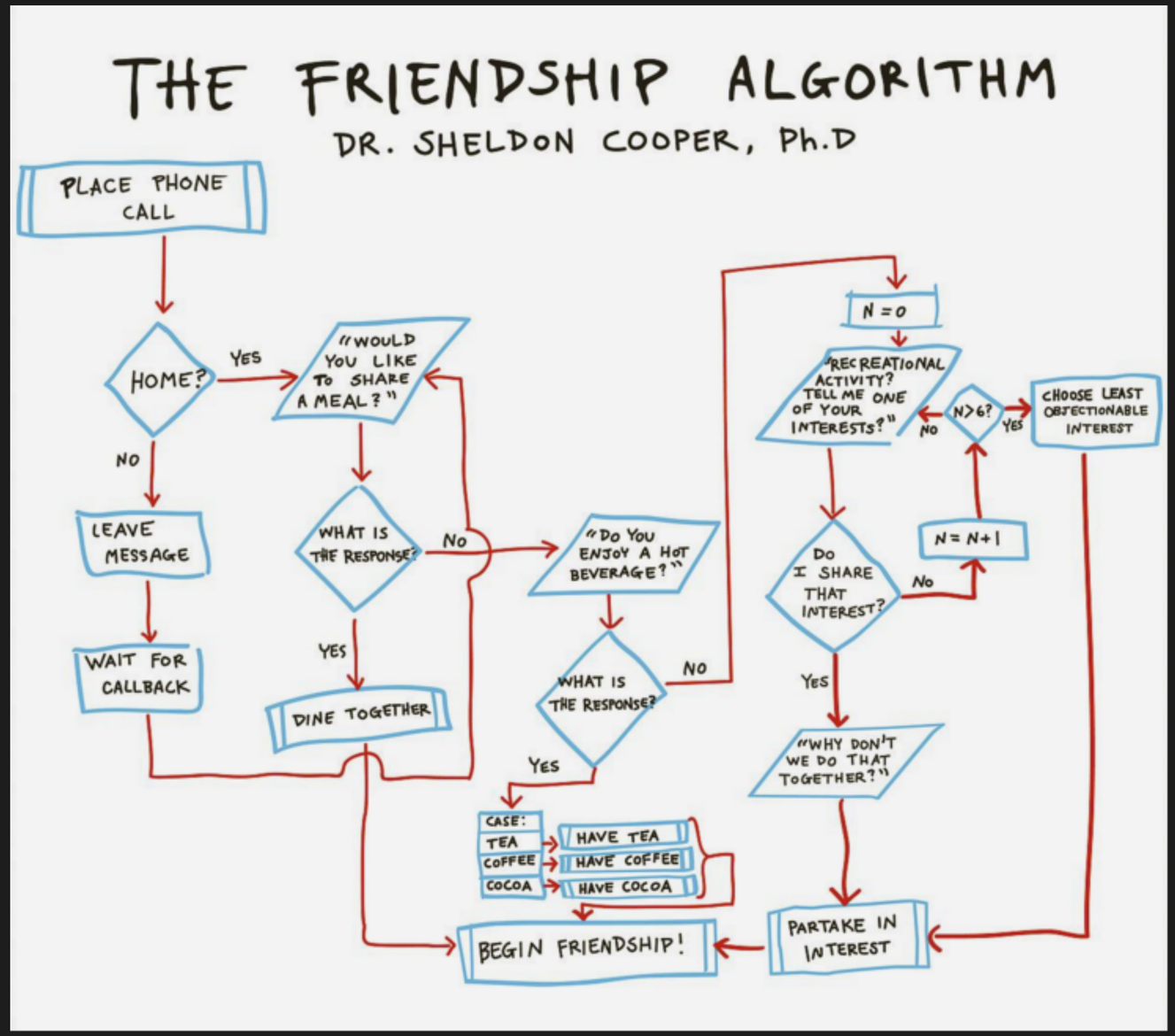


Argentinian callibrated TEC database

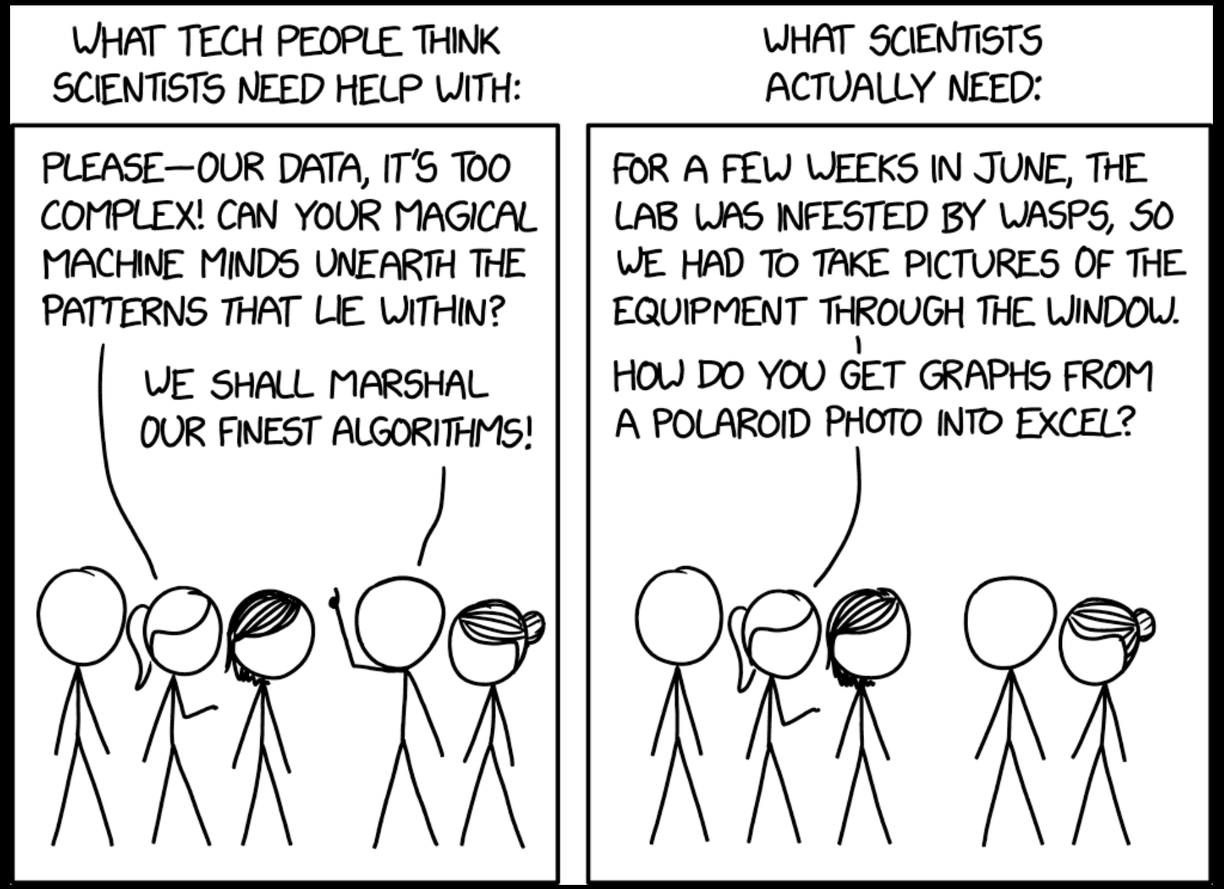


- On-line (20 min)
- On-request (>> resol)
- RAMSAC network
- Collaboration ICTP researchers

What about you?



