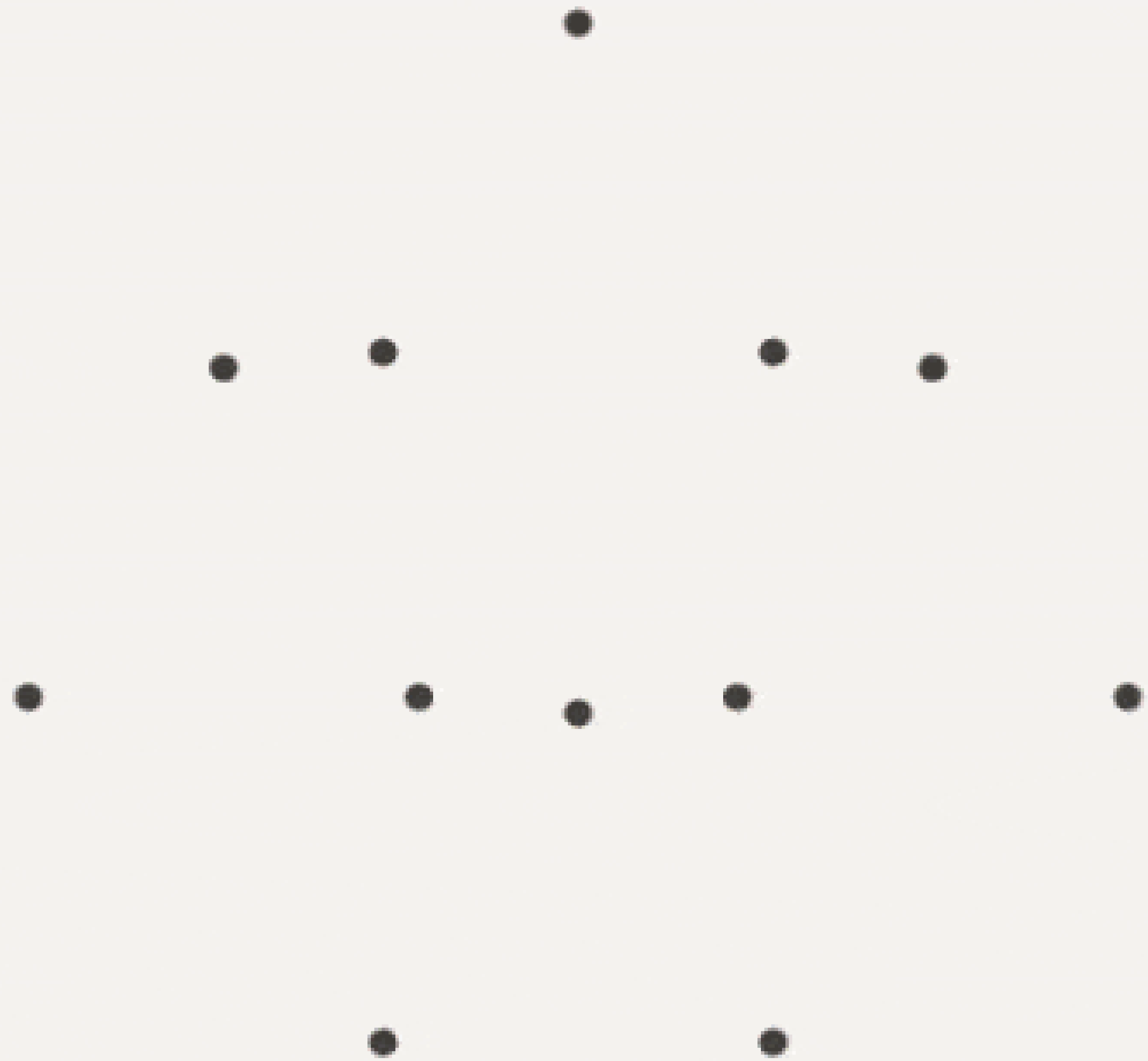
FRONTIER DEVELOPMENT LAB

A PUBLIC PRIVATE PARTNERSHIP FOR APPLIED AI FOR SPACE SCIENCE AND EXPLORATION

Lika Guhathakurta
NASA Ames Research Center
On behalf of the FDL Team

Artificial Intelligence is just fancy Statistics



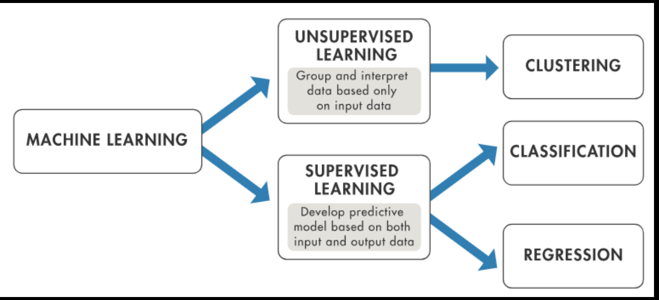
Artificial Intelligence : A Few Definitions

Artificial Intelligence (AI)

A computer which mimics cognitive functions typically associated with human intelligence.
Examples : goal seeking strategy formulation, complex image recognition, "learning", inference, and creative problem solving.

Machines Learning (ML): A branch of artificial intelligence in which a computer progressively improves its performance on a specific task by “learning” from data, without being explicitly programmed.

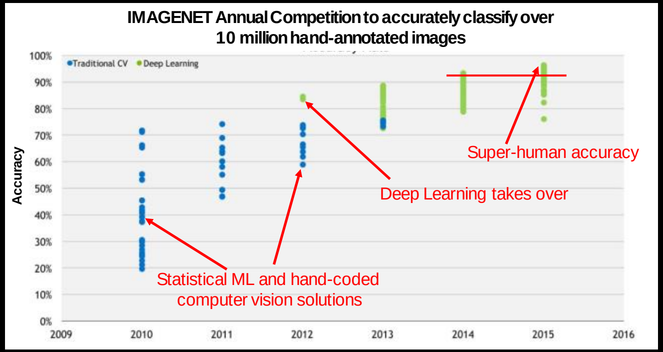
- Closely related to computational statistics, which focuses on prediction and optimization.



Data Mining: Discovering patterns in large data sets using techniques at the intersection of machine learning, statistics, and data management.

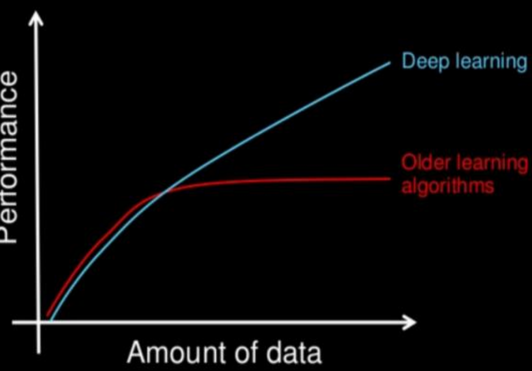
Deep Learning (DL): An extension of Machine Learning that uses the mathematical concept of a neural network (NN) to loosely simulate information processing and adaptation patterns seen in biological nervous systems.

- Many problems which have been traditionally tackled with pensive coding have been overwhelmingly superseded by neural nets that outperform the humans that trained them.
- Exponential investment (patents, publications, funding) has fueled rapid advances in DL capabilities to make predictions, to identify anomalies, and even create new content that mimics what it has previously seen.



Statistical Machine Learning vs. Deep Learning

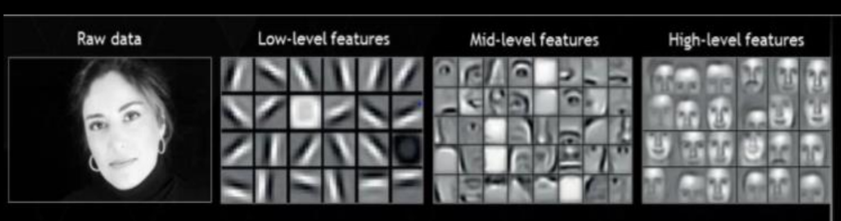
Data Scale: When properly architected, the efficacy of DL systems continue to improve with more data, long after statistical models have plateaued.



