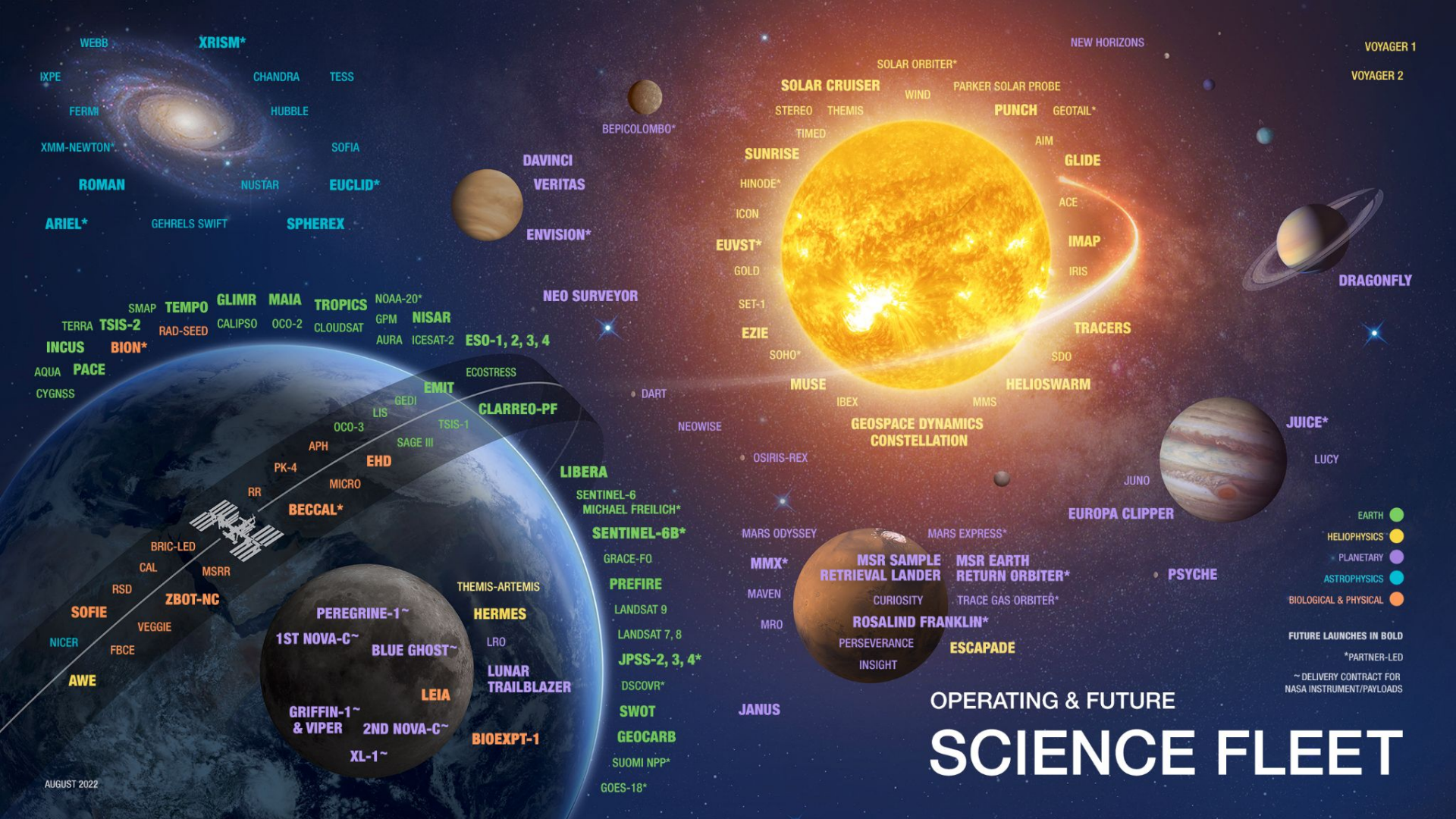
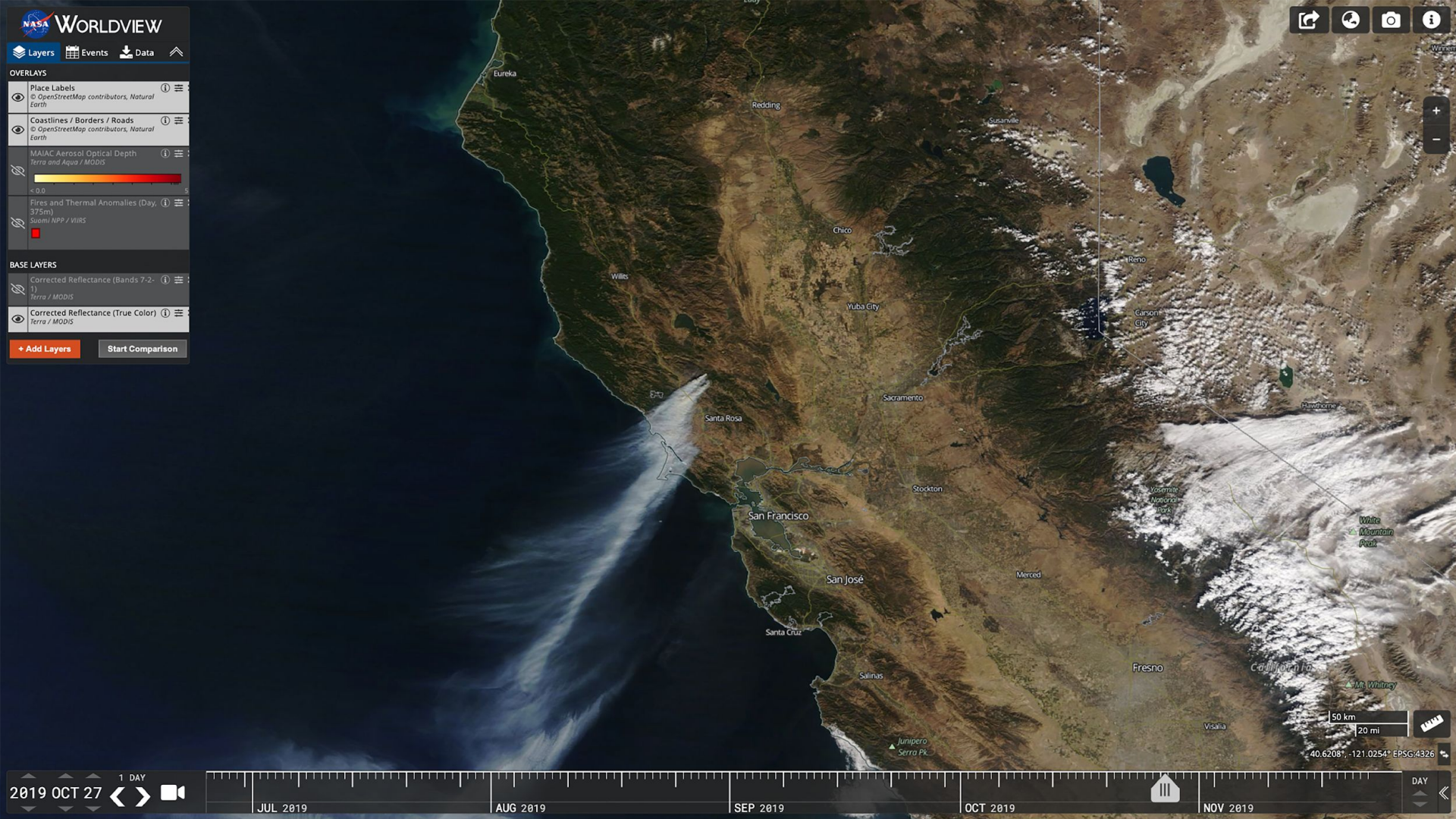