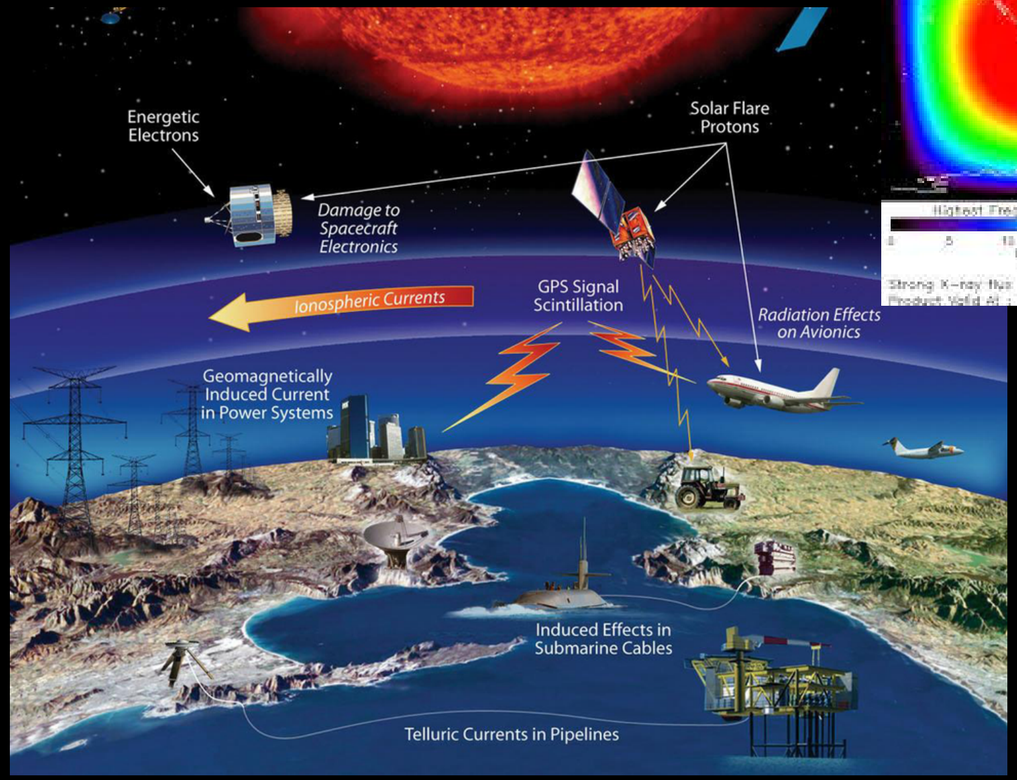
Intro

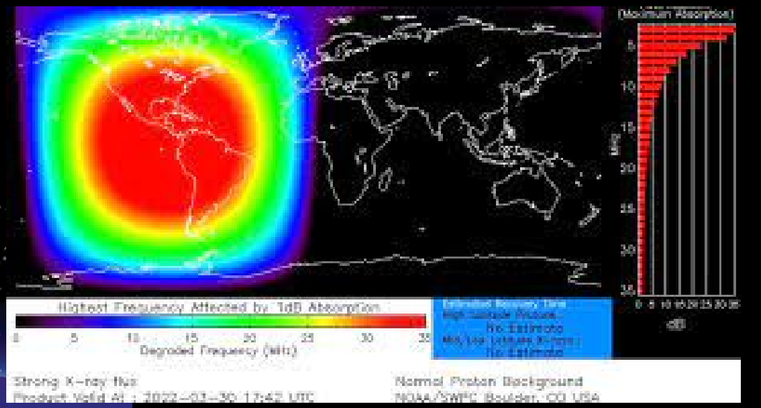

WHAT TECH PEOPLE THINK SCIENTISTS NEED HELP WITH:
PLEASE—OUR DATA, IT'S TOO COMPLEX! CAN YOUR MAGICAL MACHINE MINDS UNEARTH THE PATTERNS THAT LIE WITHIN?
WE SHALL MARSHAL OUR FINEST ALGORITHMS!
WHAT SCIENTISTS ACTUALLY NEED:
FOR A FEW WEEKS IN JUNE, THE LAB WAS INFESTED BY WASPS, SO WE HAD TO TAKE PICTURES OF THE EQUIPMENT THROUGH THE WINDOW.
HOW DO YOU GET GRAPHS FROM A POLAROID PHOTO INTO EXCEL?

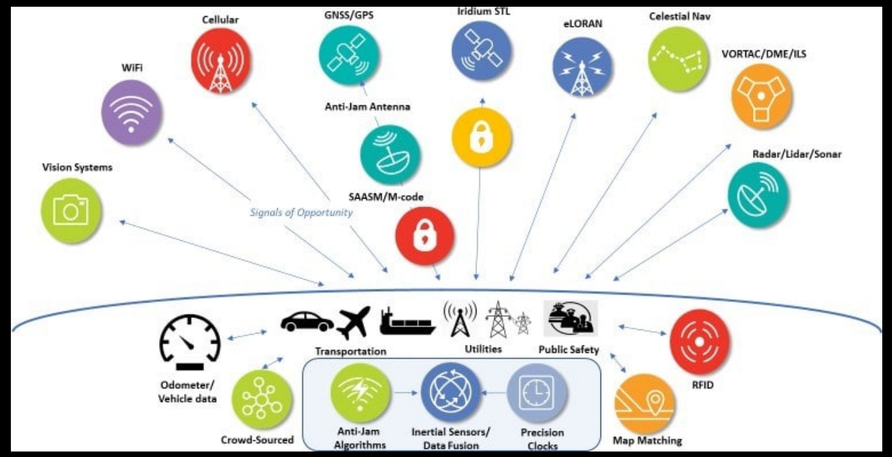
Space Weather (SWx)

Many definitions …

Space Weather is the physical and phenomenological state of natural space environments. The associated discipline aims, through observation, monitoring, analysis and modelling, at understanding and predicting the state of the Sun, the interplanetary and planetary environments, and the solar and non-solar driven perturbations that affect them, and also at forecasting and nowcasting the potential impacts on biological and technological systems. -COST Action 724 , 2009







COST (European Cooperation in Science and Technology) is a funding organisation for research and innovation networks

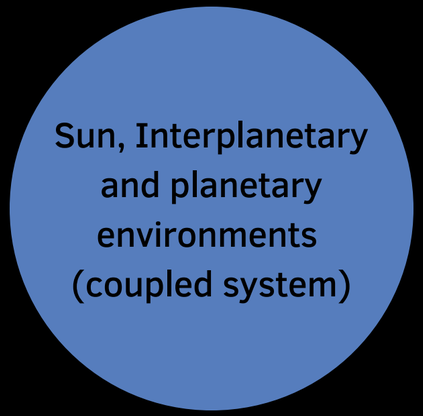
TSWC, 2022

Space Weather (SWx)

COST (European Cooperation in Science and Technology) is a funding organisation for research and innovation networks

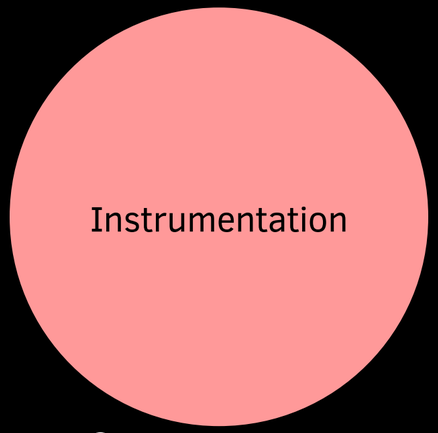
Space Weather is the physical and phenomenological state of natural space environments. The associated discipline aims, through observation, monitoring, analysis and modelling, at understanding and predicting the state of the Sun, the interplanetary and planetary environments, and the solar and non-solar driven perturbations that affect them, and also at forecasting and nowcasting the potential impacts on biological and technological systems. -COST Action 724 , 2009

Many definitions ...

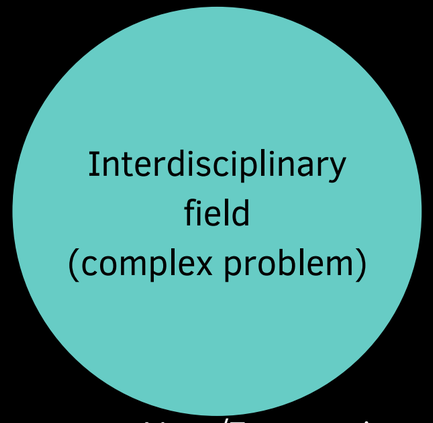


- Science

- Steakholders
- R2O
- Drives the services (e.g. ICAO - interested in HF)



- Coverage
- Partial observations
- Scientific + military + commercial
- Networks



- Now/Forecasting
- Modelling
- Analysis
- Monitoring
- Observations

TSWC, 2022

Space Weather (SWx)

COST (European Cooperation in Science and Technology) is a funding organisation for research and innovation networks

Space Weather is the physical and phenomenological state of natural space environments. The associated discipline aims, through observation, monitoring, analysis and modelling, at understanding and predicting the state of the Sun, the interplanetary and planetary environments, and the solar and non-solar driven perturbations that affect them, and also at forecasting and nowcasting the potential impacts on biological and technological systems. -COST Action 724 , 2009

Many definitions ...

Sun, Interplanetary and planetary environments (coupled system)

- Science

Impacts on biological and technological systems (health and economical impacts)

- Steakholders
- R2O
- Drives the services (e.g. ICAO - interested in HF)

DATA & PROGRAMMING

Instrumentation

- Coverage
- Partial observations
- Scientific + military + commercial
- Networks

Interdisciplinary field (complex problem)

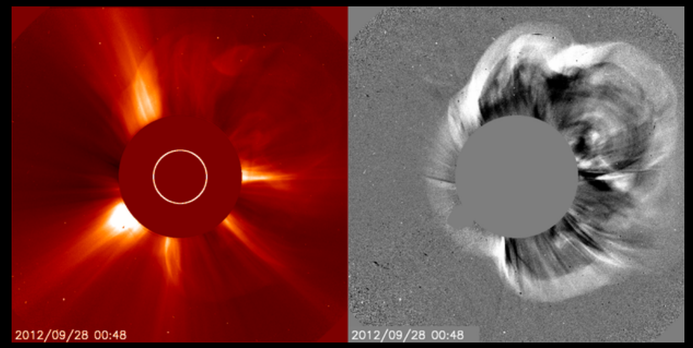
- Now/Forecasting
- Modelling
- Analysis
- Monitoring
- Observations

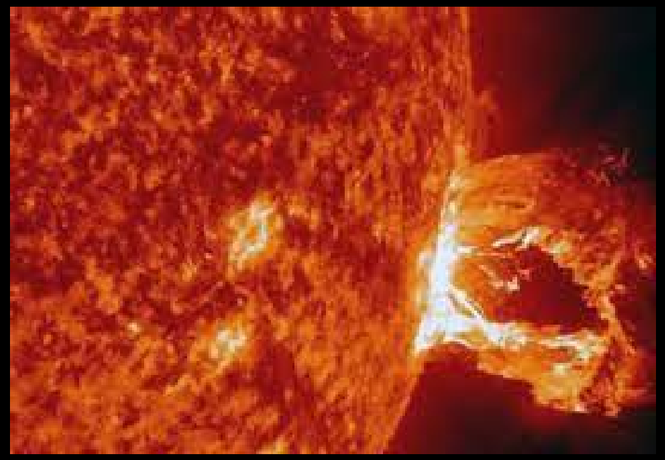
TSWC, 2022

Space Weather (SWx) Main sources & propagation - Observations - Consequences

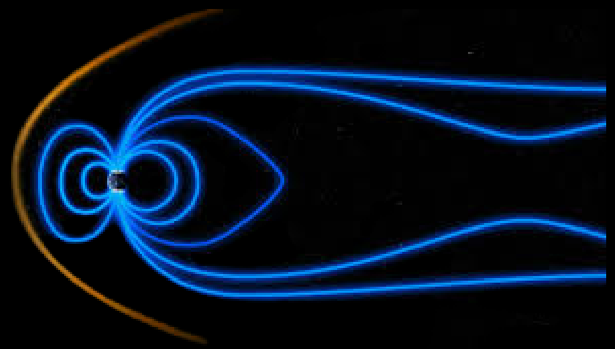
CMEs: large portions of the corona (outer atmosphere) of the Sun explosively blown (plasma) into space. Sometimes are directed toward the Earth. If CMEs reach the Earth's magnetosphere, CMEs magnetic field can interact with the Earth's magnetic fields and produce (under special conditions) the so-called geomagnetic storms

https://umbra.nascom.nasa.gov/lasco/observations/halo/2012/120928/





- Solar wind: constant outflow of electrons and protons from the Sun (interacts with Earth Mag field)



- Can reach the Earth in ~ 1.5 - 3 days (typically)
- SLOW events but difficult to predict

Geomagnetic Storms: Temporary disturbances of the Earth's magnetic field (coupling solar wind-magnetosphere through magnetic reconection)

TSWC, 2022

Space Weather (SWx) Main sources & propagation - Observations - Consequences

Solar flares: Energetic bursts of radiation (radio, white light, EUV, soft X-rays, hard X-rays and Gamma-rays) and particles triggered by the release of magnetic energy on the Sun.