Interpretation: Machine Learning systems provide "visibility" into their statistical foundations, allowing their results to be interpreted and explained. Deep Learning systems are more of a "black box", although this is improving… and in some cases this is not an impediment (e.g. AI-enhanced science discovery)

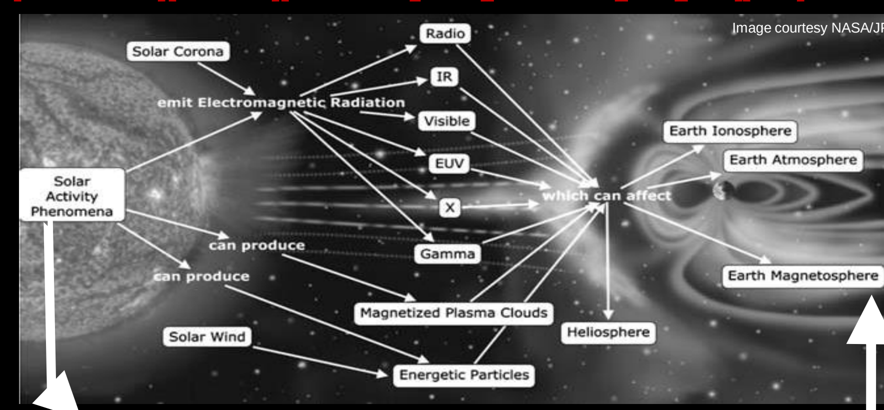
Whole System: Machine Learning typically requires that complex systems be "chunked" into trainable components that are then manually recombined. Deep Learning can often "short circuit" that process and successfully model complex systems from end-to-end

Feature Discovery: Machine Learning often requires a human expert to create "feature extractors" that enable the statistical models to learn effectively, but Deep Learning finds these high-level features for itself (often with surprisingly creative results)



Deep Learning will discover these feature abstractions for itself.
Machine Learning needs help to extract features for statistical modeling.

Multiple ML models for each component of the Solar-Terrestrial Environment



Deep Learning can often discover features to learn from the entire system

WHY NASA FDL?

AI is evolving quickly. The revolution was started by the application of neural nets to large quantities of pre-labeled data (known as '**supervised' deep learning**) in 2012. These methods are now mainstream.

The state-of-the-art is now looking towards sparse or unlabeled data (known as **'unsupervised' deep learning or machine learning**) and how to explain the uncertainty of the results. This is known as **'explainability'** and a key part of determining the veracity and usefulness of any AI.

Although deep learning is being democratized, **producing excellence** (i.e explainable, bias free and trustworthy results) is increasingly **difficult**, requiring multi-faceted and sophisticated teams and a **deep understanding** of how to use the **newest techniques. This is where FDL sees its unique value and opportunity for space science and exploration.**

"Deep Learning Big Bang"
2012

FDL Initiated
2015

NOW

Supervised Pre-labeled data (by humans)
- becoming democratized, simple workflows, clean datasets. ML is 'black box'

Unsupervised Sparse or unlabeled data

Bayesian and Probabilistic Deep Learning
Quantifying uncertainty - leadership in explainability

Data fusion Rapid increase of data size, compute cost and workflow complexity requires AI management capacity

FDL FRONTIER DEVELOPMENT LAB

"AI is changing the way we think about problems"

Massimo Mascaro, Google Cloud Technical Director of Applied AI

Executive overview

FDL is an applied **Artificial Intelligence Research** initiative that uses interdisciplinary teams at the Phd and Postdoc level to solve challenging problems for space exploration and positive impact to humankind.

- FDL was conceived by the Office of the Chief Technologist at NASA HQ in 2015, with two primary objectives:
  - Understanding how AI and Machine Learning could be leveraged to advance basic research questions of importance to NASA and accelerate discovery and understanding, and or improve research efficiency and efficacy.
  - Exploring opportunities represented by public/private partnership where companies with technology and expertise in AI/ML domains could see the benefit in supporting NASA research programs and priorities and leveraging NASA data for societal benefit.

HISTORY

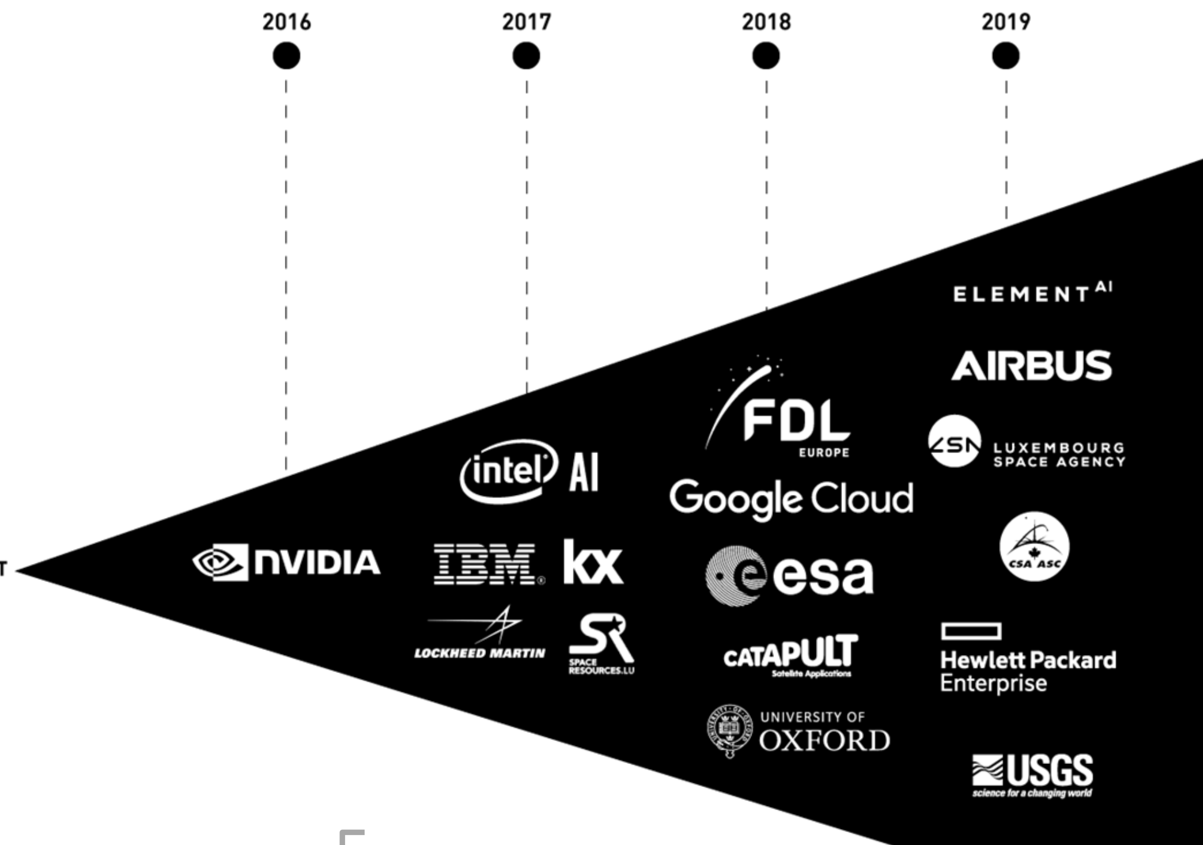
THE NASA FRONTIER DEVELOPMENT LAB (FDL) IS A PUBLIC / PRIVATE **APPLIED AI RESEARCH PARTNERSHIP** BETWEEN NASA, THE SETI INSTITUTE AND LEADERS IN COMMERCIAL AI, PRIVATE SPACE, ACADEMIA AND PARTNER SPACE AGENCIES.

- NASA FDL is four years old
- FDL has as developed a proven formula for producing excellence in applied AI research over very rapid timescales - with a focus on 'AI explainability' to match the quality expectations of the space industry.
- FDL has produced 15 peer reviewed journal papers and been accepted to 30+ scientific conferences and multiple articles in the science press. FDL results have already been deployed on NASA programs.
- The FDL brand is well respected in the research community with 450+ researcher applications in 2019. (Acceptance rate is now parity with MIT.)
- NASA FDL is currently based at NASA ARC and hosted and administered by the SETI Institute.
- The formula has attracted the attention of partner space agencies, ESA, CSA, LSA, with more to come.
- Other NASA centers are showing interest too, particularly GSFC, MSFC, JSC and Glenn.

MODEL (1): A public / private partnership

FDL has a dedicated and passionate partner community aligned behind a common vision for AI for science and space exploration for all humankind.

FRONTIER DEVELOPMENT LAB



HOW THE PARTNERSHIP IS STRUCTURED:

SETI INSTITUTE

The SETI Institute facilitates public / private partnership by acting as a hub between Space Agency, Academic and commercial partners.

TRILLIUM

Sub-contractor Trillium Technologies Inc, manages most of the commercial, academic and space agency partners and the FDL faculty and runs and co-ordinates FDL throughout the year.

