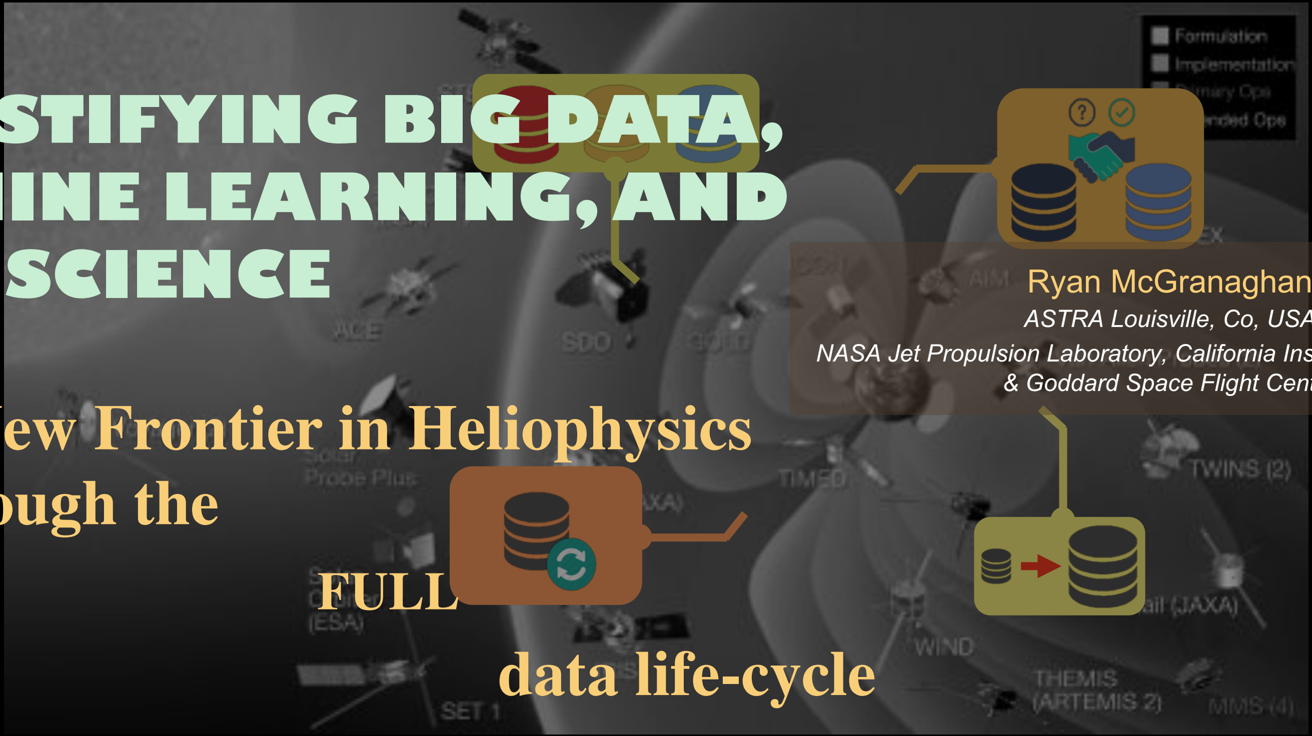
DEMYSTIFYING BIG DATA, MACHINE LEARNING, AND DATA SCIENCE

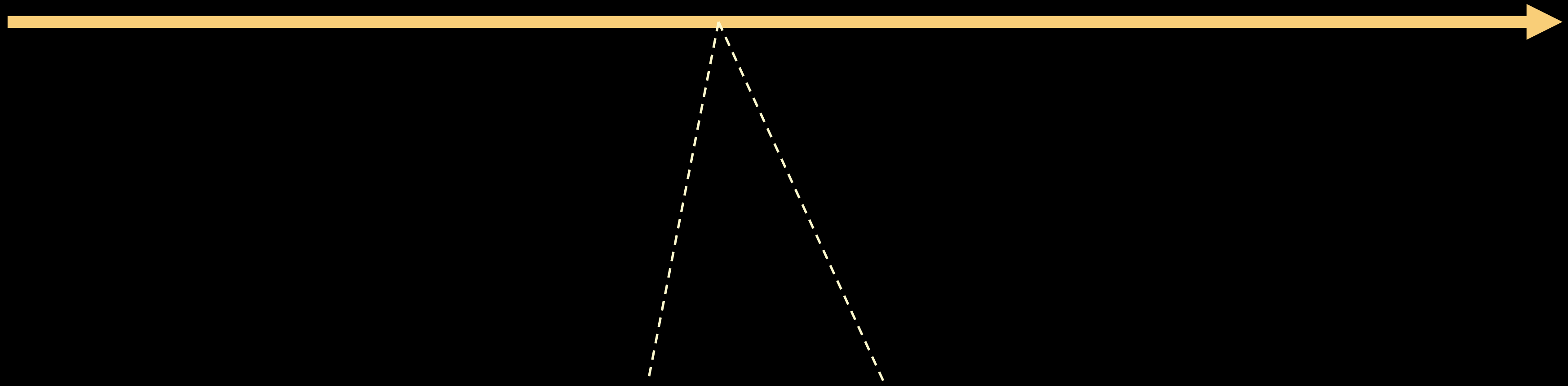
A New Frontier in Heliophysics through the FULL data life-cycle

Ryan McGranaghan
ASTRA Louisville, Co, USA
NASA Jet Propulsion Laboratory, California Institute of Technology & Goddard Space Flight Center

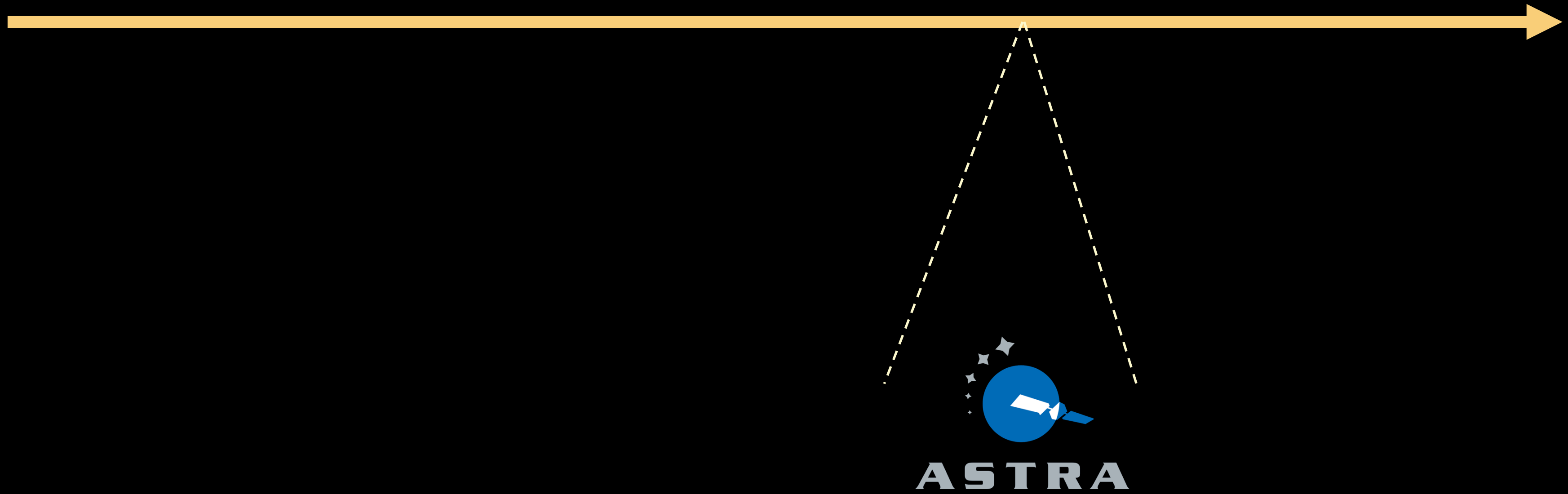


University of Colorado

DARTMOUTH



JPL


ASTRA


TEDx


Ushering in A NEW FRONTIER in HELIOPHYSICS with
Data Science

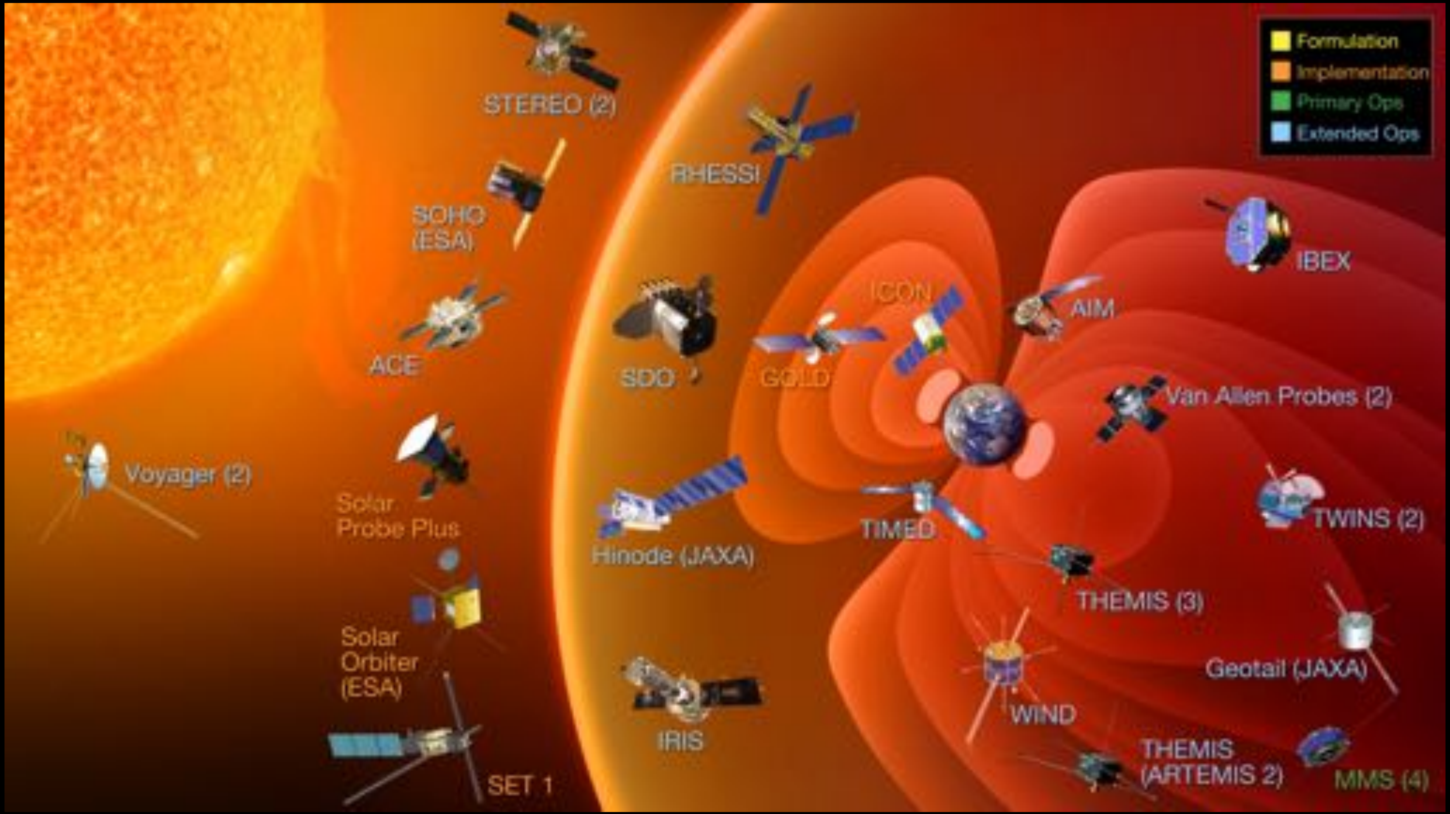

Agenda

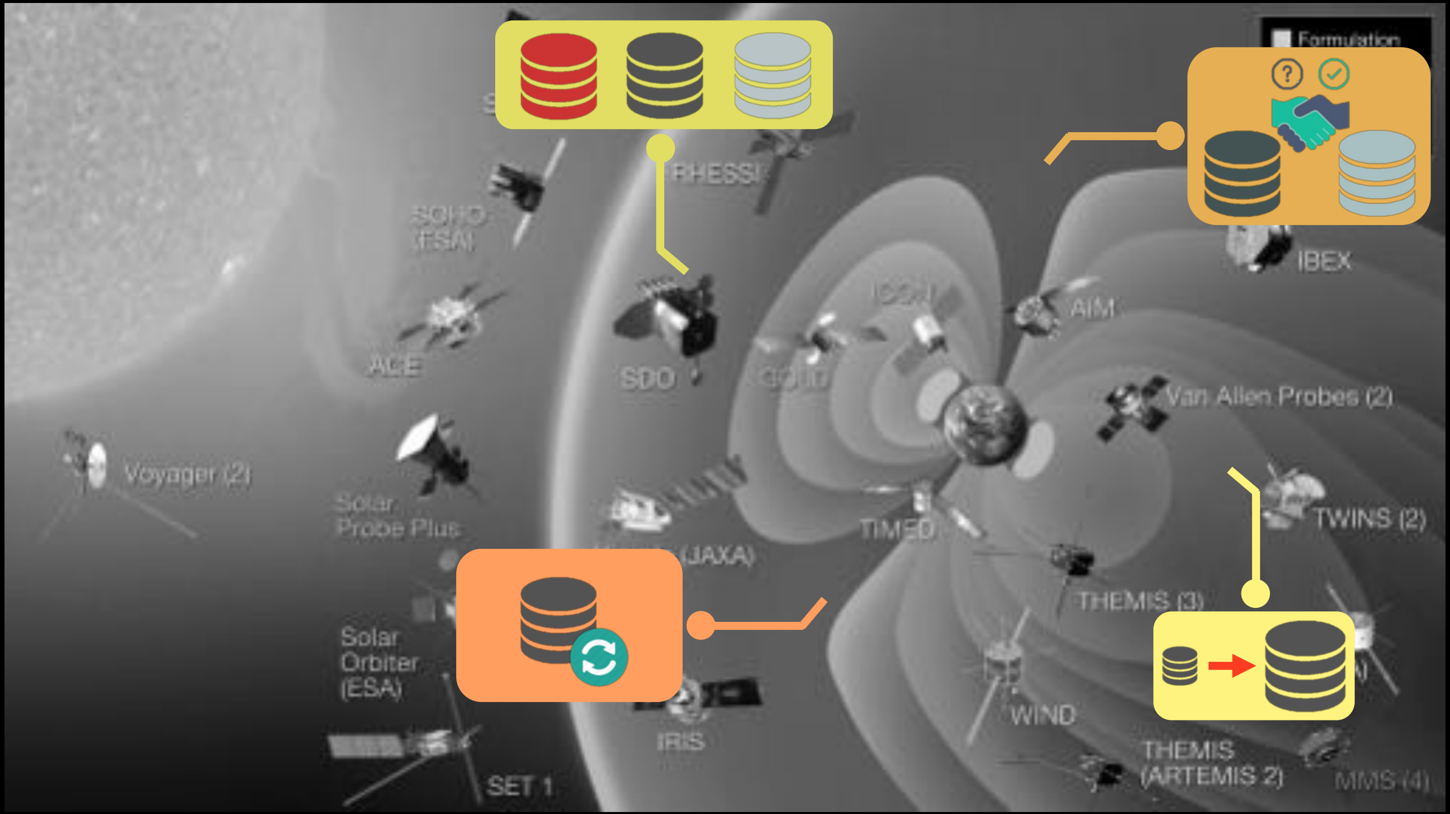
What is Data Science and why is it critical now?

What can we learn from the *fantastic* use cases across the full data lifecycle?

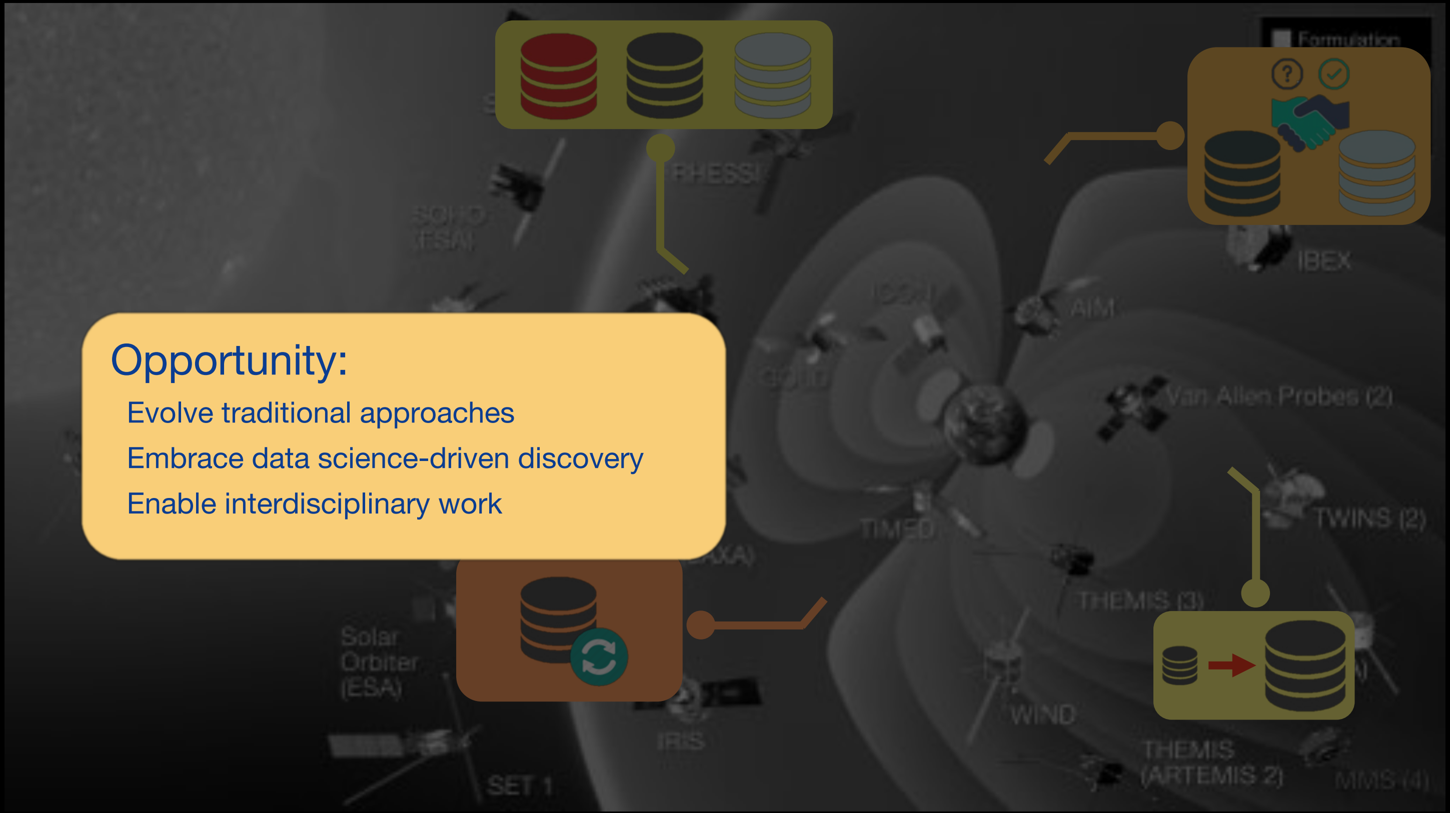
How do we shape the New Frontier?

What is Data Science and
why is it critical now?


Formulation
Implementation
Primary Ops
Extended Ops
STEREO (2)
RHESSI
SOHO (ESA)
ACE
SDO
GOLD
ICON
AIM
IBEX
Van Allen Probes (2)
Voyager (2)
Solar Probe Plus
Hinode (JAXA)
TIMED
TWINS (2)
THEMIS (3)
Solar Orbiter (ESA)
Geotail (JAXA)
WIND
IRIS
SET 1
THEMIS (ARTEMIS 2)
MMS (4)



Formulation
RHESSI
SOHO (ESA)
ACE
SDO
GOLD
ICON
AIM
IBEX
Van Allen Probes (2)
Voyager (2)
Solar Probe Plus
(JAXA)
TIMED
TWINS (2)
THEMIS (3)
Solar Orbiter (ESA)
WIND
IRIS
SET 1
THEMIS (ARTEMIS 2)
MMS (4)



Opportunity:

Evolve traditional approaches

Embrace data science-driven discovery

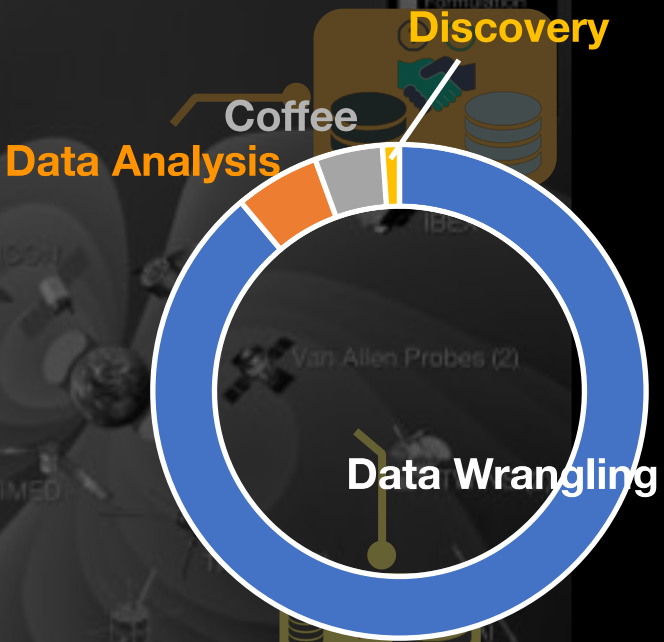
Enable interdisciplinary work

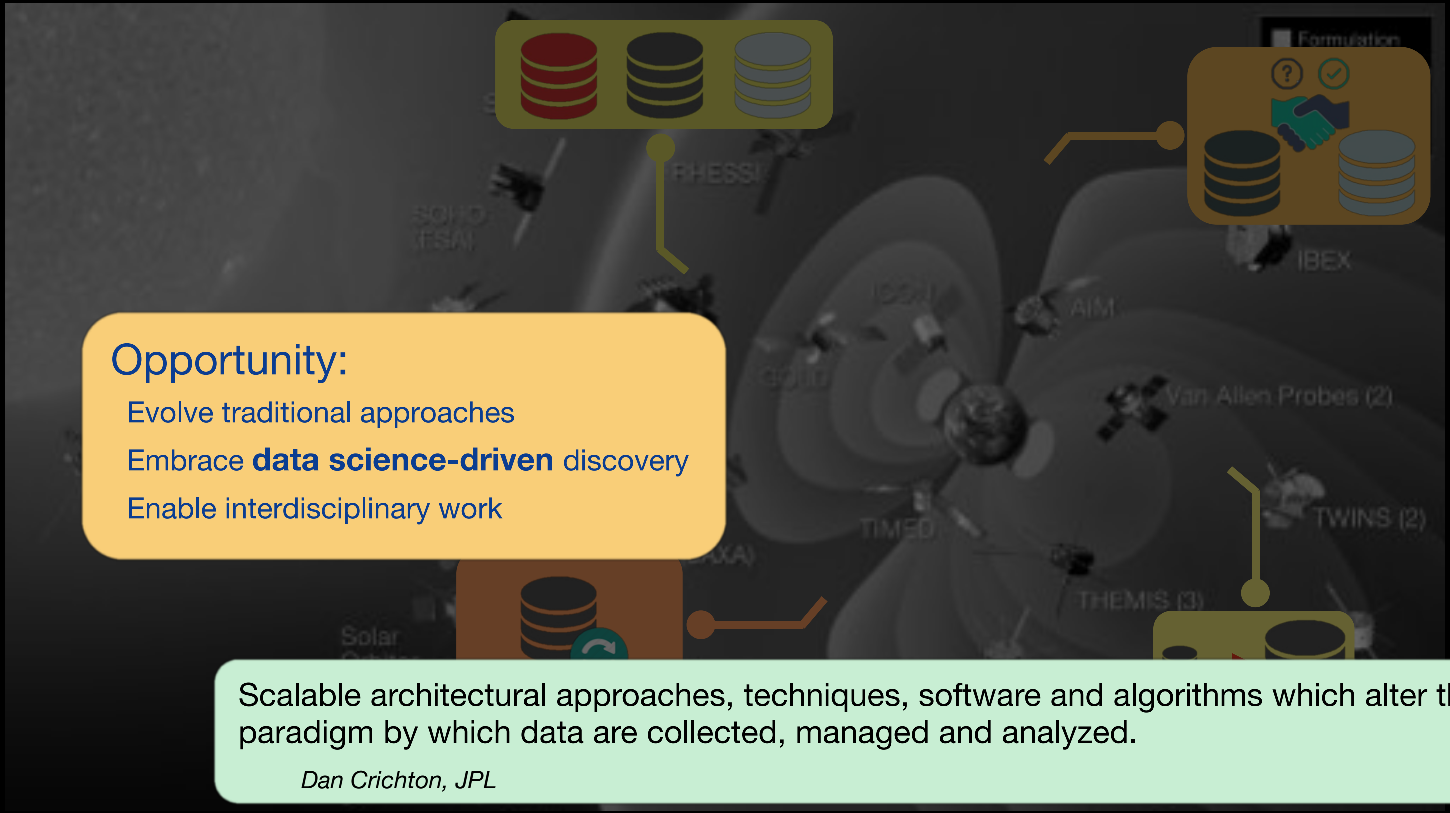
Opportunity:

Evolve traditional approaches

Embrace data science-driven discovery

Enable interdisciplinary work



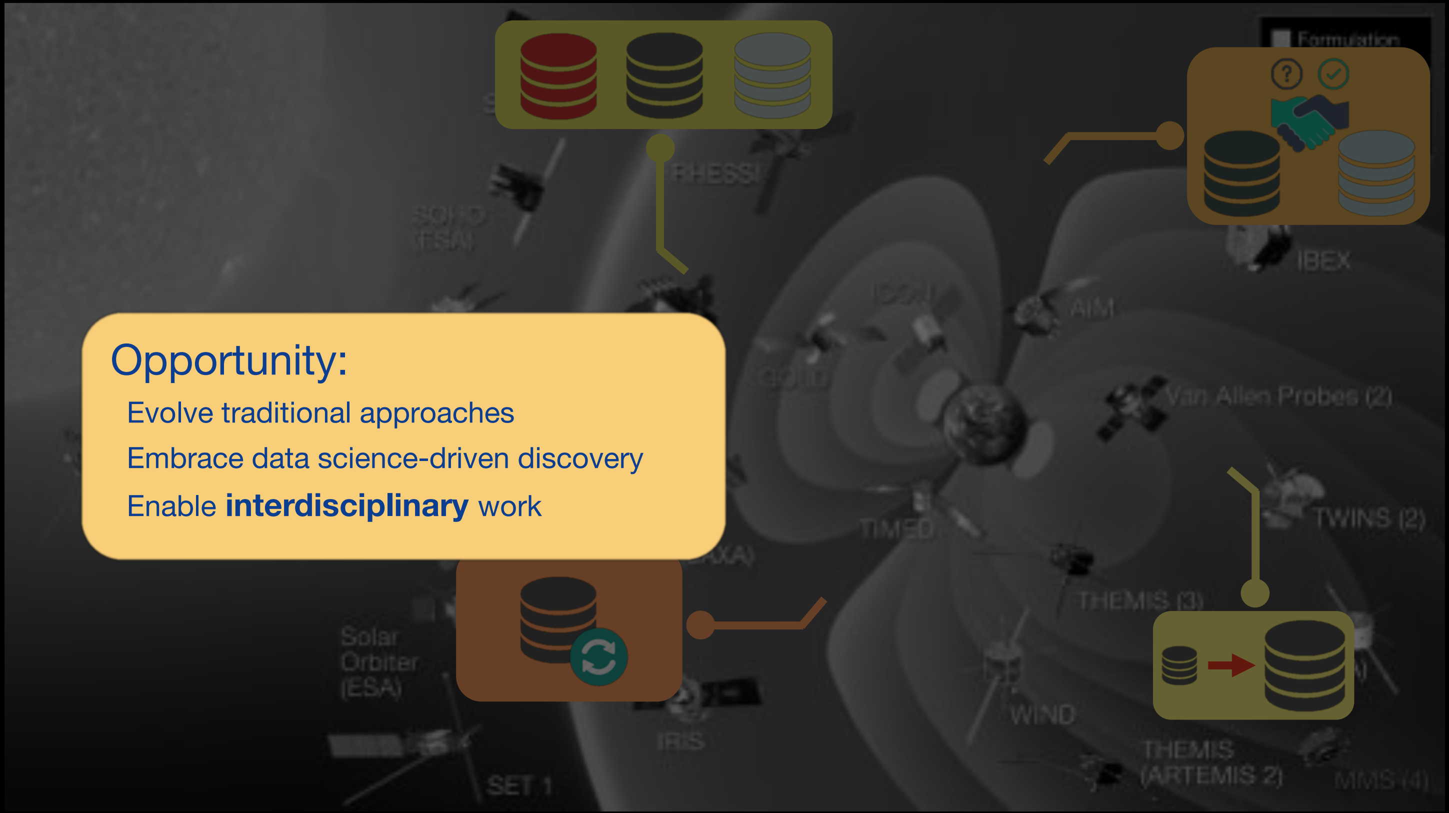


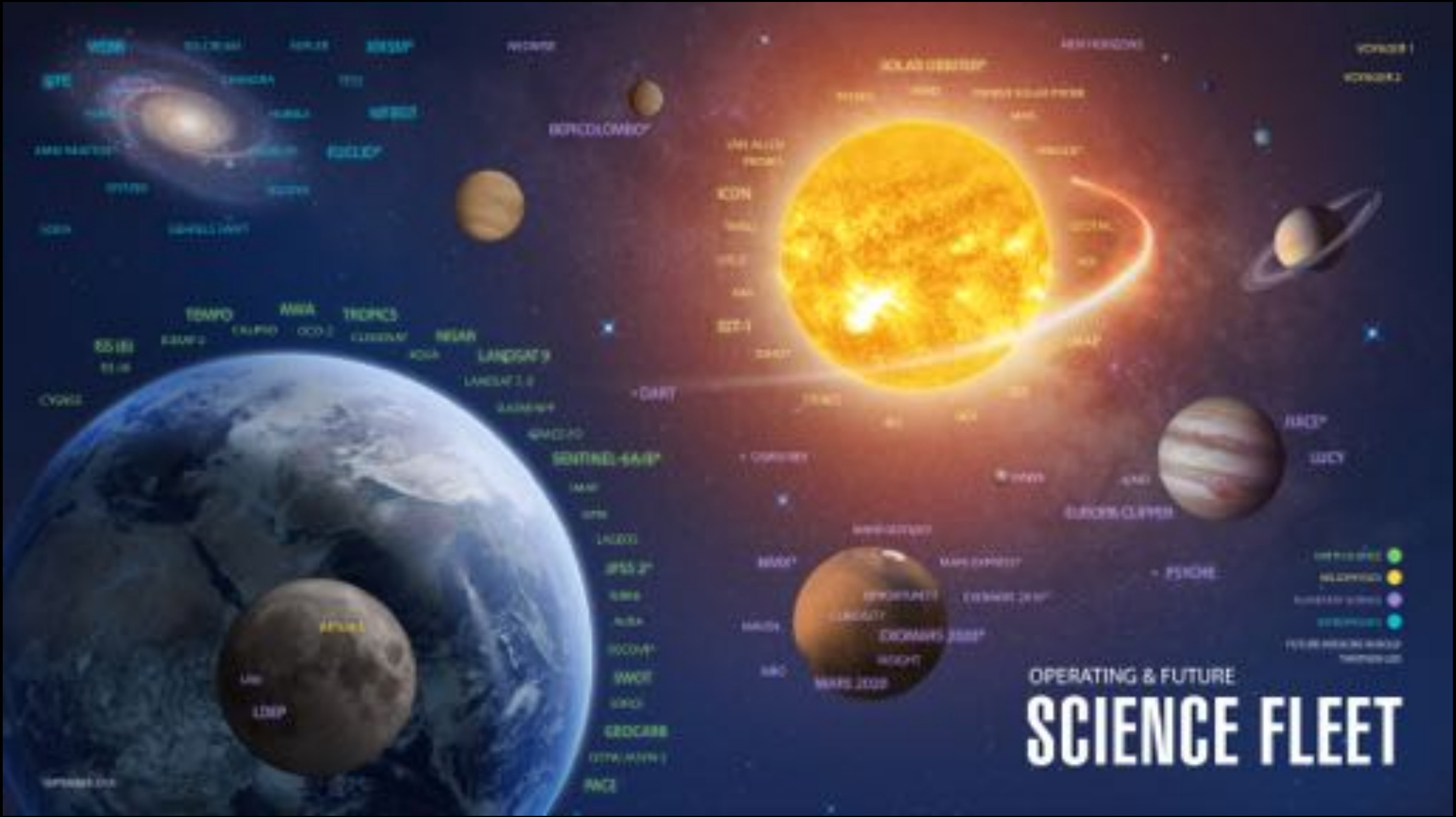
Opportunity:

- Evolve traditional approaches
- Embrace **data science-driven** discovery
- Enable interdisciplinary work

Scalable architectural approaches, techniques, software and algorithms which alter the paradigm by which data are collected, managed and analyzed.

Dan Crichton, JPL


Opportunity:
Evolve traditional approaches
Embrace data science-driven discovery
Enable **interdisciplinary** work
RHESSI
SOHO (ESA)
Formulation
IBEX
AIM
Van Allen Probes (2)
TWINS (2)
TIMED
THEMIS (3)
Solar Orbiter (ESA)
WIND
IRIS
SET 1
THEMIS (ARTEMIS 2)
MMS (4)

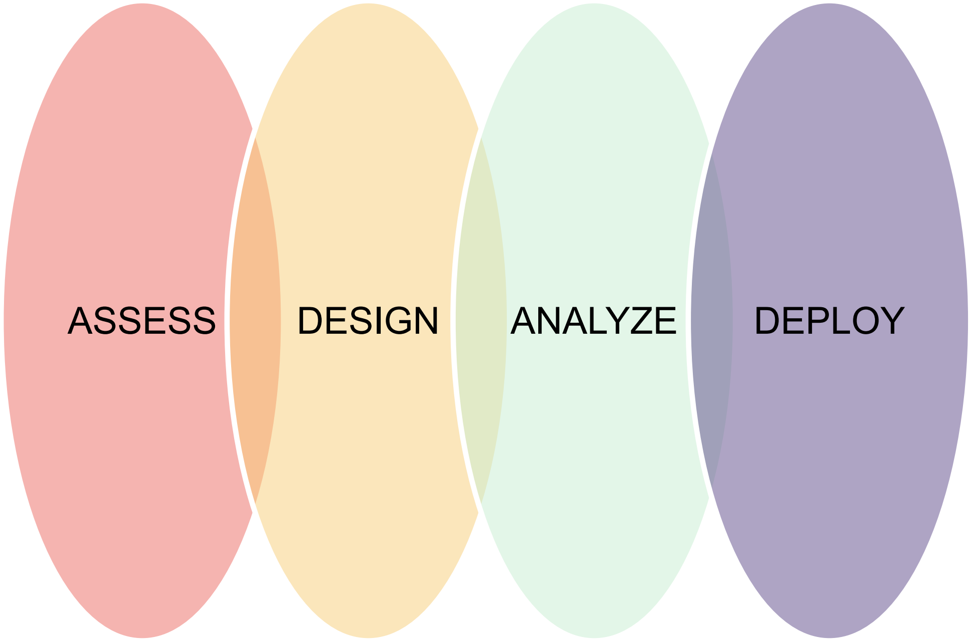


OPERATING & FUTURE
SCIENCE FLEET
TEMPO
MAIA
TROPICS
NISAR
LANDSAT 9
CYGNSS
DART
ICON
EUCLID*
BEPICOLOMBO*
SOLAR ORBITER*
LUCY
EUROPA CLIPPER
PSYCHE
MARS 2020
CURIOSITY
LRO
LDEP
SWOT
GEOCARB
PACE


OPERATING & FUTUR
SCIENCE
TEMPO
MAIA
TROPICS
NISAR
LANDSAT 9
SENTINEL-6A/6B
ICON
DART
EUROPA CLIPPER
PSYCHE
LUCY
MARS 2020

What can we learn from the *fantastic* use cases across the full data lifecycle?

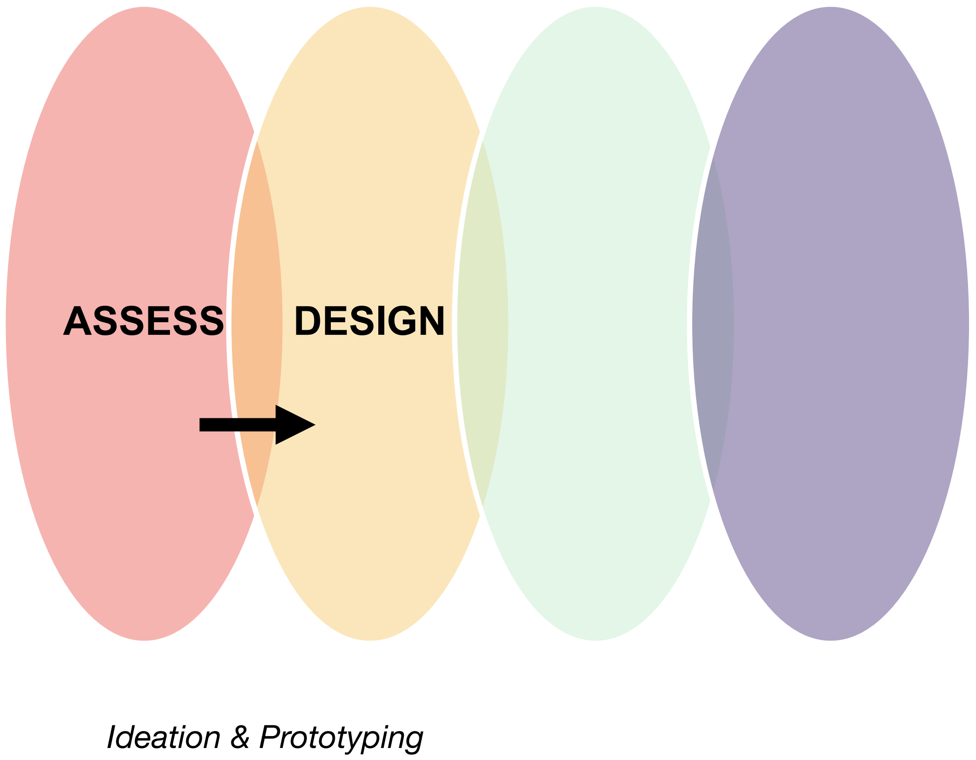
The Data Science Project Workflow



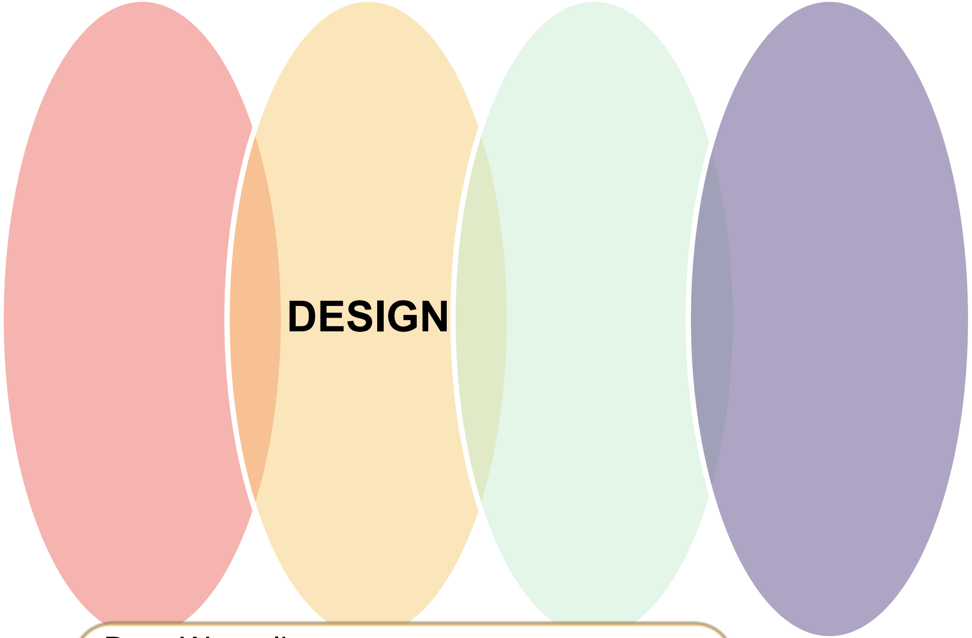
The Data Science Project Workflow

What is the explicit goal?
What data are available?
What is the state-of-the-art?
What will be the central challenges?

The Data Science Project Workflow



The Data Science Project Workflow



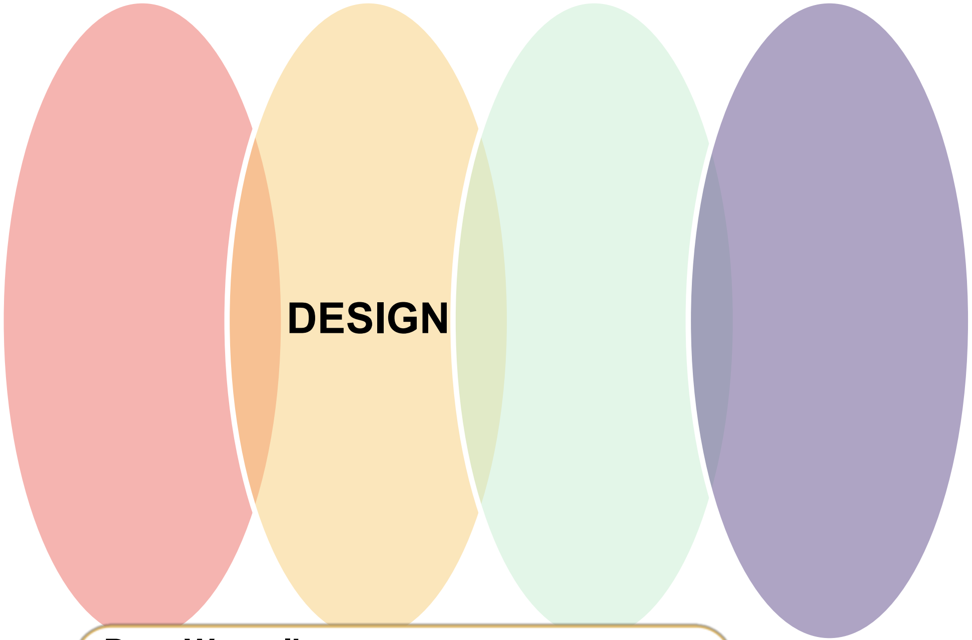
Data Wrangling

Develop potential solutions

- Experiment and evaluate

Visualize and explore the data and solutions

The Data Science Project Workflow



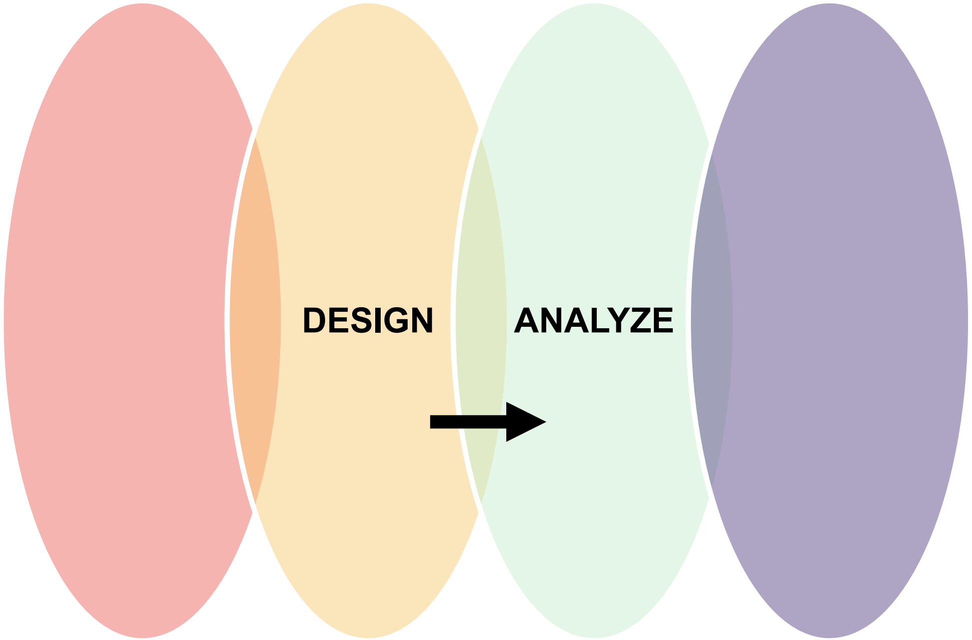
Data Wrangling

Develop potential solutions

- Experiment and evaluate

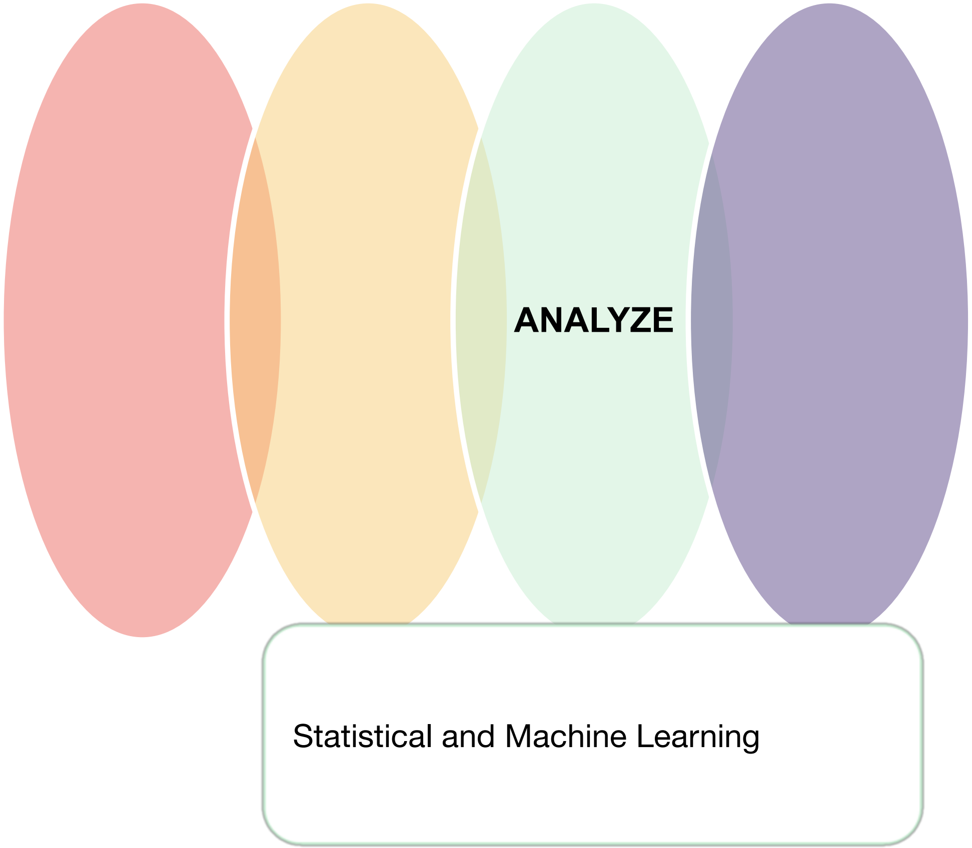
Visualize and explore the data and solutions

The Data Science Project Workflow



Analytical Innovation

The Data Science Project Workflow



The Data Science Project Workflow

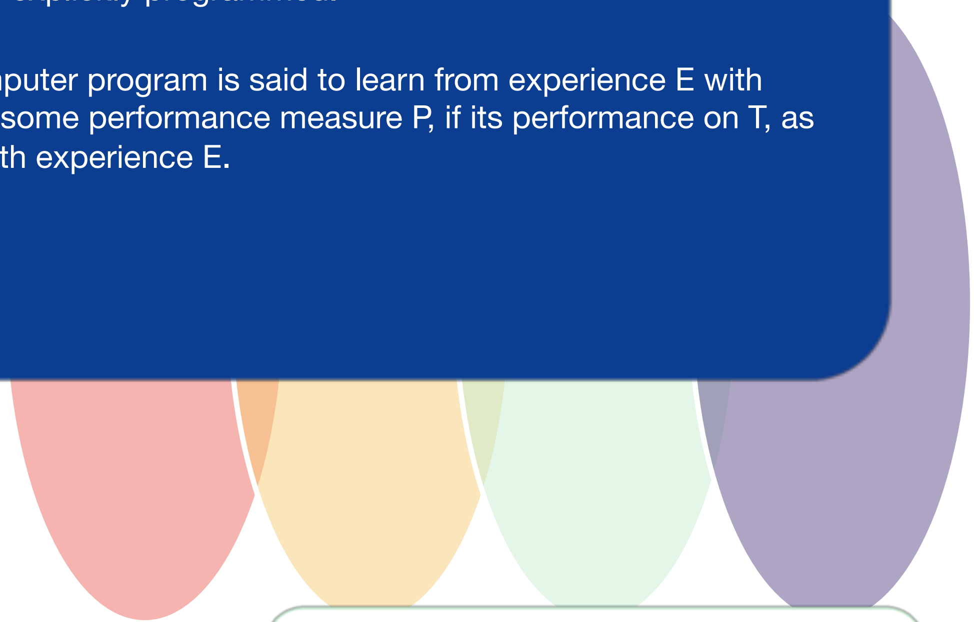
Arthur Samuel (1959) – Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed.

Statistical and **Machine Learning**

The Data Science Project Workflow

Arthur Samuel (1959) – Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed.

Tom Mitchell (1998) – A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.



Statistical and **Machine Learning**

The Data Science Project Workflow

Arthur Samuel (1959) – Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed.

Tom Mitchell (1998) – A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

Support Vector Machine (SVM) | **Decision Trees** | **Random Forests** | **Neural Networks**

Easily explainable → *Difficult to explain*

Create a narrative of new scientific understanding across spectrum of machine learning approaches

Statistical and **Machine Learning**

The Data Science Project Workflow

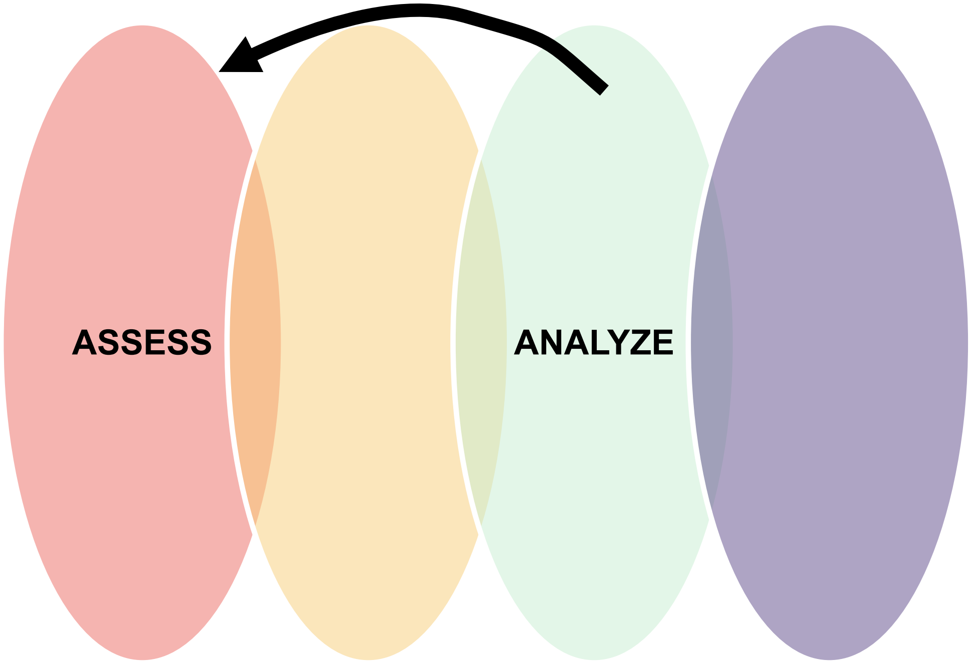
Arthur Samuel (1959) – Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed.