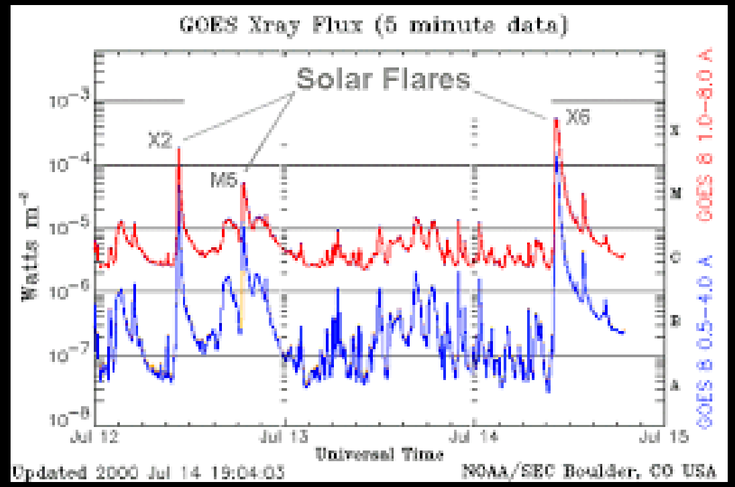
Others: e.g. filaments

CMEs

- Slow events: ~ 1.5 - 3 days (typically)
- Difficult to predict

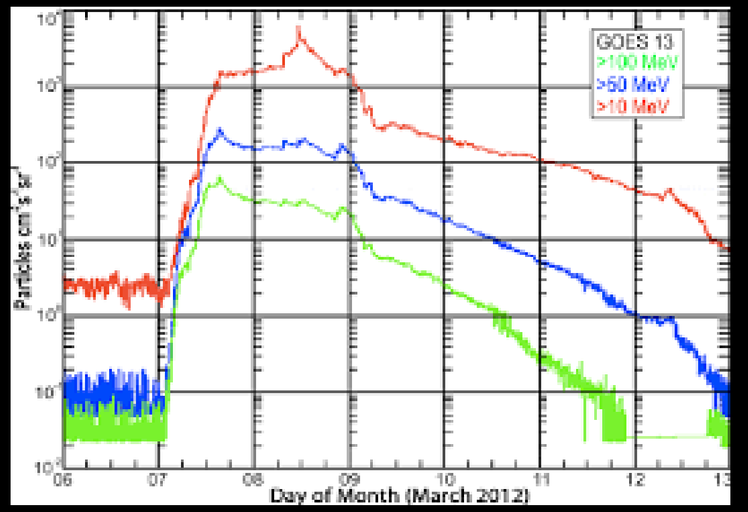
SOFT X-RAY ENHANCEMENT:

- Very Fast events: ~8 m to reach the Earth (dayside)
- Hard to predict (almost instantaneous with the observations)
- X-flares (X-ray radiation emission) classification (Near Earth): A, B, C, M or X [W/m2] of X-rays between 1- 8 Å
- Produces sudden ionization in the ionosphere



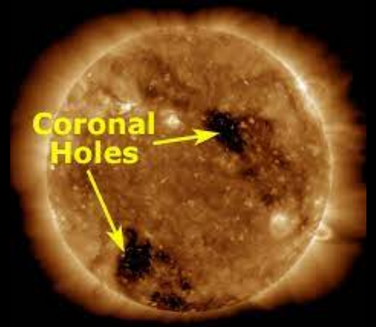
SEPs

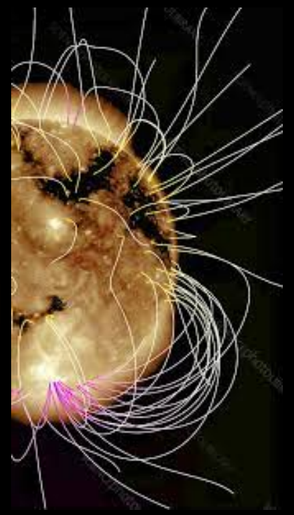
- Can reach the Earth in ~ 1.5 - 3 days (typically)
- fast events: ~20 min

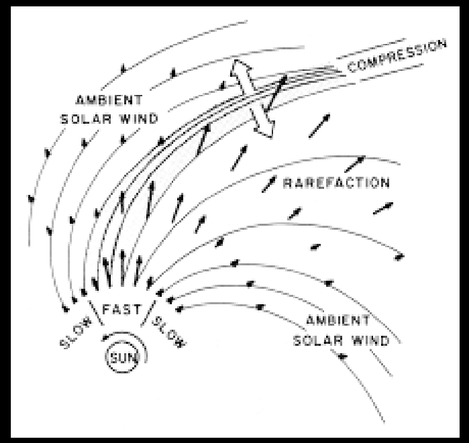


Space Weather (SWx) Main sources & propagation - Observations - Consequences

CHs: Portion of the Sun atmosphere with lower density. These are regions where the Sun magnetic field lines are connected directly with the interplanetary medium, allowing solar material to escape out in a high-speed stream of solar wing.

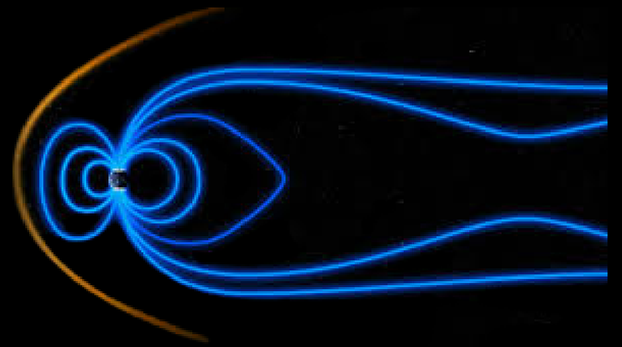






- Faster than background solar wing, but less speed than CMEs (typically)
- SLOW events - A CH can be present for more than a Sun rotation (~27 days), still difficult to predict

Geomagnetic Storms: Temporary disturbances of the Earth's magnetic field (coupling solar wind-magnetosphere through magnetic reconection)

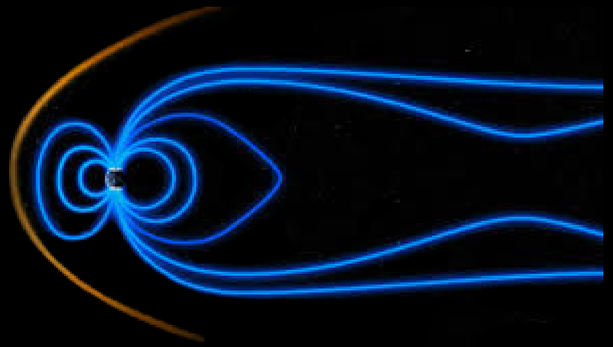


Consequences of a Solar Wind Stream Interaction Region on the Low Latitude Ionosphere: Event of 7 October 2015