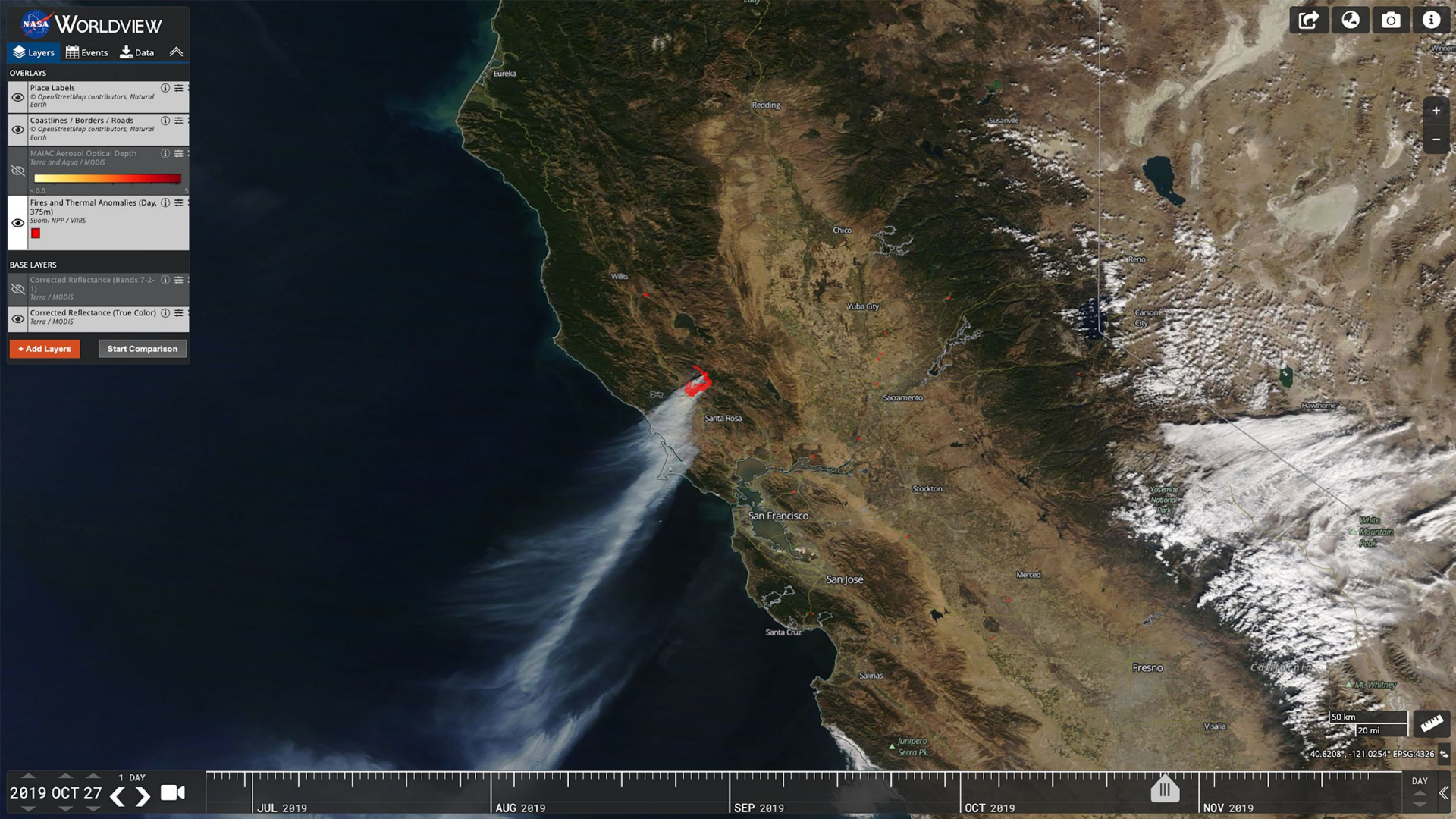
AI & Science: NASA Perspective

