

# End-to-End Machine Learning Applications Framework for Earth Science

21<sup>st</sup> Pecora Conference  
11 October, 2019

---

Brian Freitag<sup>1</sup>, A. Acharya<sup>1</sup>, M. Ramasubramanian<sup>1</sup>, D.  
Bollinger<sup>3</sup>, A. Kaulfus<sup>1</sup>, I. Gurung<sup>1</sup> M. Maskey<sup>2</sup>, R.  
Ramachandran<sup>2</sup>

1 – Univ. Alabama in Huntsville, 2 – NASA/MSFC, 3 – DevelopmentSeed



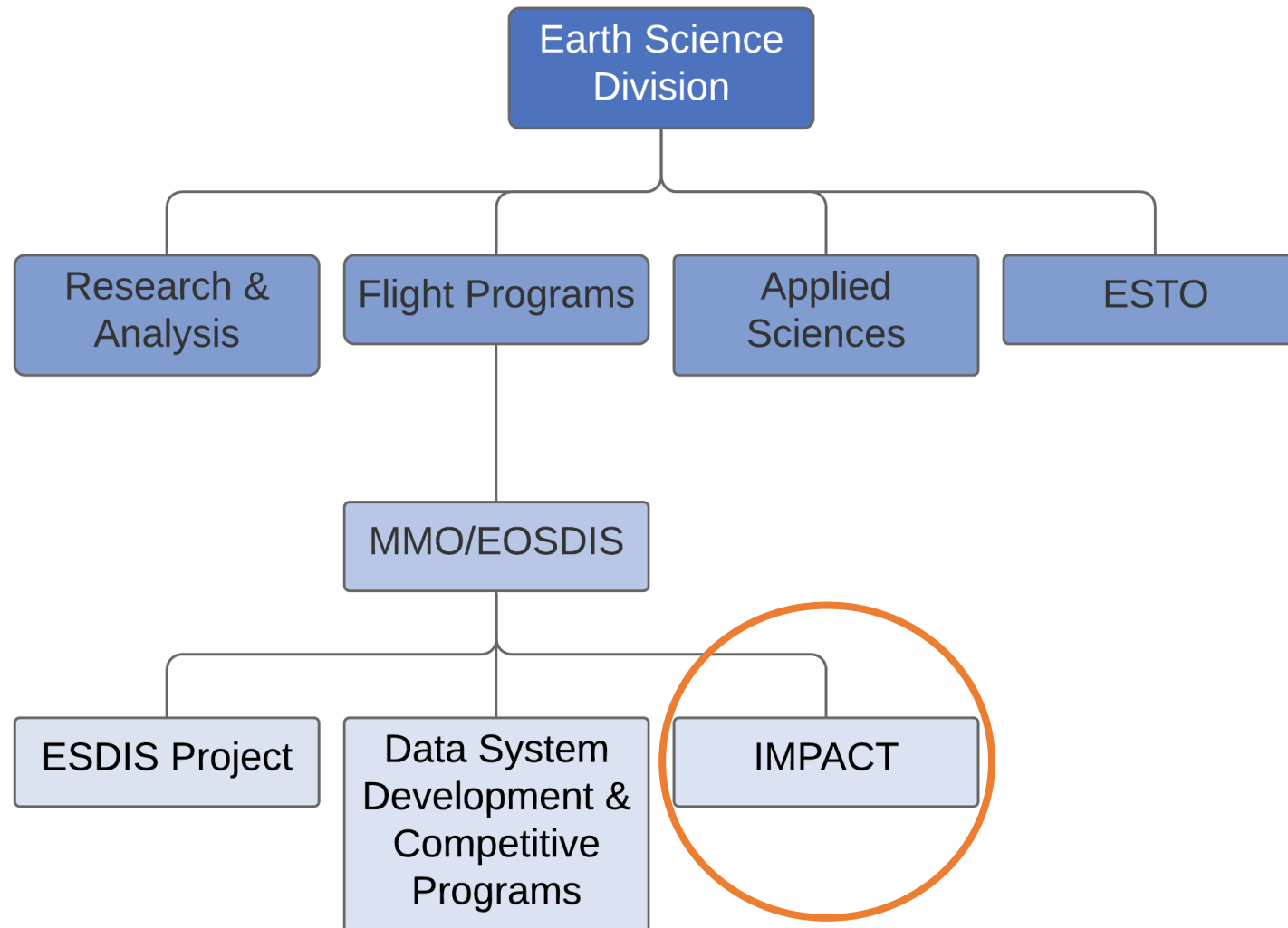
developmentSEED



# Outline

- 1) NASA IMPACT project overview
- 2) Expected challenges with NASA Earth science data
- 3) New tool for rapid development of training datasets for ML applications
- 4) User interface for visualizing model output and generating event databases for atmospheric phenomena

# NASA's Interagency Implementation and Advanced Concepts Team (IMPACT) Project



# NASA IMPACT Project

Empowering Science and Applications by integrating Data, Technology, and People together



Interagency Collaboration



Assessment &  
Evaluation

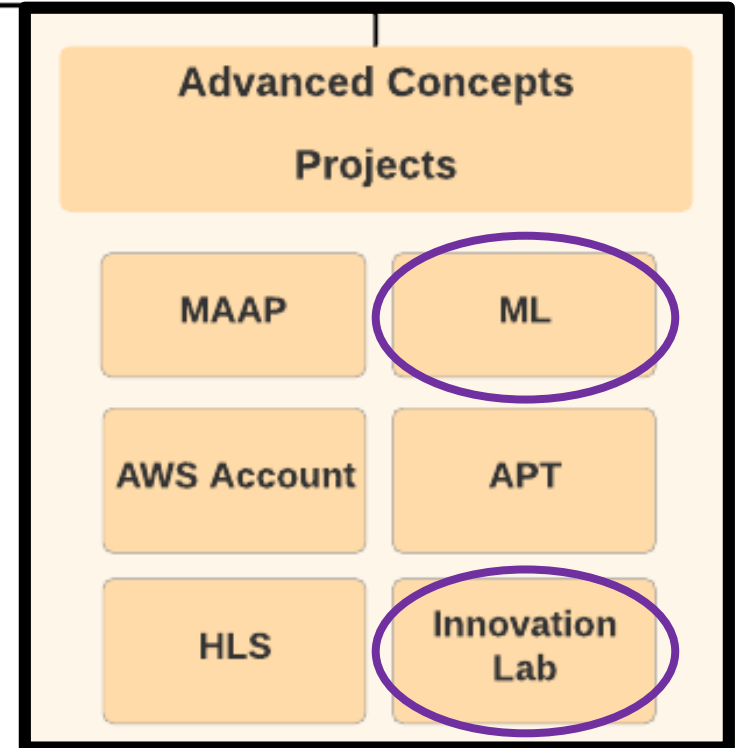
Assessment & Evaluation



Advanced  
Concepts

Advanced Concepts

# IMPACT



# NASA EOSDIS Data Systems



EOSDIS currently has over **27 Petabytes** of accessible Earth science data

Easy access and discovery of data to over **12,500 unique data products**

... of which 95% of granule searches complete in less than **1 Second**



EOSDIS delivered over **1.6 Billion** data products to over **4.1 Million** users from around the world

**33,000** Data Collections in the Common Metadata Repository (CMR)



EOSDIS also delivers near-real-time products in under **3 hours** from observation ...