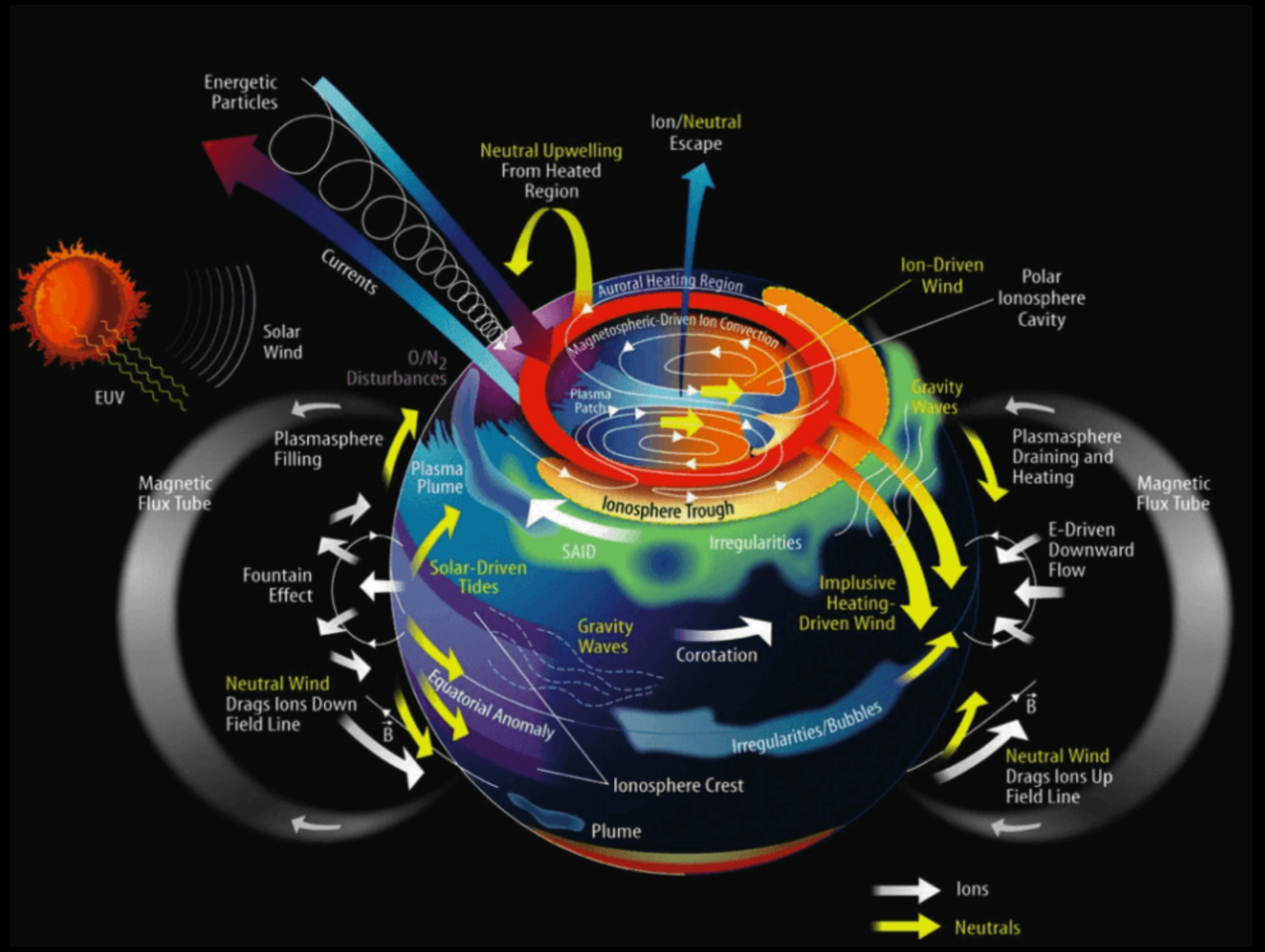
(Molina et al, 2020)

Space Weather (SWx) Main sources & propagation - Observations - Consequences

Geomagnetic Storms:
Temporary disturbances of the Earth's magnetic field (coupling solar wind-magnetosphere through magnetic reconection)



Solar Wind – Magnetosphere – Ionosphere – Thermosphere coupling



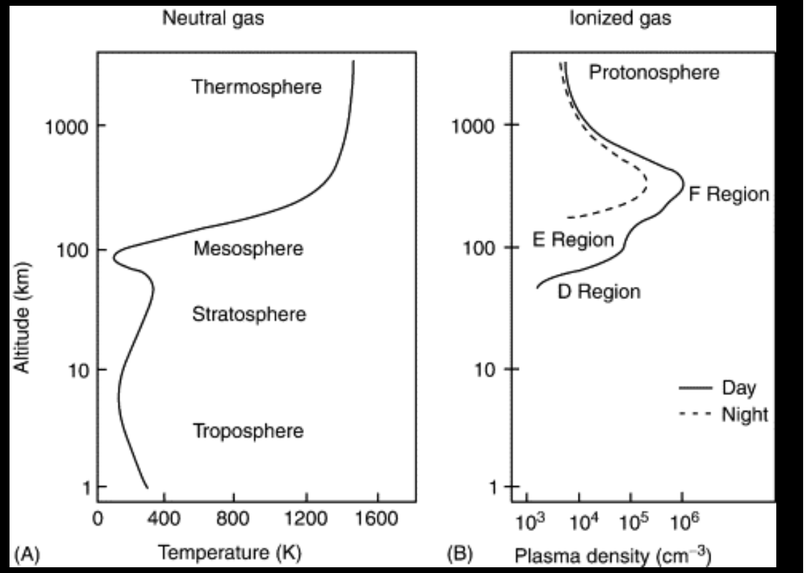
Space Weather (SWx) Main sources & propagation - Observations - Consequences

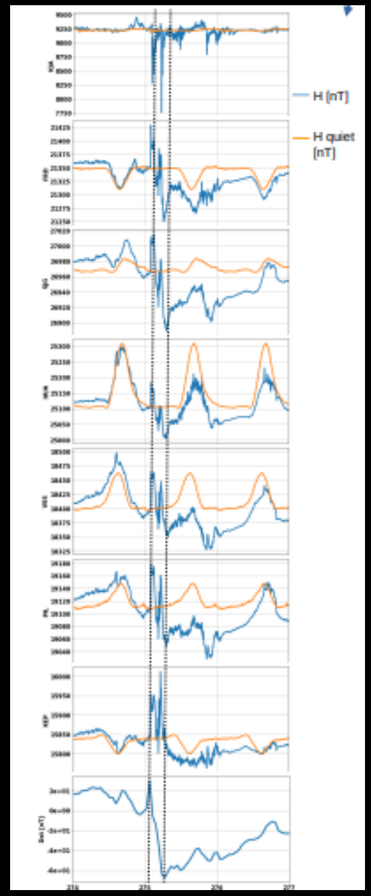
SW-M-I-T

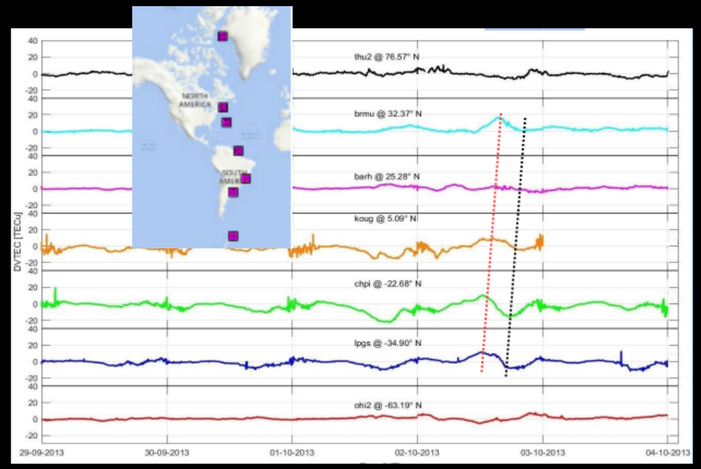
Geomagnetic field perturbations:
- Auroral lat: AU/AL/AE/AO indeces
- Sub-auroral lat: Kp
- Low lat: Dst
- Regional (South America): KSA

Particle precipitaion: enhancement at auroral latitudes

Ionosphere:
- Modification of the ionospheric current system ar high lat (auroral electroject) - PPEF
- Modification of the ionospheric current system at low lat (equatorial electroject)
- Ionospheric heating
- Travelling Ionospheric Disturbances
- (and more)







Ionospheric response to the geomagnetic storm on 2nd October 2013: Longitudinal chain analysis over the American sector. (Molina et al, 2018)

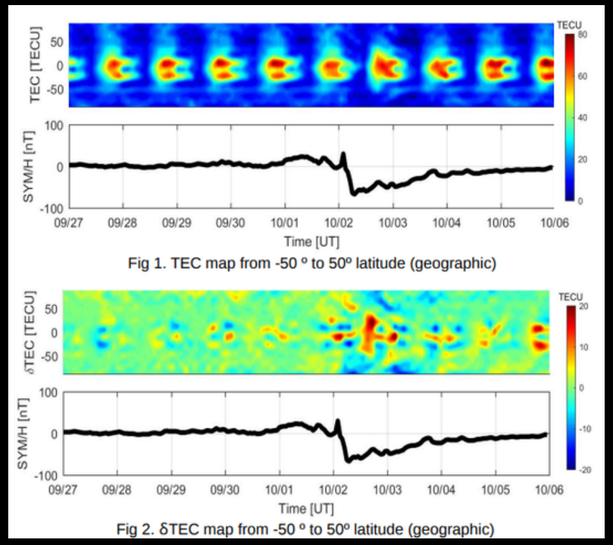


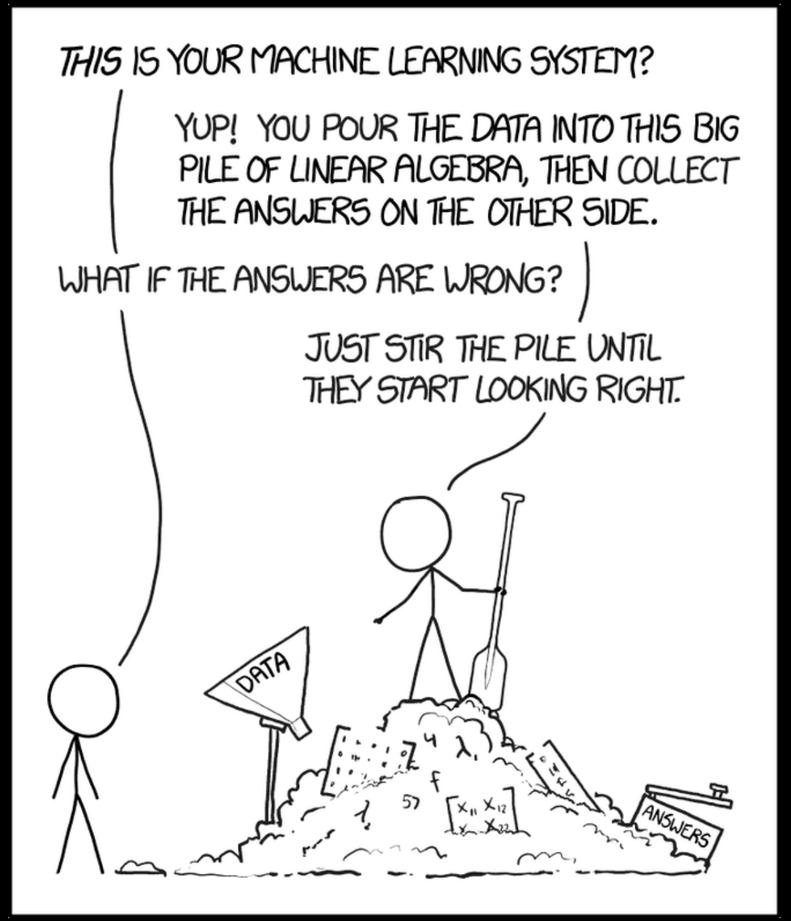
Fig 1. TEC map from -50 º to 50º latitude (geographic)

Fig 2. δTEC map from -50 º to 50º latitude (geographic)

Space Weather (SWx)

The data! just few thing to think about ...