Kevin Murphy
NASA SMD/HQ

February 22, 2024


WEBB
XRISM*
IXPE
CHANDRA
TESS
FERMI
HUBBLE
XMM-NEWTON*
SOFIA
ROMAN
NUSTAR
EUCLID*
ARIEL*
GEHRELS SWIFT
SPHEREX
BEPICOLOMBO*
DAVINCI
VERITAS
ENVISION*
NEO SURVEYOR
SMAP
TEMPO
GLIMR
MAIA
TROPICS
NOAA-20*
TERRA
TSIS-2
RAD-SEED
CALIPSO
OCO-2
CLOUDSAT
GPM
NISAR
INCUS
BION*
AURA
ICESAT-2
ESO-1, 2, 3, 4
AQUA
PACE
ECOSTRESS
CYGNSS
EMIT
GEDI
LIS
CLARREO-PF
OCO-3
TSIS-1
APH
SAGE III
PK-4
EHD
RR
MICRO
BECCAL*
BRIC-LED
CAL
MSRR
RSD
ZBOT-NC
SOFIE
VEGGIE
NICER
FBCE
AWE
PEREGRINE-1~
1ST NOVA-C~
BLUE GHOST~
GRIFFIN-1~
& VIPER
2ND NOVA-C~
XL-1~
THEMIS-ARTEMIS
HERMES
LRO
LUNAR
TRAILBLAZER
LEIA
BIOEXPT-1
AUGUST 2022
LIBERA
SENTINEL-6
MICHAEL FREILICH*
SENTINEL-6B*
GRACE-FO
PREFIRE
LANDSAT 9
LANDSAT 7, 8
JPSS-2, 3, 4*
DSCOVR*
SWOT
GEOCARB
SUOMI NPP*
GOES-18*
NEOWISE
DART
OSIRIS-REX
JANUS
SOLAR ORBITER*
SOLAR CRUISER
WIND
PARKER SOLAR PROBE
STEREO
THEMIS
PUNCH
GEOTAIL*
TIMED
AIM
SUNRISE
GLIDE
HINODE*
ACE
ICON
EUVST*
IMAP
GOLD
IRIS
SET-1
EZIE
TRACERS
SOHO*
SDO
MUSE
HELIOSWARM
IBEX
MMS
GEOSPACE DYNAMICS
CONSTELLATION
NEW HORIZONS
VOYAGER 1
VOYAGER 2
DRAGONFLY
JUICE*
LUCY
JUNO
EUROPA CLIPPER
MARS ODYSSEY
MARS EXPRESS*
MMX*
MSR SAMPLE
RETRIEVAL LANDER
MSR EARTH
RETURN ORBITER*
MAVEN
CURIOSITY
TRACE GAS ORBITER*
MRO
ROSALIND FRANKLIN*
PERSEVERANCE
INSIGHT
ESCAPADE
PSYCHE
EARTH
HELIOPHYSICS
PLANETARY
ASTROPHYSICS
BIOLOGICAL & PHYSICAL
FUTURE LAUNCHES IN BOLD
*PARTNER-LED
~ DELIVERY CONTRACT FOR
NASA INSTRUMENT/PAYLOADS
OPERATING & FUTURE
SCIENCE FLEET


WORLDVIEW
Layers
Events
Data
OVERLAYS
Place Labels
© OpenStreetMap contributors, Natural Earth
Coastlines / Borders / Roads
© OpenStreetMap contributors, Natural Earth
MAIAC Aerosol Optical Depth
Terra and Aqua / MODIS
< 0.0
5
Fires and Thermal Anomalies (Day, 375m)
Suomi NPP / VIIRS
BASE LAYERS
Corrected Reflectance (Bands 7-2-1)
Terra / MODIS
Corrected Reflectance (True Color)
Terra / MODIS
+ Add Layers
Start Comparison
Eureka
Redding
Susanville
Chico
Willits
Reno
Yuba City
Carson City
Sacramento
Santa Rosa
Hawthorne
Stockton
Yosemite National Park
San Francisco
White Mountain Peak
San José
Merced
Santa Cruz
Salinas
Fresno
Mt. Whitney
Visalia
Junipero Serra Pk.
50 km
20 mi
40.6208°, -121.0254° EPSG:4326
1 DAY
2019 OCT 27
JUL 2019
AUG 2019
SEP 2019
OCT 2019
NOV 2019
DAY