MODEL (3) PROCESS INNOVATION

HOW FDL WORKS:

FDL tackles knowledge gaps in space science by pairing machine learning experts with heliophysics, astrophysics, astrobiology, planetary science and earth science researchers for an intensive eight week research sprint, held in the summer break of the academic year - although the **journey from Challenge Definition through to finished result (Tech Memo and trained algorithm and data products) takes 12 months.**

Interdisciplinary four-person teams of PhD and postdoc level researchers address tightly defined science challenges that are **informed by knowledge of "what's possible in ML".** Mentors who are subject matter experts, provide support to the teams and drive research quality. External and partner experts, special guests, and visits to partner labs contribute to the understating of the problem and provide a community of expertise that drives excellence.

FDL's format encourages rapid iteration and prototyping to create outputs with meaningful application to the space program, with **substantial compute resources provided by FDL's commercial partners* - who have expressed ongoing commitment.** This combination of curated challenges, close mentorship, community of expertise and an emphasis on rapid prototyping has ensured a high success rate for FDL.

As such, FDL has demonstrated how structured interdisciplinary problem solving, radical collaboration methods and partnering with commercial organizations with relevant expertise can be useful to NASA's science and technology goals.

***$1.5 M USD in donated compute (2016-2019)**

PROGRAM IMPACT: DEMONSTRATED CAPACITY

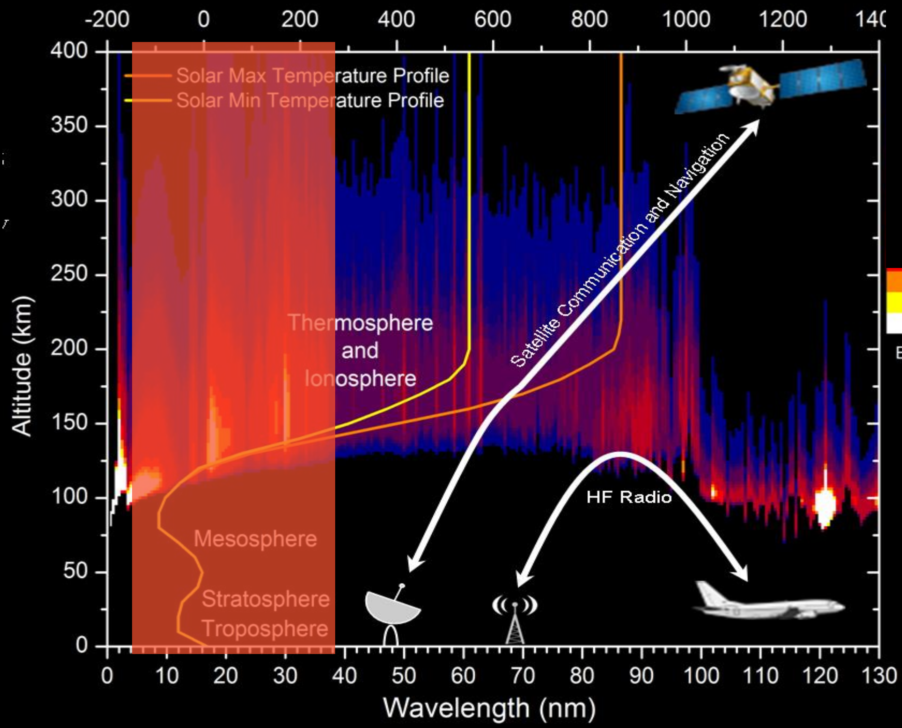
FDL has explored a broad range of AI applications for the space program.

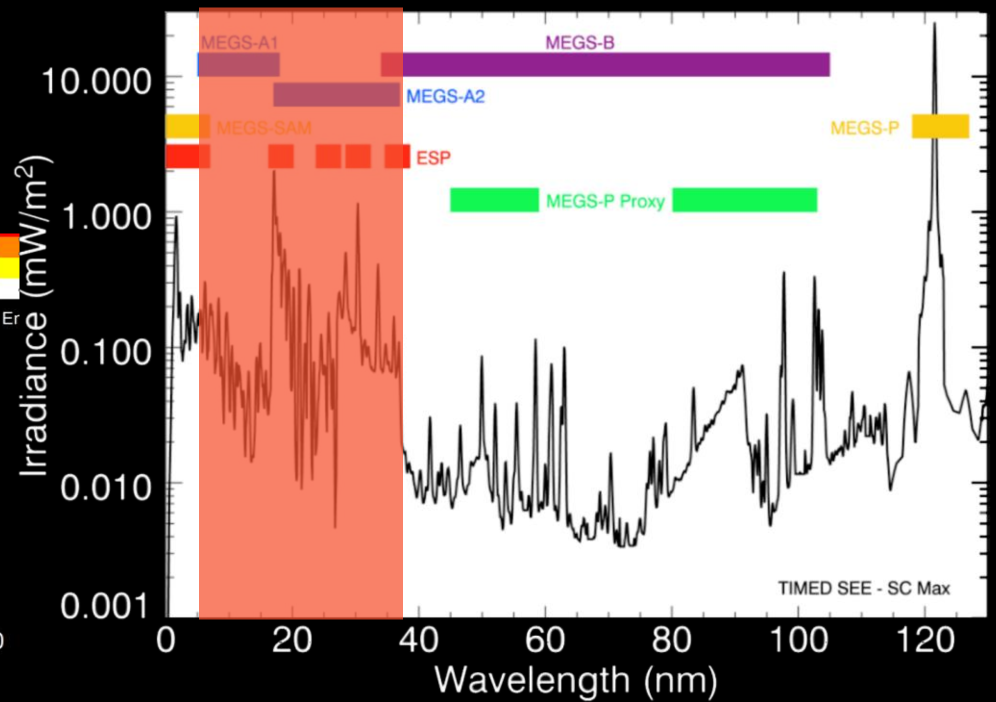
FDL is building crucial AI keystone capability for NASA

FDL RESEARCH RESULT	Discovery with multi-dimensional data	Workflow optimization	Forecasting and Prediction	Decision Intelligence	Anomaly Detection	Autonomous systems
STRATEGIC CAPABILITIES FOR NASA	Science breakthroughs	Radical cost and time savings Classification at scale	Situational Awareness / Uncertainty reduction Virtual Instruments	Rapid assessment / improved strategic decisions	Safer systems	On-board AI + Swarms Auto Calibration


FDL

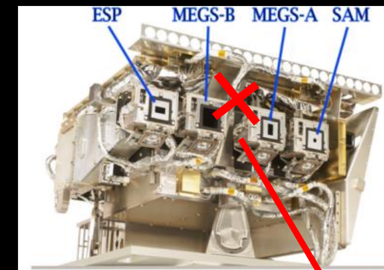
Loss of sensor in SDO/EVE left an observational gap in the most energetic part of the EUV spectrum





FDL 2018 Case Study
SYNTHESIZE SDO MEGS-A TO DATA

NASA Solar Dynamics Observatory (SDO)



Failed in 2014

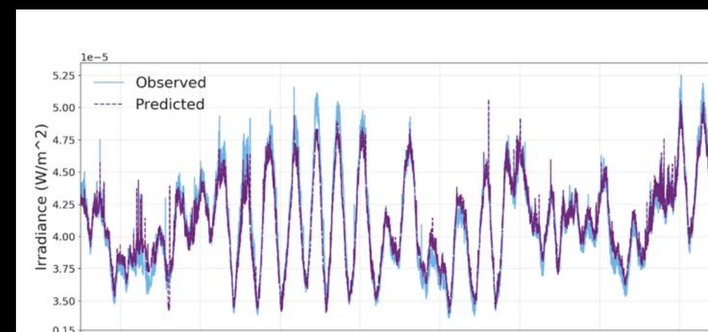
- **Need:** Measurement of solar spectral irradiance is needed for satellite orbit boost planning. Currently, this can be difficult because the MEGS-A module on SDO stopped functioning in 2014.

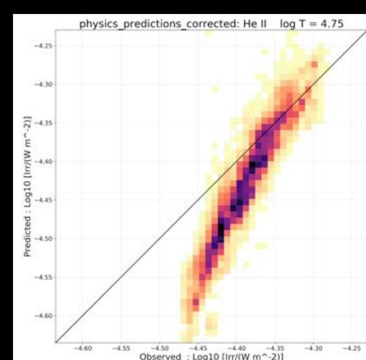
- **Goal:** The SDO AIA EUV imager co-observed with MEGS-A from 2011 to 2014 -- Can we use this data overlap to train a deep learning model to “virtually resurrect” the MEGS-A instrument and fill the observational gap left by the MEGS-A failure, thereby improving spectral irradiance prediction?