Over **330,000 users** have registered with EOSDIS to date



And Over **380 Million** data granules



American Customer Satisfaction Index (ACSI) survey scoring **79** from over **4,000** respondents

# NASA Earth Science Missions: Present through 2023

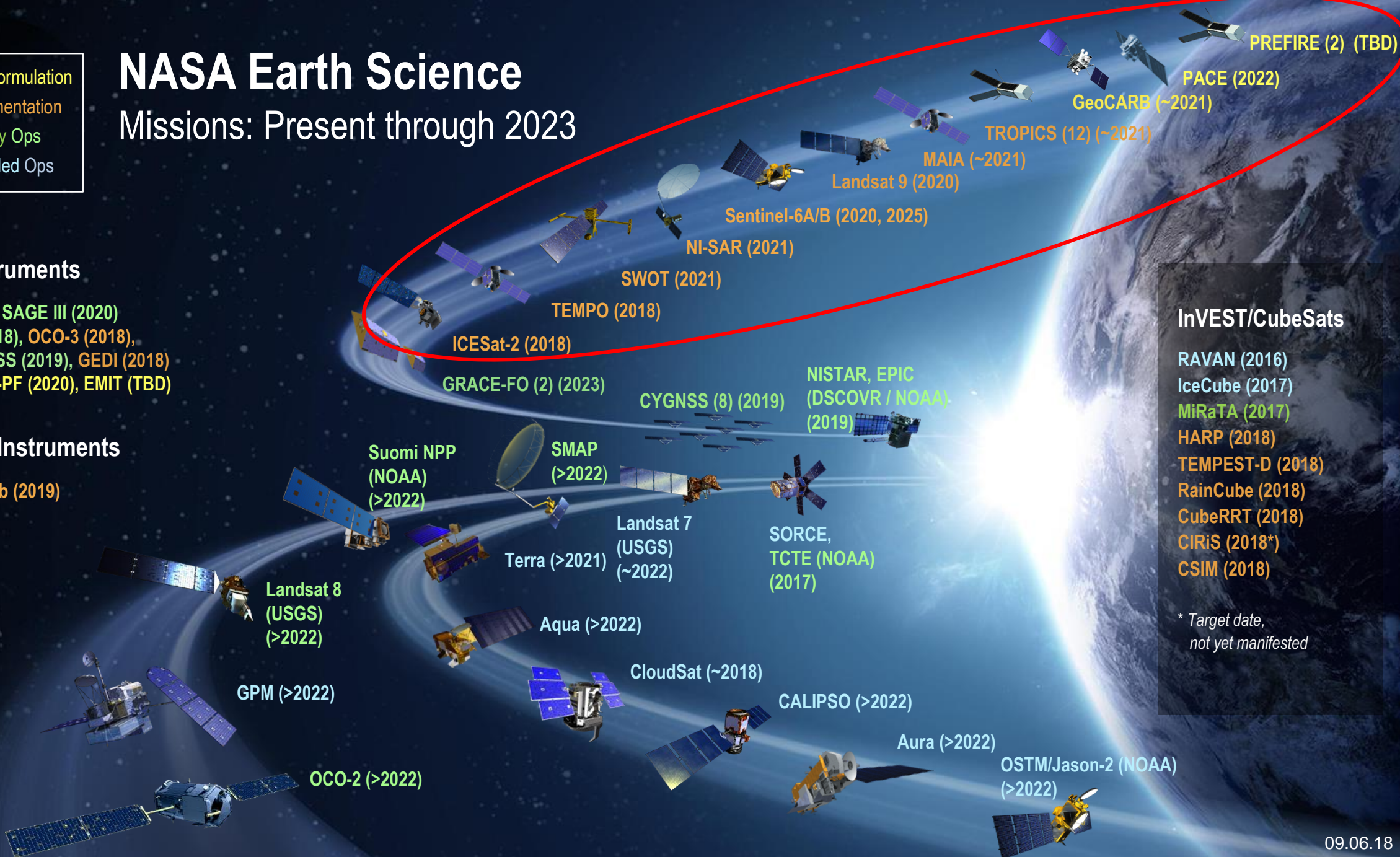
- (Pre)Formulation
- Implementation
- Primary Ops
- Extended Ops

## ISS Instruments

LIS (2020), SAGE III (2020)  
 TSIS-1 (2018), OCO-3 (2018),  
 ECOSTRESS (2019), GEDI (2018)  
 CLARREO-PF (2020), EMIT (TBD)

## JPSS-2 Instruments

OMPS-Limb (2019)



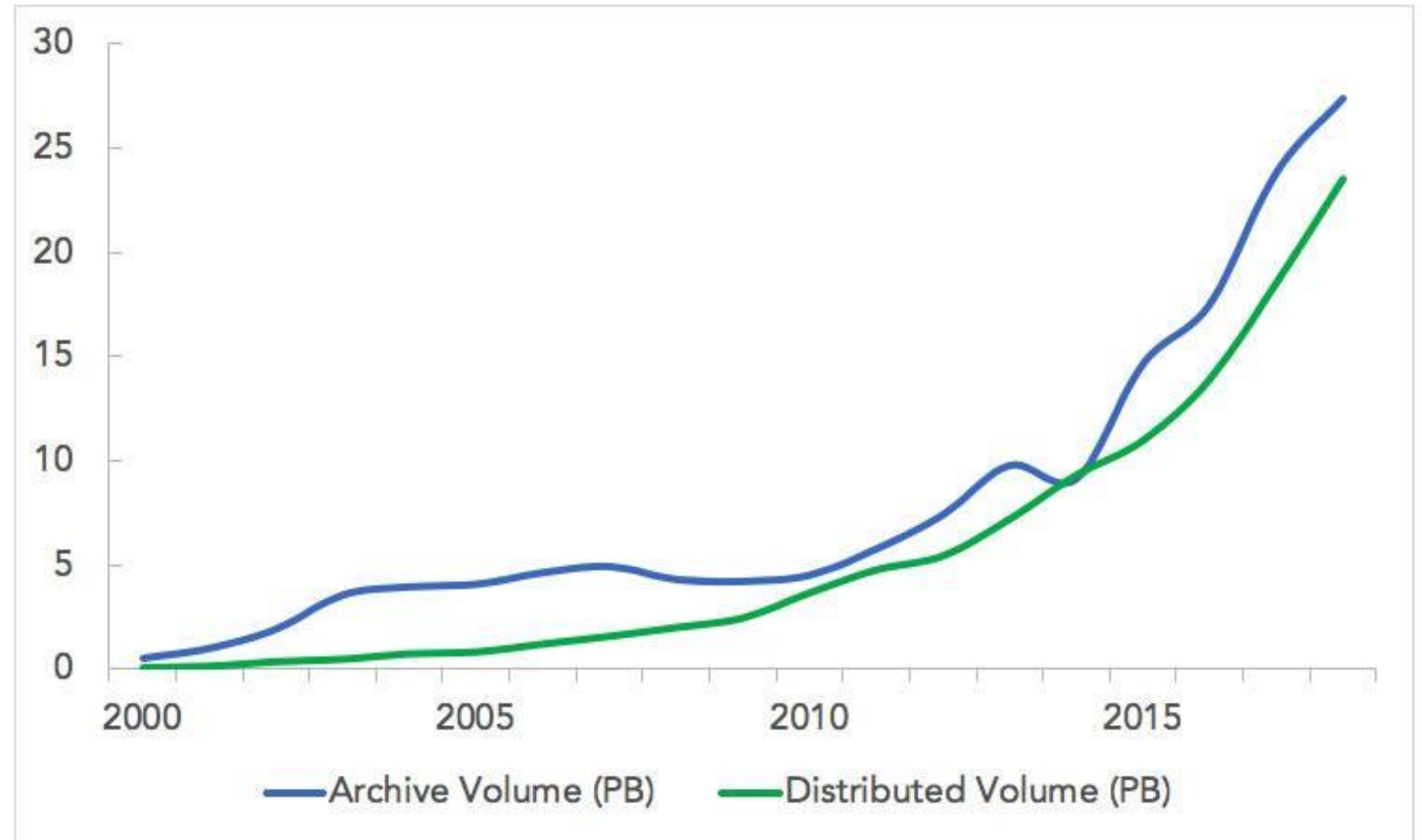
### InVEST/CubeSats

- RAVAN (2016)
- IceCube (2017)
- MiRaTA (2017)
- HARP (2018)
- TEMPEST-D (2018)
- RainCube (2018)
- CubeRRT (2018)
- CIRiS (2018\*)
- CSIM (2018)

*\* Target date, not yet manifested*

# NASA EOSDIS Data Systems

The EOSDIS data archive is growing at an extraordinary pace and the volume of data in the archive (end FY18) was more than 27.5 petabytes (PB). The current architecture will not be cost effective as the annual ingest rate increases from 4 to 50PB/year





# NASA EOSDIS Data Systems

- How is the NASA IMPACT project using machine learning to address future data concerns that fits with NASA ESD objectives?