Tom Mitchell (1998) – A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

So, your success depends on how well you:
1) define your **task**
2) understand, populate, and manage your **sample set**
3) set up the means by which your **performance** is assessed

Statistical and **Machine Learning**

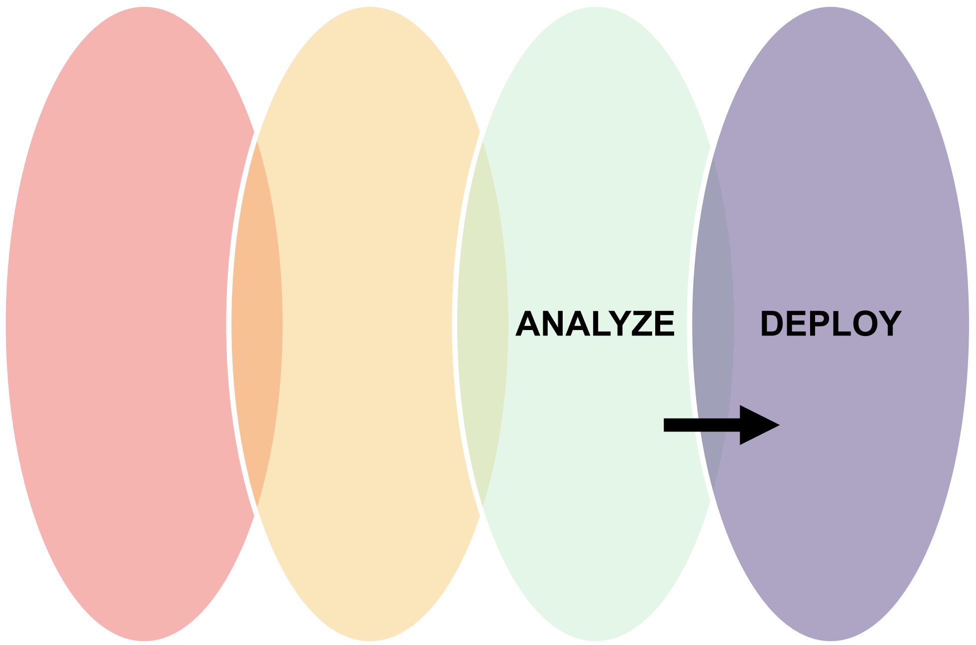
The Data Science Project Workflow



Interrogate the models

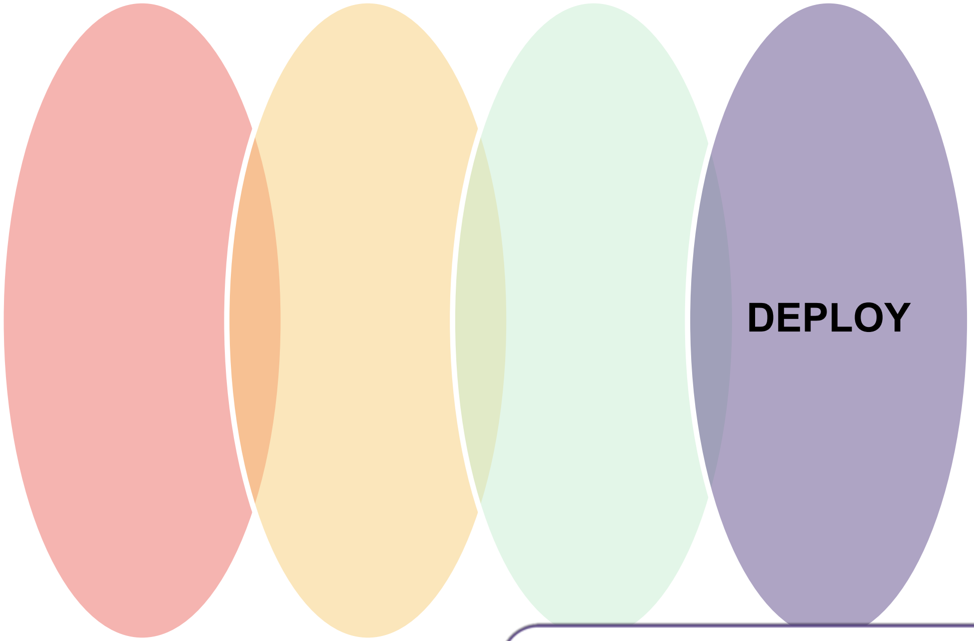
Feedback & Iteration

The Data Science Project Workflow



Add Capability & Refine

The Data Science Project Workflow

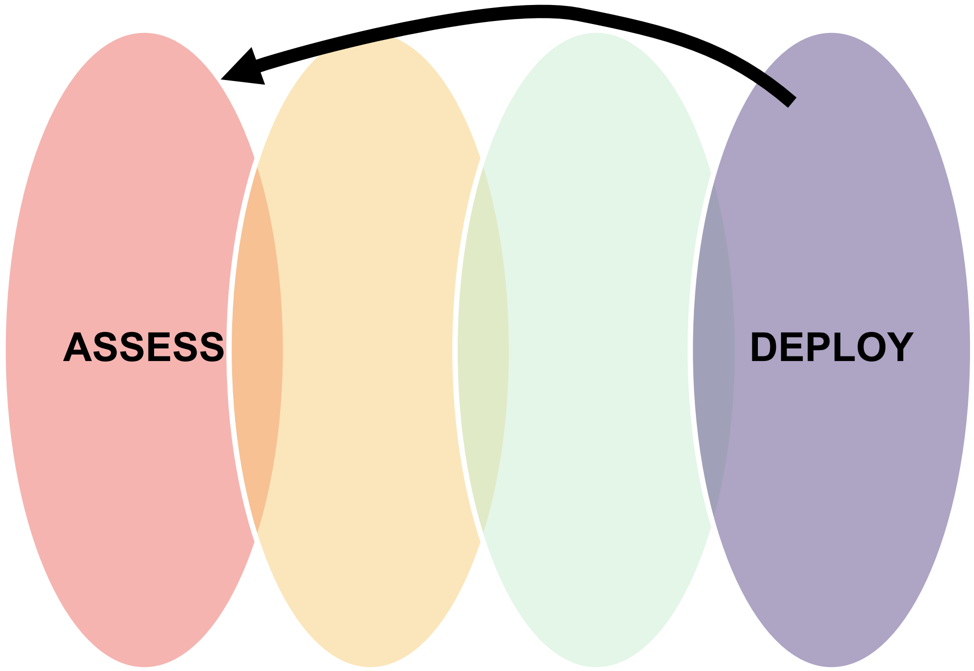


Communicate results (scientific discovery/operations)

Develop dashboards

Automate

The Data Science Project Workflow

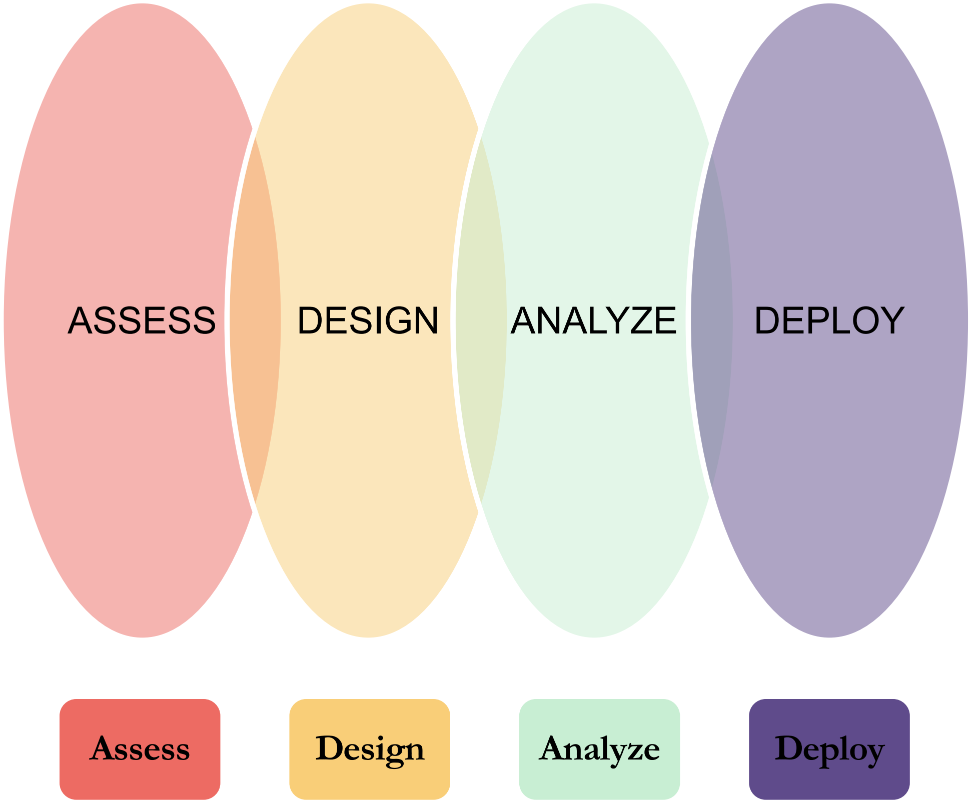


Interrogate the models

Feedback & Iteration

The Data Science Project Workflow

Applied



Assess

Design

Analyze

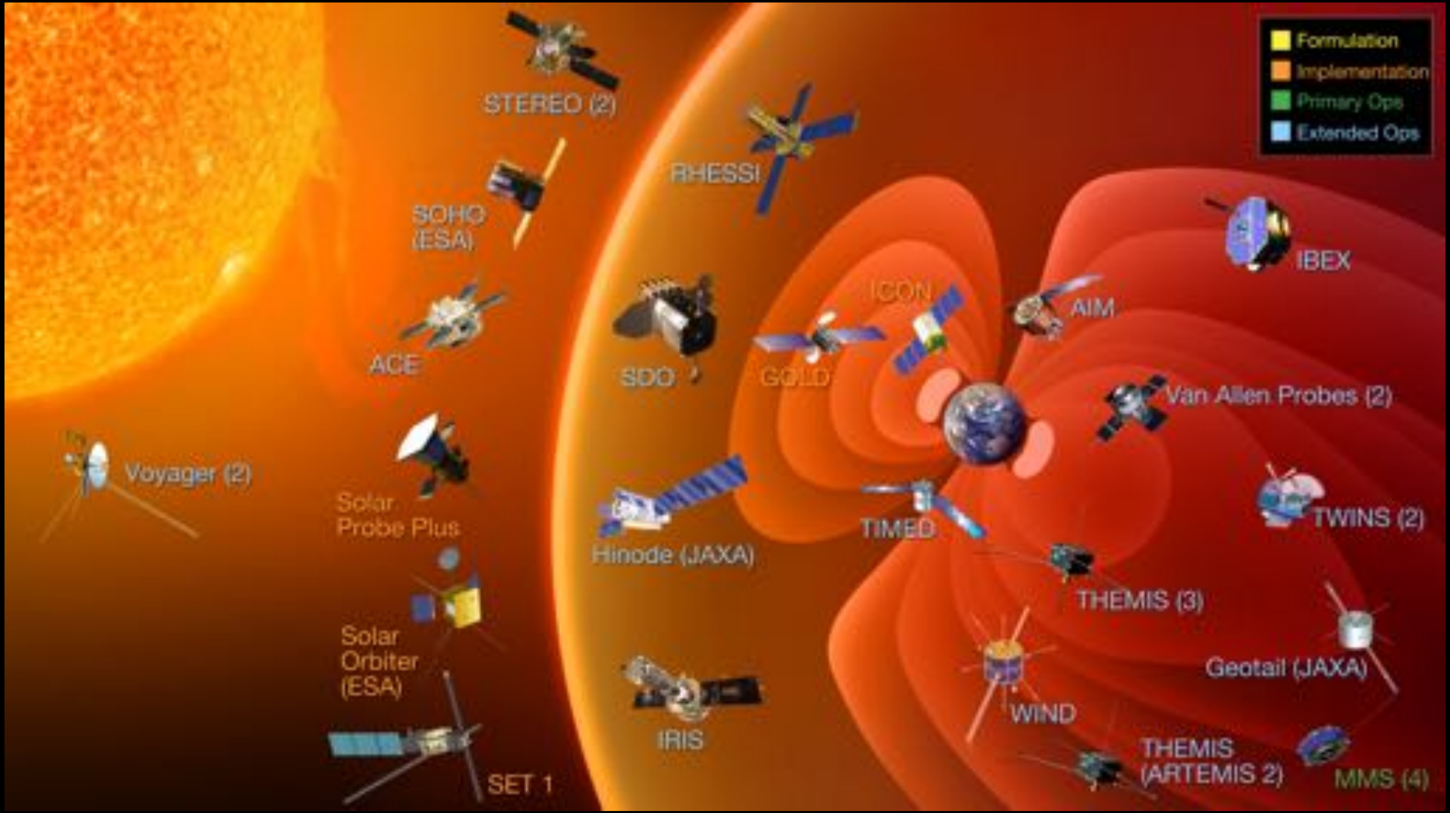
Deploy

Assess

Design

Analyze

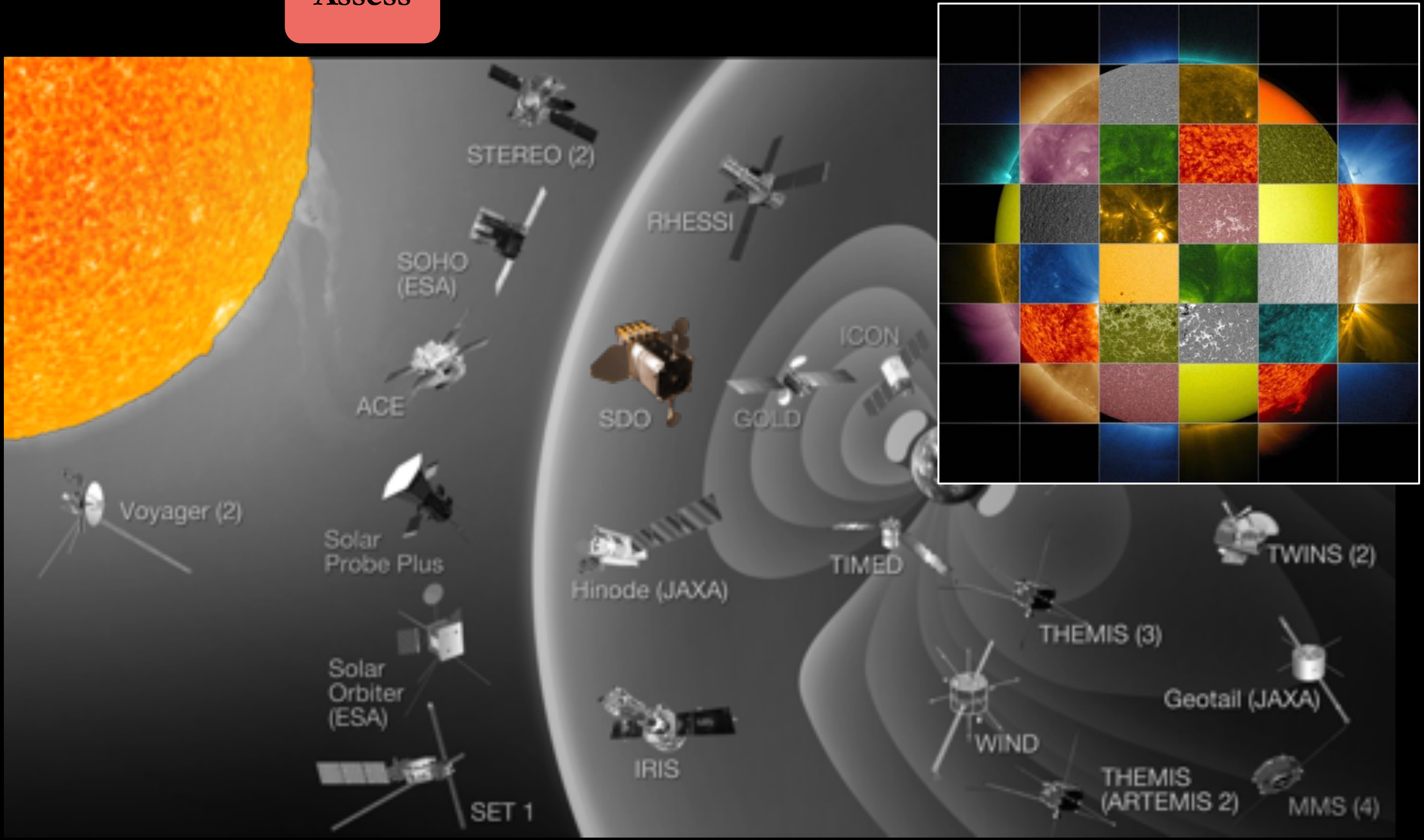
Deploy



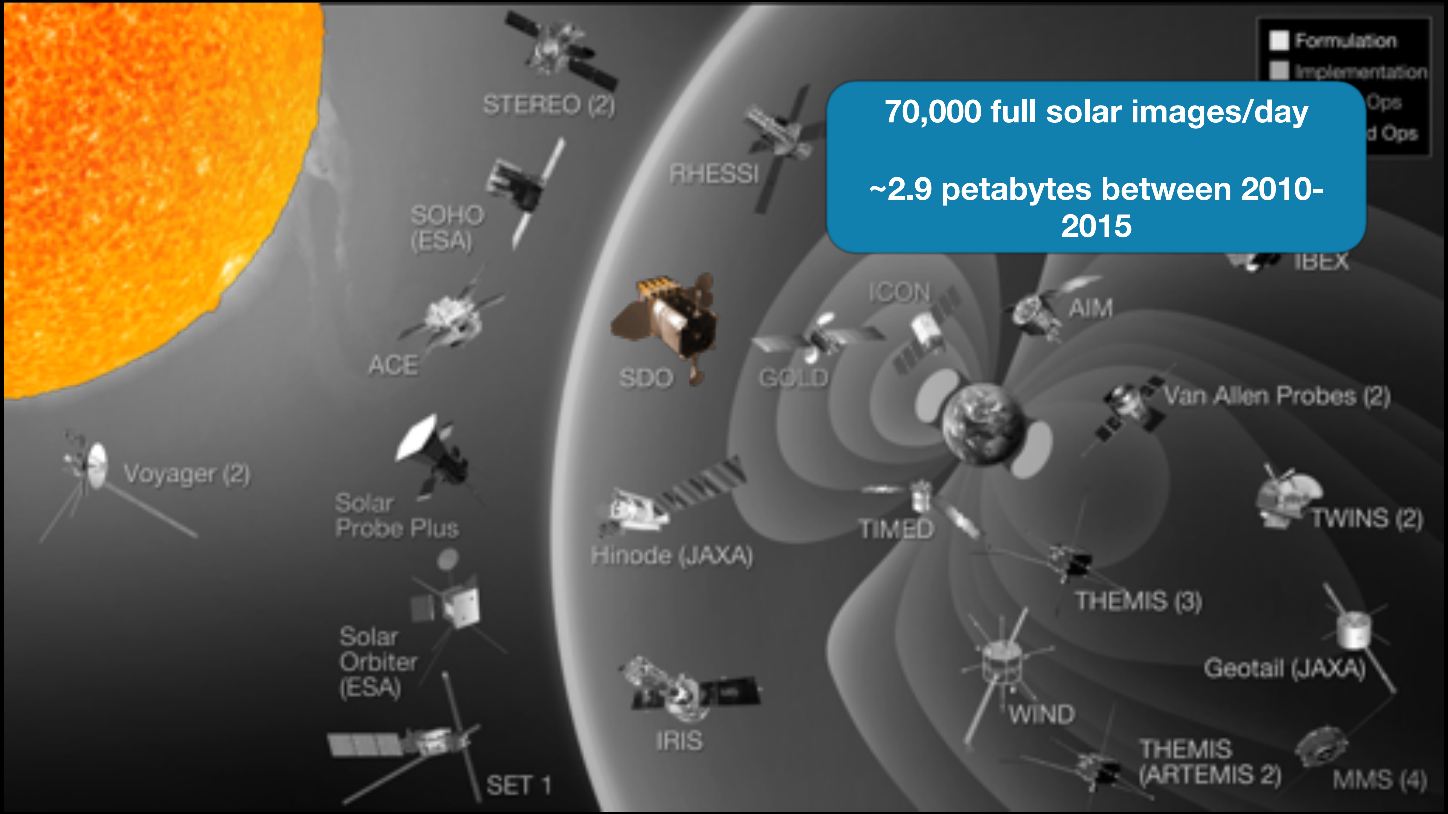
Assess


Formulation
Implementation
Primary Ops
Extended Ops
STEREO (2)
RHESSI
SOHO
(ESA)
IBEX
ICON
AIM
ACE
SDO
GOLD
Van Allen Probes (2)
Voyager (2)
Solar
Probe Plus
Hinode (JAXA)
TIMED
TWINS (2)
THEMIS (3)
Geotail (JAXA)
Solar
Orbiter
(ESA)
WIND
IRIS
THEMIS
(ARTEMIS 2)
MMS (4)
SET 1

Assess


STEREO (2)
RHESSI
SOHO (ESA)
ACE
SDO
GOLD
ICON
Voyager (2)
Solar Probe Plus
Hinode (JAXA)
TIMED
TWINS (2)
THEMIS (3)
Solar Orbiter (ESA)
Geotail (JAXA)
WIND
IRIS
SET 1
THEMIS (ARTEMIS 2)
MMS (4)

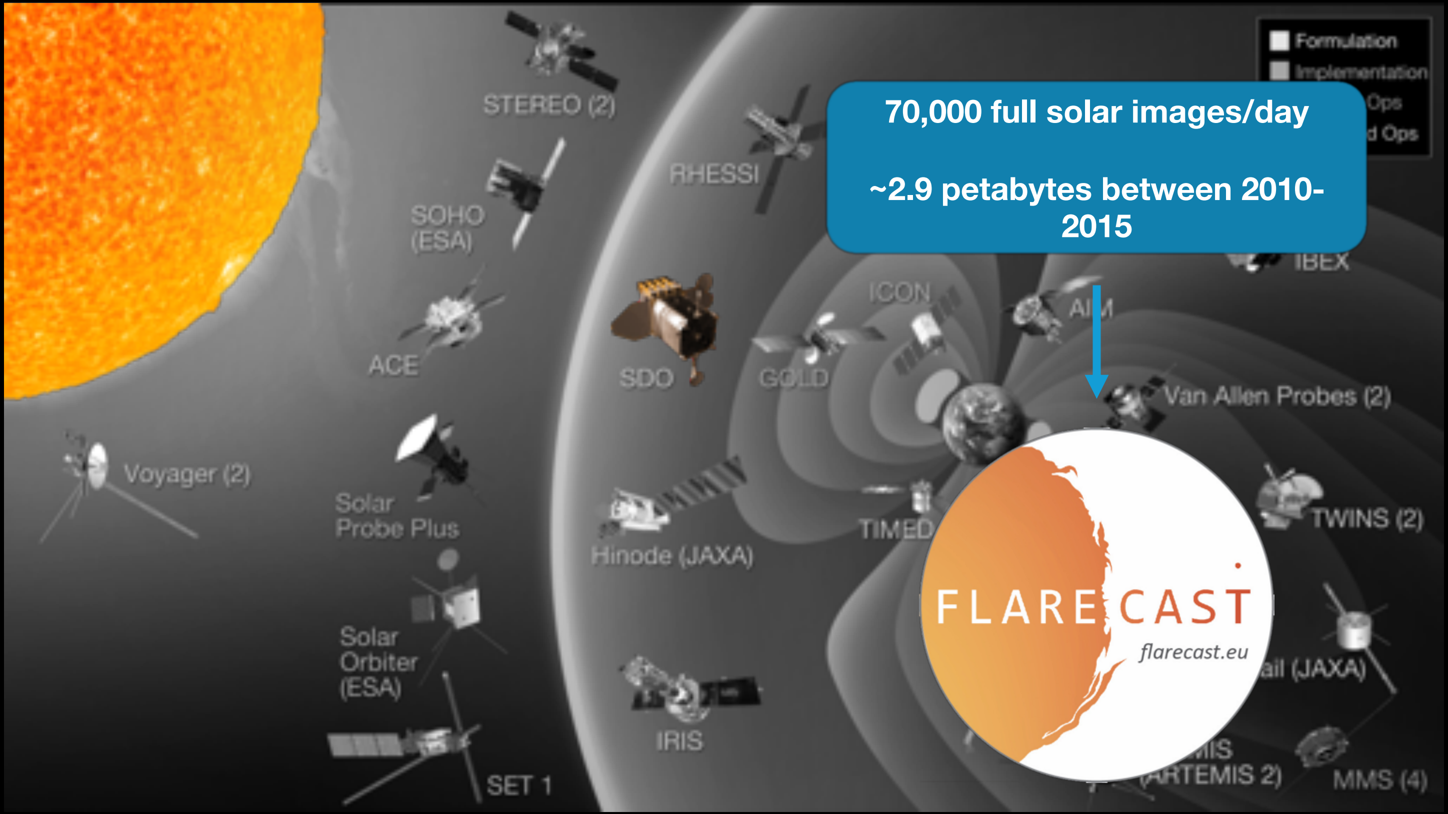
Assess



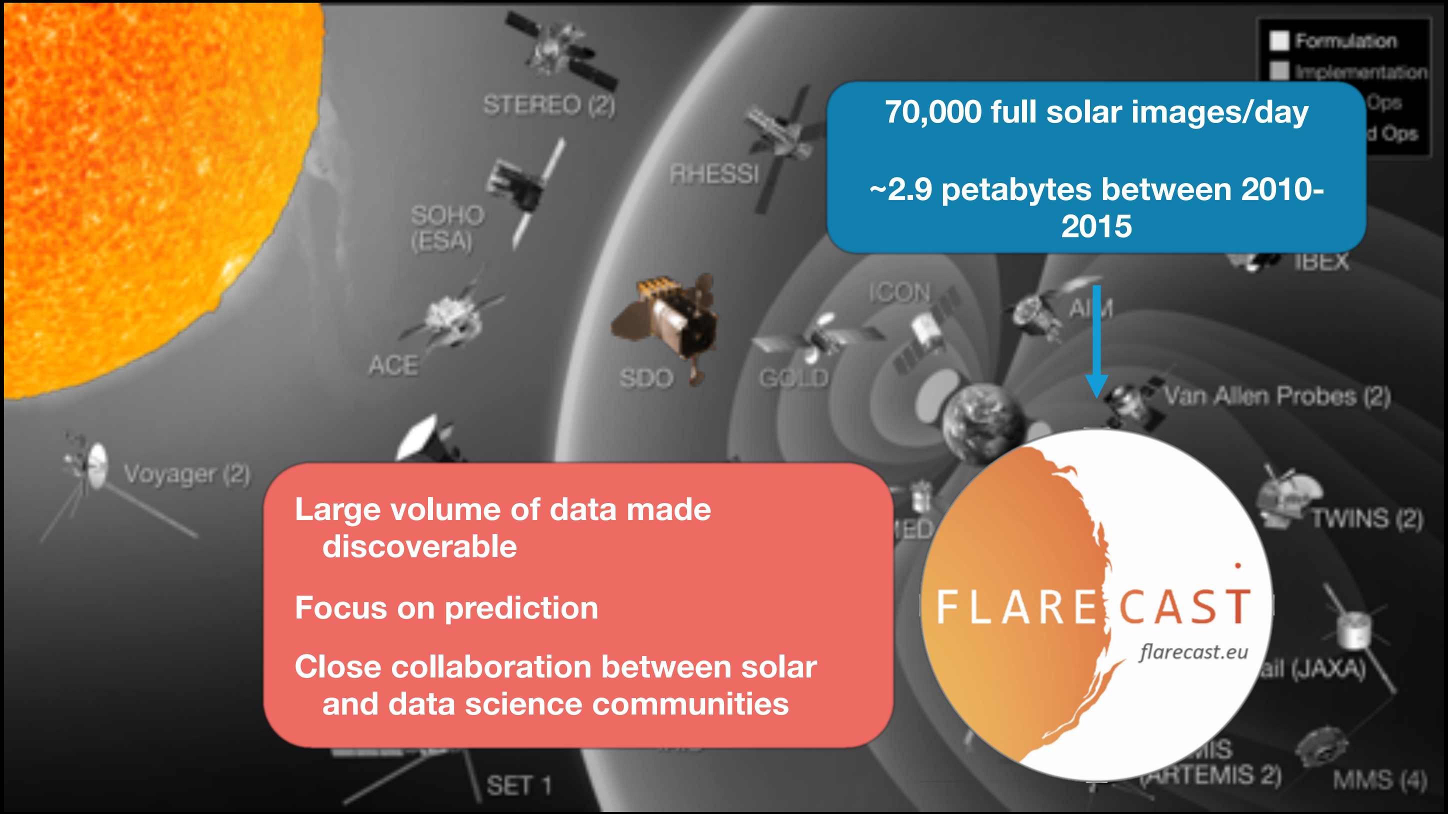
70,000 full solar images/day

~2.9 petabytes between 2010-2015

Design


STEREO (2)
SOHO (ESA)
ACE
Voyager (2)
Solar Probe Plus
Solar Orbiter (ESA)
SET 1
RHESSI
SDO
Hinode (JAXA)
IRIS
GOLD
ICON
TIMED
AIM
IBEX
Van Allen Probes (2)
TWINS (2)
MMS (4)
Formulation
Implementation
70,000 full solar images/day
~2.9 petabytes between 2010-2015
FLARECAST
flarecast.eu