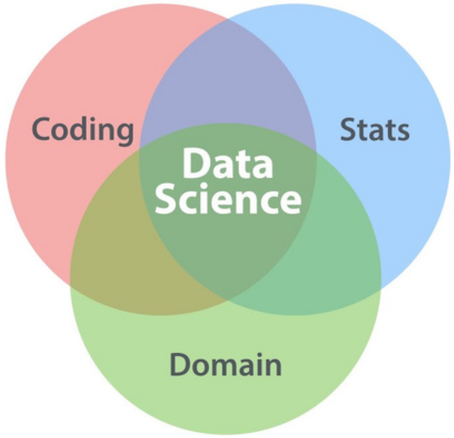
- Heterogenous Data
- Huge amount of data
- Coverage: partially observe a portion of the problem
- Data availability (remote data, military data, etc)
- Data quality: high quality = science; less quality = operations; levels of pre-processing
- Formating madness! resolution madness!
- Not straight-forward to understand
- Produced by instruments, interpreters, simulations or models, metadata
- Storage!
- Complex problems, complex data
- Data infrastructure: data model, a model for data management, hardware
- Intrinsically unbalanced data
- Scales!!!!

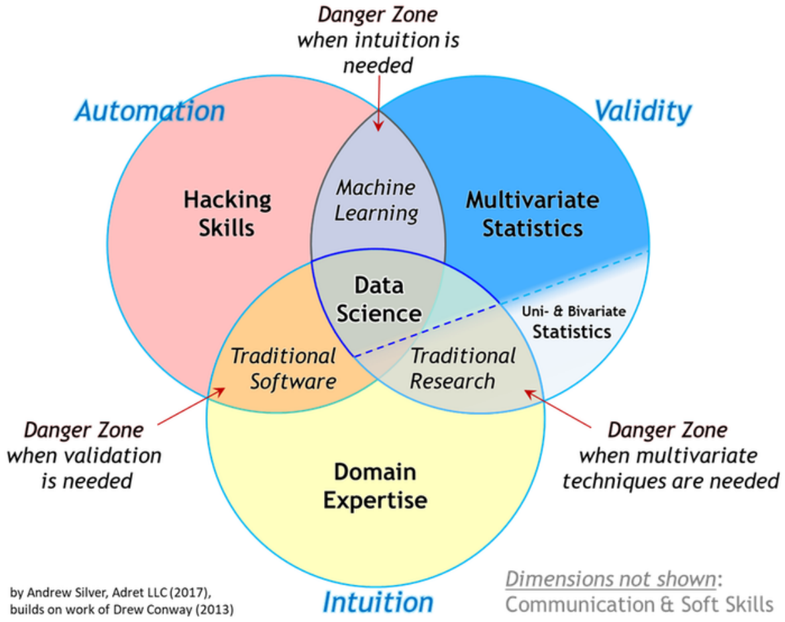


DATA SCIENCE

Data Science

- Interdisciplinary field
- Extract knowledge from data
- Transform data into knowledge/information
- Aids decision making





DATA

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE

SUMMARY SAVE SHARE 13 COMMENT TEXT SIZE PRINT $8.95 BUY COPIES

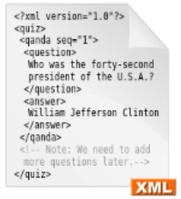
When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their

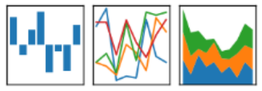
1/3 FREE ARTICLES LEFT > REGISTER FOR MORE | SUBSCRIBE + SAVE!

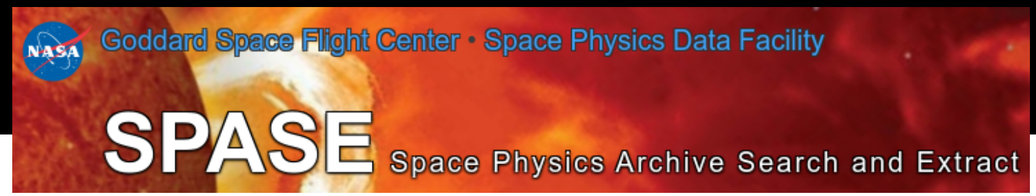
Data Science

KNOWLEDGE & SKILLS

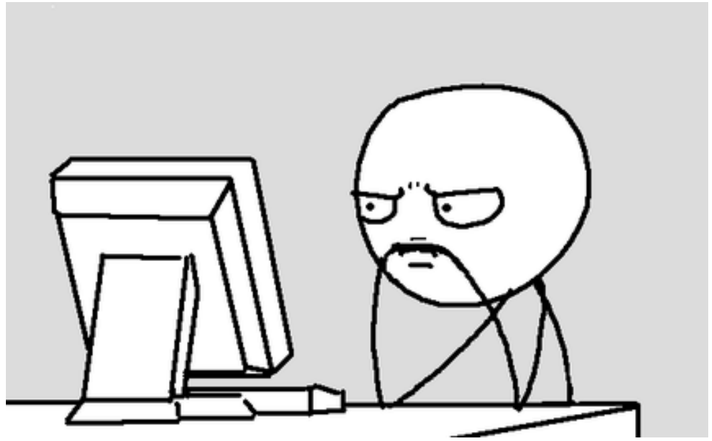
- Experience (domain)
- **Methods (e.g. ML)**
- **Programming Lenguage + tools + libraries + data formats**
- Data access / transformation
- Data storage
- Visualization tools
- Google skills!