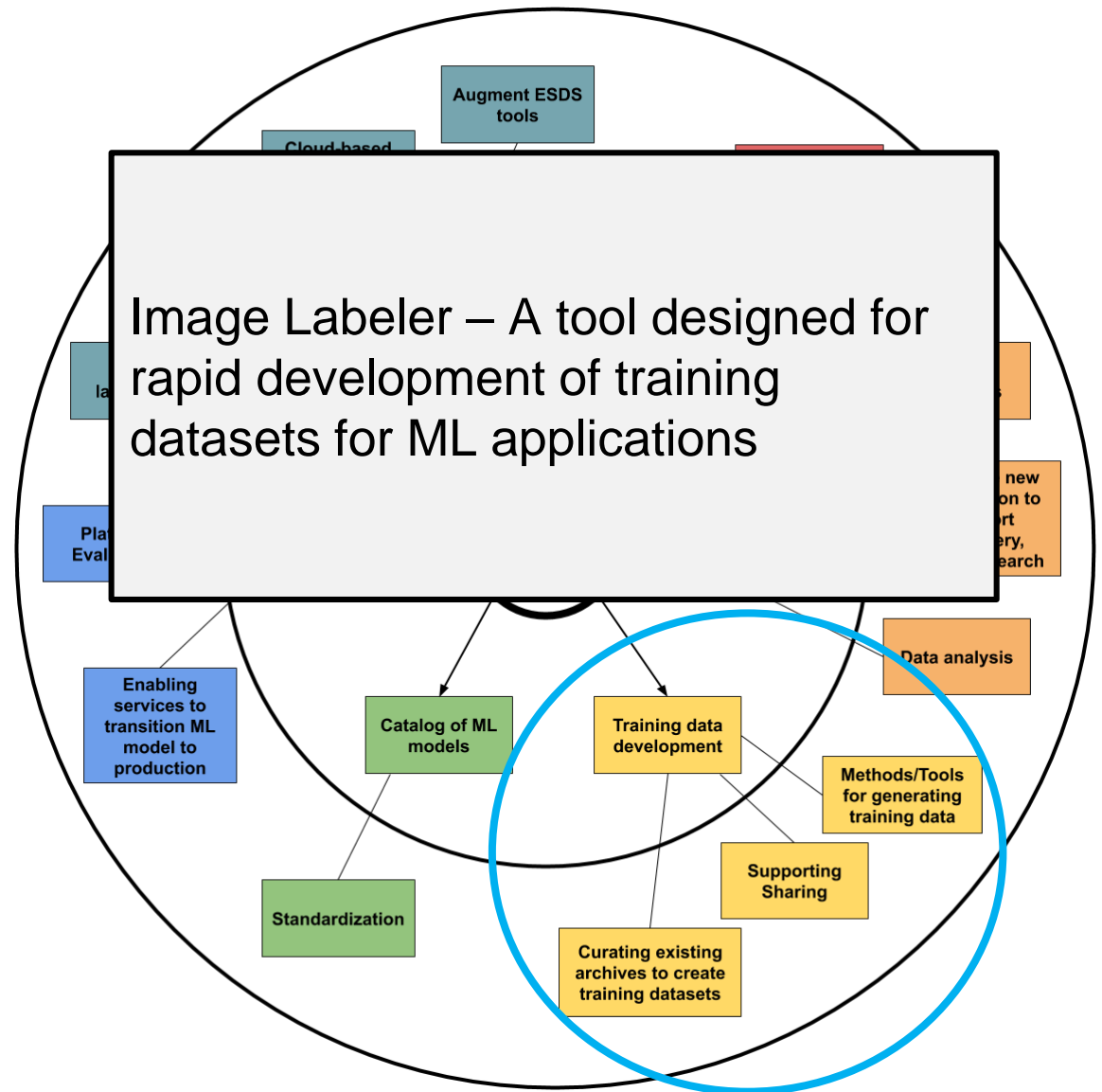
# ML Roadmap

## Need

- Provide opportunities to innovate new ways of analysis and obtain valuable insights from EO data
- Utilize ML techniques to improve search, access, and usability as well as overall operations in EOSDIS

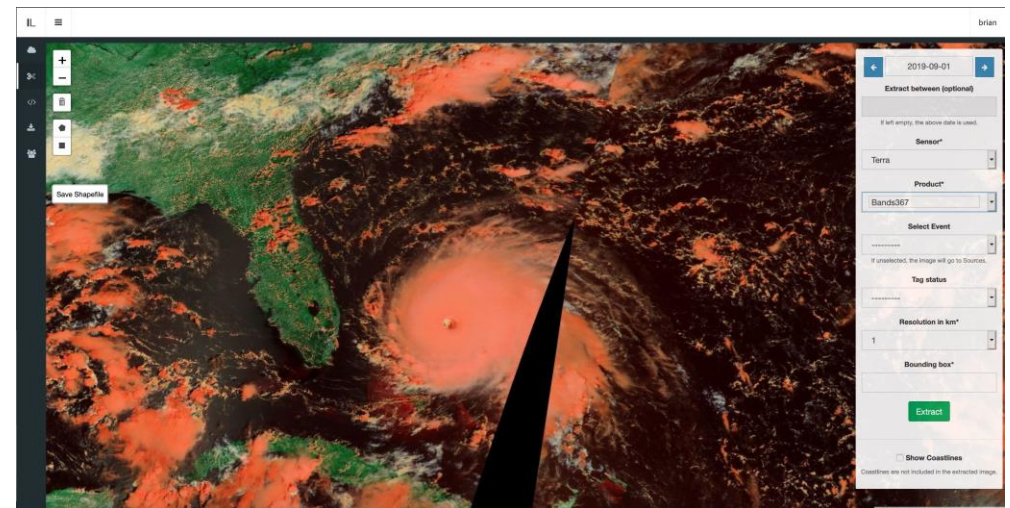
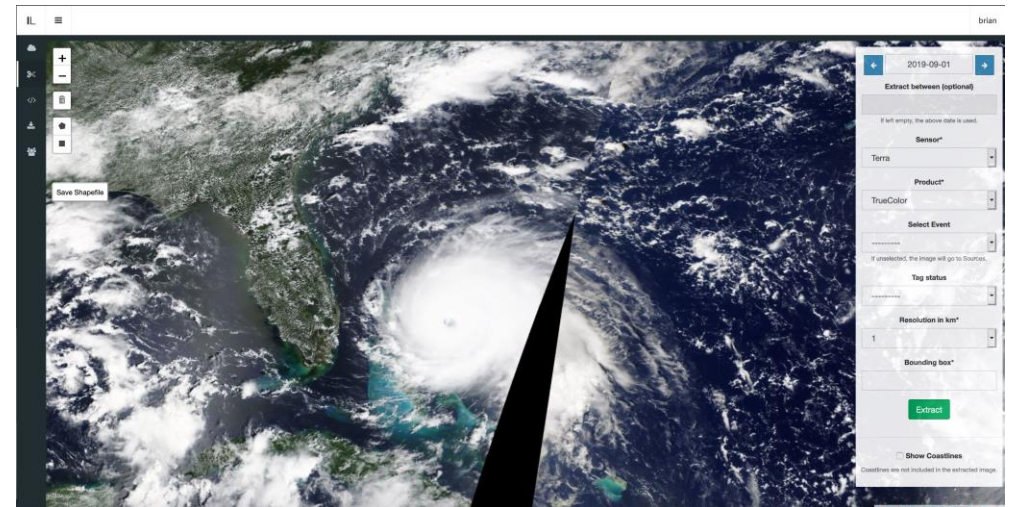
## Objectives