Analyze



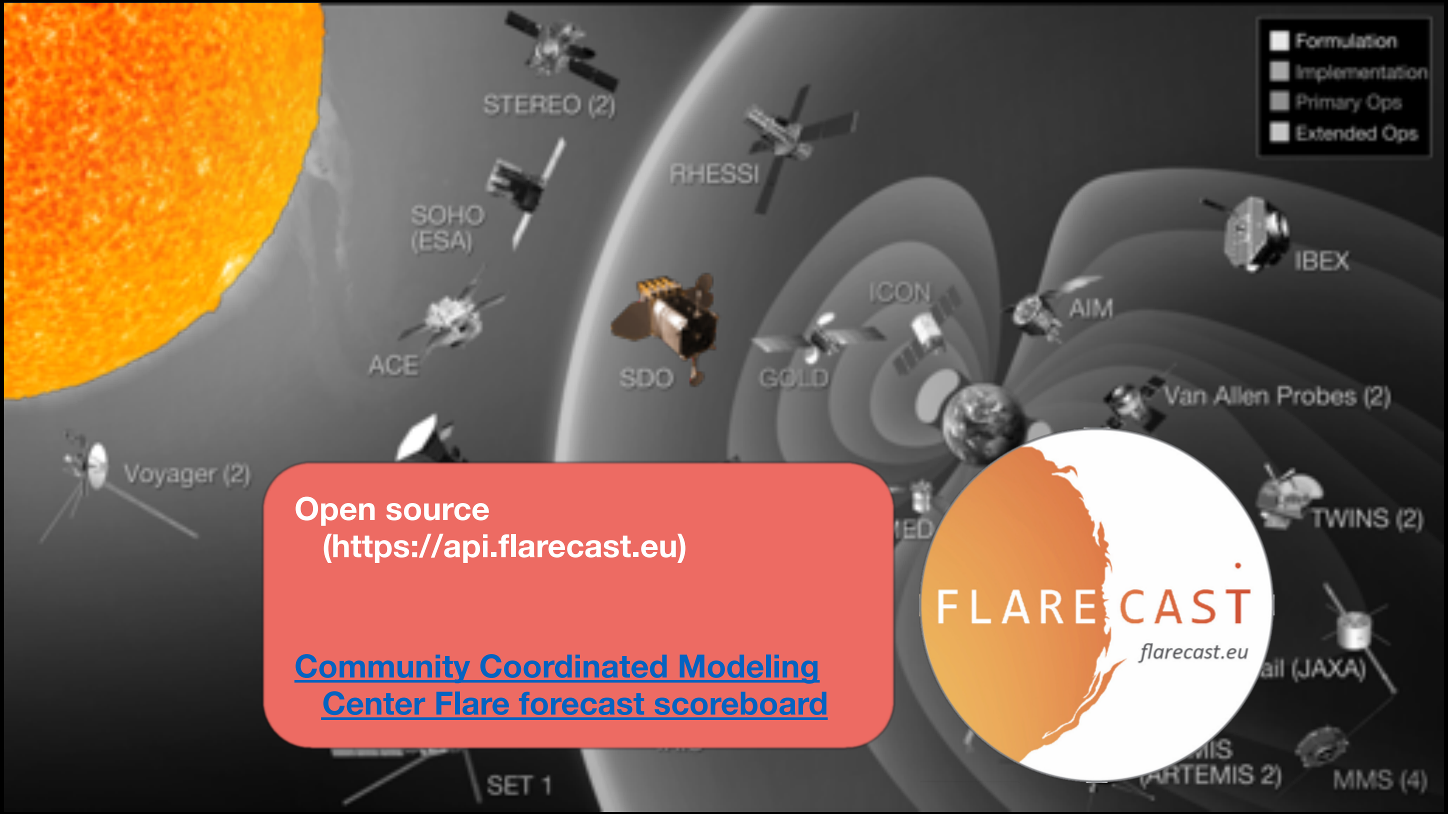
70,000 full solar images/day

~2.9 petabytes between 2010-2015

- Large volume of data made discoverable
- Focus on prediction
- Close collaboration between solar and data science communities

FLARECAST
flarecast.eu

Deploy



Open source
(https://api.flarecast.eu)

[Community Coordinated Modeling Center Flare forecast scoreboard]

FLARECAST
flarecast.eu

Assess

Design

Analyze

Deploy

Assess Design Analyze Deploy

STRETCHING GNSS SIGNALS FOR SPACE WEATHER DISCOVERY

Ryan McGranaghan, Anthony Mannucci

University Corporation for Atmospheric Research (UCAR)

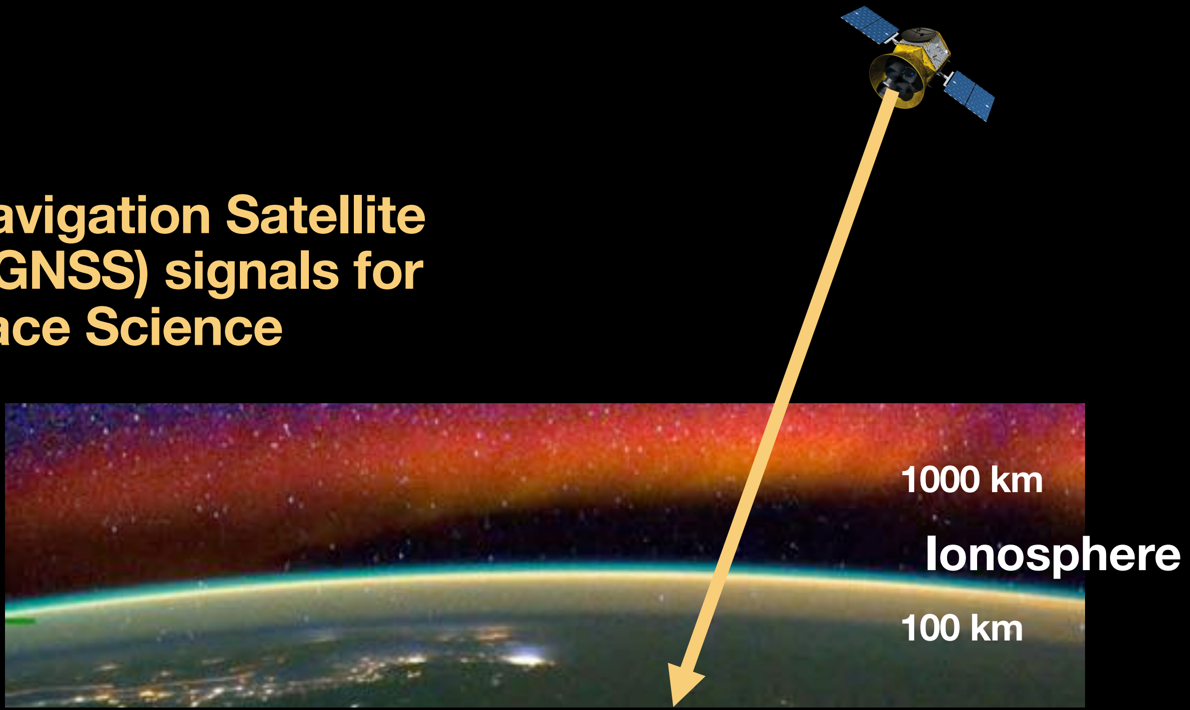
NASA Jet Propulsion Laboratory, California Institute of Technology

Brian Wilson, Chris Mattmann, Sujen Shah, Huikyo Lee

NASA Jet Propulsion Laboratory, California Institute of Technology

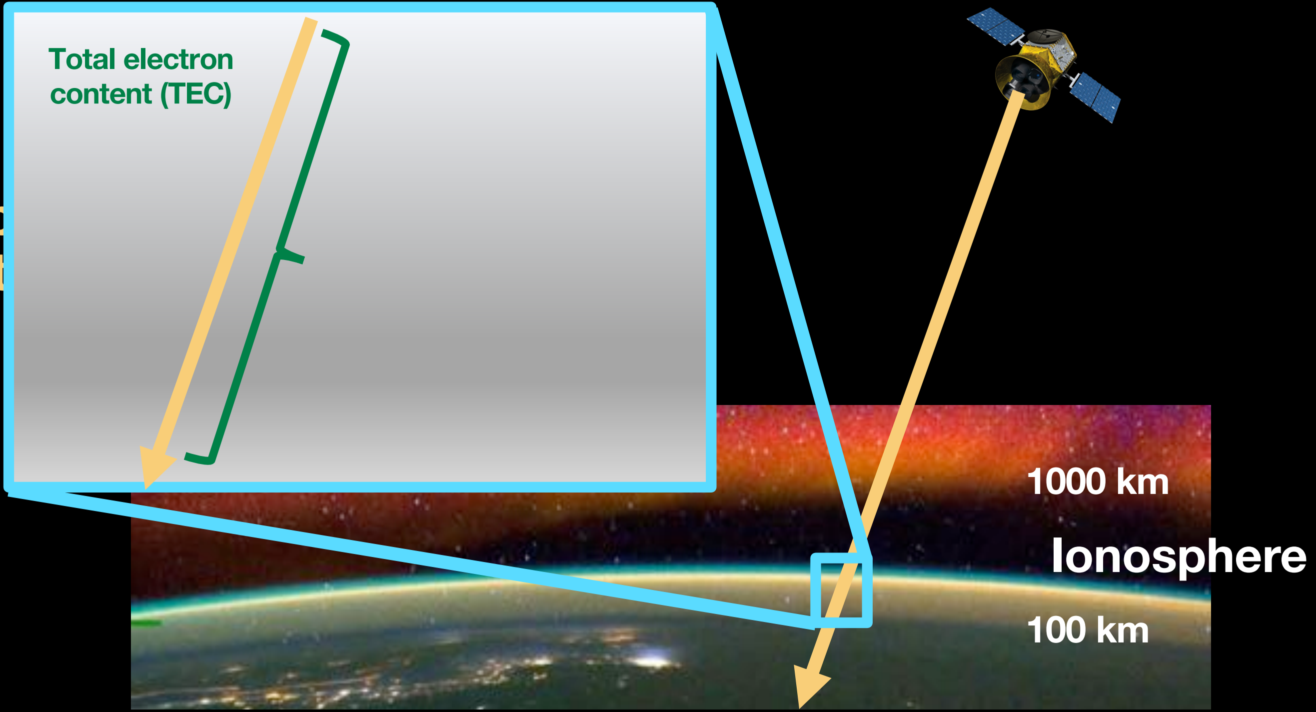
Assess

Global Navigation Satellite System (GNSS) signals for Space Science

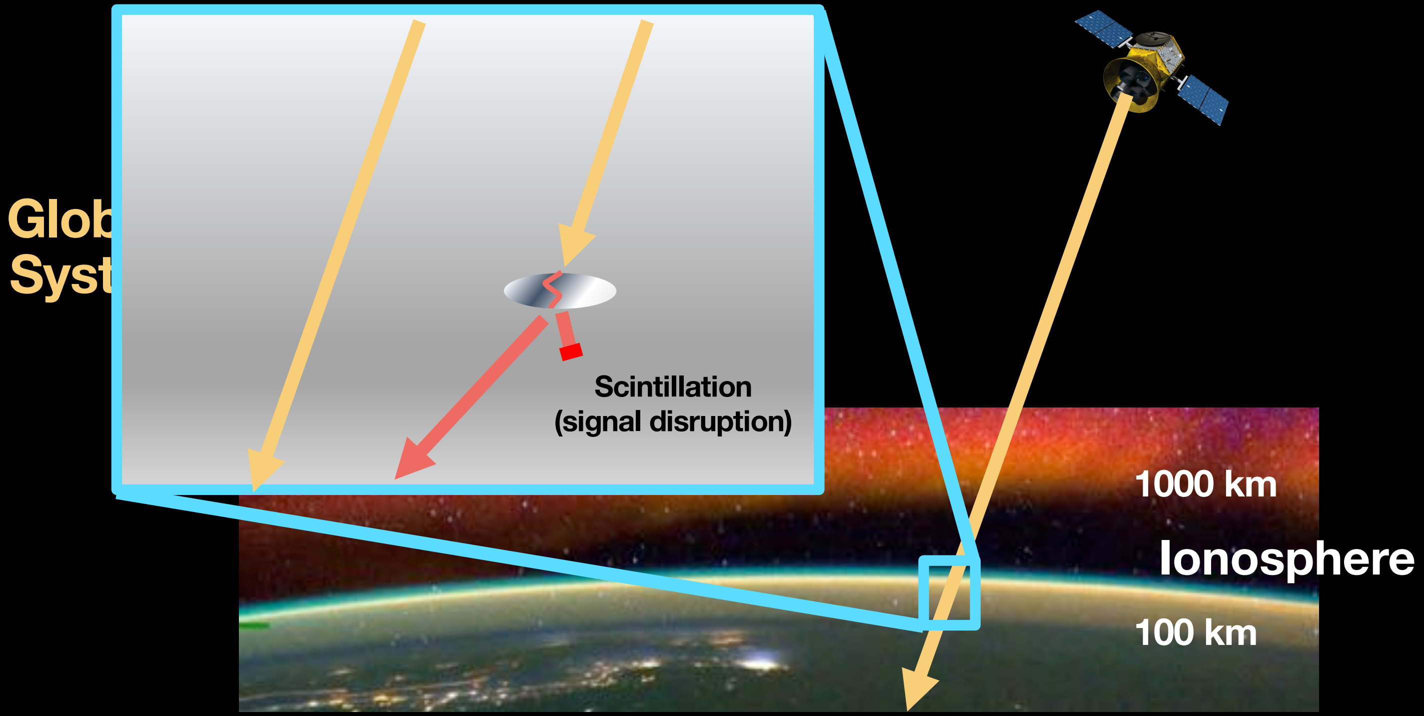


Assess

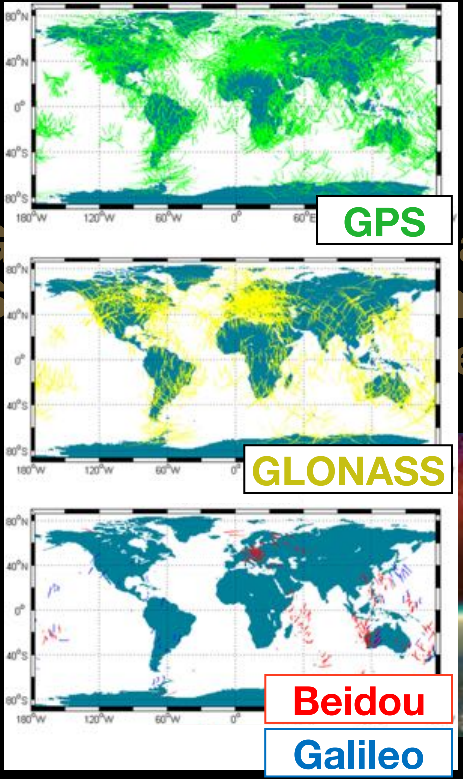
Glob
Syst

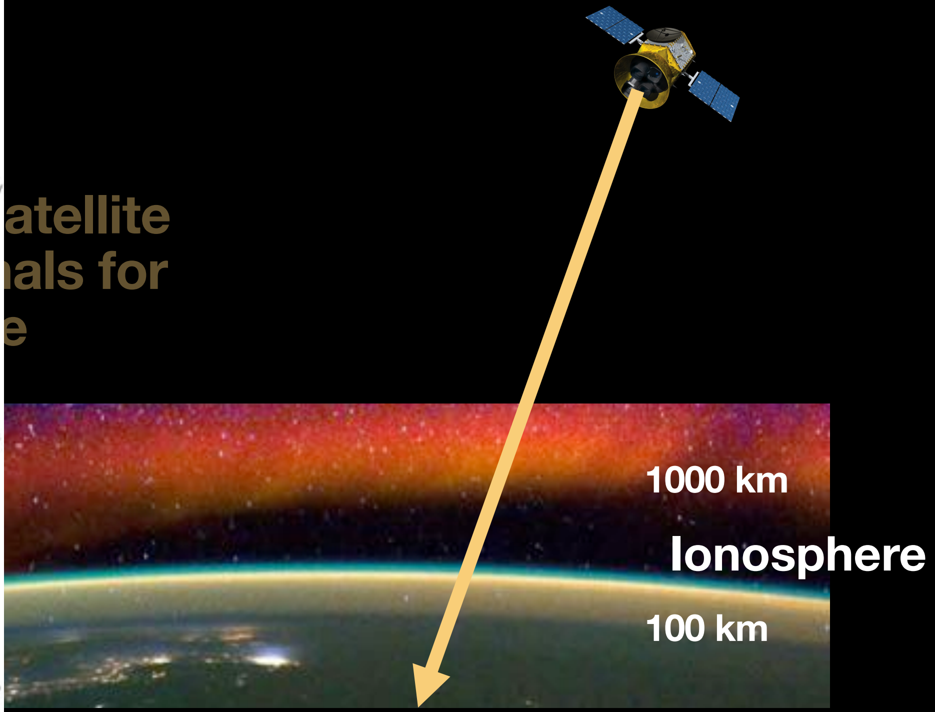


Assess


Glob
Syst
Scintillation
(signal disruption)
1000 km
Ionosphere
100 km

Assess


GPS
GLONASS
Beidou
Galileo
80°N
40°N
0°
40°S
80°S
180°W
120°W
60°W
0°
60°E


1000 km
Ionosphere
100 km

Design

Problems well-suited to machine learning

- Classification
- Event detection
- Segmentation
- Clustering
- Prediction
- Recommendation

Design

Problems well-suited to machine learning

- Classification
- Event detection
- Segmentation
- Clustering
- **Prediction**
- Recommendation

Design

Problems well-suited to machine learning

- Classification
- Event detection
- Segmentation
- Clustering
- **Prediction**
- Recommendation

Support Vector Machine (SVM) — **Decision Trees** — **Random Forests** — **Neural Networks**

Easily explainable → *Difficult to explain*

Create a narrative of new scientific understanding across spectrum of machine learning approaches

Design

Problems well-suited to machine learning

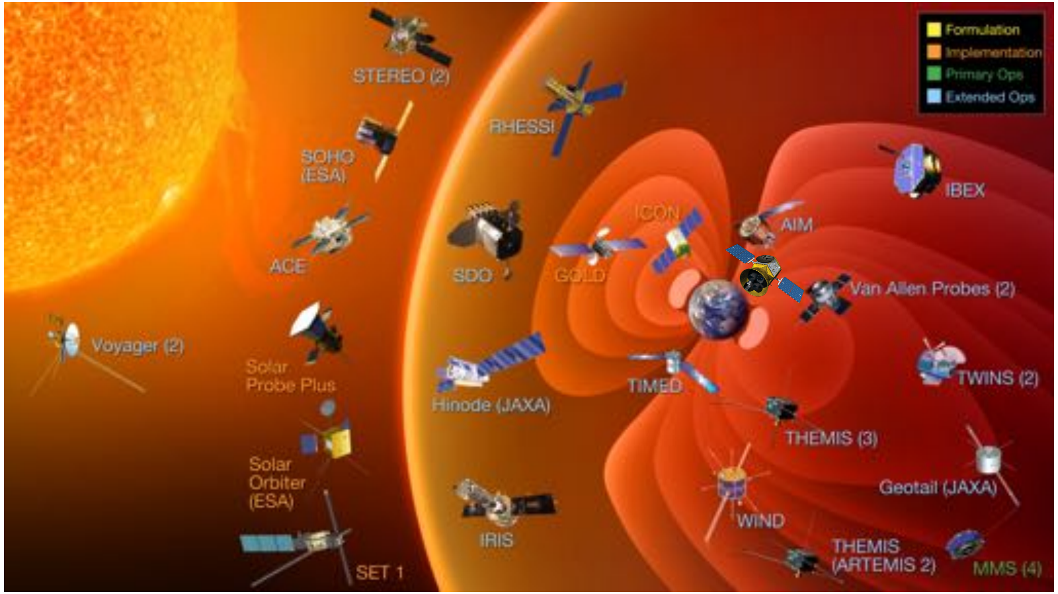
- Classification
- Event detection
- Segmentation
- Clustering
- **Prediction**
- Recommendation

Support Vector Machine (SVM) | **Decision Trees** | **Random Forests** | **Neural Networks**

Easily explainable → *Difficult to explain*

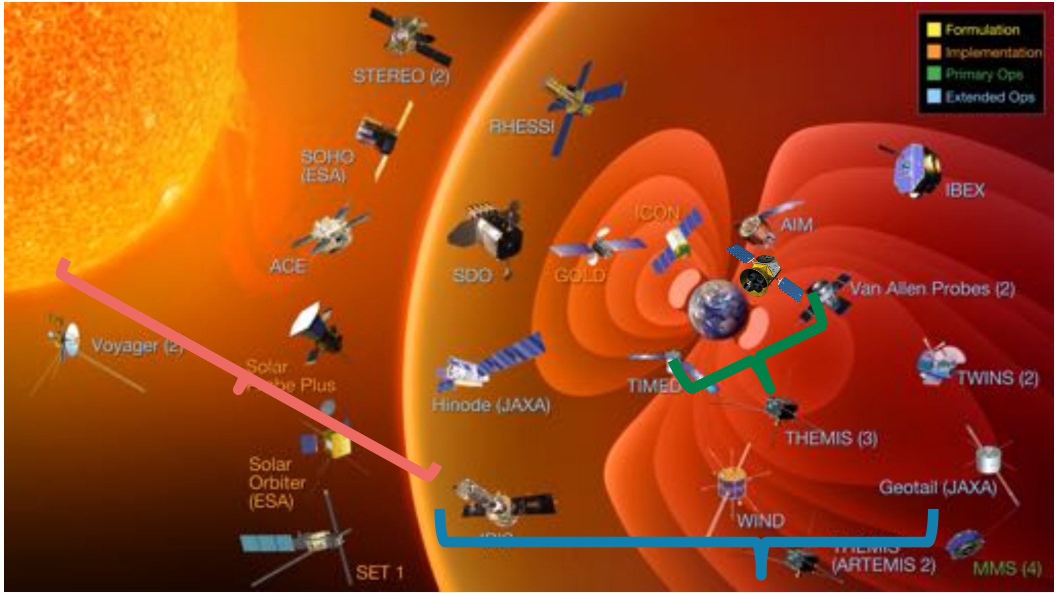
Create a narrative of new scientific understanding across spectrum of machine learning approaches