WORLDVIEW
Layers
Events
Data
OVERLAYS
Place Labels
© OpenStreetMap contributors, Natural Earth
Coastlines / Borders / Roads
© OpenStreetMap contributors, Natural Earth
MAIAC Aerosol Optical Depth
Terra and Aqua / MODIS
< 0.0
Fires and Thermal Anomalies (Day, 375m)
Suomi NPP / VIIRS
BASE LAYERS
Corrected Reflectance (Bands 7-2-1)
Terra / MODIS
Corrected Reflectance (True Color)
Terra / MODIS
+ Add Layers
Start Comparison
Eureka
Redding
Susanville
Chico
Willits
Yuba City
Reno
Carson City
Sacramento
Santa Rosa
Hawthorne
Stockton
San Francisco
Yosemite National Park
White Mountain Peak
San José
Merced
Santa Cruz
Salinas
Fresno
Mt. Whitney
Visalia
Junipero Serra Pk.
50 km
20 mi
40.6208°, -121.0254° EPSG:4326
1 DAY
2019 OCT 27
JUL 2019
AUG 2019
SEP 2019
OCT 2019
NOV 2019
DAY

EARTHDATA Search

fire

Point: 38.62916292901474,-122.71896529197693

Start: 2019-10-23 00:00:00 Stop: 2019-11-07 23:59:59

Earthdata Login

Browse Collections

Features

- Map Imagery
- Near Real Time
- Customizable

Keywords

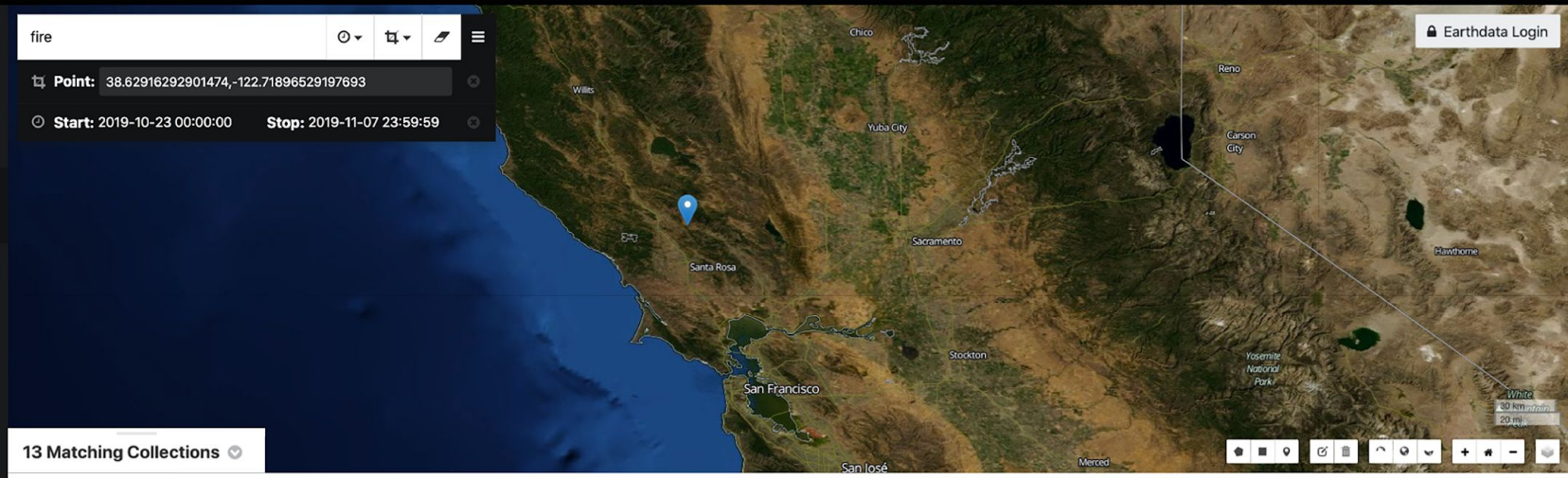
Platforms

Instruments

Organizations

Projects

Processing levels



13 Matching Collections

Sort by: Relevance ☑ Only include collections with granules ☐ Include non-EOSDIS collections

Tip: Add + collections to your project to compare and download their data.

MODIS/Terra Thermal Anomalies/Fire 8-Day L3 Global 1km SIN Grid V006

3 Granules • 2000-02-18 ongoing • The Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Thermal Anomalies and Fire 8-Day (MOD14A2) Version 6 data are generated at 1 kilometer (km) spatial resolultion as a Level 3 product. The MOD14A2 gridded composite contains the maximum value of the individual fire pixel classes detected during the eight days of acquisition. The Science Dataset (SDS) layers include the fire mask and pixel quality indicators. Improvements/Changes from Previous Versions *...

MOD14A2 v006 - LP DAAC

MODIS/Terra Thermal Anomalies/Fire Daily L3 Global 1km SIN Grid V006

3 Granules • 2000-02-18 ongoing • The Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Thermal Anomalies and Fire Daily (MOD14A1) Version 6 data are generated every eight days at 1 kilometer (km) spatial resolultion as a Level 3 product. MOD14A1 contains eight consecutive days of fire data conveniently packaged into a single file. The Science Dataset (SDS) layers include the fire mask, pixel quality indicators, maximum fire radiative power (MaxFRP), and the position of the fire pixel within the scan....