- **Methodology:** Develop a machine learning model using 2011/2014 data, test the accuracy using 2012/2013 data. After training and testing over 1000 machine learning configurations, the best implementation was found to be a Residual neural net model augmented with a Multi-Layer Perceptron.

- **Findings:** The neural net model significantly improved upon physics based models, **reducing mean error from 7.46% to 2.83%.** This improved accuracy may constitute a scientifically useful virtualization of MEGS-A.

AI model reduced mean error of spectral irradiance prediction to 2.83%

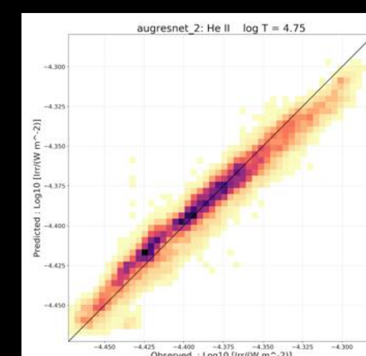




Plot of Predicted vs.Observed

Physics-based model

Improved AI model



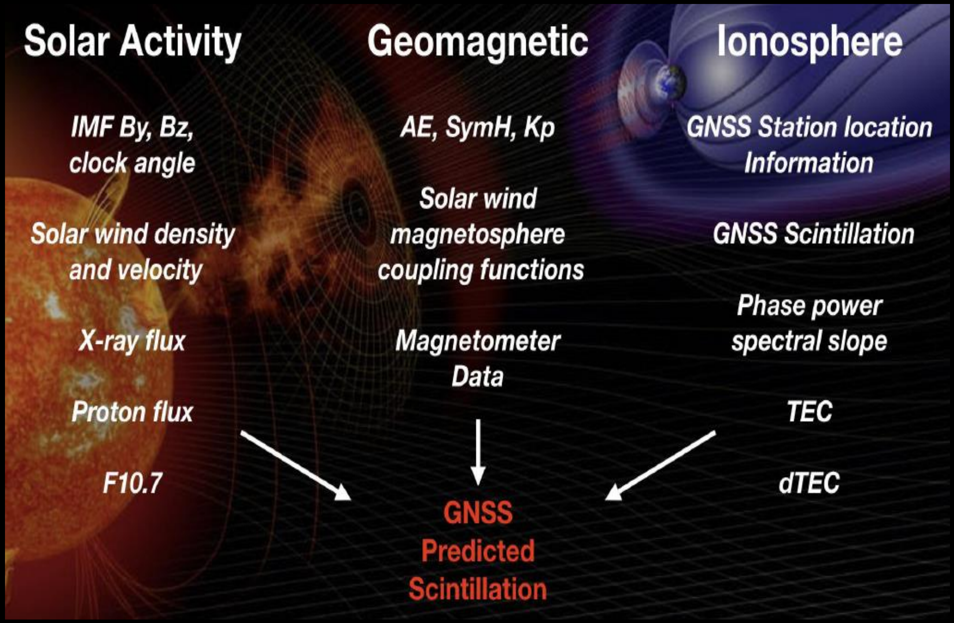
FDL 2018 Case Study
FORECASTING GNSS/GPS DISRUPTIONS

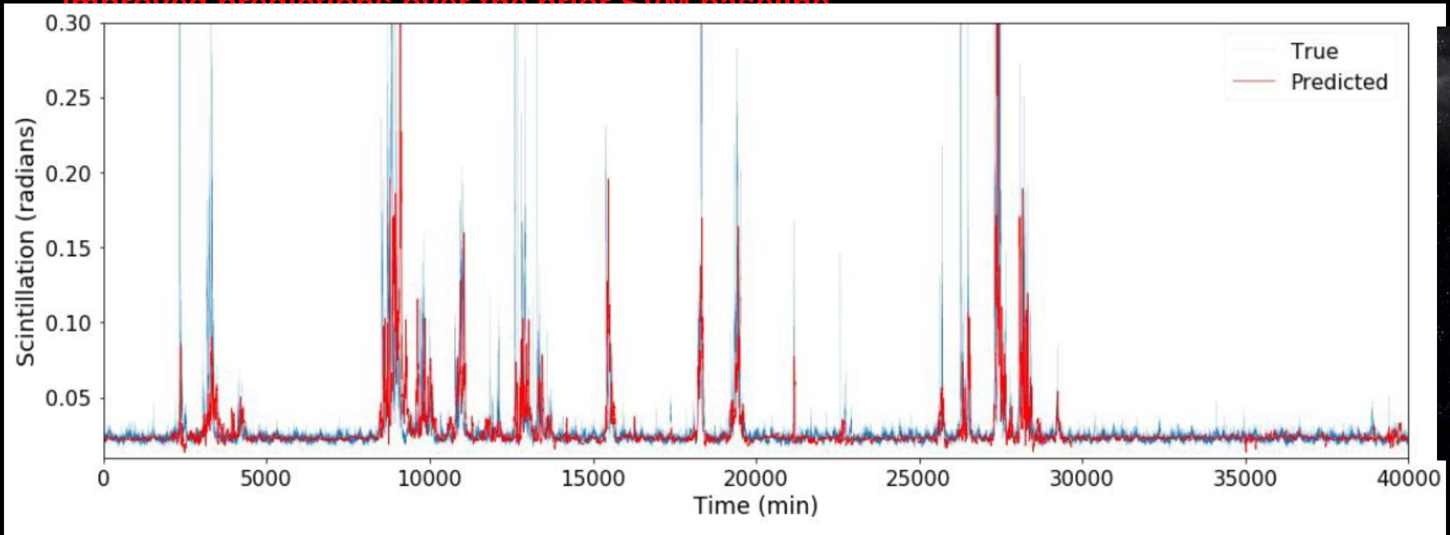
- **Need:** GNSS/GPS systems are a critical component of our global technology infrastructure. We must improve our ability to predict how space weather will degrade GNSS accuracy.

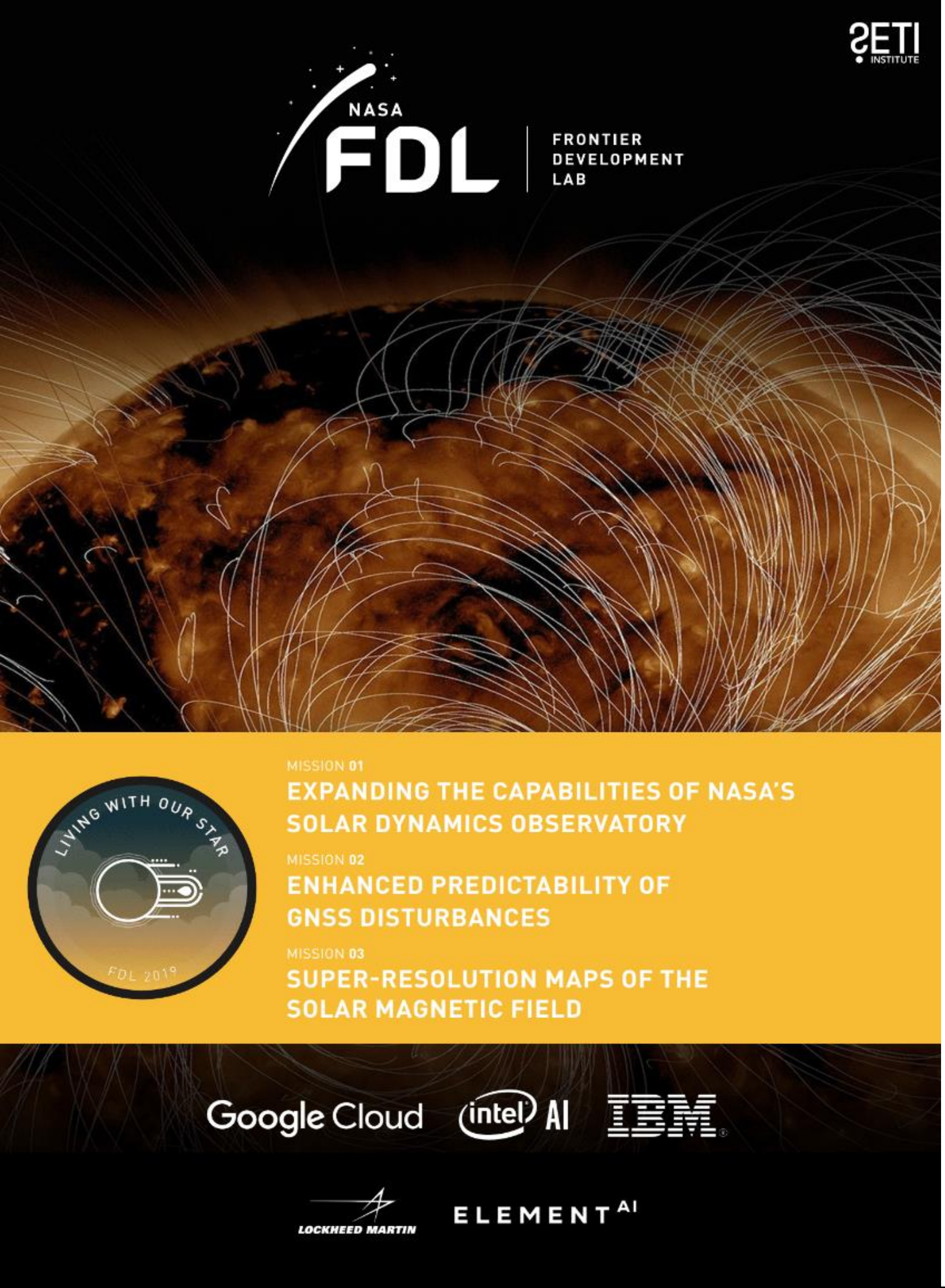
- **Goal:** Use high-latitude ionospheric and geomagnetic data in conjunction with solar data (OMNI database) to predict GNSS signal scintillations.