- Identify areas of strategic development
- Track new developments in the AI and ML community to inform the ESDS Program
- Develop and promote the concept of “OpenML”
- Evaluate new tools and platforms

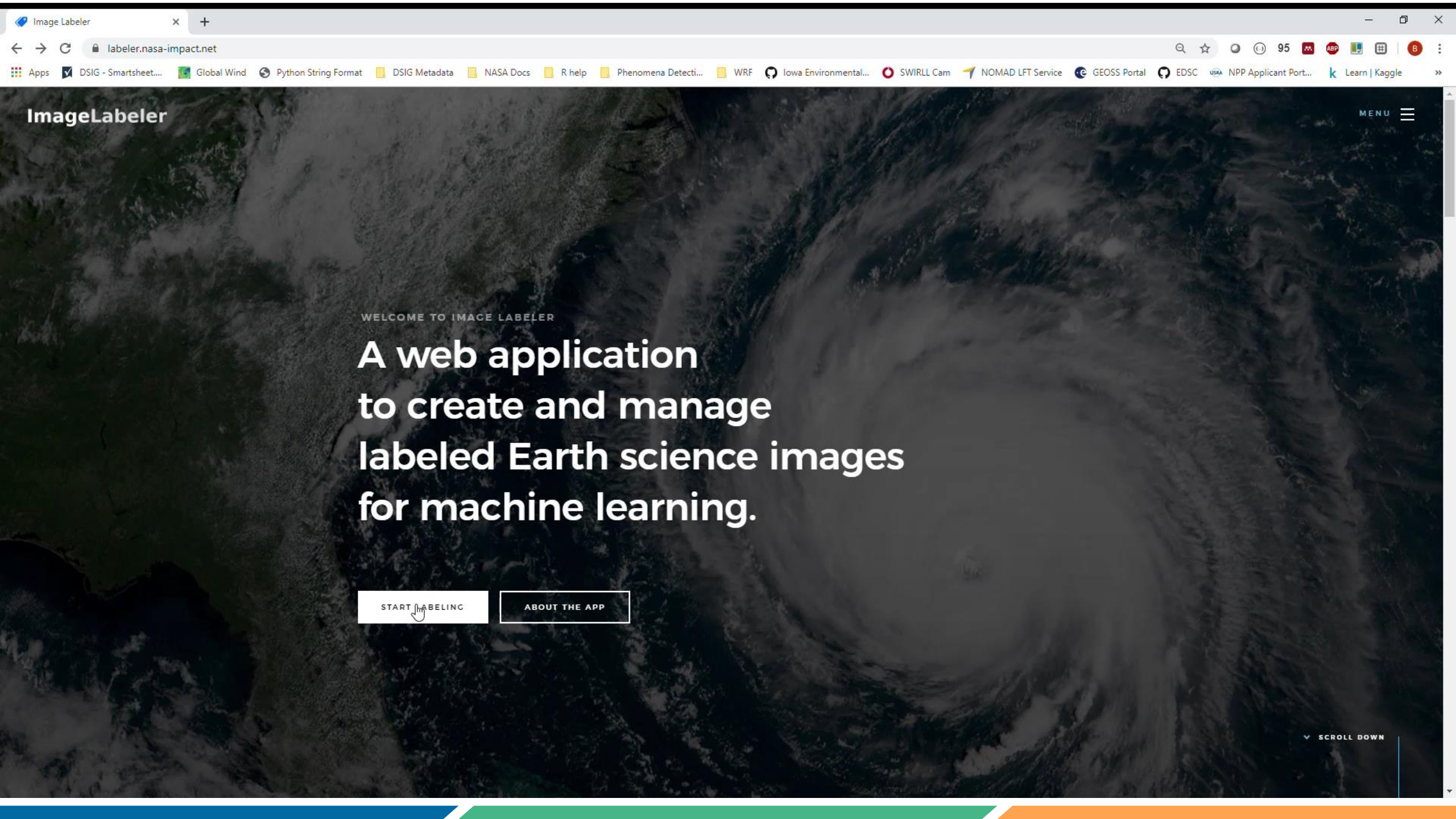


# Image Labeler

- Developed by a GRA to facilitate rapid classification of training images for ML models
- Powered by NASA web services (GIBS WMS layers)
  - MODIS true color and false color (7-2-1, 3-6-7) composites
- Supports ingest of both images and geotiffs
- Allows users to draw shapefiles on images and WMS layers for image segmentation