MOD14A1 v006 - LP DAAC

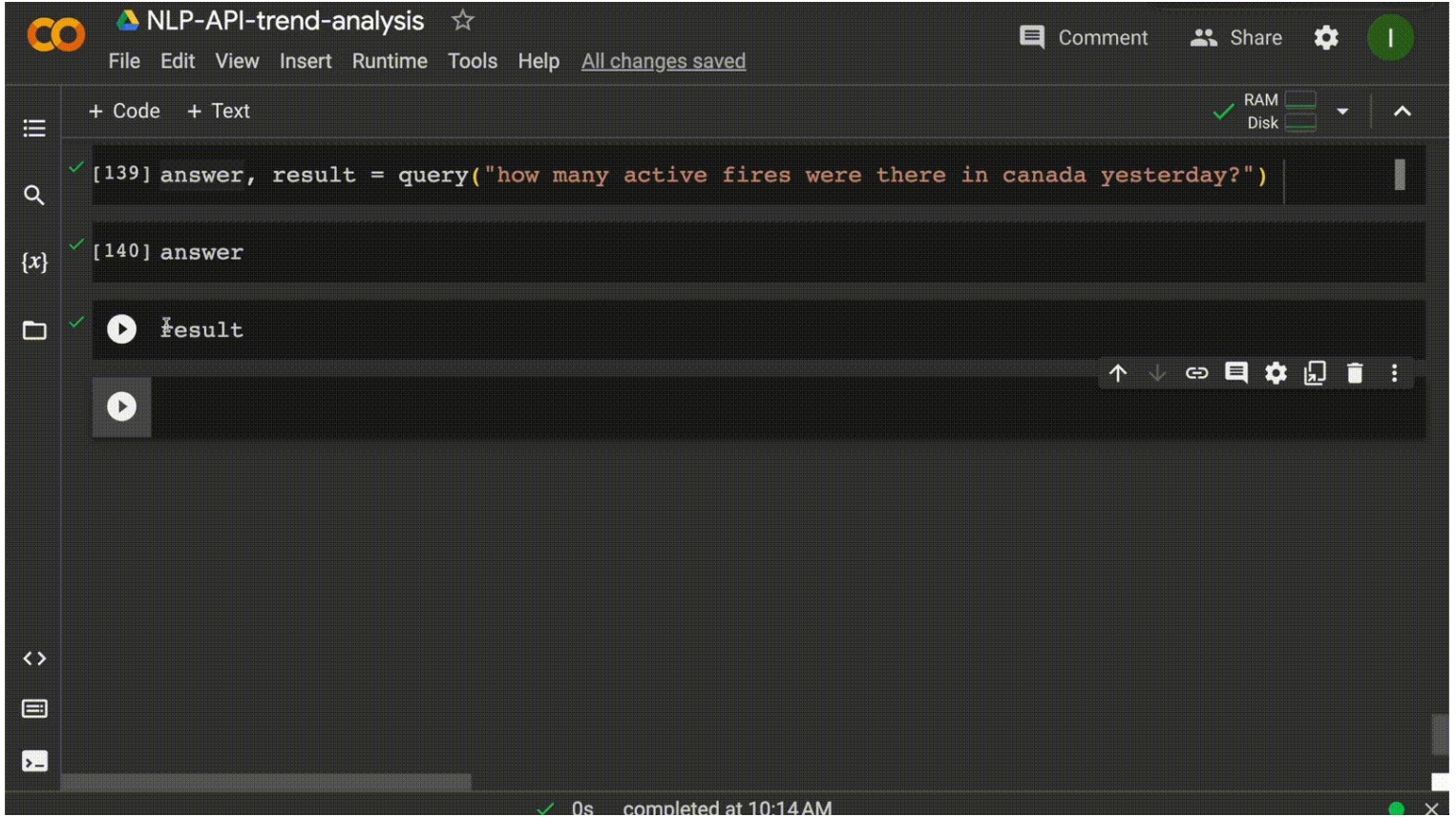
MODIS/Aqua Thermal Anomalies/Fire Daily L3 Global 1km SIN Grid V006

3 Granules • 2002-07-04 ongoing • The Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) Thermal Anomalies and Fire Daily (MYD14A1) Version 6 data are generated every eight days at 1 kilometer (km) spatial resolultion as a Level 3 product. MYD14A1 contains eight consecutive days of fire data conveniently packaged into a single file. The Science Dataset (SDS) layers include the fire mask, pixel quality indicators, maximum fire-radiative-power (MaxFRP), and the position of the fire pixel within the scan....

MYD14A1 v006 - LP DAAC

MODIS/Aqua Thermal Anomalies/Fire 8-Day L3 Global 1km SIN Grid V006

3 Granules • 2002-07-04 ongoing • The Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) Thermal Anomalies and Fire 8-Day (MYD14A2) Version 6 data are generated at 1 kilometer (km) spatial resolultion as a Level 3 product. The MYD14A2 gridded composite contains maximum value of individual fire pixel classes detected during the eight days of acquisition. The Science Dataset (SDS) layers include the fire mask and pixel quality indicators. Improvements/Changes from Previous Versions * Refinements to...


NLP-API-trend-analysis
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
[139] answer, result = query("how many active fires were there in canada yesterday?")
[140] answer
result
0s completed at 10:14 AM