- **Methodology:** Curate over 350GB of data (2015-2018) to extract over 100 features for model training. Compare a baseline Support Vector Machine (SVM) model with a Multi-layer perceptron (MLP) neural net implementation.

- **Findings:** The neural net model significantly **improved predictions over the prior SVM baseline**







EXPANDING THE CAPABILITIES OF NASA SOLAR DYNAMICS OBSERVATORY

Challenge:

- By using a prepared “AI-ready” SDO dataset, this challenge aims to transform multiple EUV channels data into extreme ultraviolet (EUV) images.

  This will help the reduced instrumentation strategy that will be central to the success of future SmallSat missions.

  Using the same dataset, this challenge will also identify spatial patterns on the Sun to determine the calibration factor that would correct for SDO EUV instrument degradation, which would help to avoid the cost of regular suborbital launches to obtain calibration data.

2019 **Auto-cal+ Virtual Telescope**

EXPANDING THE CAPABILITIES OF NASA'S SOLAR DYNAMICS OBSERVATORY

Back to **AI Portfolio**

NEED > CHALLENGE

1. UV and EUV instruments in orbit suffer time-dependent degradation which reduces instrument sensitivity. Accurate calibration for EUV instruments currently depend on sounding rockets (e.g. for SDO/EVE, and SDO/AIA), which are costly and infrequent. Furthermore, such calibration experiments are not practical for missions in deep space (e.g. STEREO satellites). Using the SDO data, we propose to exploit spatial patterns in multi-wavelength observations to arrive at a **auto-calibration** of (E)UV imaging instruments.

2. The capabilities of Heliophysics missions are limited by the cost of launch, of instrument development and of telemetry. We propose the development of a **virtual telescope** that can generate desired science data products using fewer measurements (e.g. fewer EUV channels) as a possible solution to mitigate these challenges.

RESULTS

- The team devised a Convolutional Neural Network (CNN) that takes multi-channel EUV images as input, and outputs per-channel degradation factors. The CNN solution outperforms a baseline method which uses pixel intensity histogram analysis.
- The team trained a Deep Neural Network (DNN) with a U-net architecture to synthetically generate AIA 211 Å channel images from three other (94, 171, 193 Å) EUV images. The synthetically generated image has good correspondence with ground truth images over three orders of magnitude dynamic range.
- These techniques demonstrate how we can enhance the scientific return of space missions (especially deep space missions like STEREO), and paves the way for an autonomous space weather constellation.

DATA / METHOD

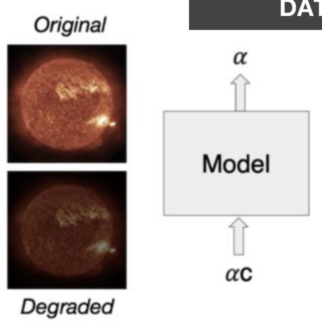


Fig.1 A schematic of the auto-calibration problem.

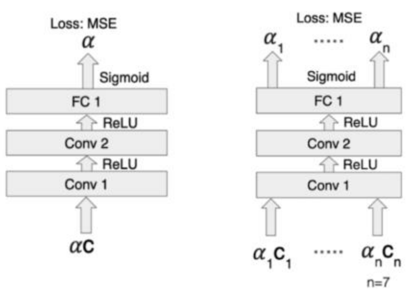
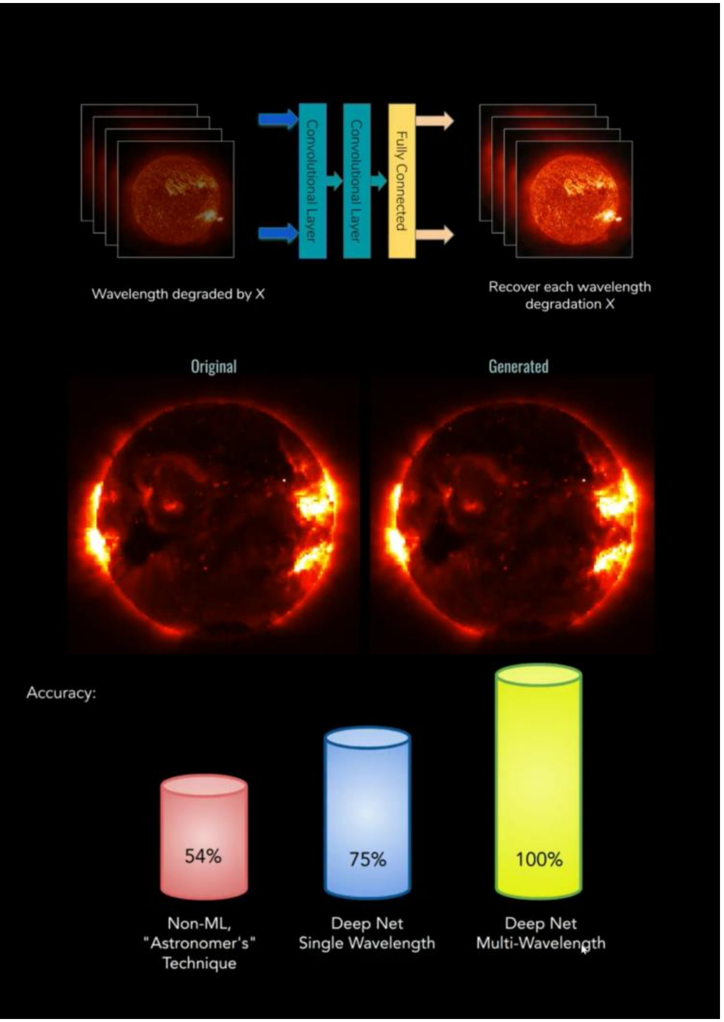


Fig.2 CNN taking in one channel (left) and multiple channels (right).

The machine learning-ready SDO dataset prepared by Galvez et al. (ApJ 2019) was used for both challenges. The dataset consists of a subset of the original SDO data dating from 2010 to 2018, is comprised of 7 EUV channels + 2 UV channels + HMI vector magnetograms.