Design



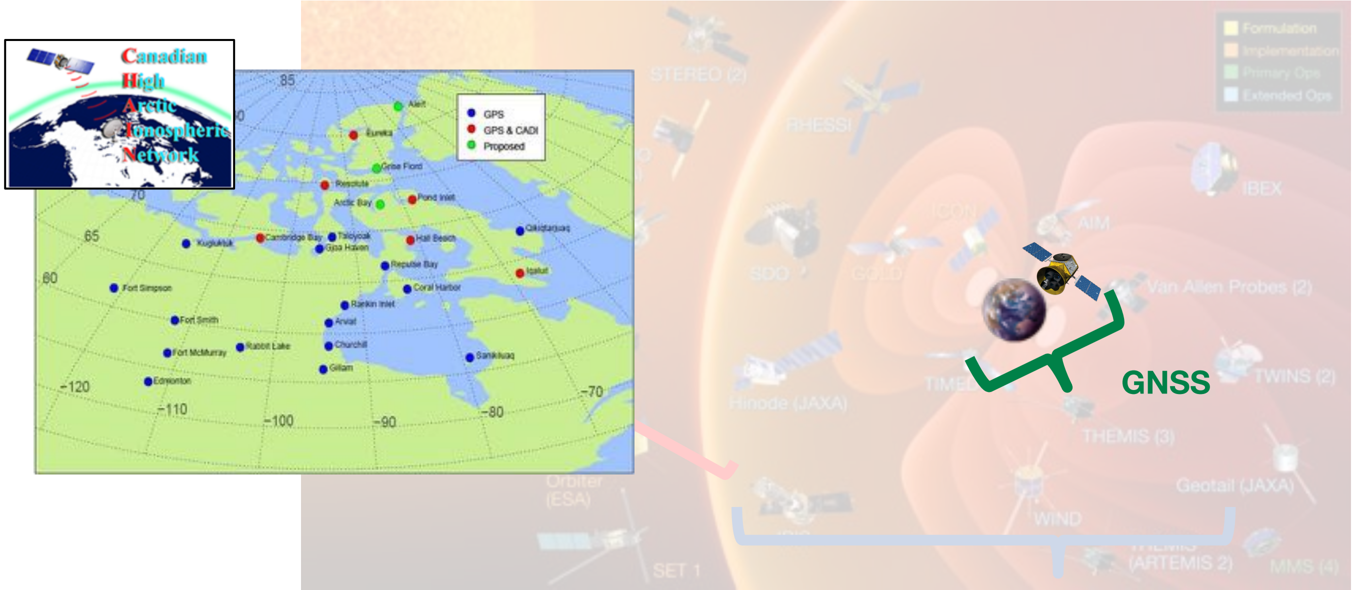
Step 1:
Obtain solar, geomagnetic, and ionospheric data

Design



Step 1:
Obtain solar, geomagnetic, and ionospheric data

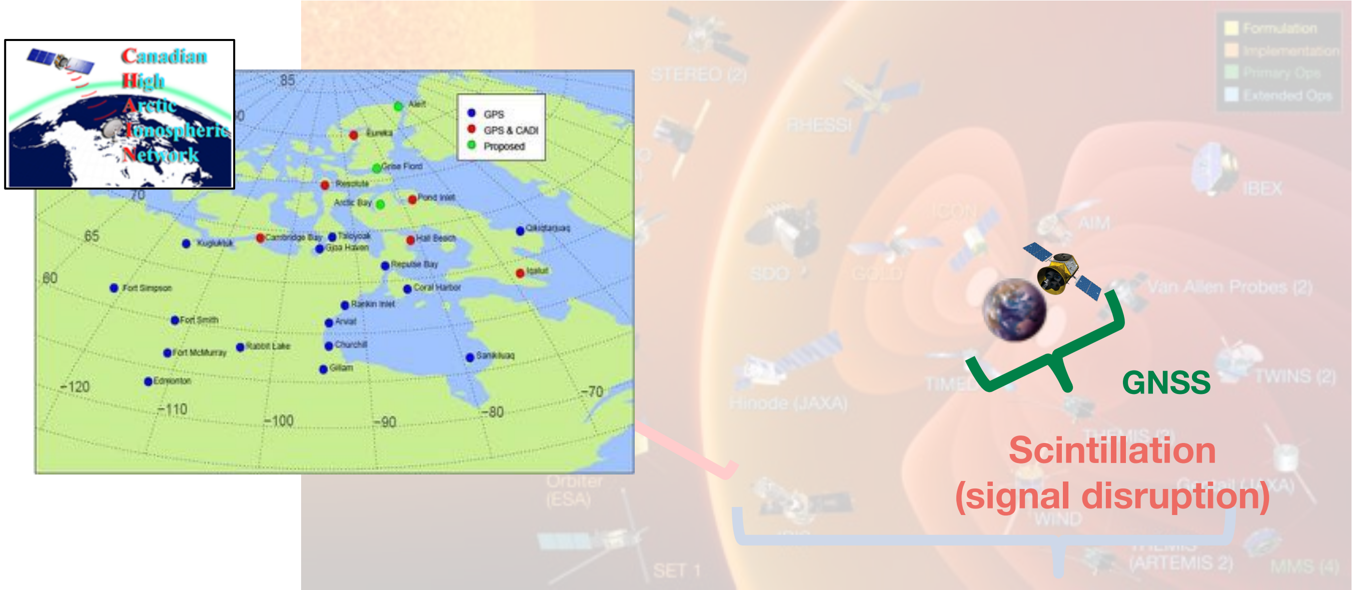
Design



Step 1:

Obtain solar, geomagnetic, and ionospheric data

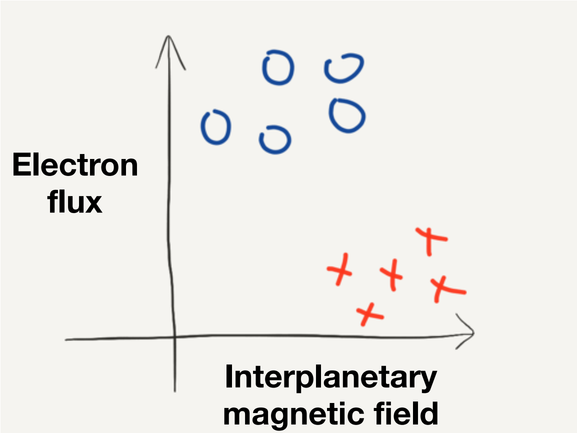
Design



Step 2:
Define the predictive task

Design

Support Vector Machine

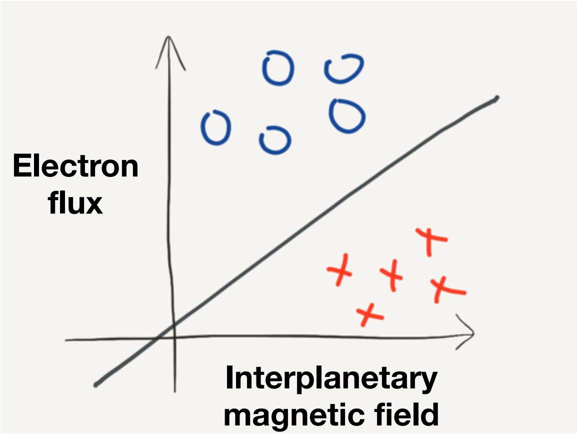


Cortes and Vapnik (1995)

Step 3:
Machine learning algorithm for prediction

Design

Support Vector Machine

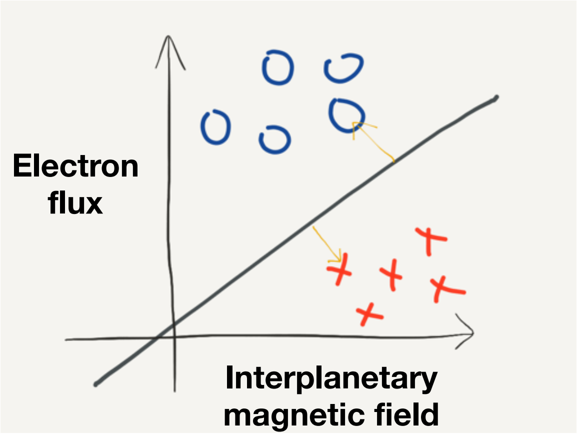


Cortes and Vapnik (1995)

Step 3:
Machine learning algorithm for prediction

Design

Support Vector Machine

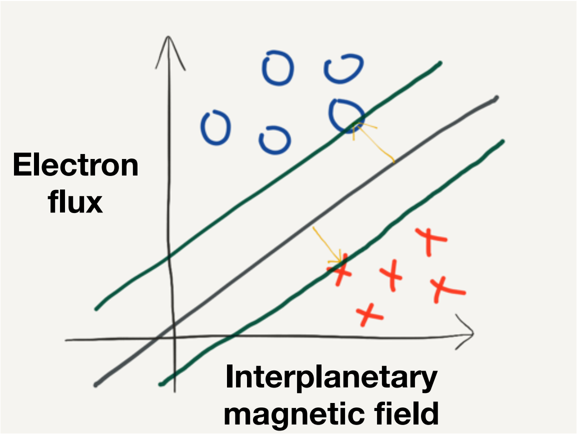


Cortes and Vapnik (1995)

Step 3:
Machine learning algorithm for prediction

Design

Support Vector Machine



Cortes and Vapnik (1995)

Step 3:

Machine learning algorithm for prediction

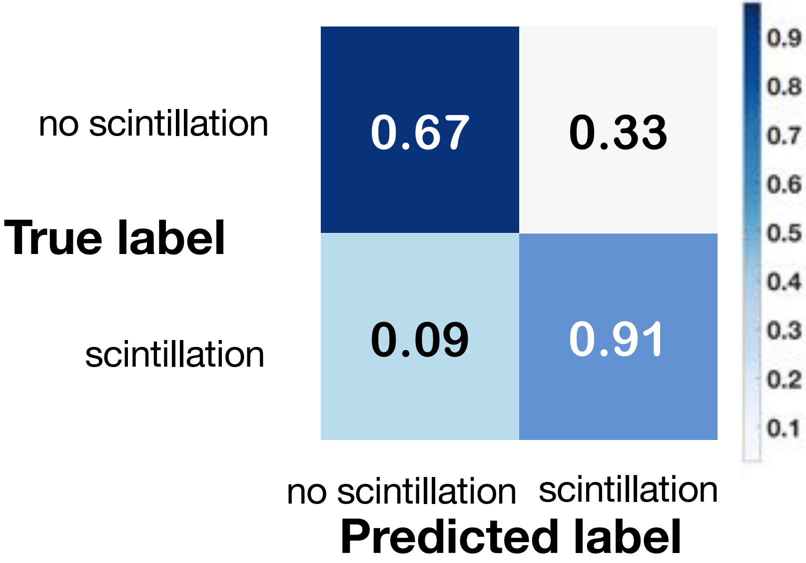
Analyze

True label	Predicted label: no scintillation	Predicted label: scintillation
no scintillation	True negative	False positive
scintillation	False negative	True positive

McGranaghan et al., (2018)

Step 4:
Evaluate and interrogate the model

Analyze



McGranaghan et al., (2018)

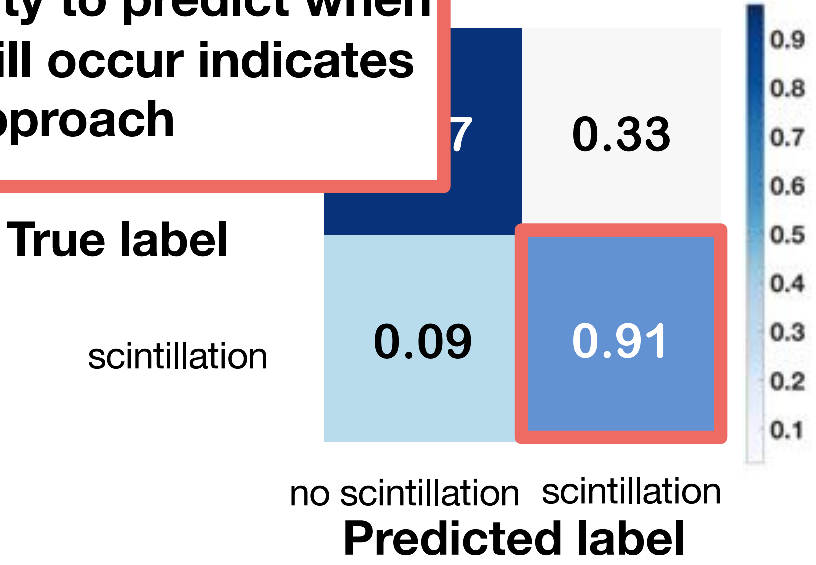
Step 4:

Evaluate and interrogate the model

Analyze

91%

Improved ability to predict when scintillation will occur indicates potential of approach

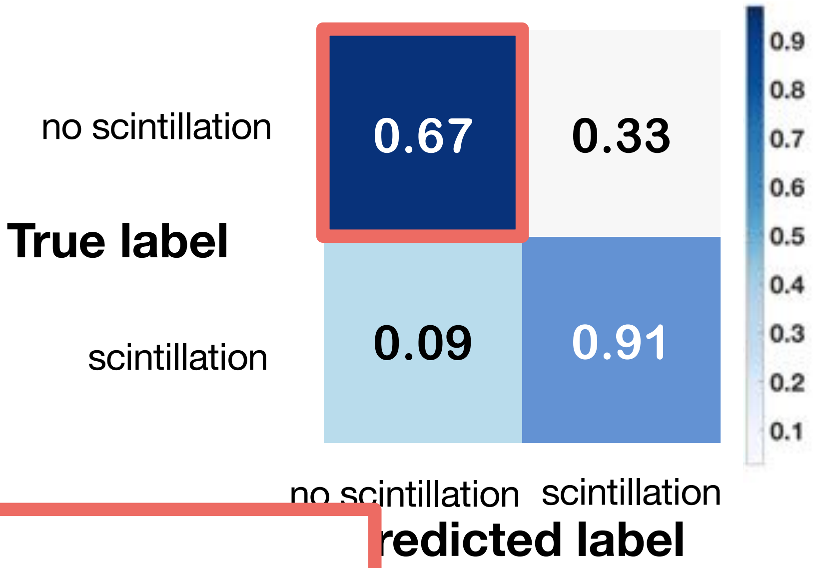


McGranaghan et al., (2018)

Step 4:

Evaluate and interrogate the model

Analyze



67%

High accuracy predicting when scintillation would not occur

Step 4:

d interrogate the model

Analyze

Step 4:

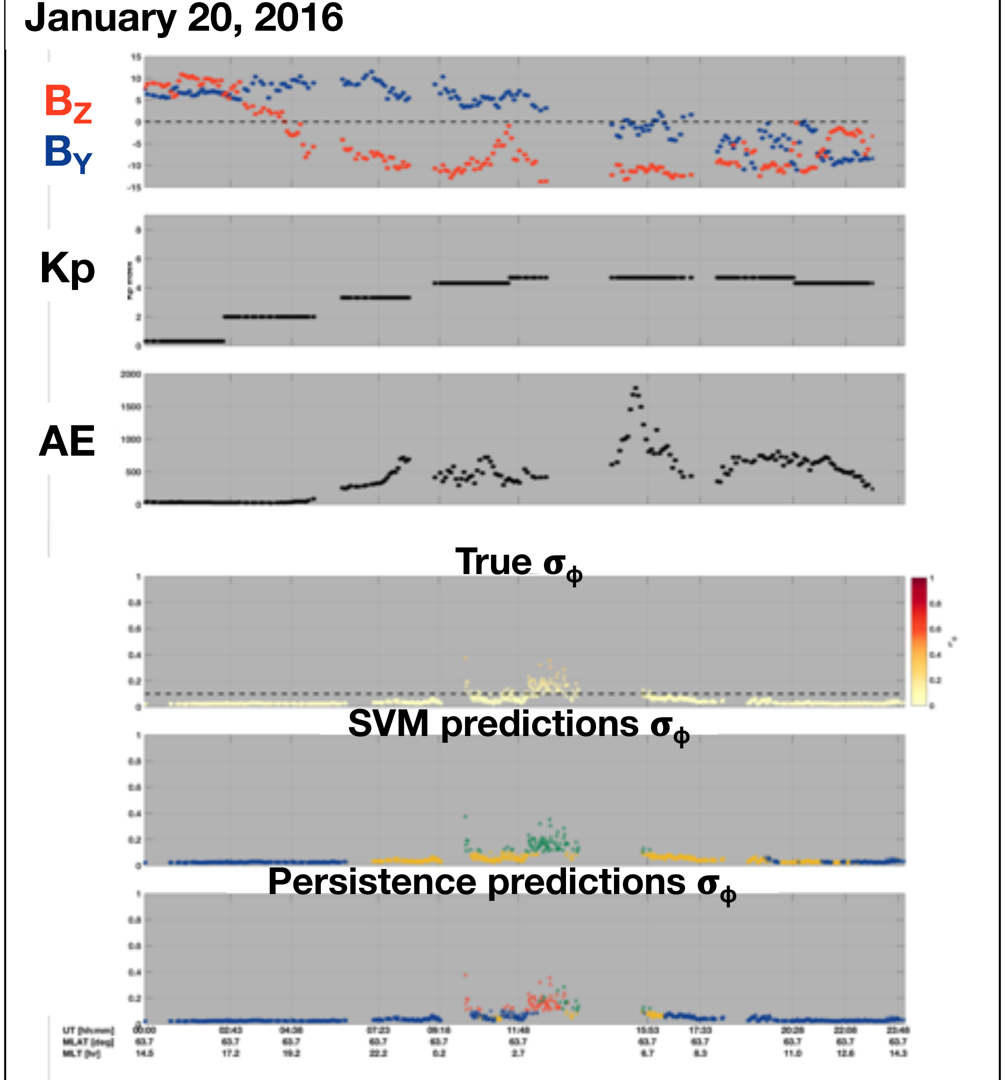
Evaluate and **<u>interrogate</u>** the model

Analyze

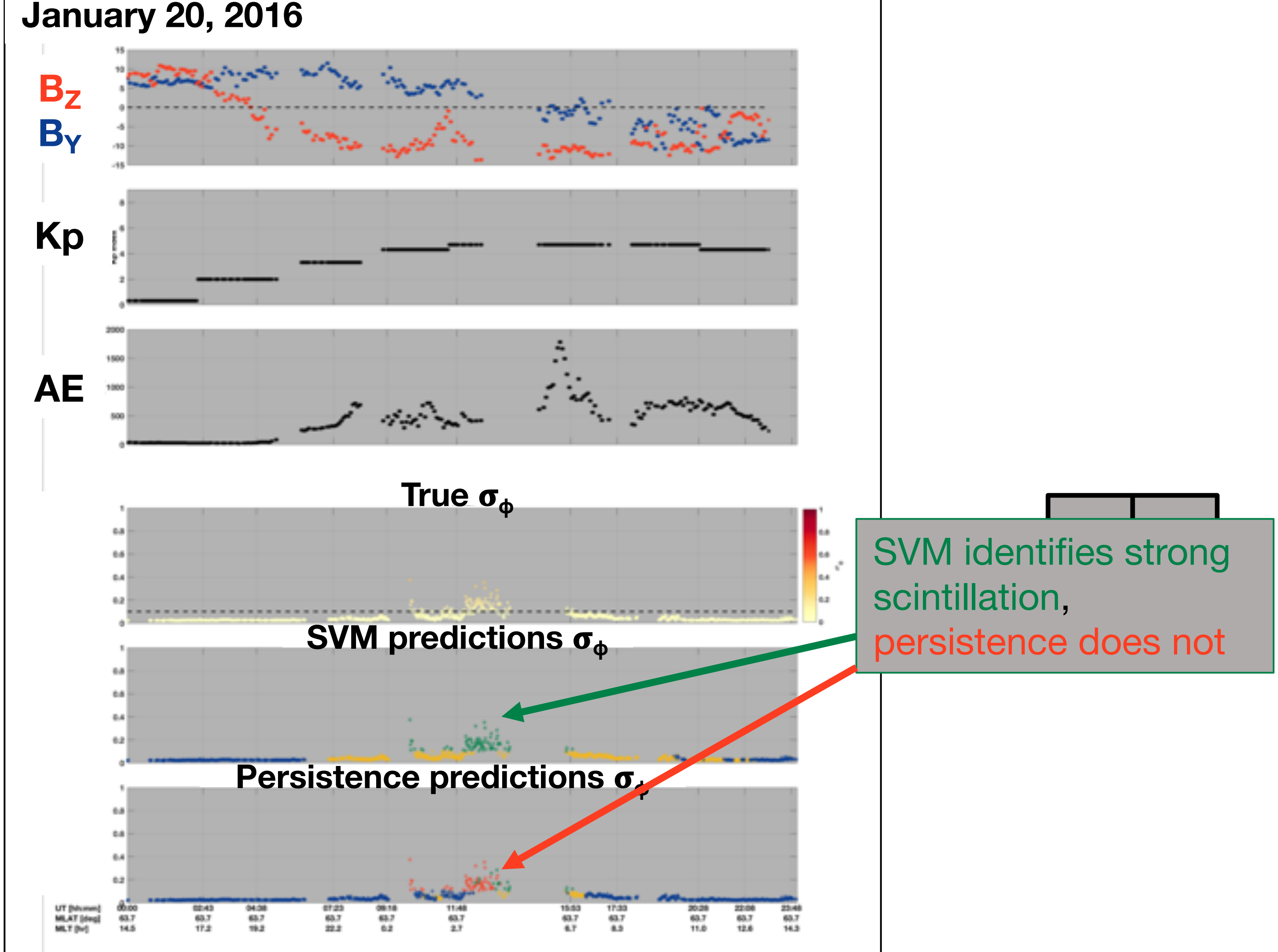
Interrogation

Step 4:
Evaluate and **<u>interrogate</u>** the model

Interrogation

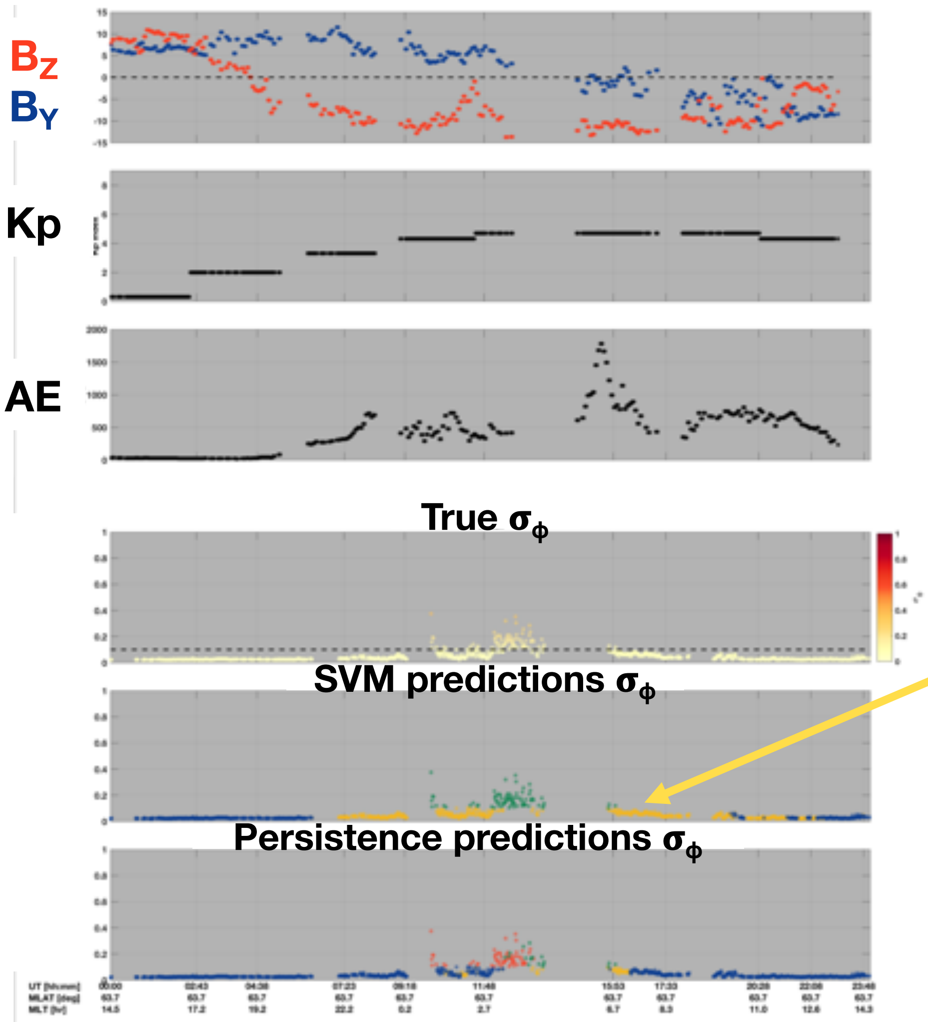


Interrogation



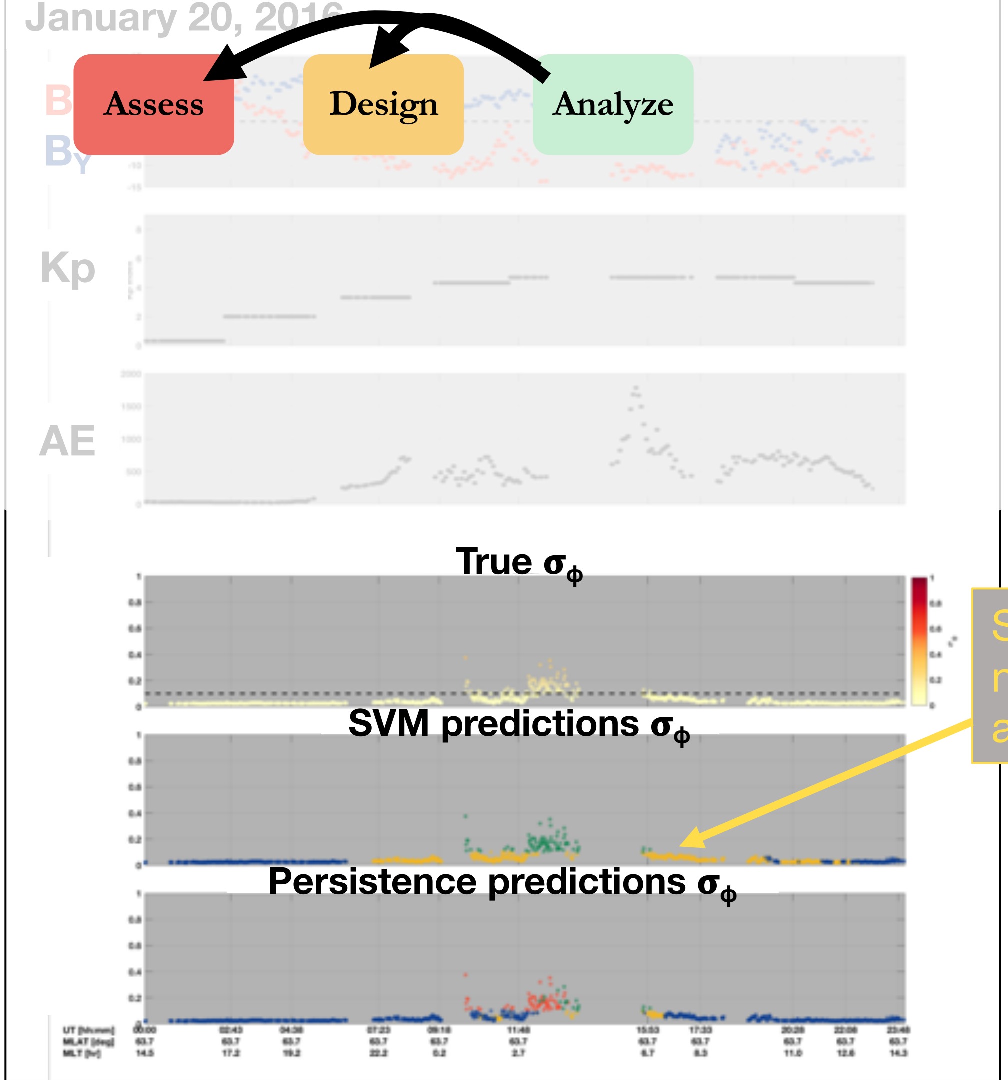
Interrogation

January 20, 2016



SVM contains high number of 'false alarms'

Explanation



SVM contains high number of ‘false alarms’

Deploy

STRETCHING GNSS SIGNALS FOR SPACE WEATHER DISCOVERY

Ryan McGranaghan, Anthony Mannucci
University Corporation for Atmospheric Research (UCAR)
NASA Jet Propulsion Laboratory, California Institute of Technology

Brian Wilson, Chris Mattmann, Sujen Shah, Huikyo Lee
NASA Jet Propulsion Laboratory, California Institute of Technology

Open source & Challenge data set
(https://doi.org/10.6084/m9.figshare.6813143)

NASA Frontier Development Lab

How do we shape
the New Frontier?