Scientific process is fundamentally changing due to AI

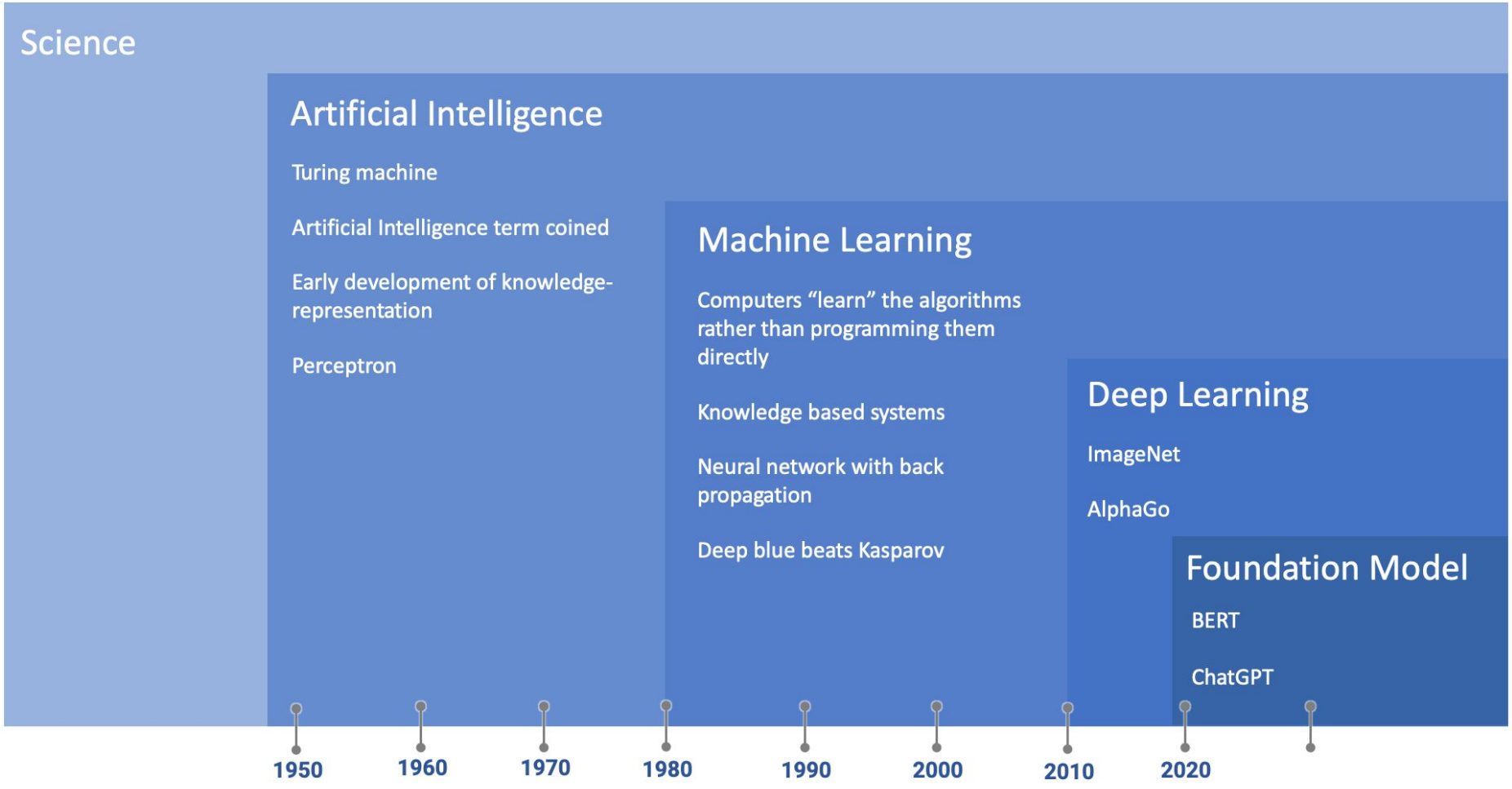
How can we best utilize AI tools in science?

AI has the potential to change every step of the science research lifecycle

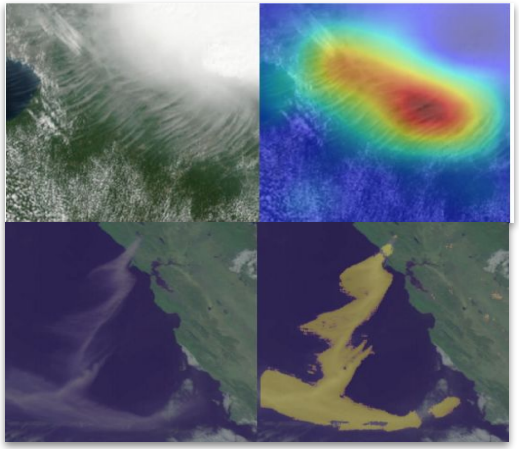
AI can speed up scientific discovery process - no repetitive tasks

Completely new AI generated hypothesis?

AI use should be disclosed


Science
Artificial Intelligence
Turing machine
Artificial Intelligence term coined
Early development of knowledge-representation
Perceptron
Machine Learning
Computers “learn” the algorithms rather than programming them directly
Knowledge based systems
Neural network with back propagation
Deep blue beats Kasparov
Deep Learning
ImageNet
AlphaGo
Foundation Model
BERT
ChatGPT
1950
1960
1970
1980
1990
2000
2010
2020

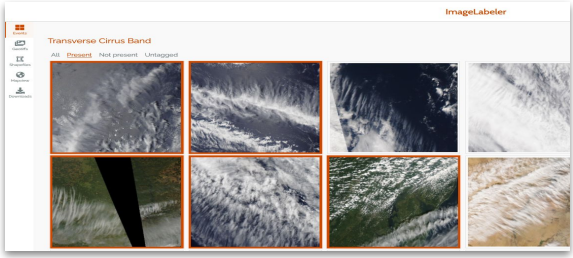
Supervised learning over the years...