STATUS > FUTURE WORK

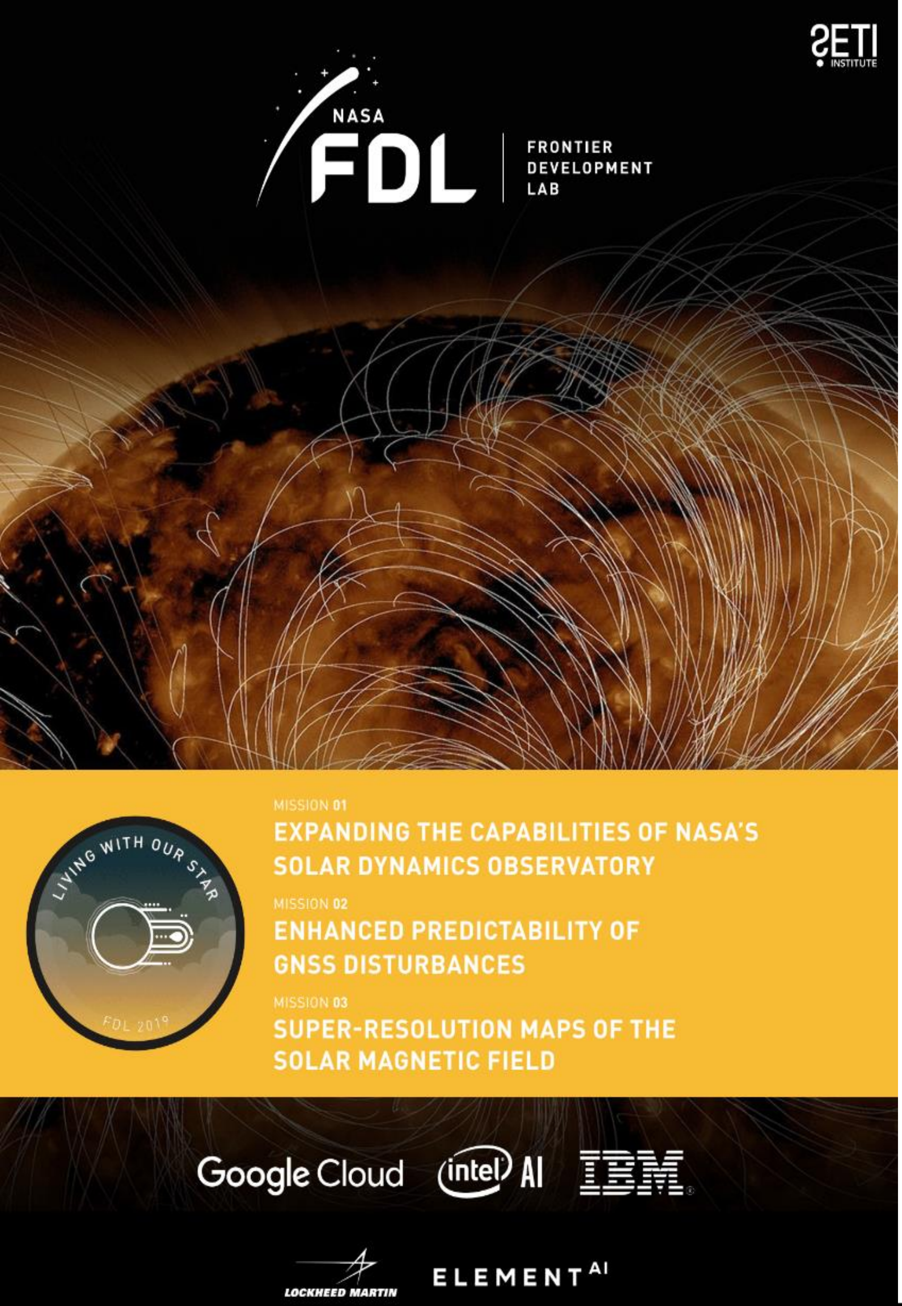
- The team "plug-and-play" pipeline that allows feeding input from the SDO dataset as well as plugging in either the auto-calibration or the virtual telescope experiments via a configuration file. The source code, as well as documents on how to use the code and explanations will be made available on GitLab after the first publication.
- The team is transitioning from the IBM Cloud platform to the Google Cloud platform to refine model development and deployment.
- Two abstracts have been submitted to the Machine Learning for Physical Sciences Workshop at NeurIPS 2019.



EXPANDING THE CAPABILITIES OF NASA'S SOLAR DYNAMICS OBSERVATORY

Results overview:

- The team devised a way for solar extreme UV telescopes to self-calibrate, improving our capability to monitor space weather.
- Furthermore, the team created a synthetic telescope to image the Sun's corona.
- These techniques enhance the scientific return of space missions, and paves the way for an autonomous space weather constellation.

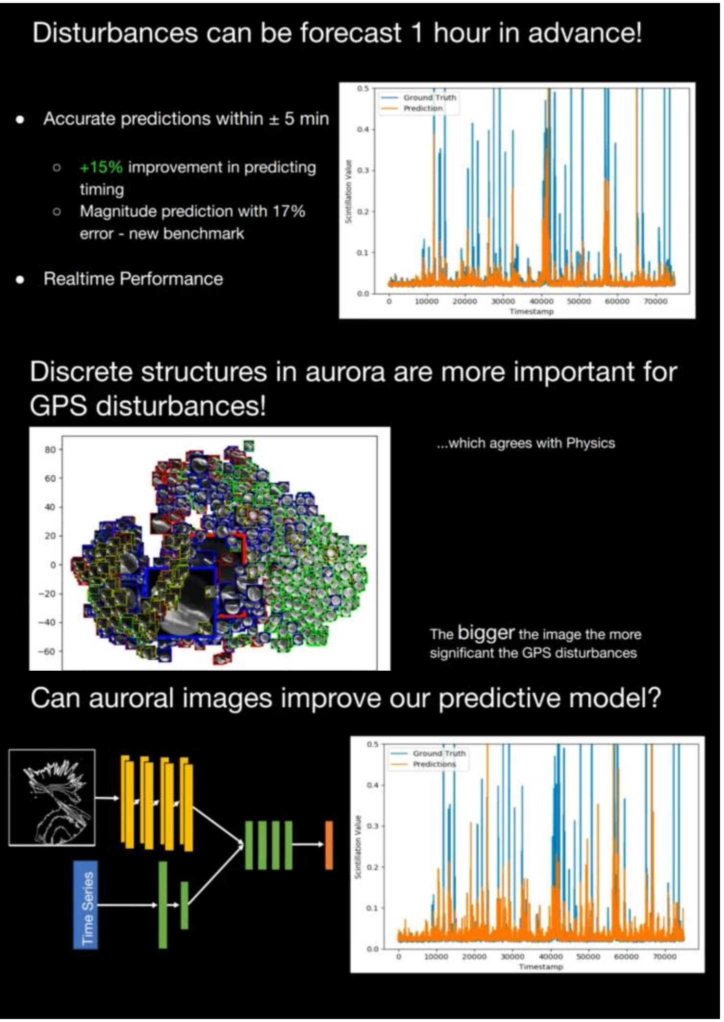


ENHANCED PREDICTABILITY OF GNSS DISTURBANCES

Challenge:

- Solar output can range from low-velocity solar wind to episodic eruptions like CMEs that have the potential to negatively impact communication/ navigation systems and other critical elements of our techno-social infrastructure.

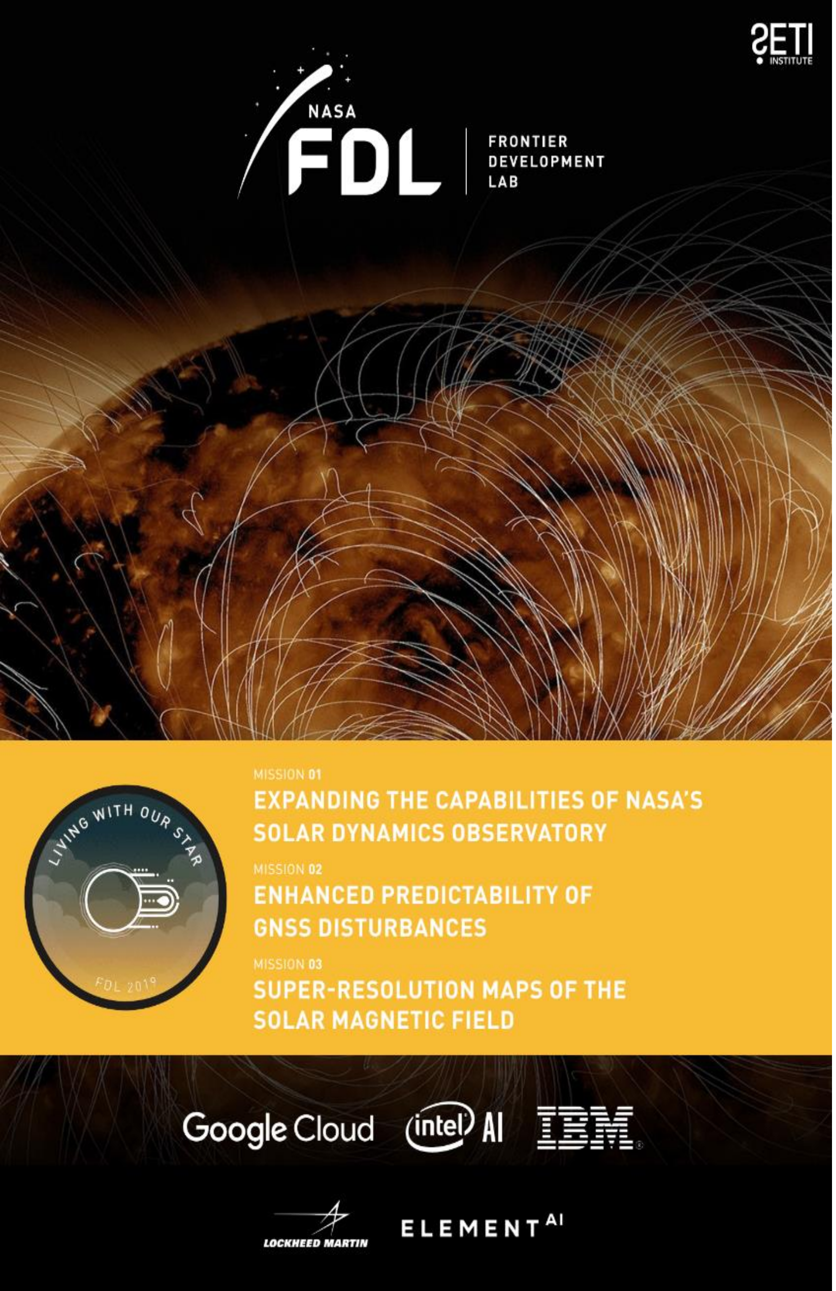
- An unresolved question is whether certain solar outputs will be 'geoeffective' – meaning effective in generating disruptive effects in the solar-terrestrial system.