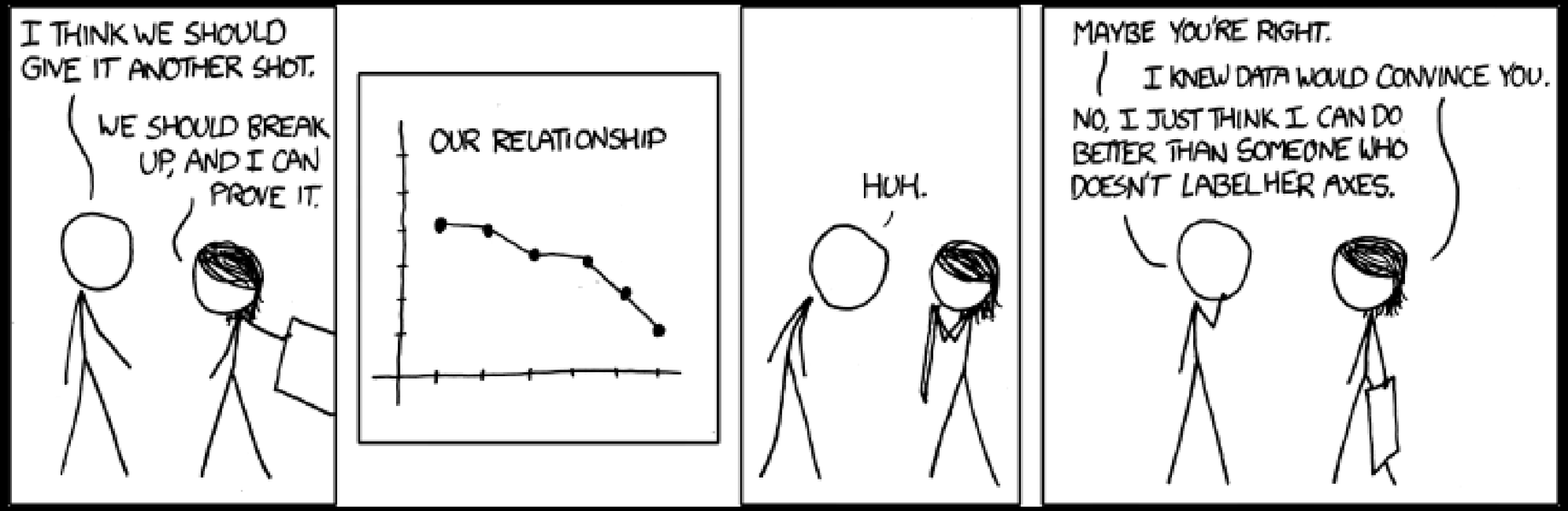
https://spase-group.org/data/



TSWC, 2022

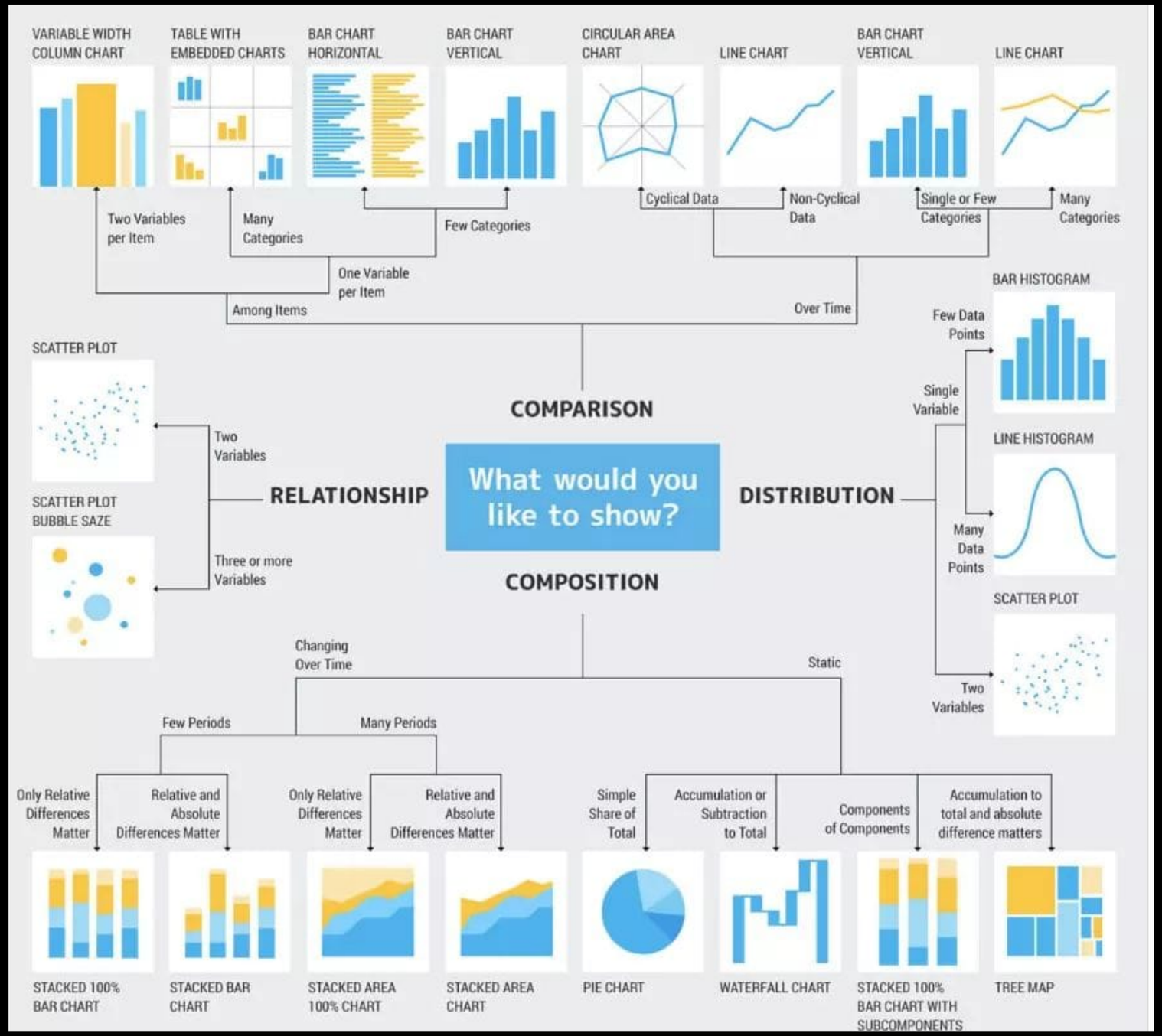
KNOWLEDGE & SKILLS

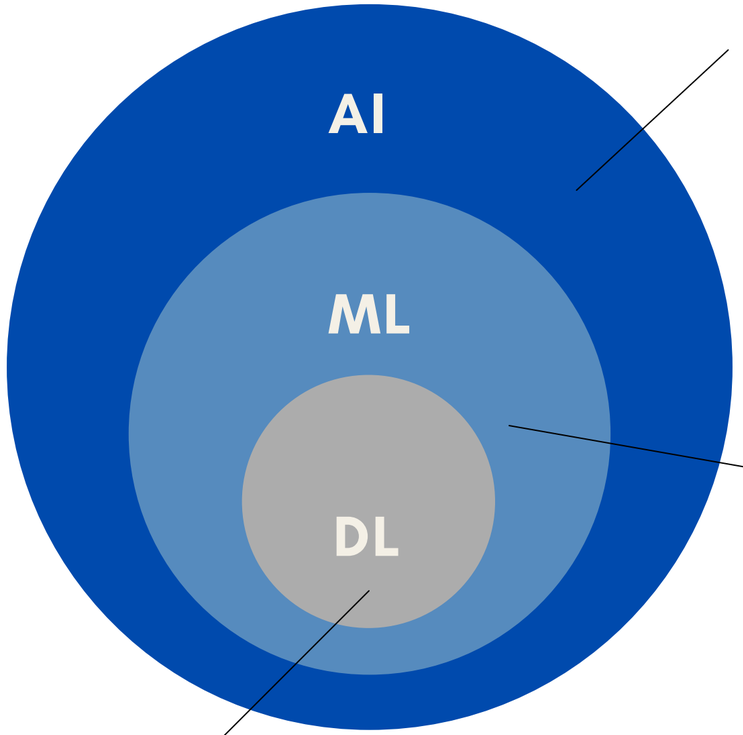
- **Visualization tools**
- **Google skills!**
  - https://www.kaggle.com/
  - https://es.stackoverflow.com/



KNOWLEDGE & SKILLS

- **Visualization tools**
- **Google skills!**
  - https://www.kaggle.com/
  - https://es.stackoverflow.com/

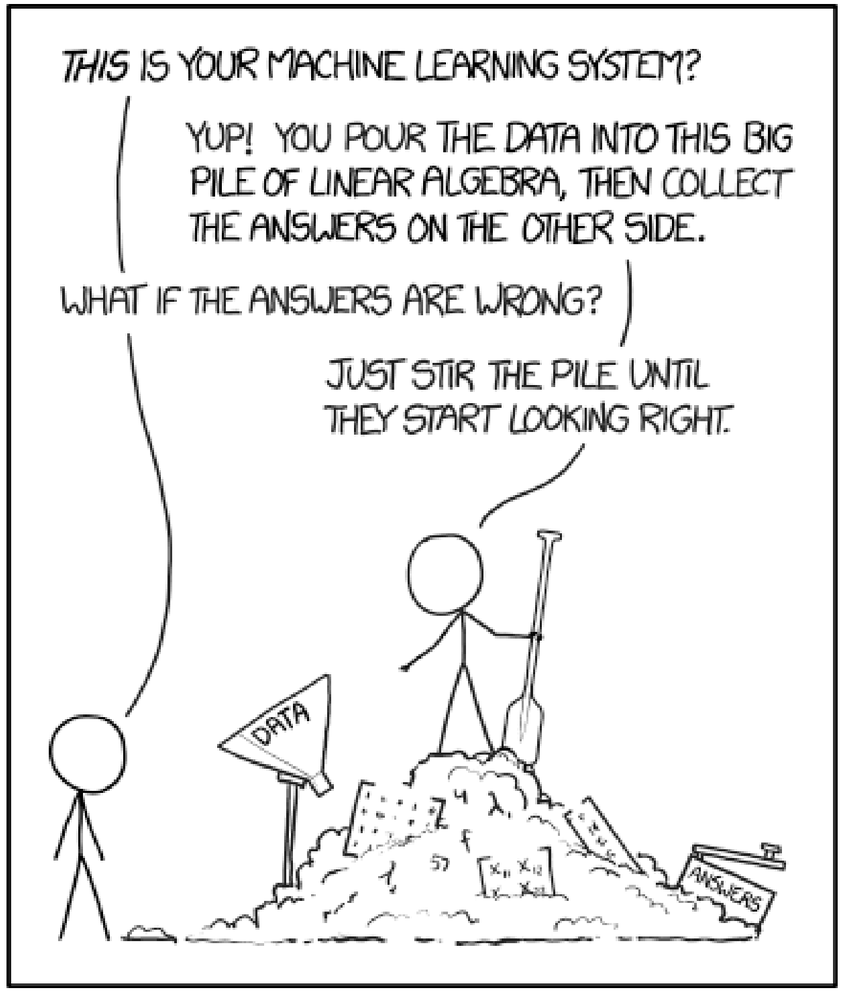




AI = area of computer science that emphasizes the creation of intelligent machines that work and react like humans.

ML = method of data analysis that automates data model building. ML uses algorithms that learn from data and can find insights without explicit programming.

DL = a subset of ML and refers to artificial neural networks that are composed of many layers.



ML (DL) WHY NOW?

Neural Networks date back decades, so why the resurgence?

1952 Stochastic Gradient Descent

1958 Perceptron
- Learnable Weights

1986 Backpropagation
- Multi-Layer Perceptron

1995 Deep Convolutional NN
- Digit Recognition

1. Big Data
- Larger Datasets
- Easier Collection & Storage

IMAGENET

WIKIPEDIA The Free Encyclopedia

2. Hardware
- Graphics Processing Units (GPUs)
- Massively Parallelizable

3. Software
- Improved Techniques
- New Models
- Toolboxes

TensorFlow

+ more mature algorithms
+ applications
+ bigger community

DOI: 10.1214/AOMS/1177729586 • Corpus ID: 16945044

A Stochastic Approximation Method

H. Robbins • Published 1 September 1951 • Mathematics • Annals of Mathematical Statistics

Let M(x) denote the expected value at level x of the response to a certain experiment. M(x) is assumed to be a monotone function of x but is unknown tot he experiment, and it is desire to find the solution x=0 of the equation M(x) = a, where x is a given constant. we give a method for making successive experiments at levels x1, x2,... in such a way that x, will tend to 0 in probability.