Atmospheric phenomena identification

COVID 19 indicators

Marine debris segmentation



Labeling tools

Limitations of Supervised Learning

Advancing Application of Machine Learning Tools for NASA's Earth Observation Data

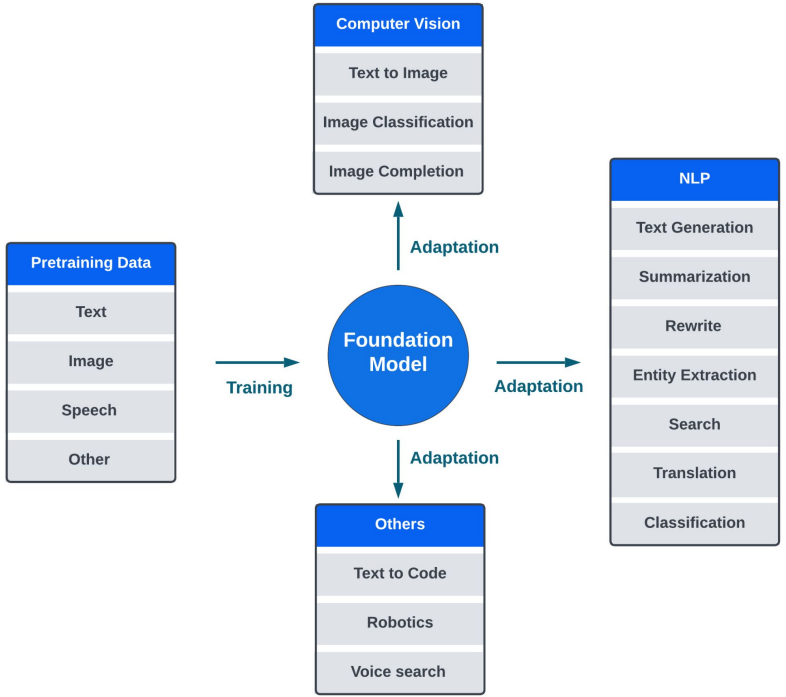
Jan. 21-23, 2020 | Washington, D.C.
Workshop Report

- **Training data** is the main component of supervised machine learning techniques and is increasingly becoming the **main bottleneck to advance applications of machine learning** techniques in Earth science.
- Geoscience models must **generalize across space and time**; however, for supervised learning one needs large training datasets to build generalizable models.

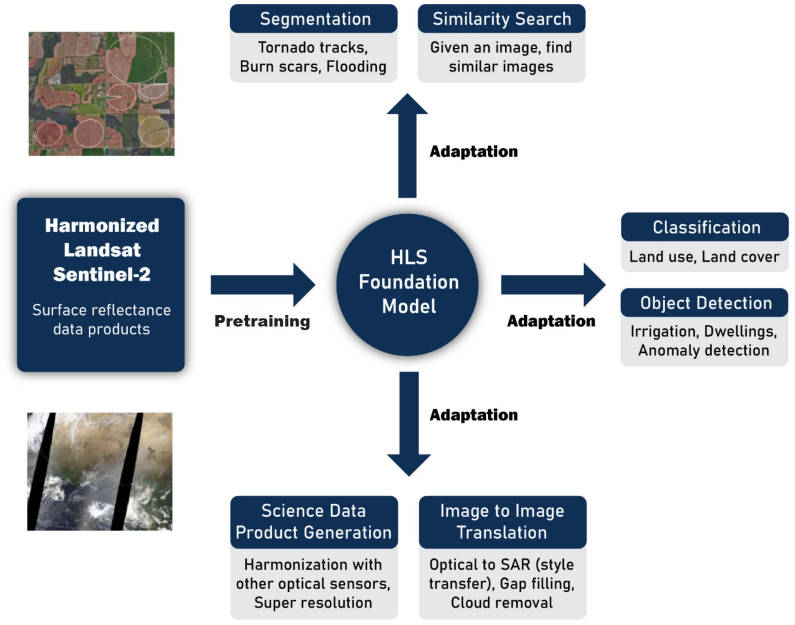
Maskey et al. "Advancing AI for Earth Science: A Data Systems Perspective," AGU Eos 2020

AI foundation models

- Large-scale models pre-trained on vast amounts of data, serving as a starting point for fine-tuning on specific tasks
- Unlike traditional models FMs are pre-trained on general data and then adapted to specialized tasks
- Pre-training captures broad knowledge, allowing for versatility across multiple applications
- Substantially reduce the downstream effort for building AI applications, including the need for large labeled training datasets

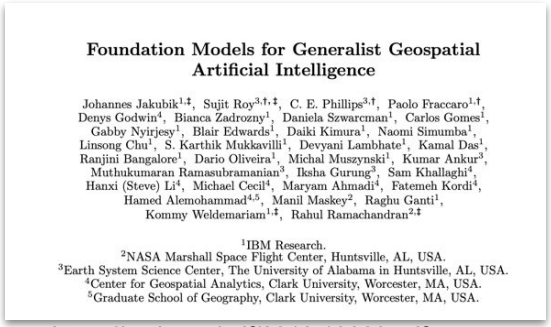


Geospatial foundation model with Harmonized Landsat Sentinel-2: Prithvi



- Build with collaboration with IBM Research
- Initial version released are 100M and 300M parameter models
- Masked Autoencoder where attention mechanism is extended in space and time
- Being evaluated for adaptation for different categories of downstream tasks