<https://labeler.nasa-impact.net>



WELCOME TO IMAGE LABELER

# A web application to create and manage labeled Earth science images for machine learning.

START LABELING

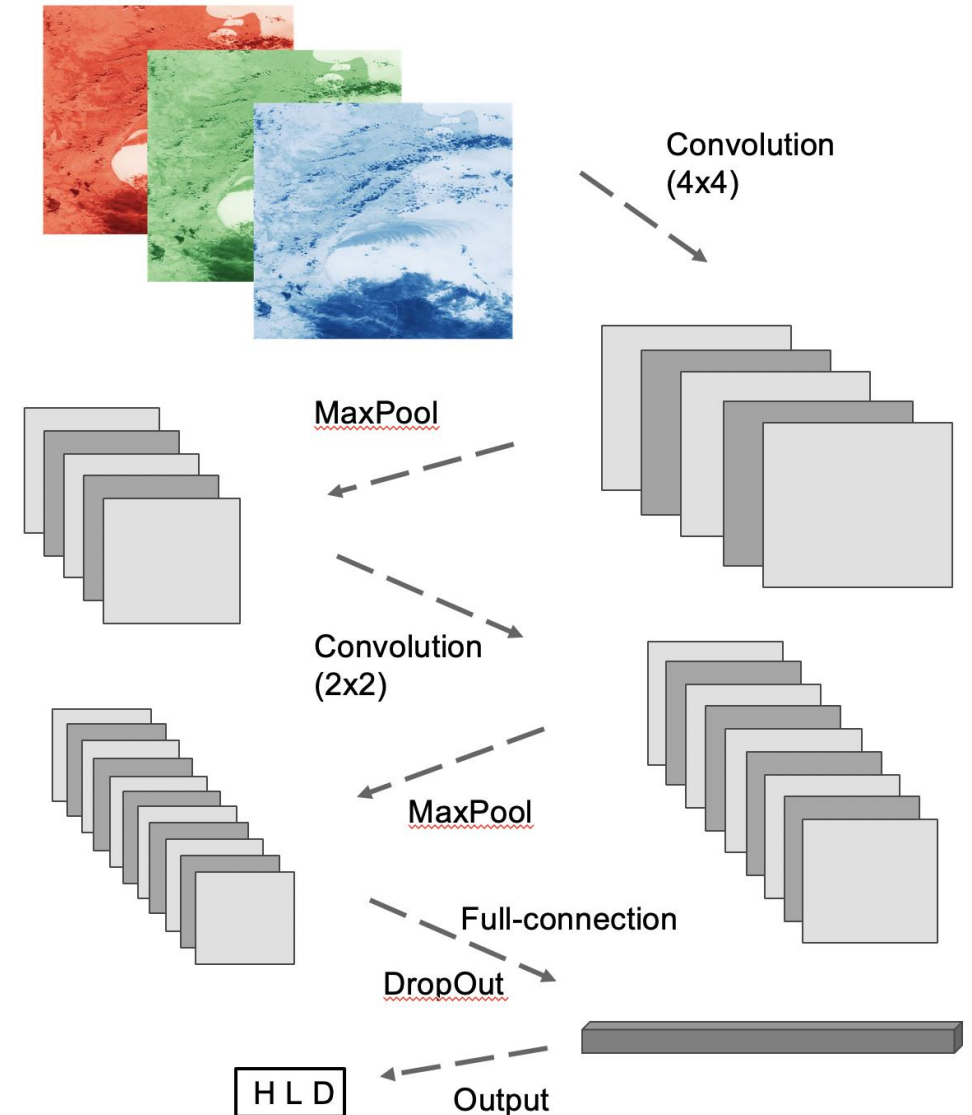
ABOUT THE APP

SCROLL DOWN

# High Latitude Dust ML Model

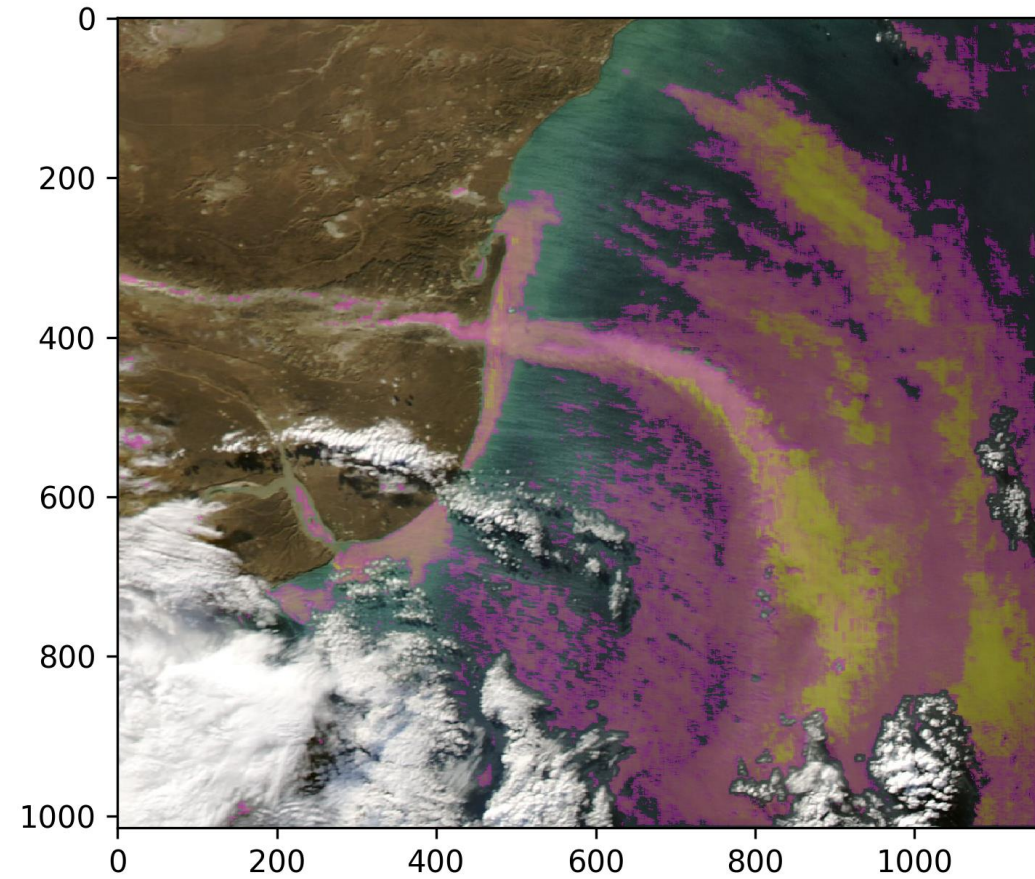
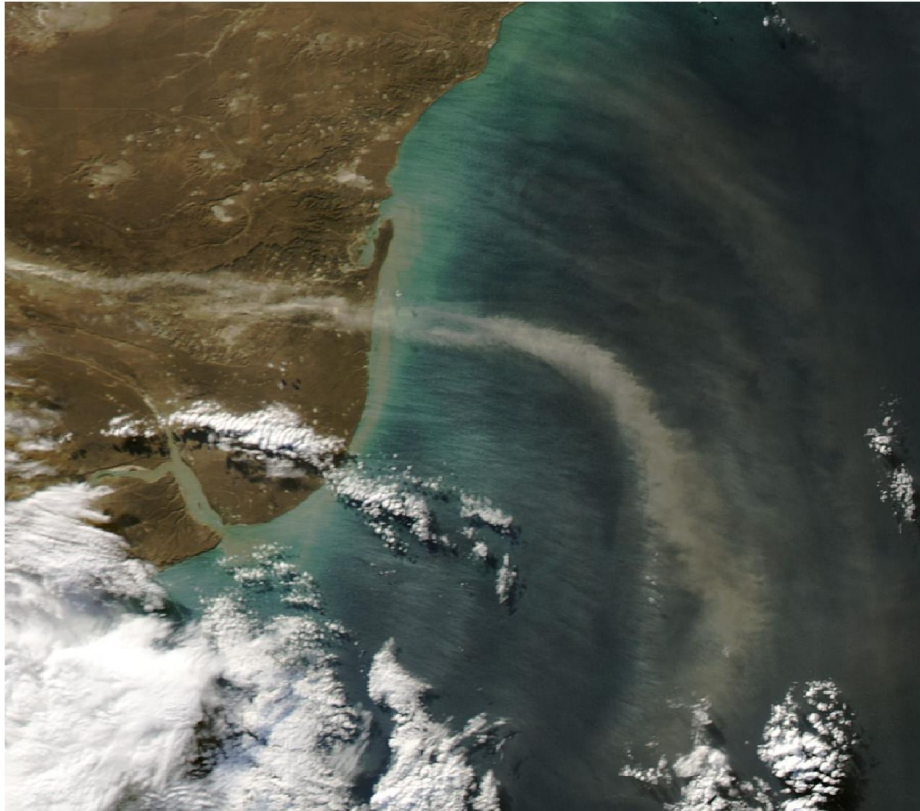
- Image labeler used for tagging and segmentation of high latitude dust (HLD) from MODIS true color imagery
- Training dataset used to detect pixel level model for automated detection of HLD
  - Pixels are split 80/20 for training and validation