View via Publisher | [PDF] projecteuclid.org | Save to Library | Create Alert | Cite

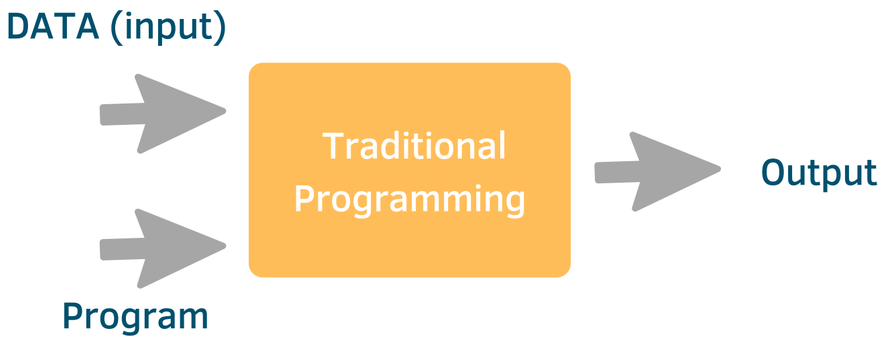
Share This Paper

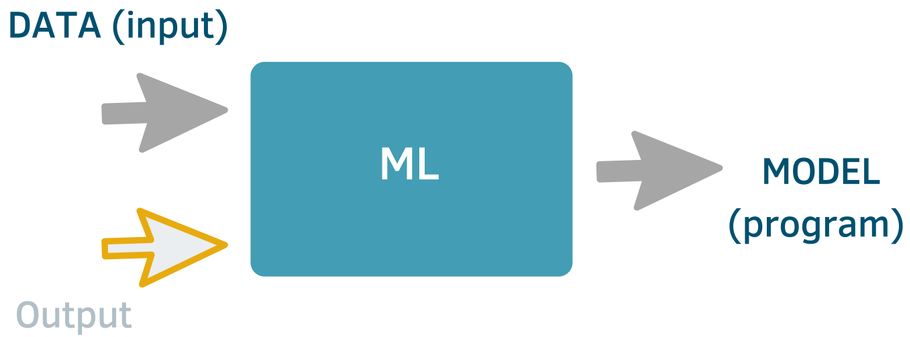
6,527 Citations

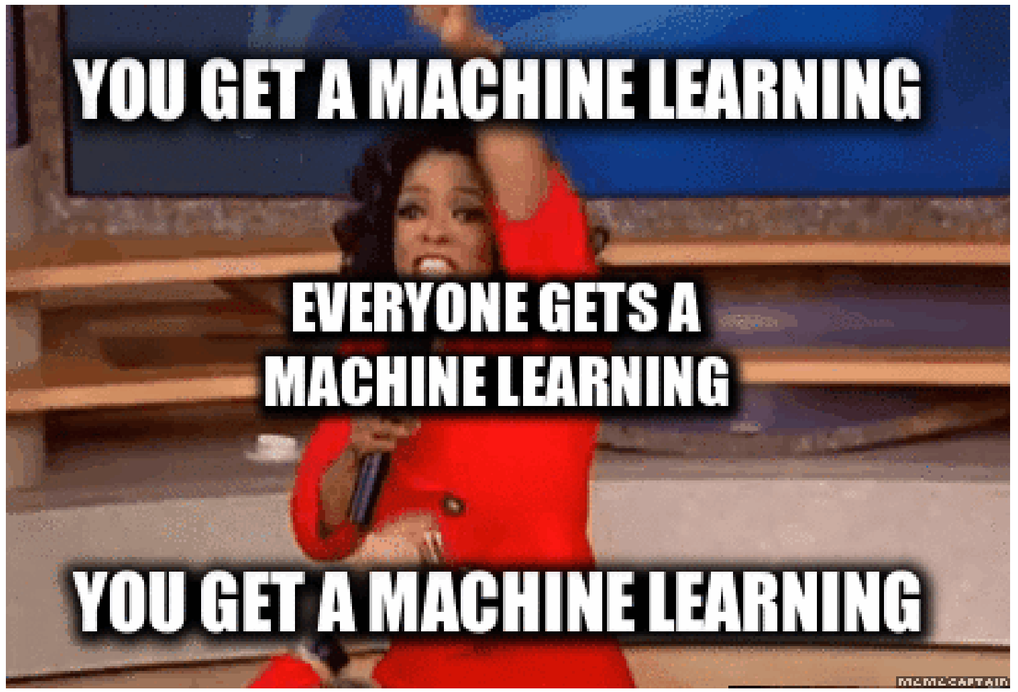
Highly Influential Citations	764
Background Citations	1,848
Methods Citations	2,663
Results Citations	45

View All

ML (DL) WHY NOW?

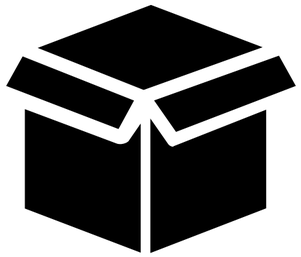

DATA (input)
Traditional Programming
Output
Program


DATA (input)
ML
MODEL (program)
Output


YOU GET A MACHINE LEARNING
EVERYONE GETS A MACHINE LEARNING
YOU GET A MACHINE LEARNING

Data-driven modeling

data
+
hyperparameters

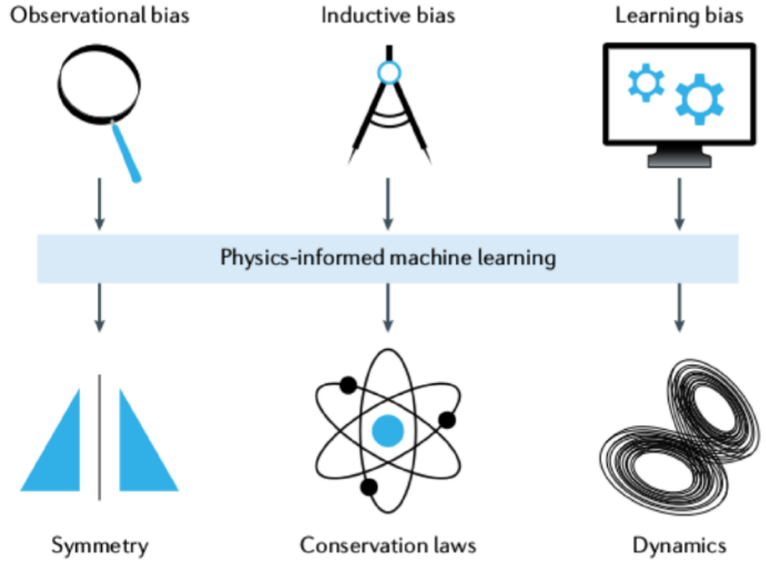


model (parameters)

data
+
hyperparameters
+
physics/constraints/domain knowledge

model
(parameters)

- Physical consistency (definitions, conservation laws…)
- Ability to generalize outside of the training set
- Interpretability
- Stability
- Data limitations



goal = enhance generalization

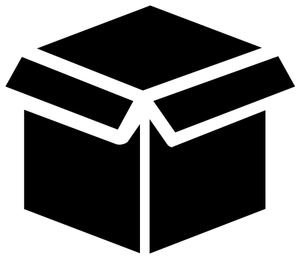
Physics-informed learning = process by which prior knowledge stemming from our observational, empirical, physical or mathematical understanding of the world can be leveraged to improve the performance of a learning algorithm

Karniadakiset al. Physics-informed machine learning. Nat Rev Phys (2021).

Data-driven modeling

- Should we give a chance to the black box?

data
+
hyperparameters

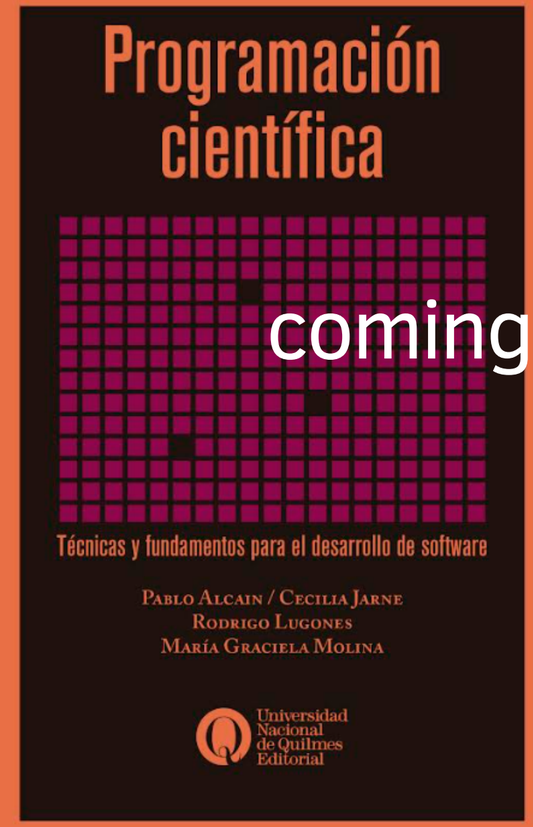


model
(parameters)

TSWC, 2022

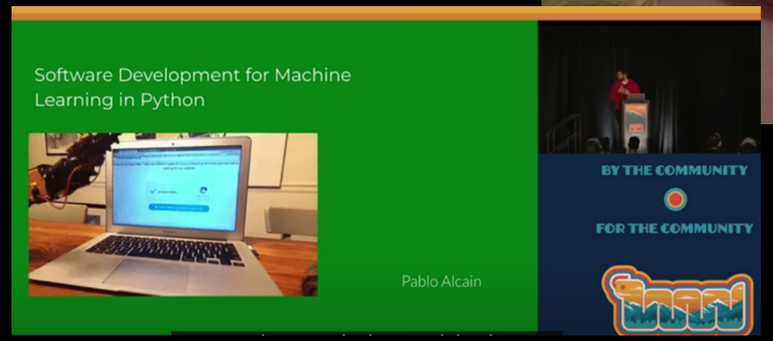
Scientific programming

WTPC | WORKSHOP EN TÉCNICAS DE PROGRAMACIÓN CIENTÍFICA



https://wtpc.github.io/

coming soon!