ENHANCED PREDICTABILITY OF GNSS DISTURBANCES

Results overview:

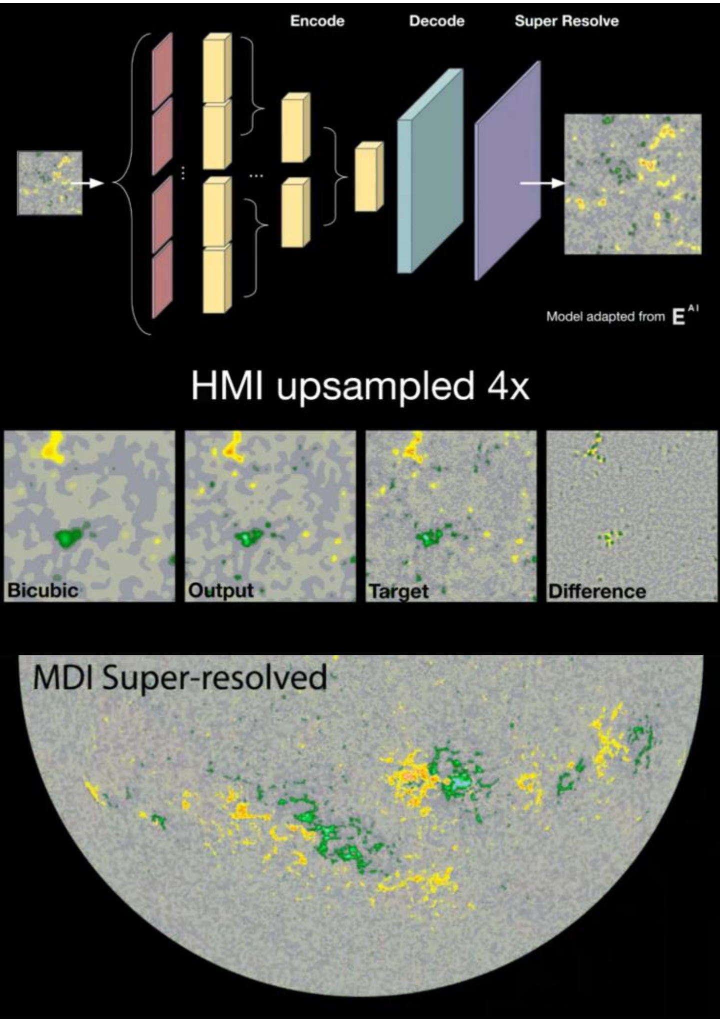
- The team used a novel machine learning approach of bringing together auroral imagery and solar-magnetosphere-ionosphere observations to improve the predictability of GPS/GNSS signal disruptions.
- By using ML techniques to understand auroral structures, they achieved 15% improvement over the state of the art and instantaneous results.



SUPER-RESOLUTION MAPS OF THE SOLAR MAGNETIC FIELD COVERING 40 YEARS OF SPACE WEATHER EVENTS

Challenge:

- Predicting geo-effective space-weather events is challenged by the time-limited coverage of SDO data (2010-present).
- This challenge proposes to address this problem by using deep learning solutions to upscale lower resolution images from earlier missions, thereby allowing for a second neural net to normalize and combine a much longer temporally-composited data product from multiple solar observation missions.



SUPER-RESOLUTION MAPS OF THE SOLAR MAGNETIC FIELD COVERING 40 YEARS OF SPACE WEATHER EVENTS

Results overview:

- Used state of the art deep neural networks to calibrate and super-resolve historical maps of the solar magnetic field.
- This addresses a problem that the heliophysics community has been unable to solve in 50 years and enables the study of both space weather and space climate evolution.

AI Capability Portfolio

2016 - 2019

EARTH SCIENCE

2019 **Cloud Classification**

EARTH SCIENCE - DISASTERS

2018 **Informal Settlements**

2018 **Disaster Response**

2019 **Flood Mapping**

2019 **Edge Inference: Floods**

HELIOPHYSICS - SPACE WEATHER

2019 **GNSS Disturbances**

2018 **Ionospheric Scintillations**

HELIOPHYSICS

2019 **Auto Calibration**

2019 **Virtual Telescopes**

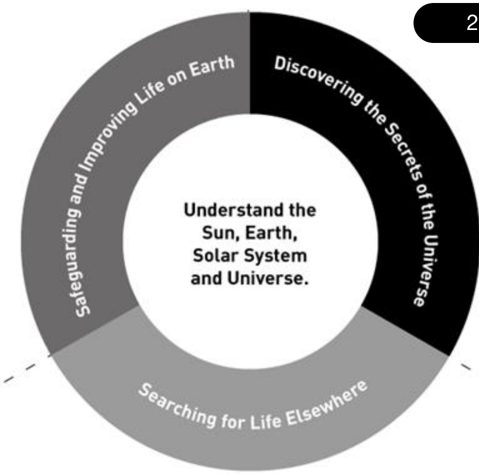
HELIOPHYSICS - SPACE WEATHER

2017 **C-Class Solar Flares**

2017 **Predict KP Index**

2018 **SDO UV Irradiance**

2019 **Super Res. Mag. Fields**



PLANETARY DEFENSE

2016 **Meteorite Hunting Drone**

2016 **Long-period Comet Detection**

2016 **Asteroid Deflection Selector**

2017 **NEO Shape Modeling**

ARTEMIS + BEYOND

2019 **Lunar Resource Mapping (Metals)**

2017 **Crater Identification**

ASTROBIOLOGY

2017 **Exoplanet Detection**

2017 **Atmospheric Retrieval from Spectra**

2017 **Biosignatures**

ARTEMIS + MARS

2018 **Rover Localization**

2018 **Autonomous Route Planning**

= with ESA

NASA FDL

PROGRAM IMPACT: DEMONSTRATED CAPACITY

EXPAND HUMAN KNOWLEDGE THROUGH NEW SCIENTIFIC DISCOVERIES
Understand the Sun, Earth, Solar System, and Universe. SMD / HEOMD

FDL RESEARCH RESULT	Earth Science	Planetary Science	Heliophysics
Discovery with Multi-dimensional data		2019 Lunar Resource Mapping (Metals) 2016 Long-period Comet Detection	2019 GNSS Disturbances 2017 **KP Index**
Workflow optimization	2019 **Edge Inference: Floods** 2019 **Cloud Classification**	2016 **Meteorite Hunting Drone** 2017 Exoplanet Detection 2017 **Crater Identification**	2019 **Virtual Telescopes** 2019 Super Res. Mag. Fields
Forecasting and Prediction	2019 Flood Mapping	2017 NEO Shape Modeling 2017 Atmospheric Retrieval from Spectra	2018 **SDO UV Irradiance** 2017 **KP Index** 2017 **C-Class Solar Flares** 2018 **Ionospheric Scintillations**
Decision Intelligence	2018 **Disaster Response**	2016 **Asteroid Deflection Selector**	
Anomaly Detection		2017 **Biosignatures**	
Autonomous Systems		2018 Rover Localization 2018 **Autonomous Route Planning**	2019 Auto Calibration

FDL IN NUMBERS

- **4** years, **6** research sprints (**4** NASA, **2** ESA)
- **4** Space Agency Partners / **12** Commercial
- **8** Big Thinks
- **843** applicants (436 in 2019)
- **11%** acceptance rate (‘18 / ‘19)
- **108** Researchers (Phd and Post-Doc)
- **138** mentors and guest experts
- **633** Partner reviewer community, **25+** Universities
- **$1.5m** compute (partner in-kind)
- **26** Research Projects, **15** Publications* / **30+** Scientific and AI conferences
- **50/50** US / International split (NASA FDL)