TEMPO
MAIA
TROPICS
LANDSAT 9
ICON
DART
SENTINEL-6A/B*
EUROPA CLIPPER
LUCY
PSYCHE
MARS 2020
SWOT
PACE
OPERATING & FUTUR
SCIENCE

Convergence entails integrating knowledge, methods, and expertise from different disciplines and forming novel frameworks to catalyze scientific discovery and innovation

National Science Foundation, 10 Big Ideas

Opportunity:

- Evolve traditional approaches
- Embrace data science-driven discovery
- Enable **interdisciplinary** work

Opportunity:

- Evolve traditional approaches
- Embrace data science-driven discovery
- Enable **interdisciplinary** work

Someone or something that doesn't fit within traditional academic discipline-a field of study with its own particular words, frameworks, and methods

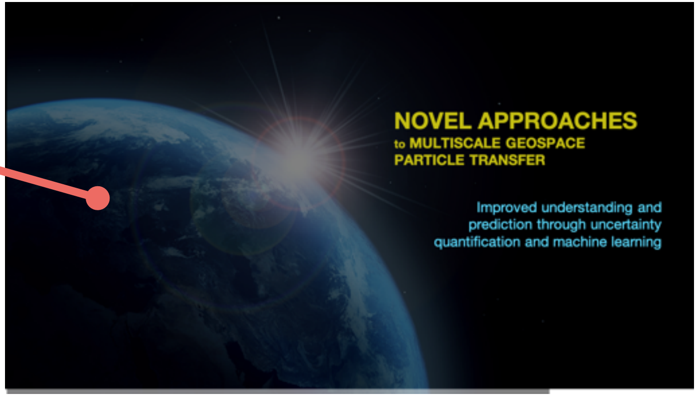
Joi Ito, MIT Media Lab, “Antidisciplinary”

Toward antidisciplinary

Non-traditional collaborations and approaches

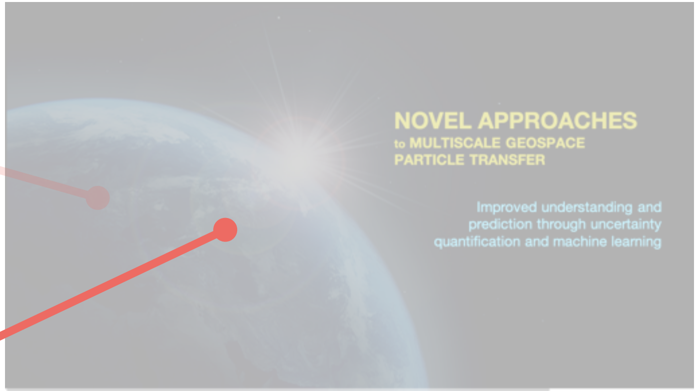
Toward antidisciplinary

Non-traditional collaborations and approaches



Toward antidisciplinary

Non-traditional collaborations and approaches



Toward antidisciplinary

Non-traditional collaborations and approaches

HelioML

Unique opportunity:

Be pioneers of the New Frontier in your networks

NASA FRONTIER DEVELOPMENT LAB

JPL

ASTRA

NASA

Goddard SPACE FLIGHT CENTER

?

Shaping the New Frontier

Understand the models

Be open by default

Be radically interdisciplinary (i.e., *antidisciplinary*)

@AeroSciengineer

ryan.mcgranaghan@gmail.com

RyanMcGranaghan.com

McGranaghan, R. M., Bhatt, A., Matsuo, T., Mannucci, A. J., Semeter, J. L., & Datta-Barua, S. (2017). Ushering in a new frontier in geospace through data science. Journal of Geophysical Research: Space Physics, 122, 12,586–12,590. https://doi.org/10.1002/2017JA024835

McGranaghan, R. M., A.J. Mannucci, B.D Wilson, C.A. Mattmann, and R. Chadwick. (2018), New capabilities for prediction of high-latitude ionospheric scintillation: A novel approach with machine learning, Space Weather, 16. https://doi.org/10.1029/2018SW002018

Backup slides

Resources

- HelioAnalytics – Goddard Space Flight Initiative to "*harness advanced statistics, informatics and computer science methods to achieve science*"
- Thought-leaders:
  - Kirk Borne and on Twitter
  - Joi Ito
  - Cesar Hidalgo
  - Andrew Ng
  - Barbara Thompson
  - Naval Ravikant
  - Hilary Mason
  - *Expand your horizons with the papers that you read, the fields to which you pay attention, and the thinkers that you choose to learn from
- Compilations of resources
  - Non-traditional funding resources
  - Data science tools and resources
- Being 'antidisciplinary'
  - MIT Media Lab
  - Fall AGU Town Hall 2018: "Data Science and a New Scientific Frontier in Space Science"
  - Fall AGU Town Hall 2019: "Antidisciplinary: Science and engineering in the digital age"
- Podcasts
  - Origins
  - Microsoft Research Podcast
  - Grey Mirror Podcast
  - Voices from DARPA
  - Artificial Intelligence Podcast
  - Data Skeptic

Resources (cont'd)

- Camporeale et al., [2019]
- AGU Earth and Space Science Informatics (ESSI)
- National Research Council "Enhancing the Effectiveness of Team Science"
- Meetups, hackathons, and unconferences
- Open source communities (e.g., Open Source Initiative)
- Citizen Science
- Many resources to discover based on your own passions and search!

Toward antidisciplinary

Innovation through communication and interaction

Thought-leaders:

NASA Headquarters

Industry

NSF

National Academy of Sciences

Toward antidisciplinary

Innovation through communication and interaction

Thought-leaders:

NASA Headquarters

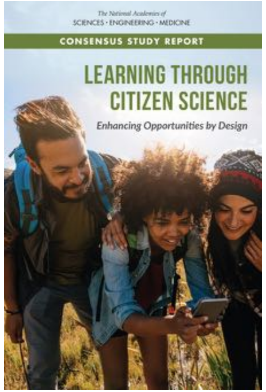
Industry

NSF

National Academy of Sciences

Toward antidisciplinary

Innovation through communication and interaction



sci foo

SCIENCE on TAP

Origins

Misconceptions

- Two things that typically turn people off to data science and machine learning are the jargon and the hype
    - These are detrimental misconceptions and this morning I want to address them in two direct ways:
        1. By explicitly defining what I mean with the term ‘data science’
        2. Crystallize the capabilities (and incapabilities!) of machine learning by looking at a fantastic use case – ionospheric scintillation

Key takeaways

Data science is much more than just machine learning

Assessing the problem and data wrangling are vastly the most time consuming component of the data science workflow

Concept of ‘Analysis Ready Data’ (ARD)

Machine learning are techniques that cover a broad spectrum

The New Frontier that we face must focus on generating new scientific discovery from innovation *in cooperation* with existing knowledge and traditional approaches

- Interrogate your models

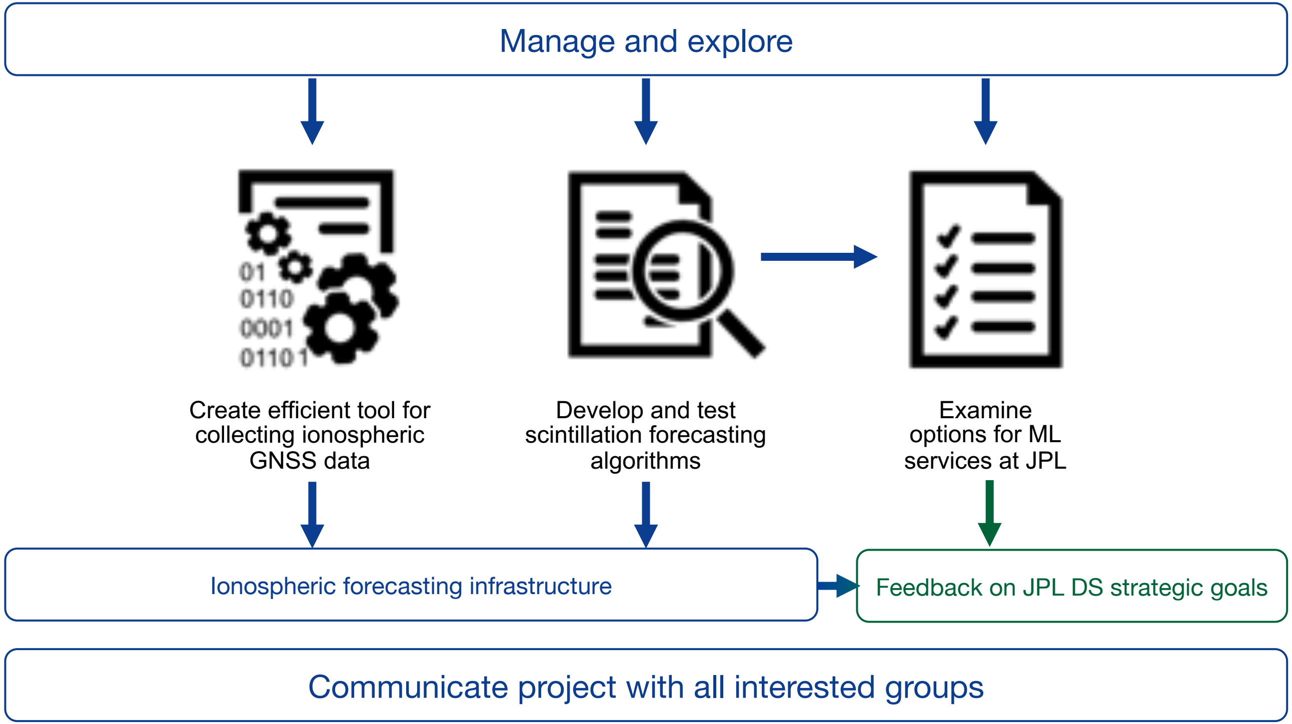
Data science requires non-traditional composition and coordination – need to be *radically interdisciplinary*

- Be open by default

Advanced information technologies

A solution:

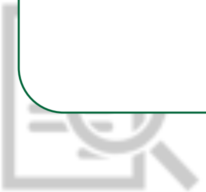
Data science solutions for space weather



A solution:
Data science solutions for space weather

Manag

Python library to gather, process, and analyze data from the Canadian High Arctic Ionospheric Network (CHAIN)

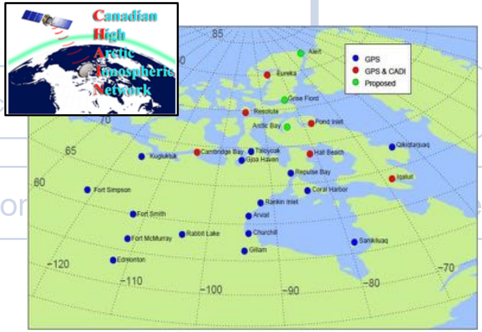


Create efficient tool for collecting ionospheric GNSS data

Develop and test scintillation forecasting algorithms

Examine options for ML services at JPL

Ionospheric



Feedback on JPL DS strategic goals

Cor

sted groups

A solution:

Data science solutions for space weather

Manag

Create efficient tool for collecting ionospheric GNSS data

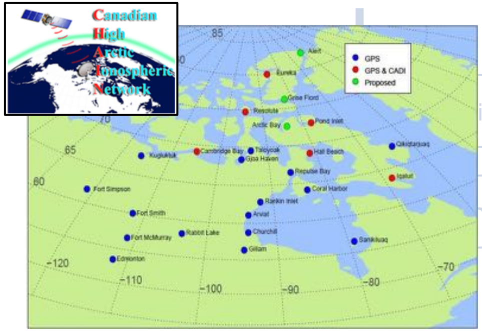
Develop
scintillation
algori

ing infrastructure

icate project wi

Python library to gather, process, and analyze data from the Canadian High Arctic Ionospheric Network (CHAIN)

Scalable to global network



INTERNATIONAL GNSS SERVICE

A solution:

DSWG strategic challenge: How can we use advanced data science methods to systematically derive scientific inferences from massive, distributed science measurements and models?

Manag

Python library to gather, process, and analyze data from the Canadian High Arctic Ionospheric Network (CHAIN)

Scalable to global network

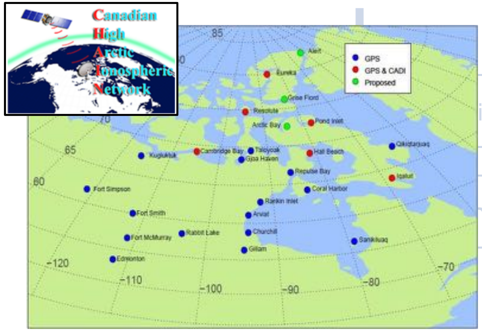
Create efficient tool for collecting ionospheric GNSS data

Develop scintillation algori

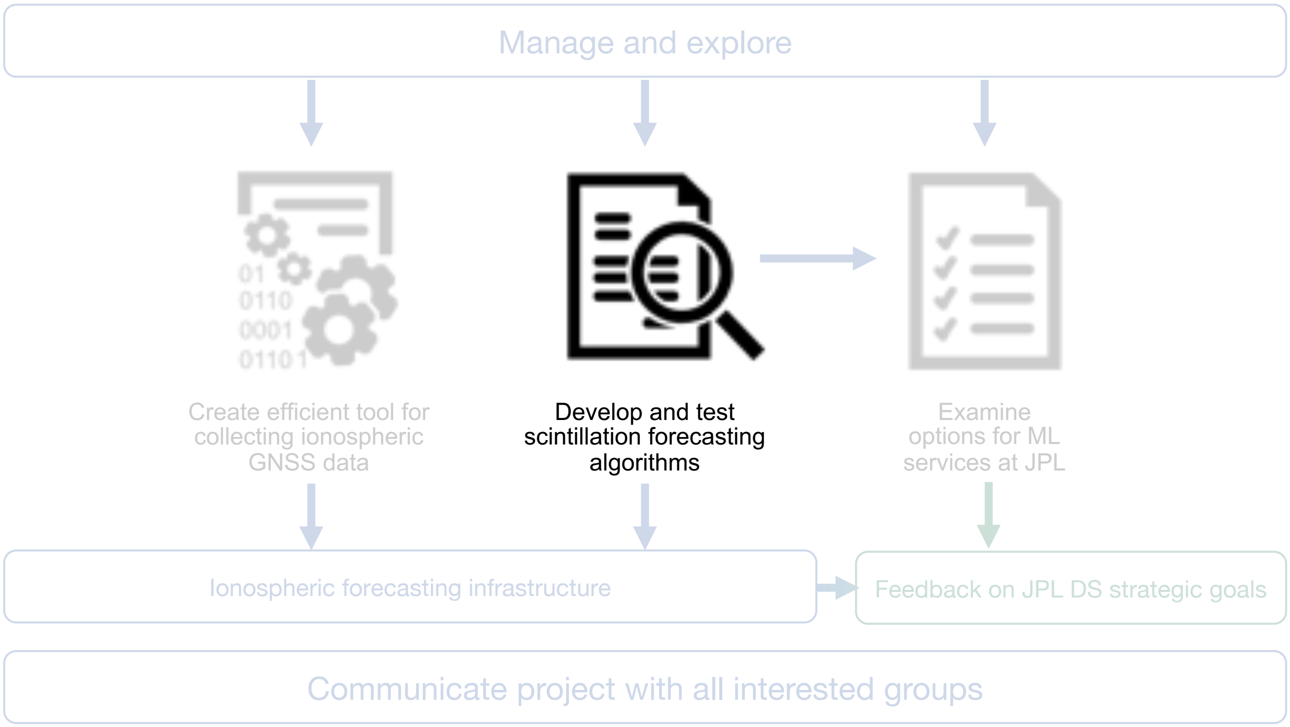
ing infrastructure

nicate project wi

INTERNATIONAL GNSS SERVICE

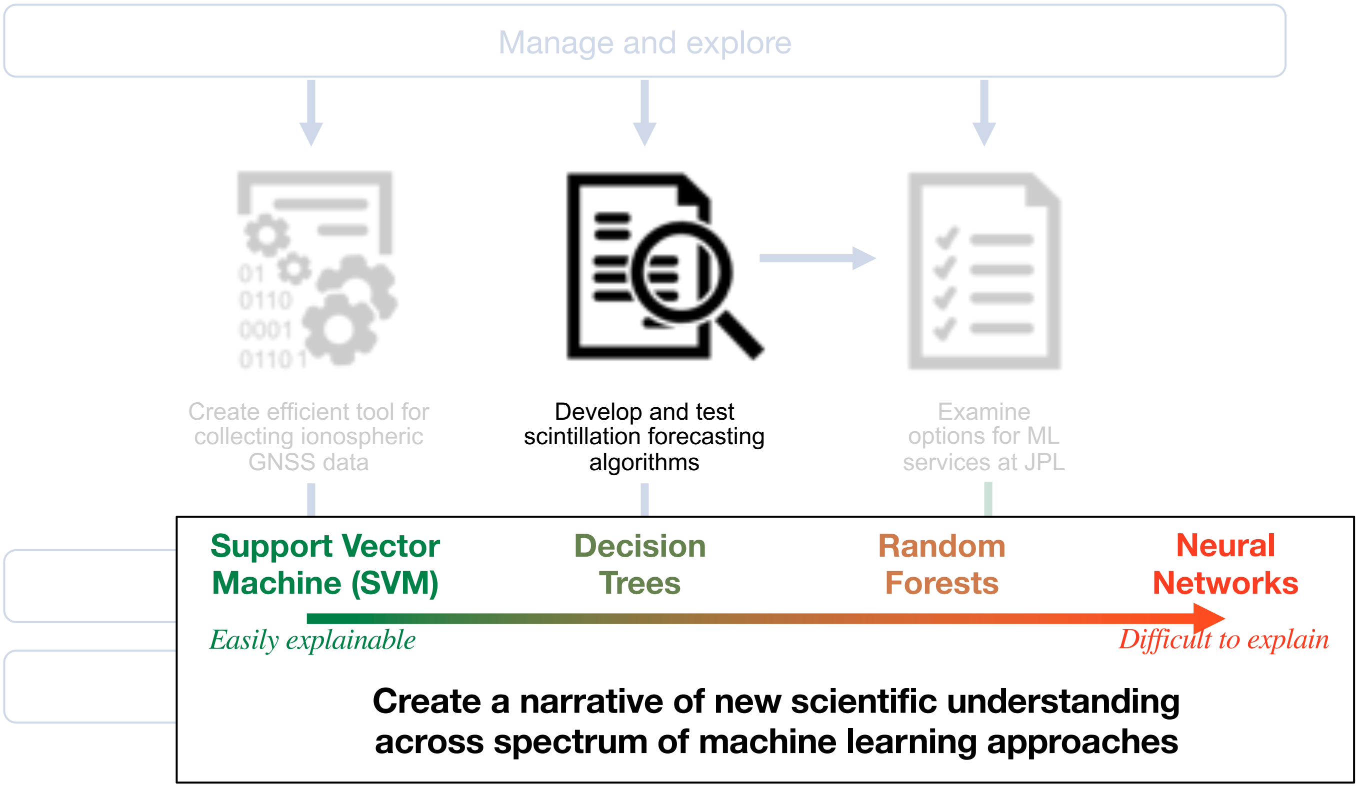


A solution:
Data science solutions for space weather



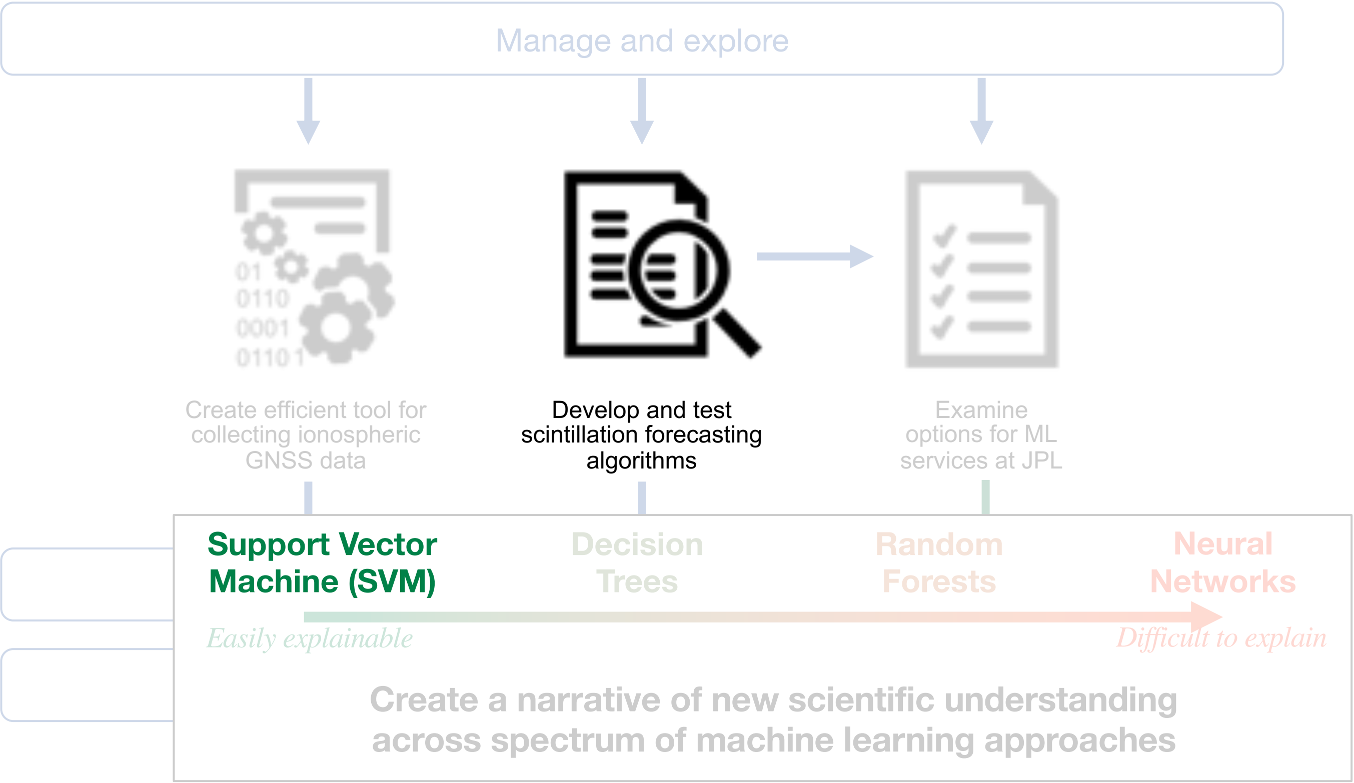
A solution:

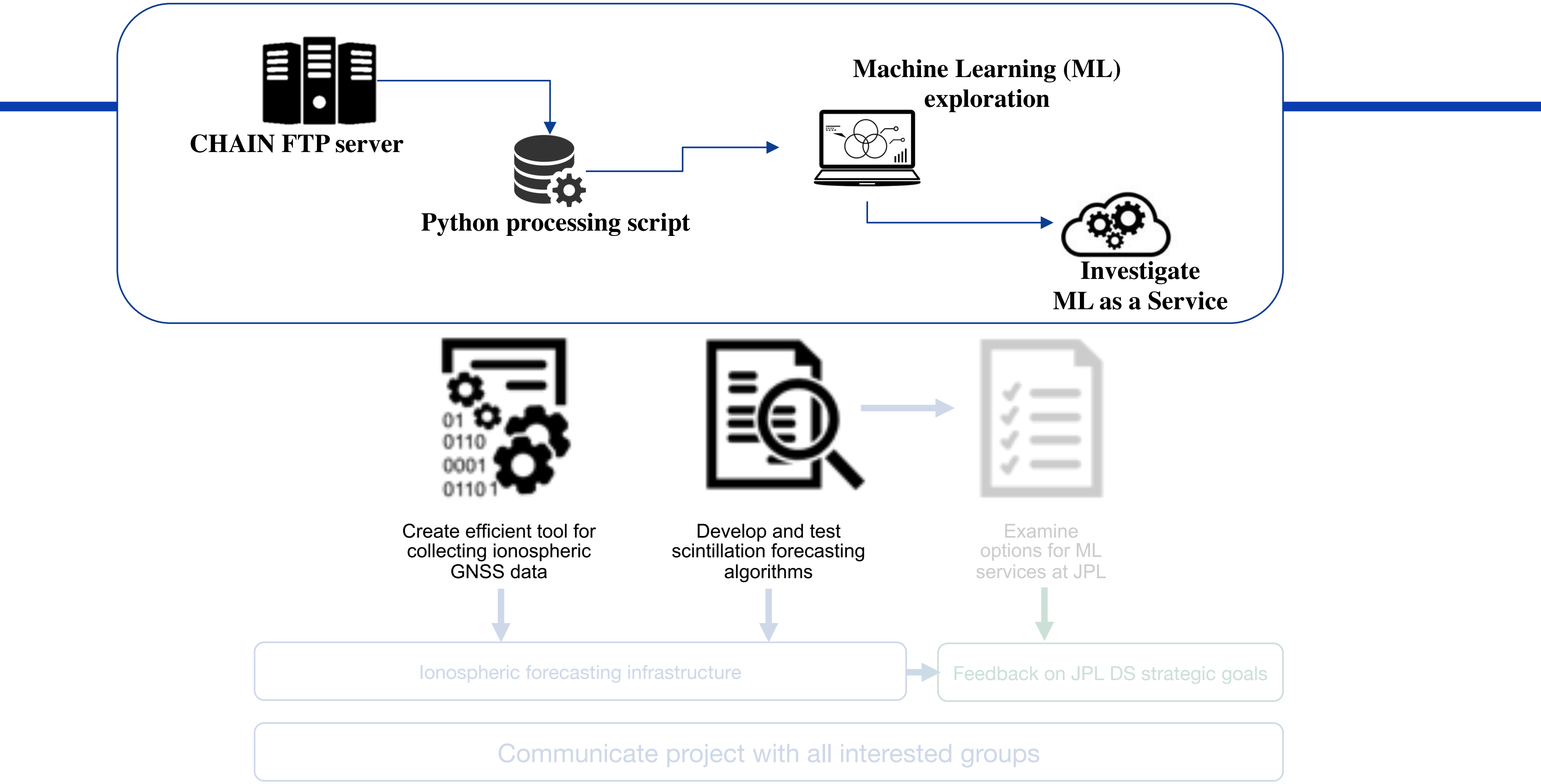
Data science solutions for space weather



A solution:

Data science solutions for space weather




CHAIN FTP server
Python processing script
Machine Learning (ML) exploration
Investigate ML as a Service
Create efficient tool for collecting ionospheric GNSS data
Develop and test scintillation forecasting algorithms
Examine options for ML services at JPL
Ionospheric forecasting infrastructure
Feedback on JPL DS strategic goals
Communicate project with all interested groups

Explainable Machine Learning

Step 1:

Obtain solar, geomagnetic, and ionospheric data

Ionosphere:

location information
station
precipitation (E_{flux}, N_{flux})
az
el
TEC
dTEC
spectral slope
S4
σ_ϕ

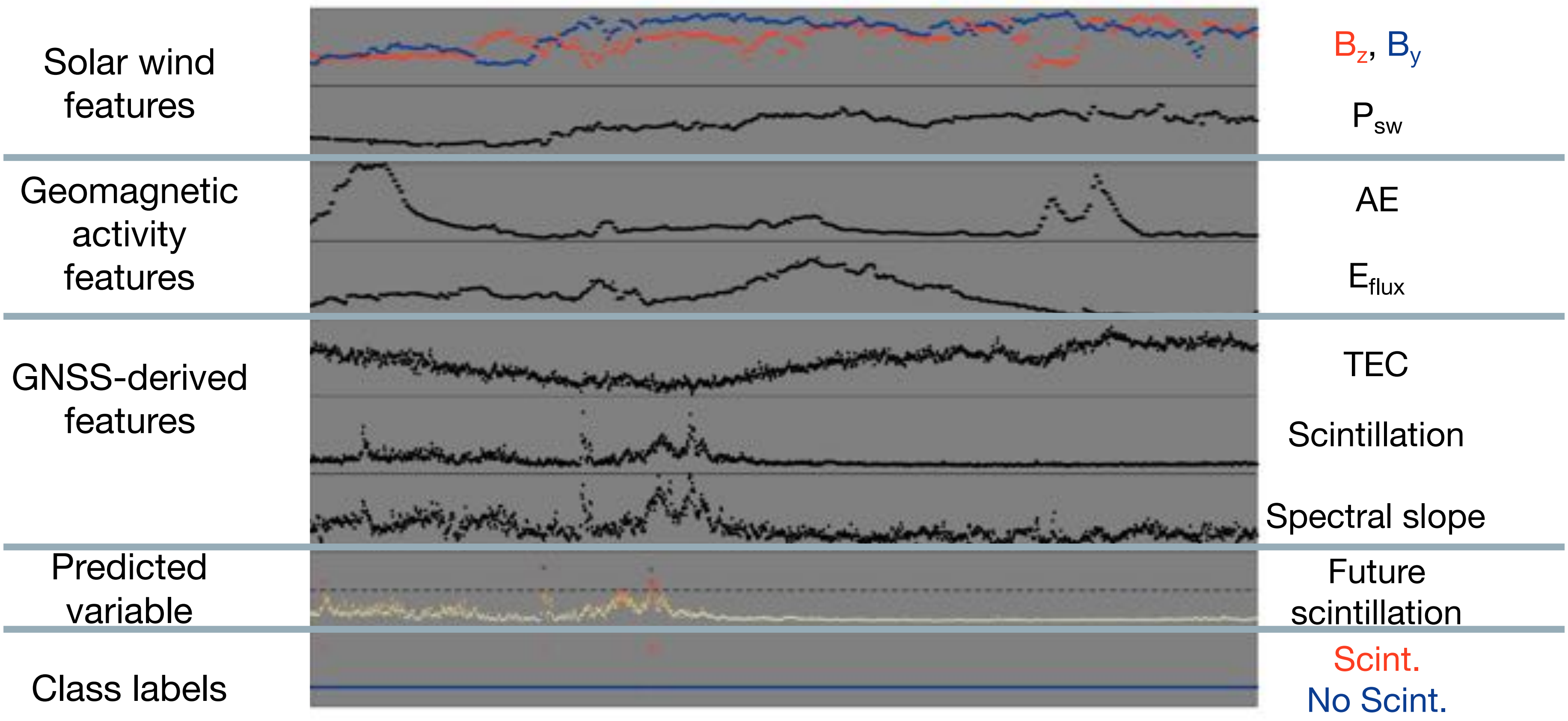
Solar:

Bz, By
clock angle
V_{sw}
P_{sw}
$F_{10.7}$

Geomagnetic:

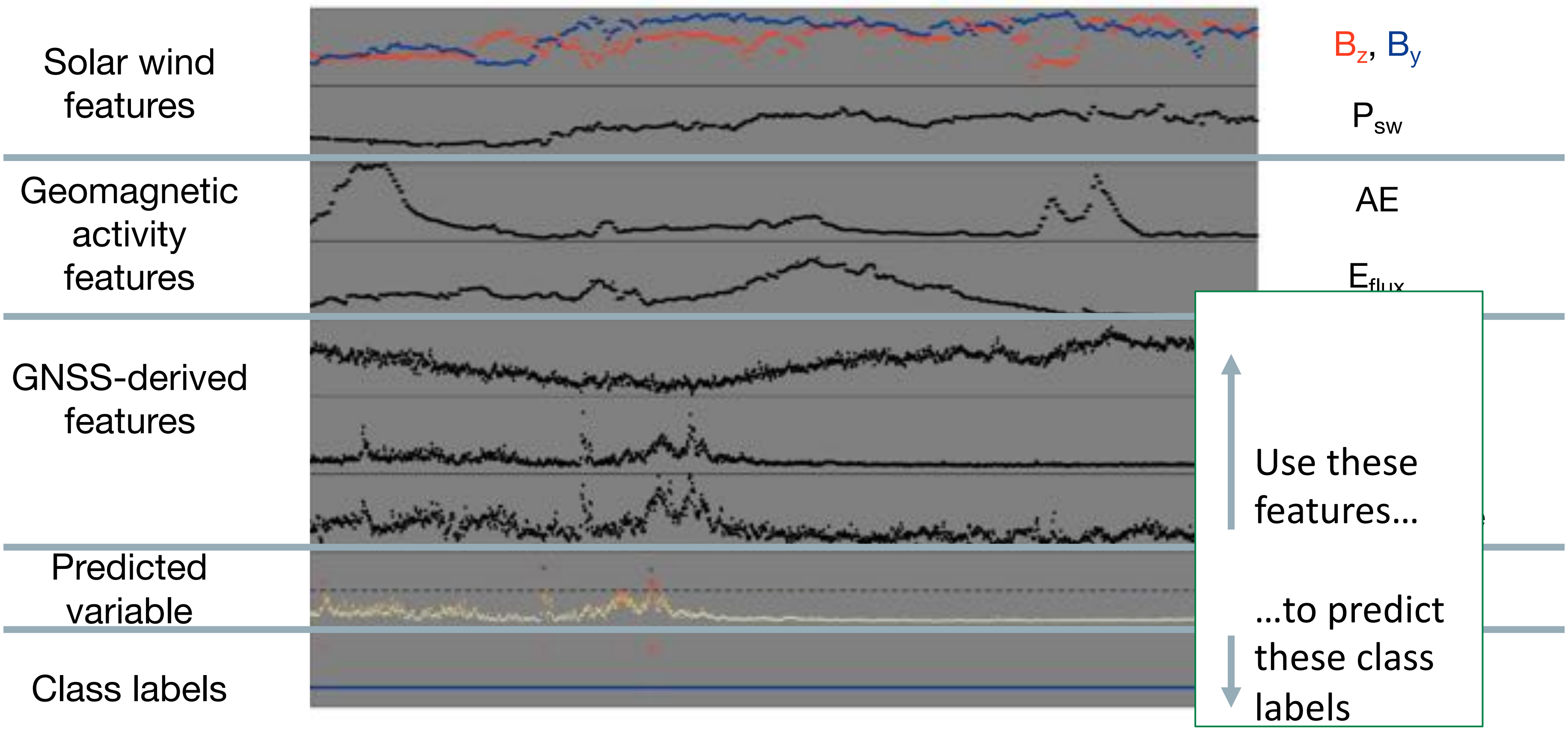
Sym-H
AE
Kp
Solar wind-magnetosphere coupling functions

Step 2:
Define the predictive task

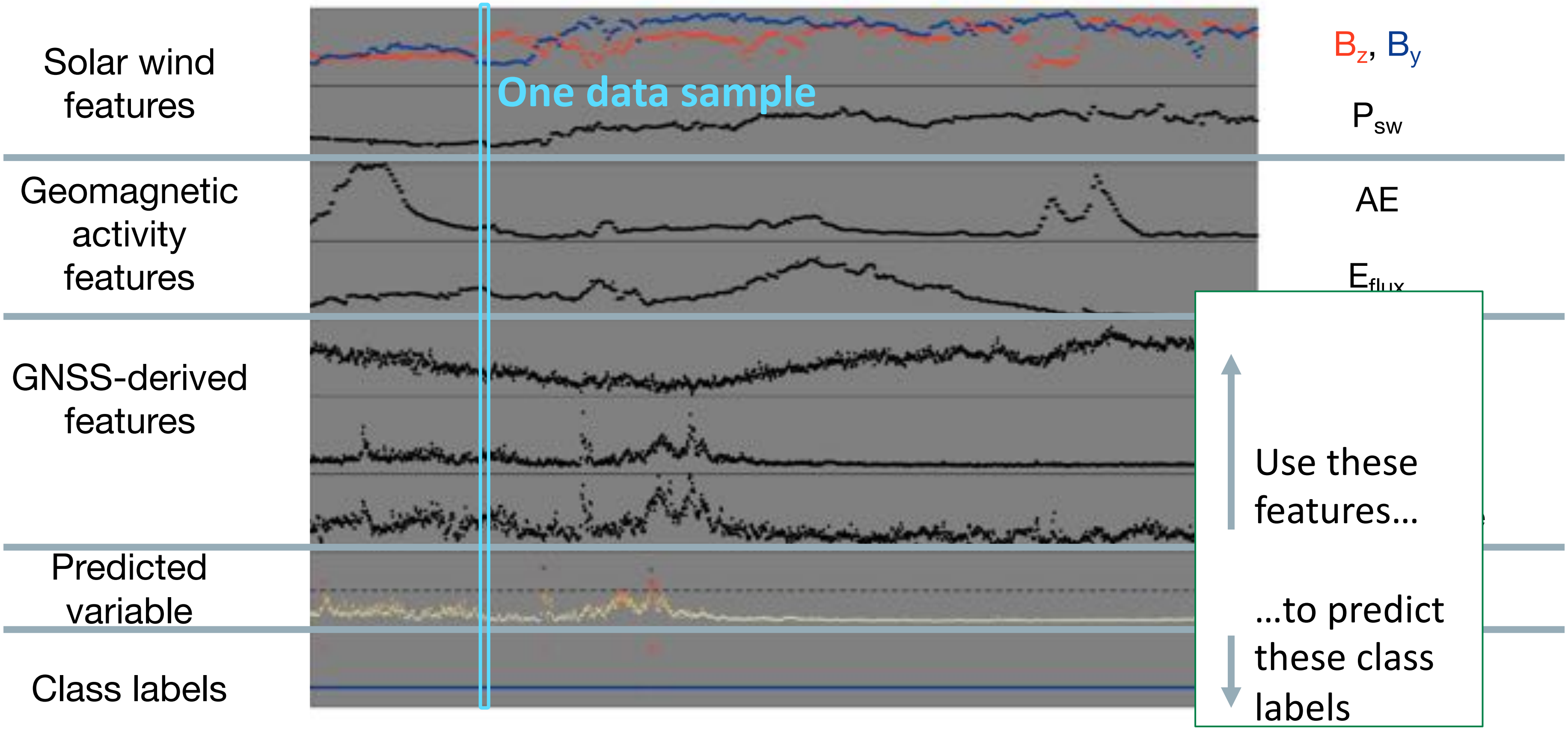


Step 2:

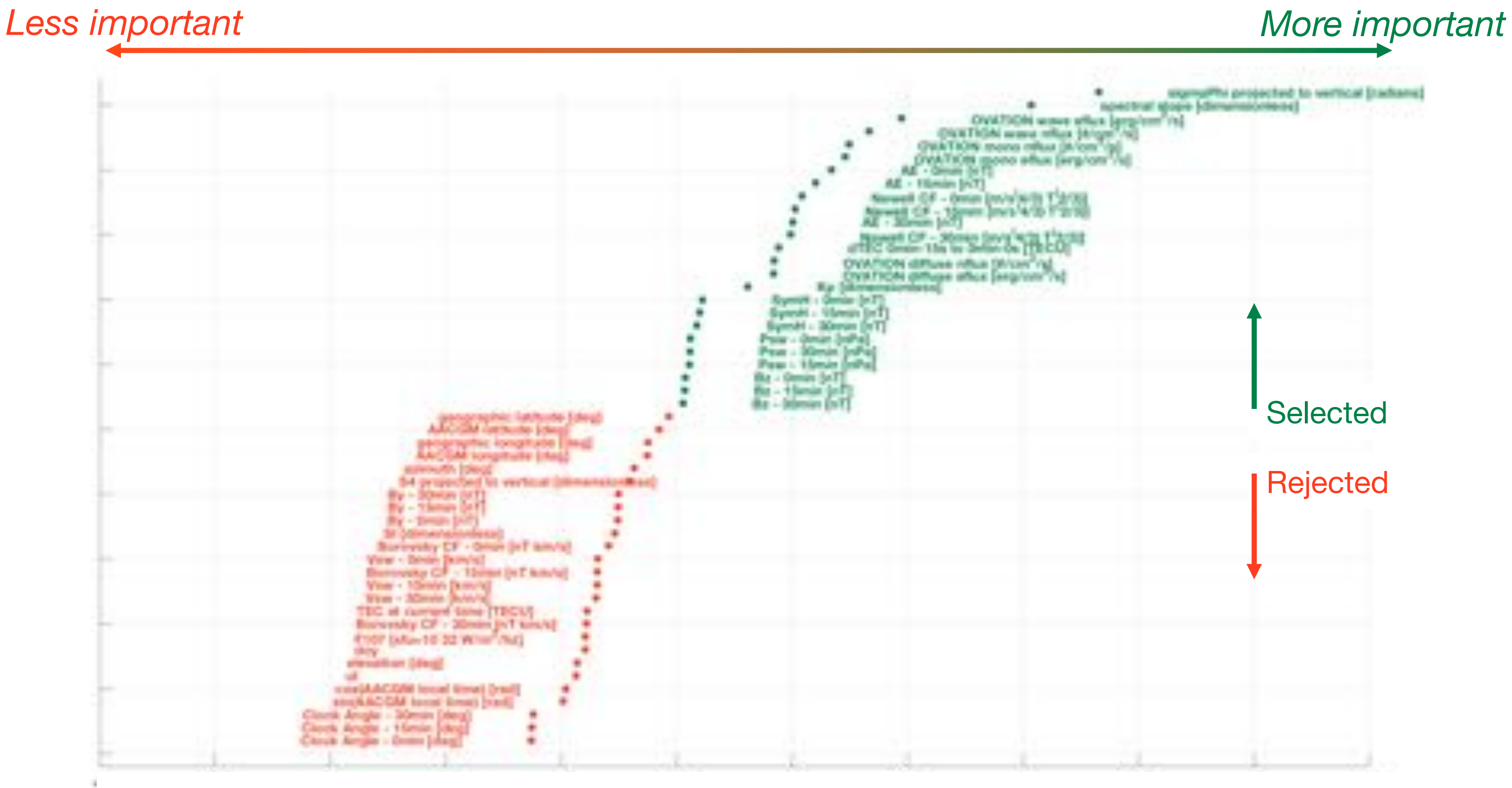
Define the predictive task



Step 2:
Define the predictive task



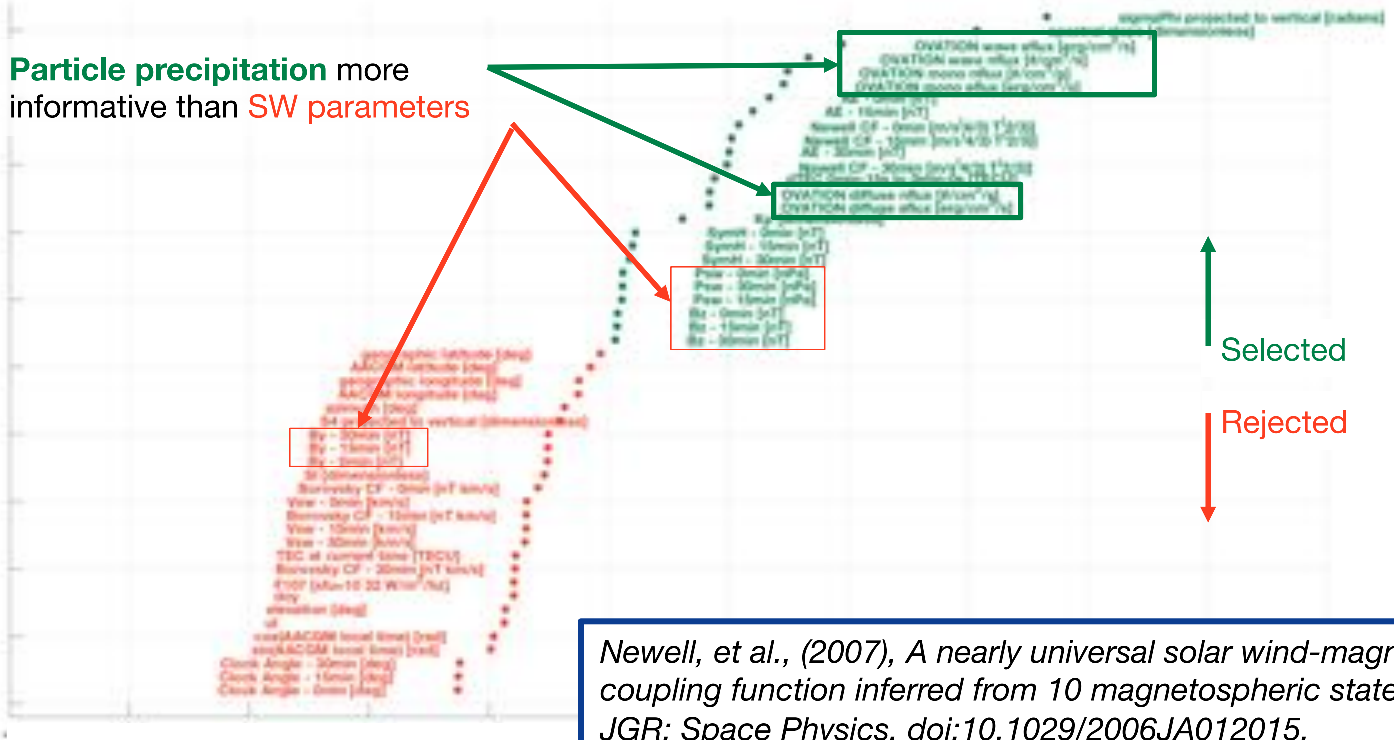
Step 4:
Interrogate the model



F-Score (i.e., importance)

Step 4:
Interrogate the model

Particle precipitation more informative than SW parameters



Selected

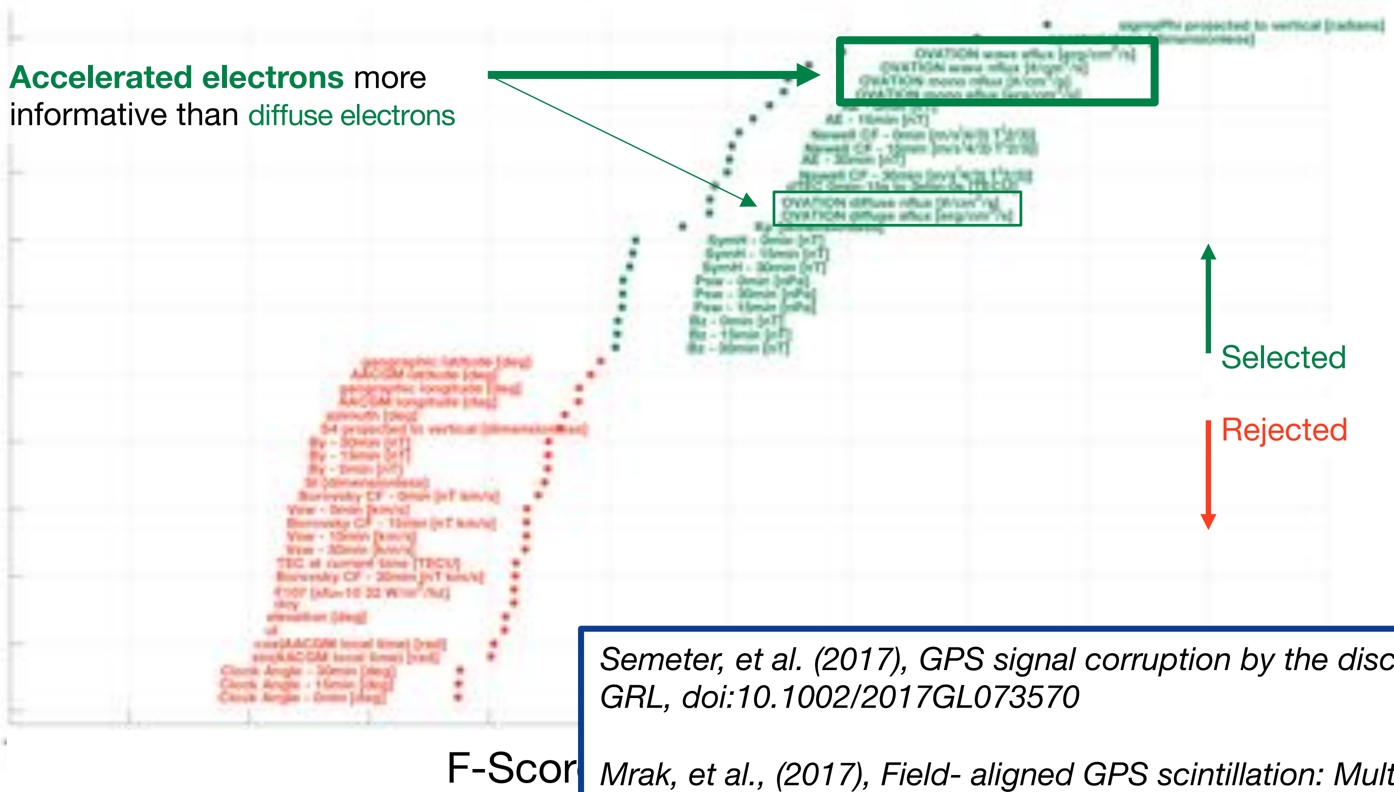
Rejected

F-Score (i.e., importance)

Newell, et al., (2007), A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables, JGR: Space Physics, doi:10.1029/2006JA012015.