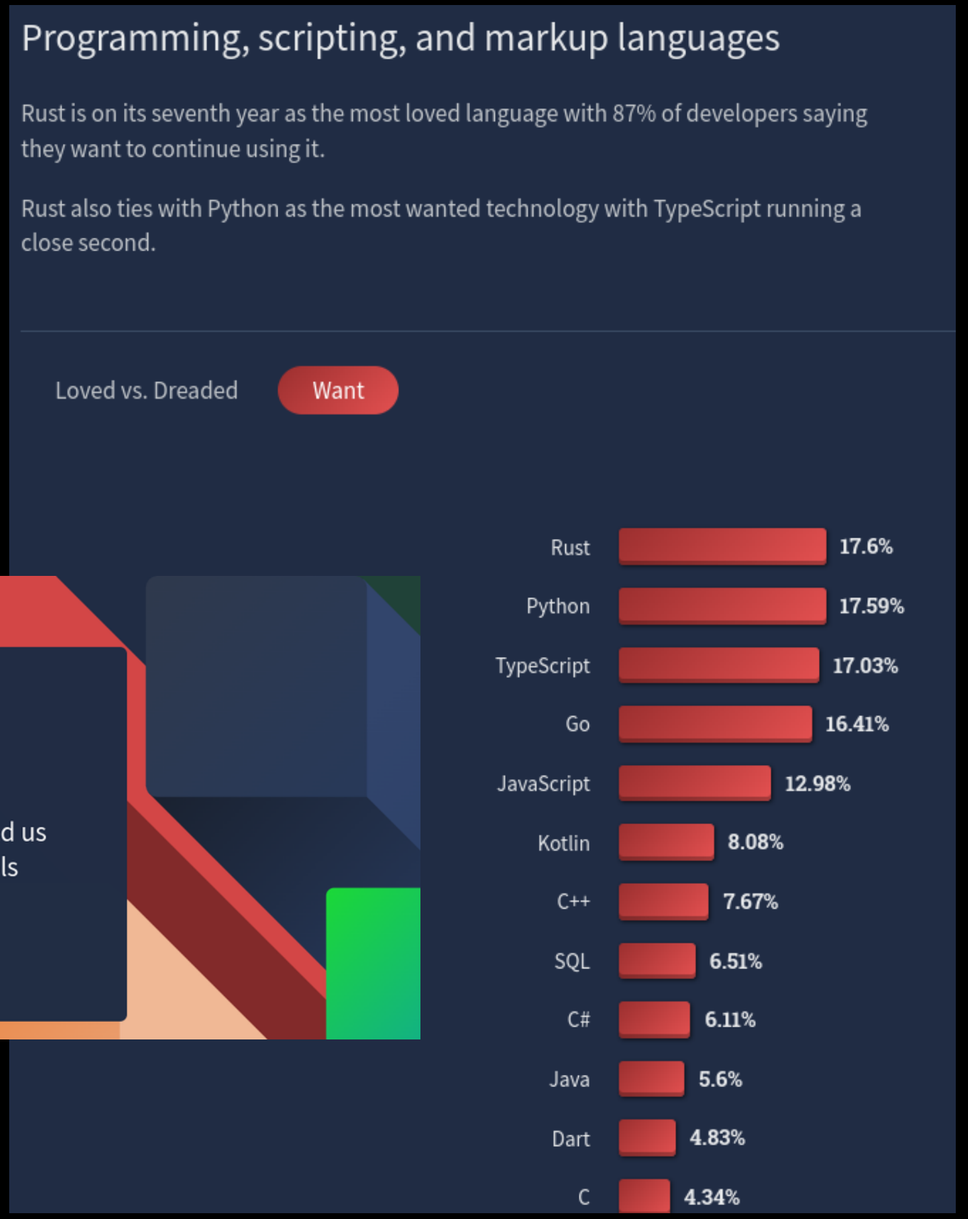


Python ?)

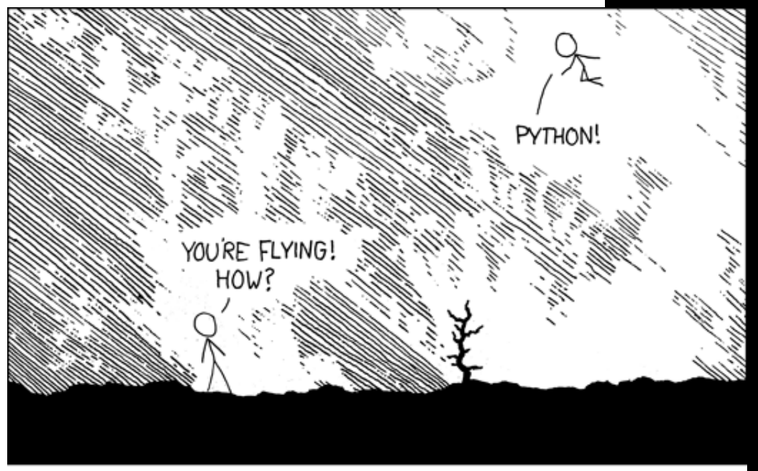
Programming Language Hall of Fame

Year	Winner
2021	Python
2020	Python
2019	C
2018	Python
2017	C
2016	Go
2015	Java
2014	JavaScript
2013	Transact-SQL
2012	Objective-C
2011	Objective-C
2010	Python

https://www.tiobe.com/tiobe-index/
(July 2022)



TSWC, 2022

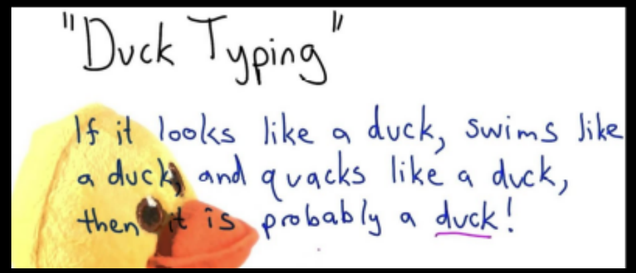


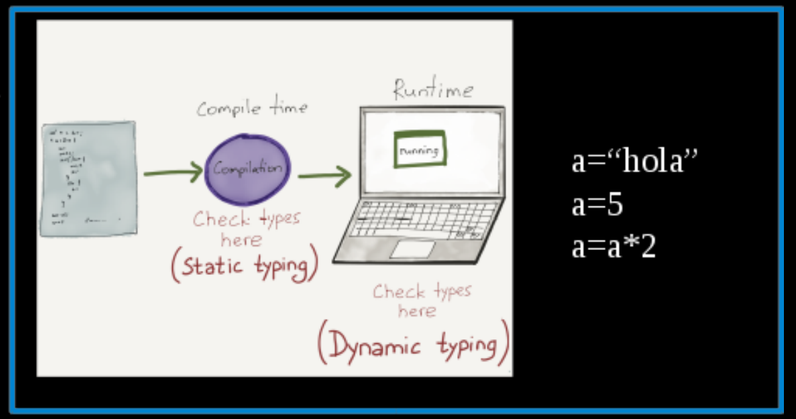
- Todo son objetos (de alguna clase)
- Fuertemente tipado
- Tipado dinámico
- Duck typing

Objeto	.	atributo
Objeto	.	método

"hola" + 5 → da error

```
def suma(a,b):
    return a+b
```



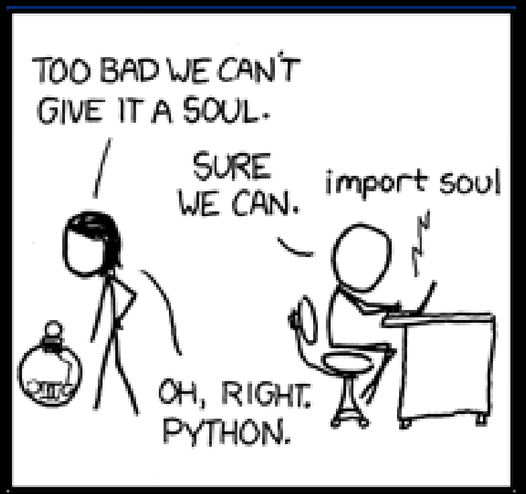



```
a="hola"
a=5
a=a*2
```

Python

NumPy pandas

Espacios de nombres (namespaces)
y Módulos



TSWC, 2022

HANDS-ON LAB

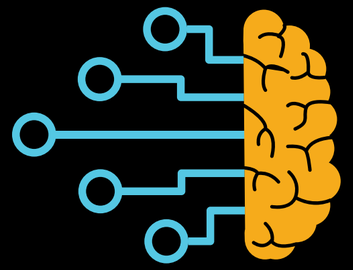
TP - 0

- Google Colabs - Jupyter Notebooks
- Codes & Data access
- Python refresh (numpy + Pandas)
- sklearn - Playing with a dataset (splitting and modeling)

APRENDIZAJE AUTOMÁTICO.

Fundamentos y Aplicaciones en Meteorología del Espacio

DCAO, UBA
08-12 AGOSTO 2022



Dra María Graciela Molina
FACET-UNT / CONICET
Tucumán Space Weather Center - TSWC

https://spaceweather.facet.unt.edu.ar/
IG -> @spaceweatherargentina

gmolina@herrera.unt.edu.ar