Accuracy	0.7
Precision	0.91
F1_score	0.57
Recall	0.41



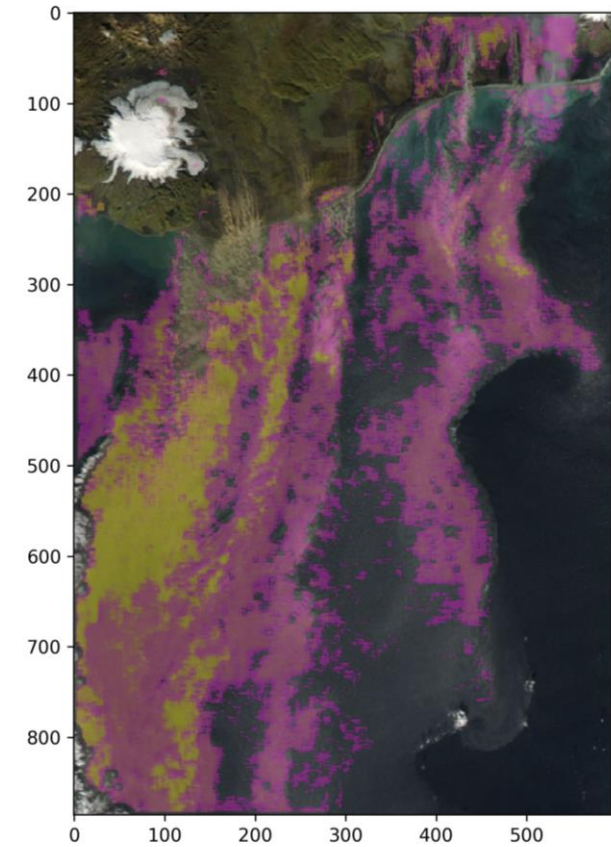
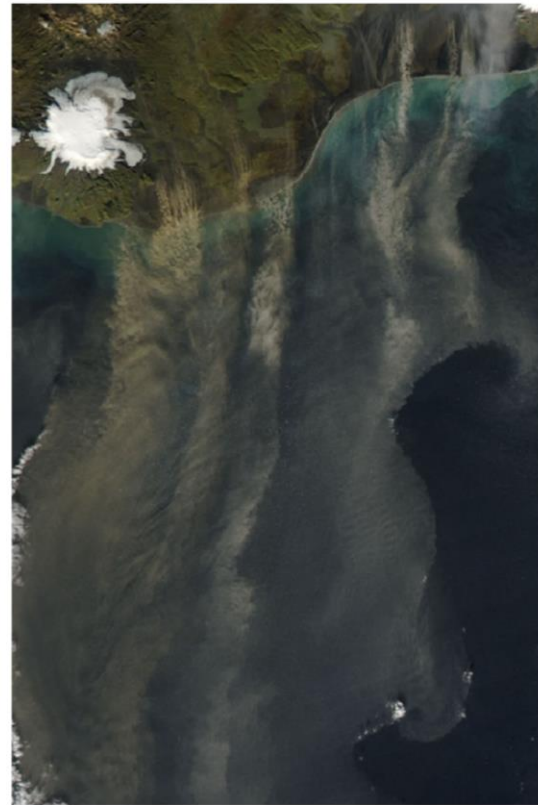
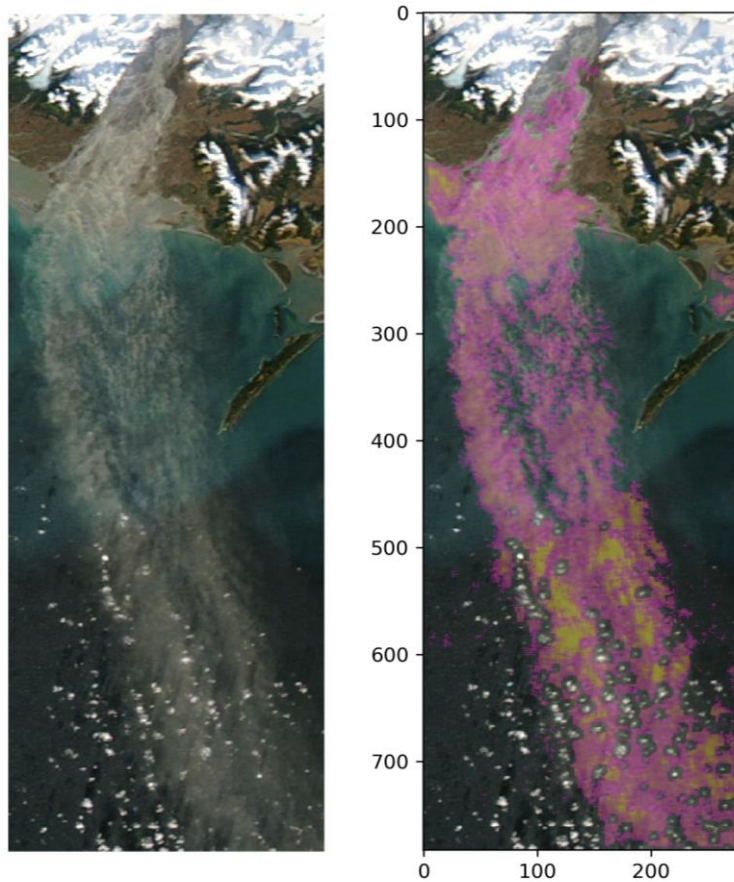
# High Latitude Dust ML Model