Step 4:
Interrogate the model

Accelerated electrons more informative than diffuse electrons

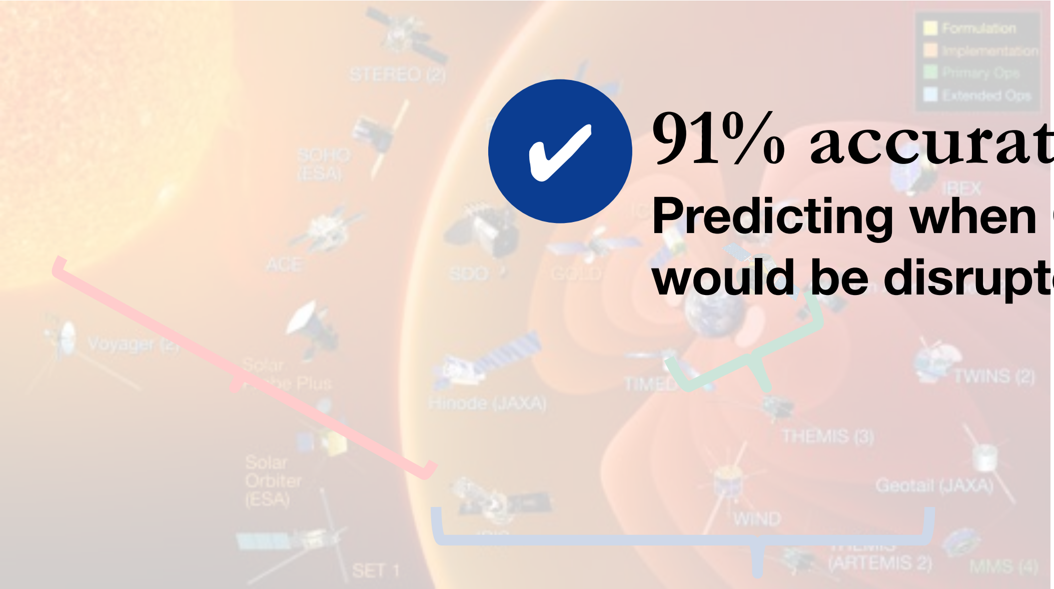


Semeter, et al. (2017), GPS signal corruption by the discrete aurora. GRL, doi:10.1002/2017GL073570

Mrak, et al., (2017), Field- aligned GPS scintillation: Multisensor data fusion, JGR: Space Physics, doi:10.1002/2017JA024557.

Data-driven space weather:

Machine learning



91% accurate

Predicting when GNSS signal would be disrupted

Support Vector Machine (SVM)

Decision Trees

Random Forests

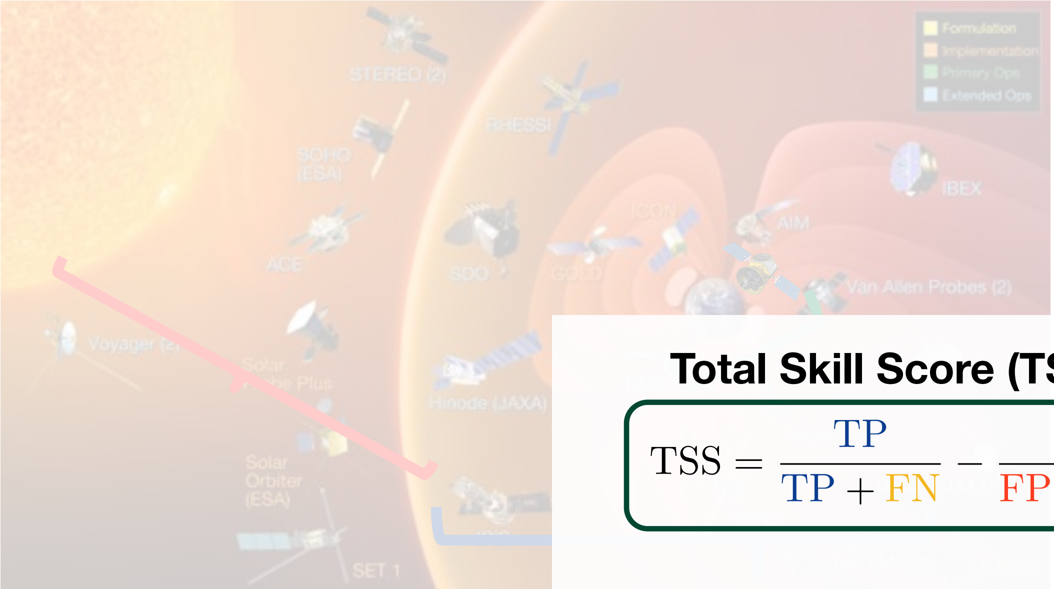
Neural Networks

Easily explainable → *Difficult to explain*

Create a narrative of new scientific understanding across spectrum of machine learning approaches

Data-driven space weather:

Machine learning



Total Skill Score (TSS)

$$\mathrm{TSS} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} - \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$

TSS

-1
Worst

+1
Perfect

Support Vector Machine (SVM)

Decision Trees

Random Forests

Neural Networks

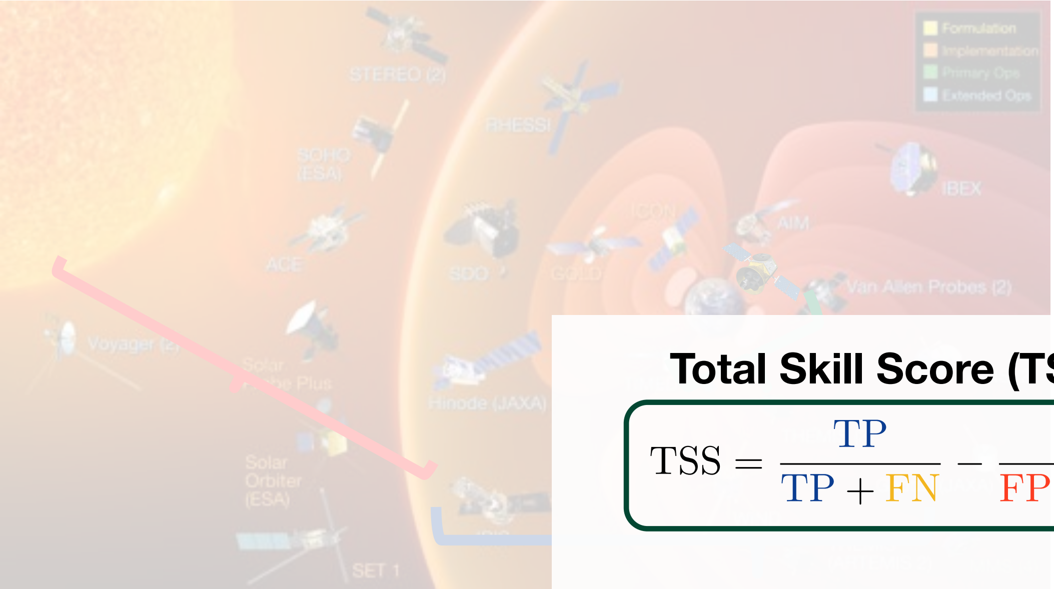
Easily explainable

Difficult to explain

Create a narrative of new scientific understanding across spectrum of machine learning approaches

Data-driven space weather:

Machine learning



Total Skill Score (TSS)

$$\boxed{\mathrm{TSS} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} - \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}}}$$

TSS

-1
Worst

+1
Perfect

Previous state-of-the-art

SVM benchmark

Support Vector Machine (SVM)

Decision Trees

Random Forests

Neural Networks

Easily explainable

Create a narrative of new scientific understanding across spectrum of machine learning approaches

Step 4:

Interrogate the model

Evaluation

True Skill Statistic (TSS)

$$\mathrm{TSS} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} - \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$

True label / **Predicted label**	no scintillation	scintillation
no scintillation	True negative	False positive
scintillation	False negative	True positive

Step 4:
Interrogate the model

Evaluation

True Skill Statistic (TSS)

$$\mathrm{TSS} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} - \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}}$$

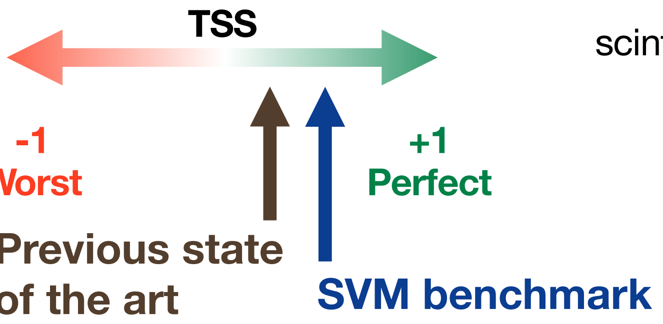
True label \ Predicted label	no scintillation	scintillation
no scintillation	True negative	False positive
scintillation	False negative	True positive

Step 4:
Interrogate the model

Evaluation

True Skill Statistic (TSS)

$$\mathrm{TSS} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} - \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$



True label \ Predicted label	no scintillation	scintillation
no scintillation	True negative	False positive
scintillation	False negative	True positive

- What have we learned?
  - Be radically interdisciplinary
  - Find the data science and be open by default: Leverage the advances in information technology to make data discoverable (i.e., develop agile and completely open data tools)
  - Understand the models: Interrogate models and bring domain knowledge (e.g., traditional approaches) together with new data-driven advances (e.g., ML)
    - Evaluation dashboard

Mapping ‘What’s next?’

- FDL

- MLaaS

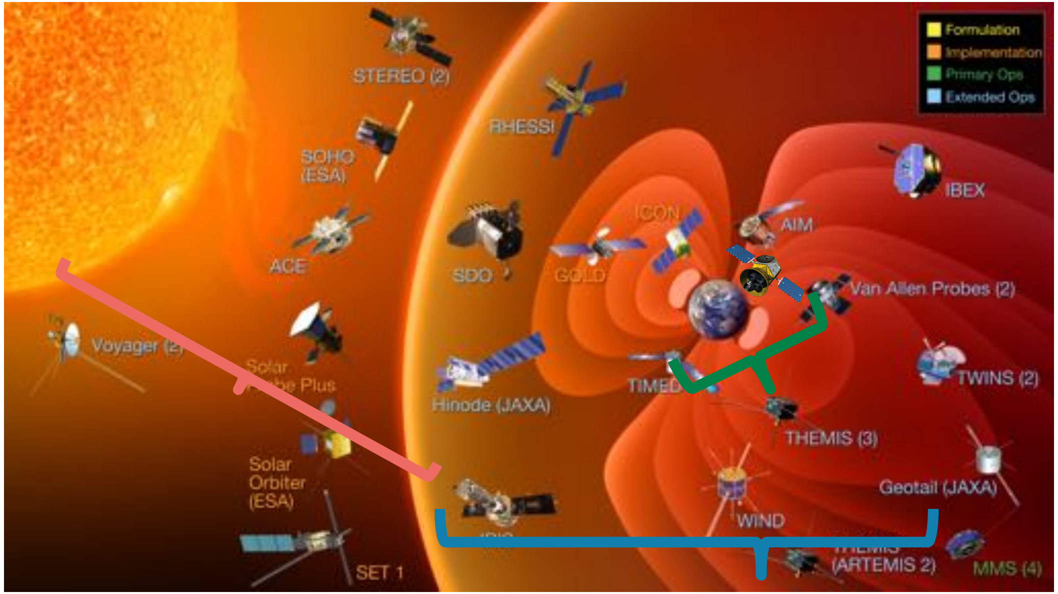
- Openness
  - Main point: Openness engenders new progress from radically different directions
    - Citizen Science
    - Successes from open source

- Data Viz
  - Main point: innovation can illuminate new connections and directions
    - Find an example to share from the sciences? Maybe something NSF-related?
    - Create my own example from simple text-based study and visualization

- Communication and broader impact
  - Main point: Innovation in how we communicate our science can enable us to overcome disciplinary boundaries
    - Podcast
    - Science on Tap seminar series

What's next?

Machine learning as a Service (MLaas)



Support Vector Machine (SVM)

Decision Trees

Random Forests

Neural Networks

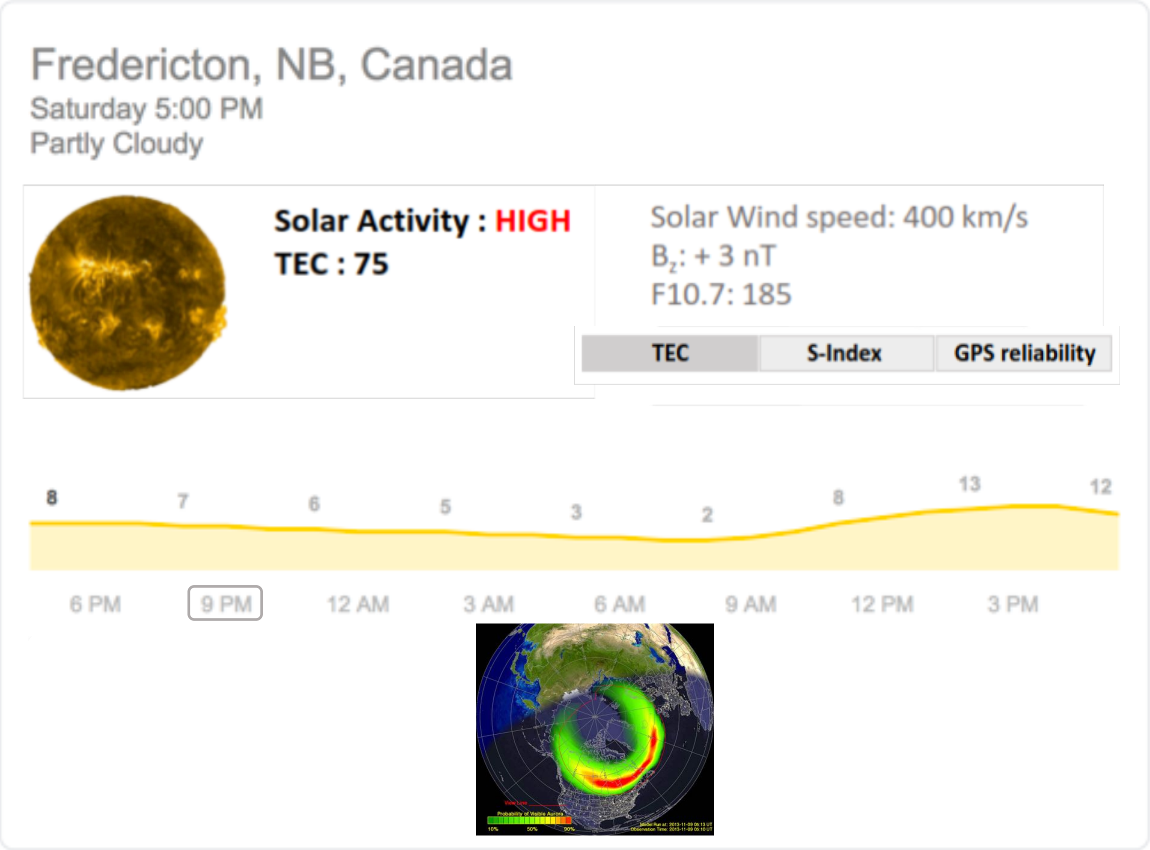
Easily explainable

Difficult to explain

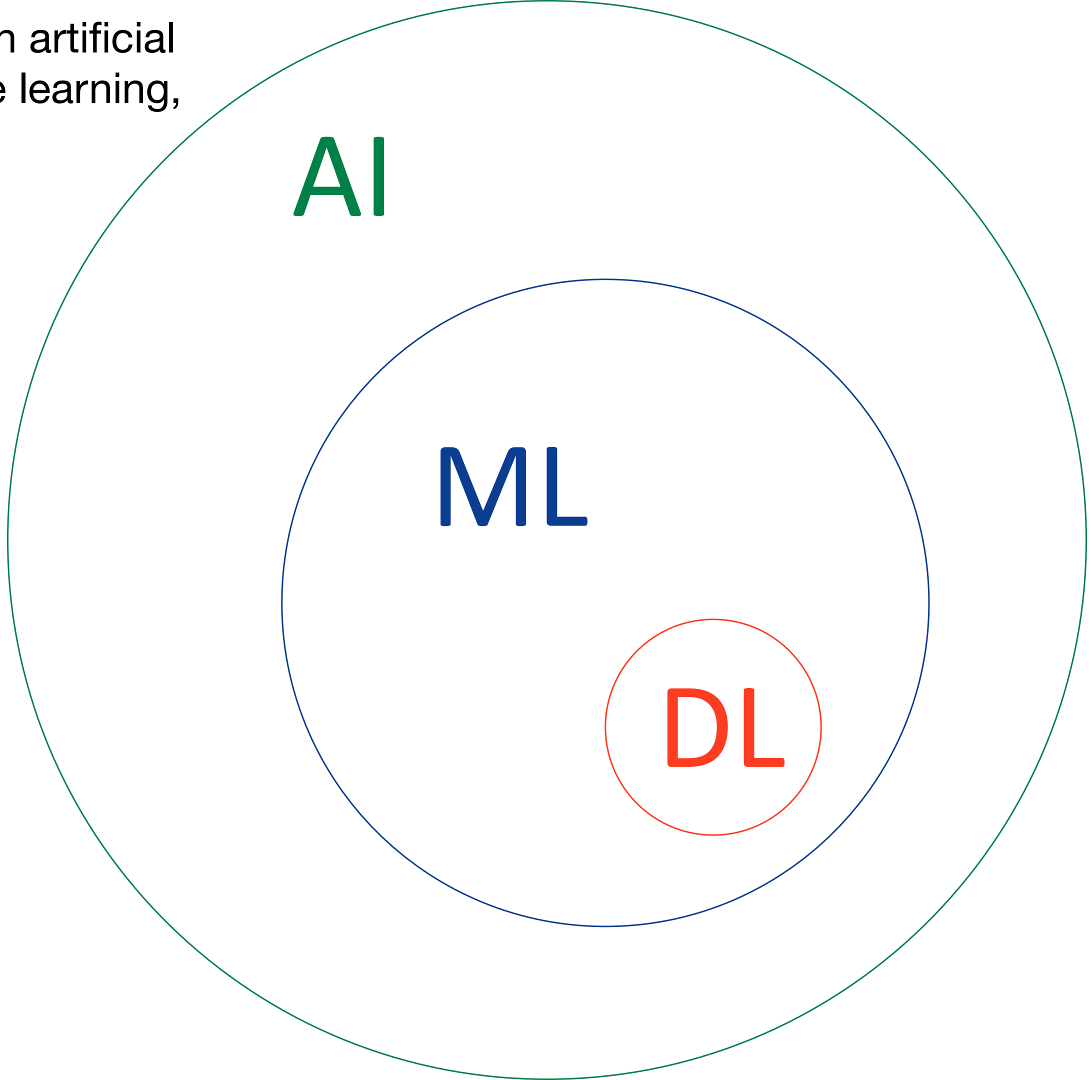
Machine Learning as a Service (MLaaS)

“Hey Google, what’s the weather like today?”

“Hey Google, what’s the ***space*** weather like today?”



Relationship between artificial intelligence, machine learning, and deep learning

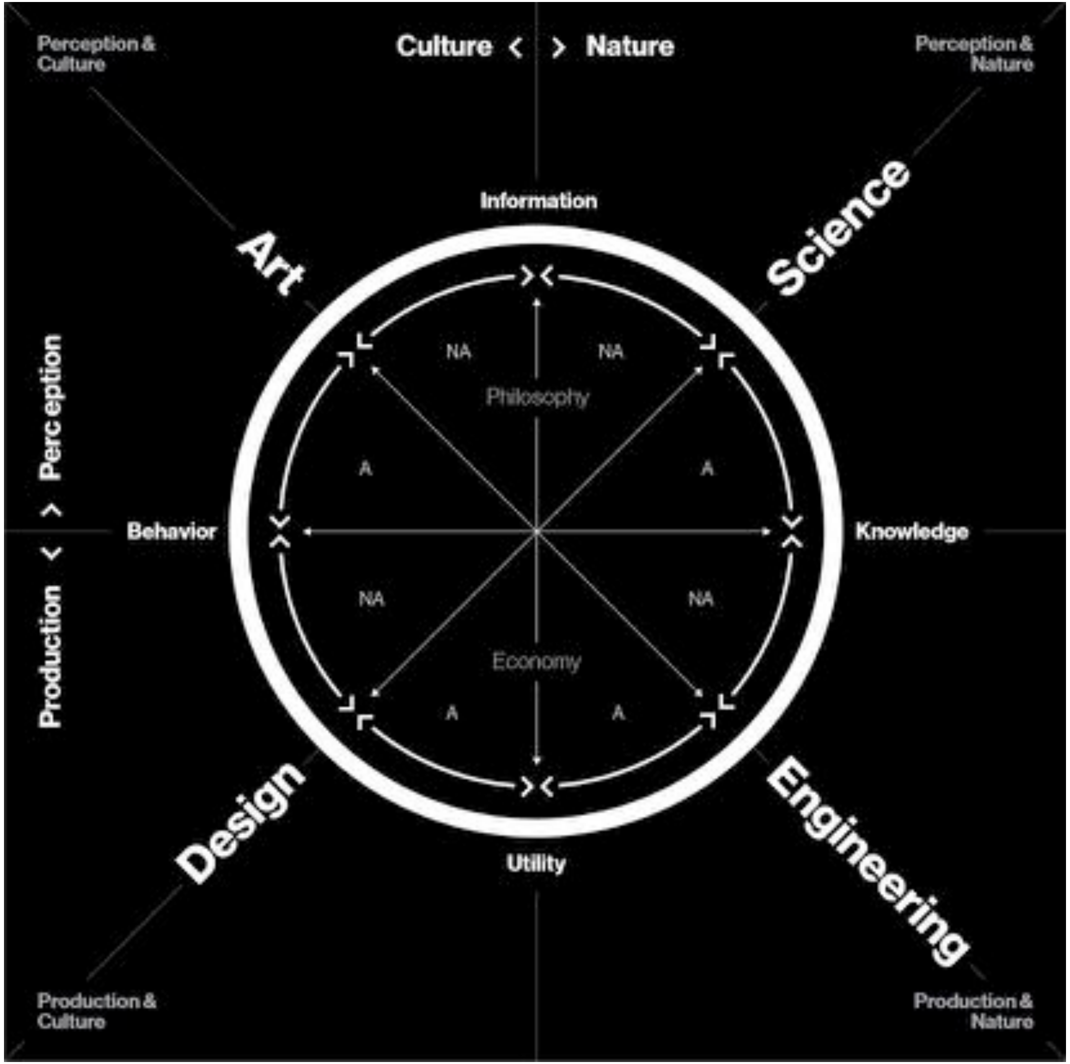


Toward convergence

Innovation

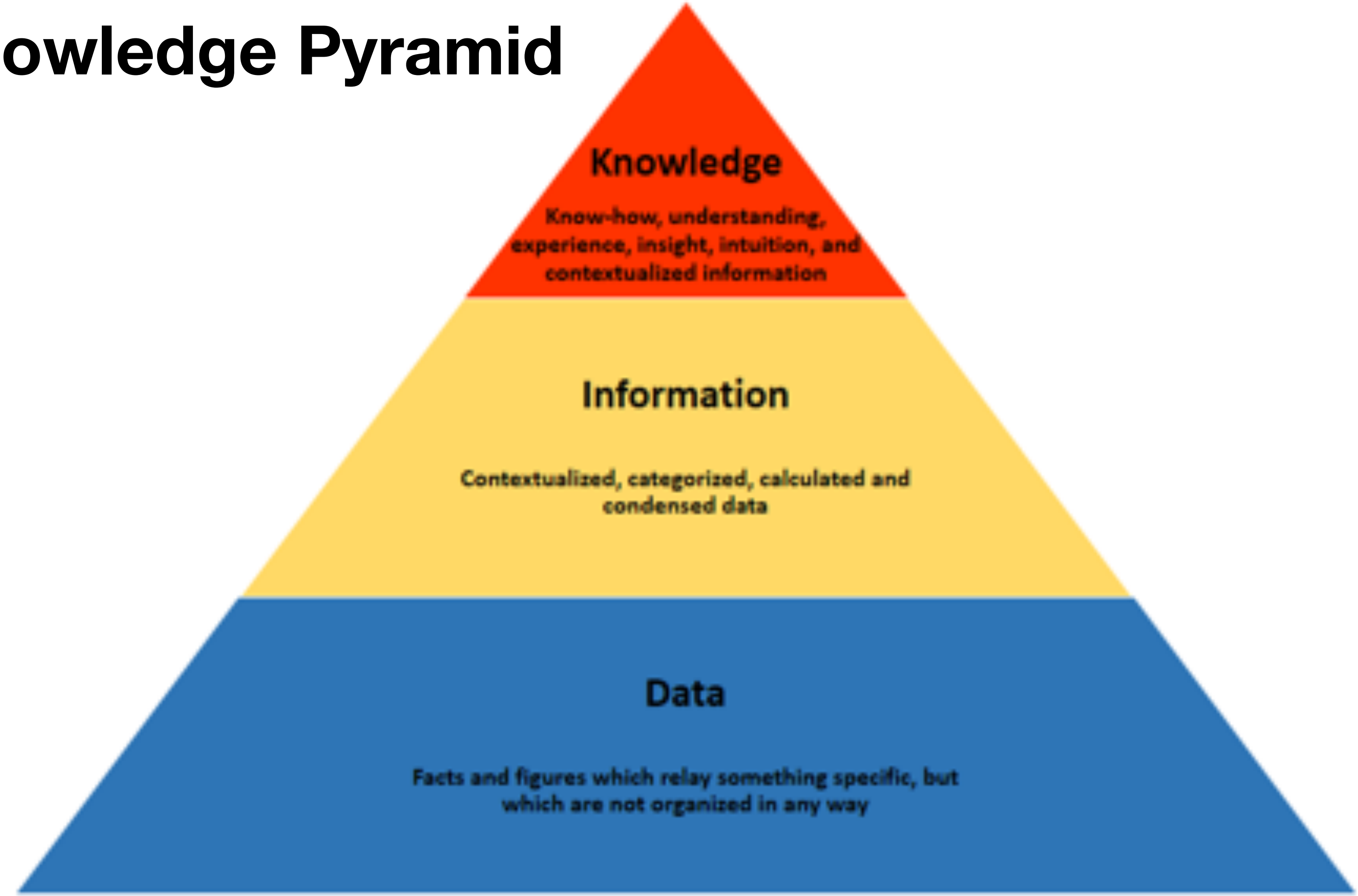
NASA Living With a Star Targeted Research and Technology 10-year vision report

Two NSF convergence HTF projects



Culture
Nature
Perception & Culture
Perception & Nature
Information
Art
Science
NA
NA
Philosophy
A
A
Perception
Production
Behavior
Knowledge
NA
NA
Economy
A
A
Utility
Design
Engineering
Production & Culture
Production & Nature

Knowledge Pyramid



BEPICOLOMBO*
EUCLID*
TEMPO
TROPICS
LANDSAT 9
ICON
SOLAR ORBITER*
LUCY
EUROPA CLIPPER
PSYCHE
SWOT
MARS 2020
OPERATING & FUTURE
SCIENCE FLEET


LSST
Large Synoptic Survey Telescope
OPERATING & FUTURE
SCIENCE FLEET