* **As of August 2019**

Papers

A Machine-learning Data Set Prepared from the NASA *Solar Dynamics Observatory* Mission (Galvez et al. 2019), Astrophysical Journal Supplement Series

An Ensemble of Bayesian Neural Networks for Exoplanetary Atmospheric Retrieval, Astronomical Journal

Scientific Domain Knowledge Improves Exoplanet Transit Classification with Deep Learning, Astrophysical Journal Letters

Rapid Classification of TESS Planet Candidates with Convolutional Neural Networks, Astronomy & Astrophysics

A survey of southern hemisphere meteor showers, Planetary and Space Science Journal 2018

Artificial Intelligence Techniques applied to Automating Meteor Validation and Trajectory Quality Control to Direct the Search for Long Period Comets, International Meteor Conference 2017

A Deep Learning Virtual Instrument for Monitoring Extreme UV Solar Spectral Irradiance (Szenicer, Fouhey et al.), Science Advances (accepted)

The NASA FDL Exoplanet Challenge: Transit Classification with Convolutional Neural Networks, AbSciCon 2019

INARA: Intelligent exoplaNet Atmospheric RetrievAl A Machine Learning Retrieval Framework with a Data Set of 3 Million Simulated Exoplanet Atmospheric Spectra, AbSciCon 2019

EXO-ATMOS: A Scalable Grid of Hypothetical Planetary Atmospheres, AbSciCon 2019

NASA Frontier Development Lab 2018 Using machine learning to study E.T. biospheres, CiML at NeurIPS 2018

Bayesian Deep Learning for Exoplanet Atmospheric Retrieval Bayesian Deep Learning Workshop, NeurIPS 2018

Absolute Localization Through Orbital Maps and Surface Perspective Imagery: A Synthetic Lunar Dataset and Neural Network Approach, 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) *Coming Soon*

FDL's Code of Practice:

https://docs.google.com/document/d/1sXApoOJEbfR32p8gzBq2Ks17osat-V2cbs_AI5L9Vms/edit#

TIMELINE

2018			2019												2020		
oct	nov	dec	jan	feb	mar	apr	may	jun	jul	aug	sep	oct	nov	dec	jan	feb	mar

Development Phase (2019)

- Stakeholder Engagement
- “Big Think” Sessions
- Challenge Definition
- Initial Recruiting

Set Up and Kick Off

- Researcher interviews
- On-boarding
- Data preparation

Sprint Phase

- Research and development
- Stakeholder checkpoints
- Tech Review Panels
- Presentations

2019 CYCLE ENDS

FDL 2020 DEVELOPMENT PHASE BEGINS

TECH MEMOS DELIVERED, CODE AND DATA PRODUCTS COMPLETE

Deployment

- Publications
- Conference Presentations
- Technical Reports

Communicating results

- NeurIPS

MODEL (1) A public / private partnership

WHAT FDL ENABLES FOR PARTNERS

1. MARKETING AND BRAND RECOGNITION
Leadership association with AI application and big data, "for the good of humankind" - PARTICULARLY "FOR GOOD"

2. PHYSICAL CASE STUDIES
Showcasing stretch use cases and engaging talking points (e.g. "we're using AI to detect solar flares")

3. POSITIONING IN SPACE COMMUNITY
Burgeoning 'new space race' / inspirational narratives as we look to Moon and Mars.

4. INTRODUCTIONS TO OTHER PARTNERS & NETWORKS

Sitting on FDL committees / steering group
(e.g. AI technical committee)

5. B2B RELATIONSHIPS BETWEEN PARTNERS
Bilateral relationships between partners have been brokered

6. WHITE PAPERS AND ARTICLES.
Specific use cases written up.

7. CONFERENCES
Demos, Keynotes, booths

8. TALENT ACQUISITION
A number of FDL researchers have been offered roles post FDL

The FDL program lead by SETI in collaboration with NASA has become a powerful catalyst for innovation in the areas of Space Technology, Space Weather and Astronaut Health. With this partnership, researchers, developers, and data scientists have the opportunity to access IBM's most advanced Cognitive System (Power AI) to revolutionize AI innovation and solve the challenges of tomorrow. Also via the partnership with the FDL program lead by SETI, IBM gains valuable insight into next generation AI requirements so that we can advance our AI services and offerings.

Mac Devine (IBM Fellow) & Naeem Altaf (IBM Distinguished Engineer, CTO Space Tech)

SETI INSTITUTE

NASA FDL

FRONTIER DEVELOPMENT LAB

AI RESEARCH FOR SPACE EXPLORATION AND ALL HUMANKIND

TECHNOLOGY AND RESEARCH PARTNERS

Google Cloud

intel AI

CHALLENGE PARTNERS

LSA LUXEMBOURG SPACE AGENCY

Hewlett Packard Enterprise

IBM

KX

LOCKHEED MARTIN

CSA ASC

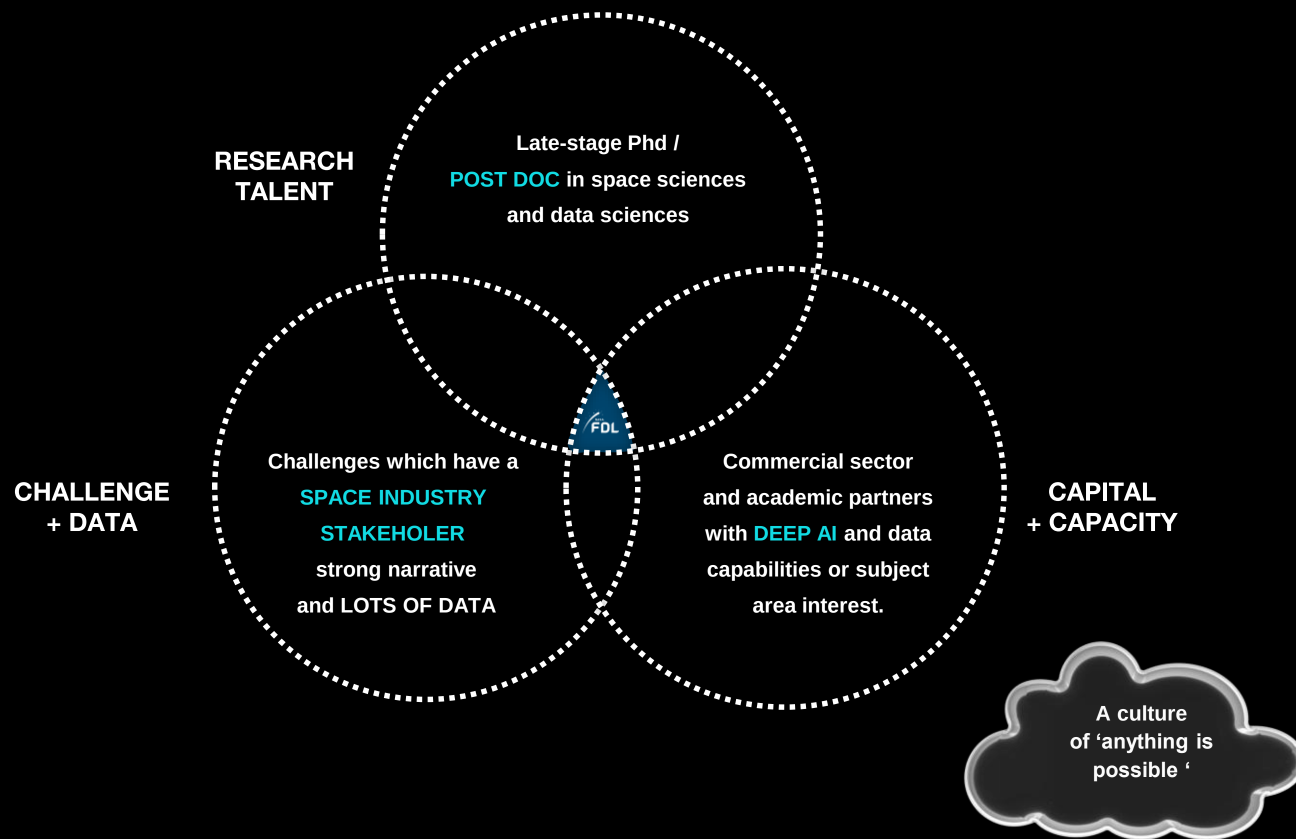
RESEARCH PARTNERS

NVIDIA

ELEMENT AI

USGS science for a changing world

NASA FRONTIER DEVELOPMENT LAB - FORMULA



https://vimeo.com/362856162

https://vimeo.com/362856096

https://vimeo.com/362856096