## Patagonia



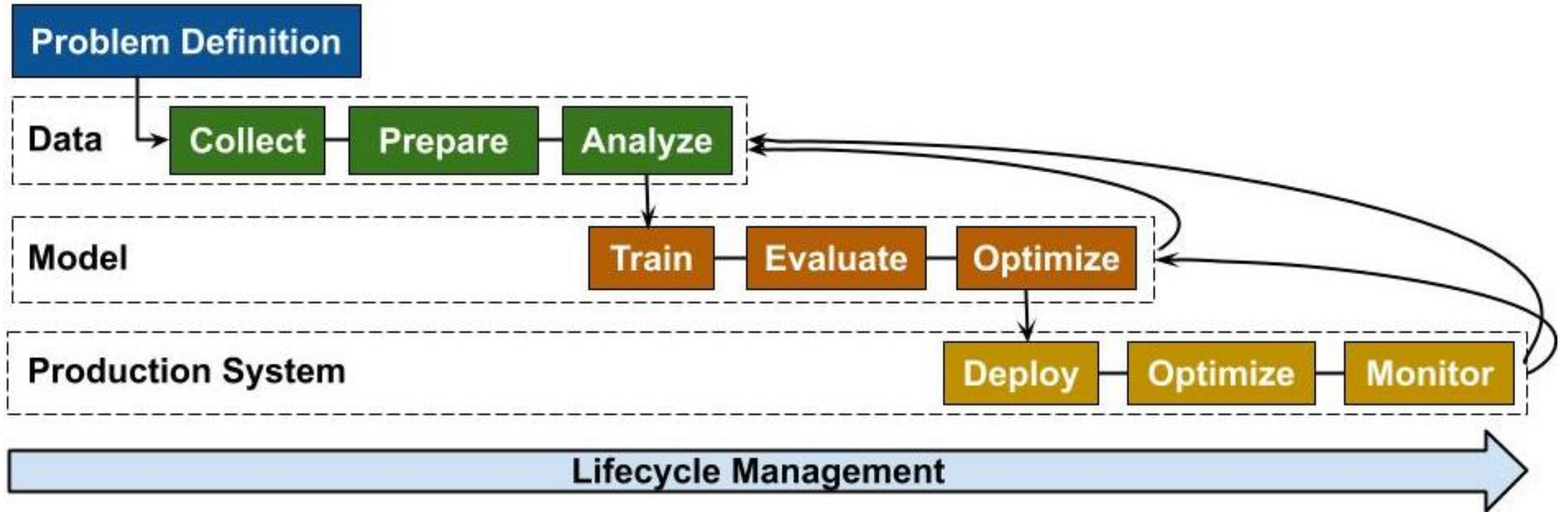
# High Latitude Dust ML Model

## Copper River Valley, AK



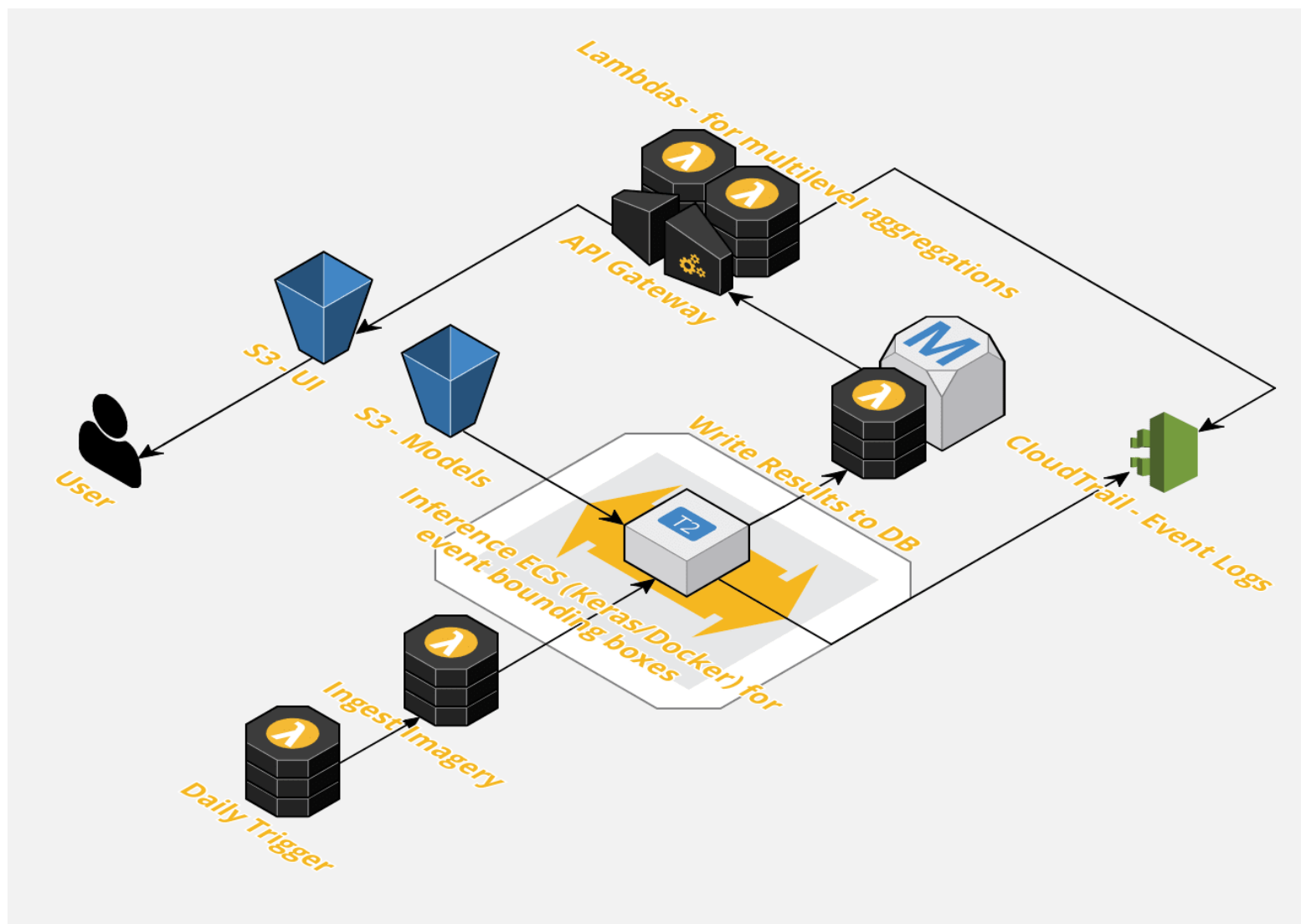
## South Iceland coastline

# Research to Operation



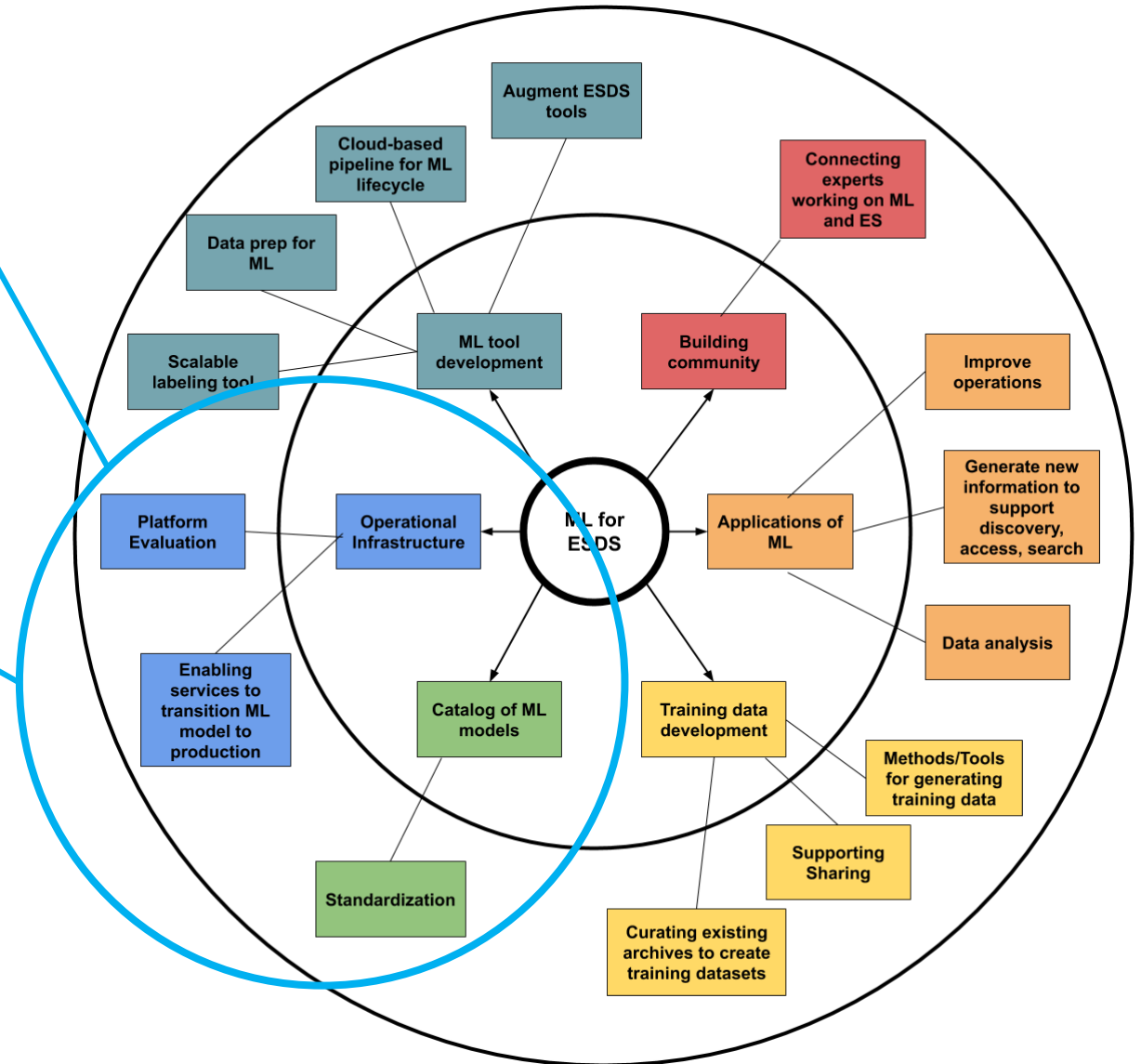


# Research to Operation



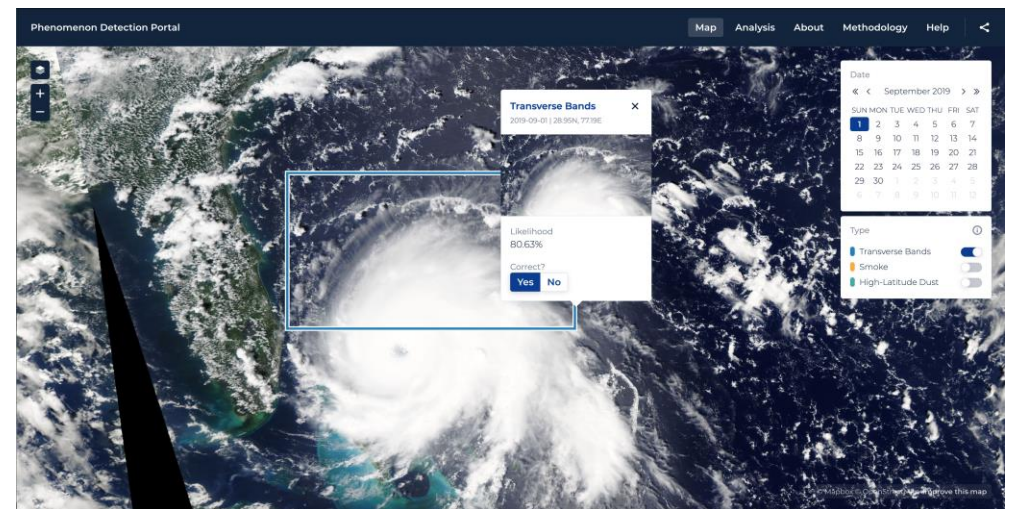
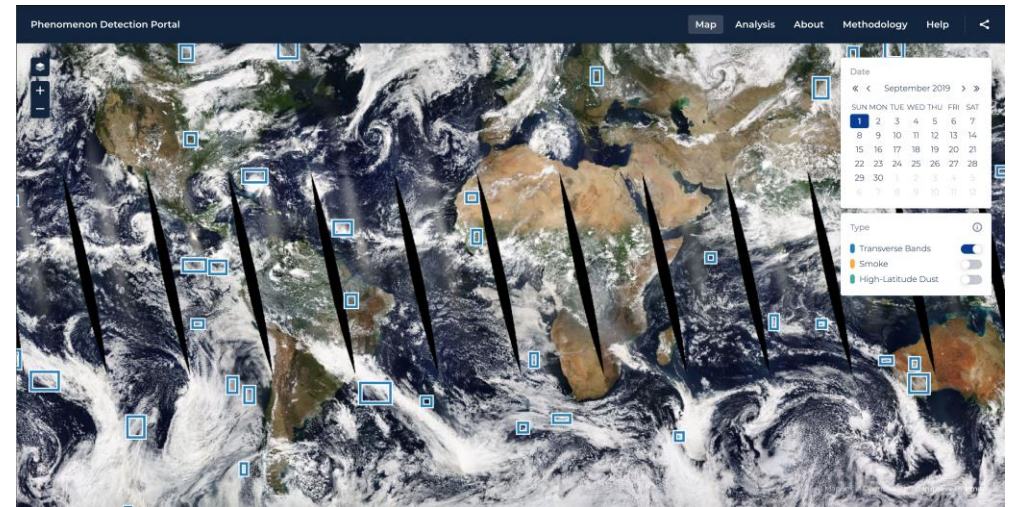
# ML Roadmap

Phenomena Portal – A user interface designed to visualize and analyze a dynamic event catalog of atmospheric phenomena

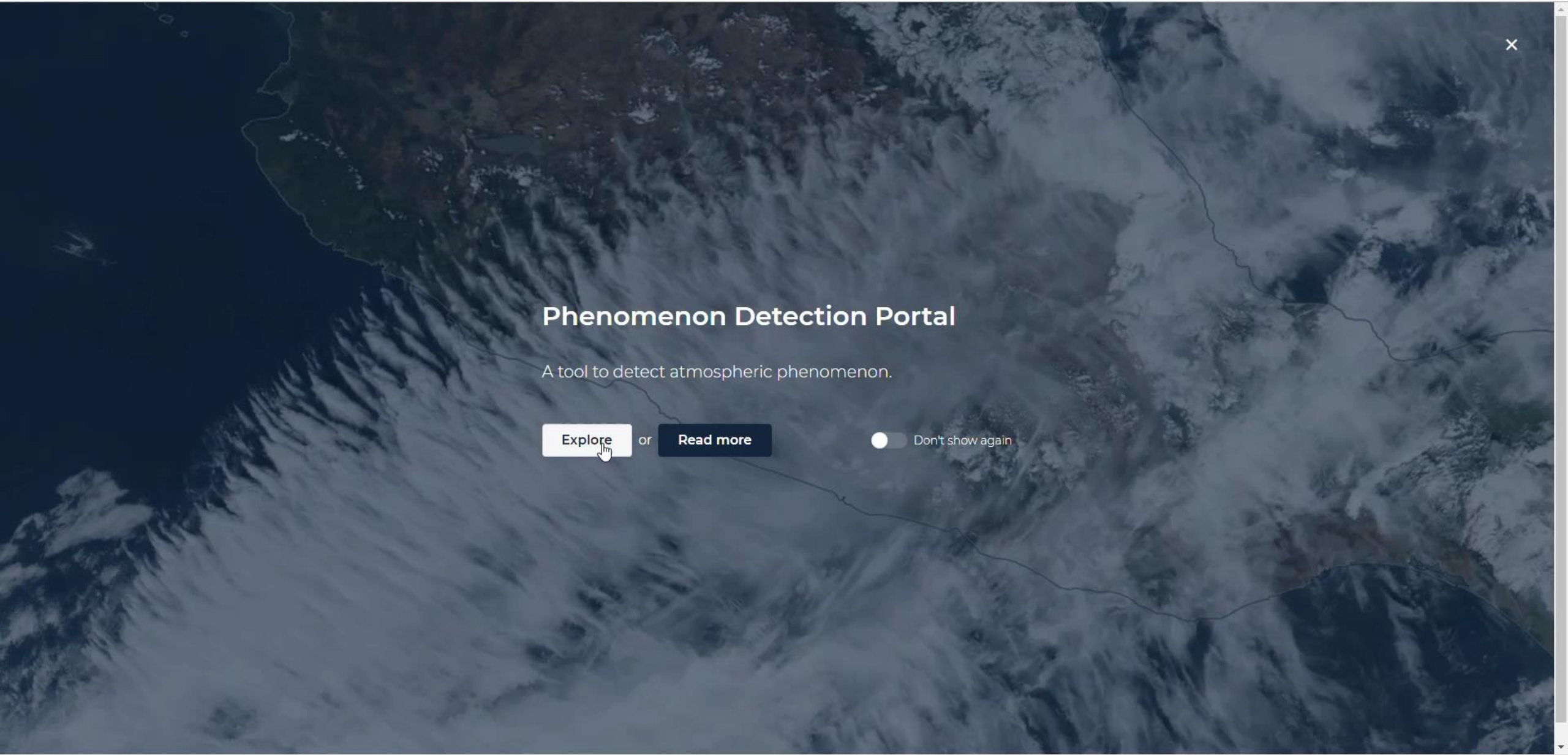


# Phenomena Portal

- Incorporate ML models with manually labeled data
- ML models run daily to detect atmospheric phenomena
  - Transverse Cirrus Bands
  - Smoke
  - High Latitude Dust
- Detections stored in event database
  - Visualize model output
  - Analyze events in space and time
- API for accessing/parsing event database



<https://phenomena.surge.sh>



# Phenomenon Detection Portal

A tool to detect atmospheric phenomenon.

Explore

or

Read more



Don't show again

# Summary

- NASA IMPACT project uses ML to support ESDS challenges
  - Exponential growth in data volume with launch of NISAR and SWOT missions
- Designed image labeler tool for rapid development of training datasets
  - Supports image classification and segmentation
  - Allows for team collaboration
- Move trained ML models to production using phenomena portal
  - Currently supports 3 ML models running daily on MODIS true color imagery
  - Includes analysis tab for quick assessment of spatiotemporal patterns in detected events

# Future Improvements

- Image Labeler
  - Fully support geotiff ingest and labeling
  - Support additional data layers from NASA web services
  - General front-end and back-end enhancements
- Phenomena Portal
  - Streamline model pipeline from training to production
  - Incorporate additional phenomena
  - Add functionality to support more detailed analysis

# Questions

Email: [brian.freitag@nsstc.uah.edu](mailto:brian.freitag@nsstc.uah.edu)

IMPACT: <https://earthdata.nasa.gov/esds/impact>

Image Labeler: <https://labeler.nasa-impact.net>

Phenomena Portal: <https://phenomena.surge.sh>