Collaborators: IBM, UAH, Clark University, ORNL, Hugging Face



Foundation Models for Generalist Geospatial Artificial Intelligence

Johannes Jakubik[1,‡], Sujit Roy[3,†,‡], C. E. Phillips[3,†], Paolo Fraccaro[1,†], Denys Godwin[4], Bianca Zadrozny[1], Daniela Szwarcman[1], Carlos Gomes[1], Gabby Nyirjesy[1], Blair Edwards[1], Daiki Kimura[1], Naomi Simumba[1], Linsong Chu[1], S. Karthik Mukkavilli[1], Devyani Lambhate[1], Kamal Das[1], Ranjini Bangalore[1], Dario Oliveira[1], Michal Muszynski[1], Kumar Ankur[3], Muthukumaran Ramasubramanian[3], Iksha Gurung[3], Sam Khallaghi[4], Hanxi (Steve) Li[4], Michael Cecil[4], Maryam Ahmadi[4], Fatemeh Kordi[4], Hamed Alemohammad[4,5], Manil Maskey[2], Raghu Ganti[1], Kommy Weldemariam[1,‡], Rahul Ramachandran[2,‡]

[1]IBM Research.
[2]NASA Marshall Space Flight Center, Huntsville, AL, USA.
[3]Earth System Science Center, The University of Alabama in Huntsville, AL, USA.
[4]Center for Geospatial Analytics, Clark University, Worcester, MA, USA.
[5]Graduate School of Geography, Clark University, Worcester, MA, USA.

https://arxiv.org/pdf/2310.18660.pdf

EARTHDATA
AI-powered Earth Insights BETA

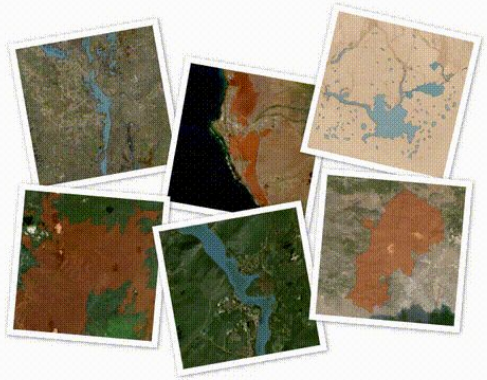
Welcome Explore About

Welcome to

AI-powered Earth Insights

AI-powered Earth Insights is a system that leverages the first of its kind open-source geospatial AI foundation model developed by NASA and IBM Research. It uses the Harmonized Landsat Sentinel-2 Foundation (HLS) data and models that are fine-tuned on Flood mapping and Burn scar segmentation tasks. It allows users to inference on the fine-tuned models and visualizes the results.

ABOUT START EXPLORING →



Understanding of foundational knowledge

Knowledge of AI is critical to develop scientific applications

Scientific knowledge is critical to evaluating AI results

Source: WIRED

Ethical issues

Bias and Fairness

Transparency and Explainability

Safety and Security

Human-AI Collaboration

nature

Explore content | About the journal | Publish with us | Subscribe

nature > comment > article

COMMENT | 31 October 2023

Garbage in, garbage out: mitigating risks and maximizing benefits of AI in research

Artificial-intelligence tools are transforming data-driven science — better ethical standards and more robust data curation are needed to fuel the boom and prevent a bust.

What AI advances are needed to realize science goals?

AI = Algorithm + Data (lots of data):
Algorithms that can learn from less data?

Novel data management

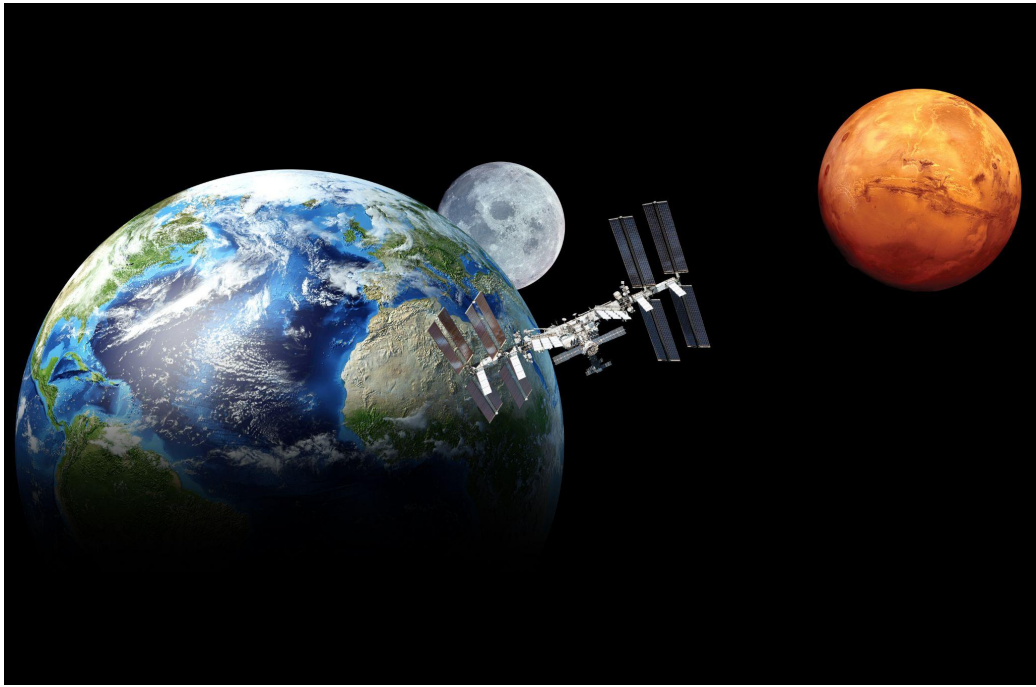
X-disciplinary tools

Optimization techniques

Ethical and explainable AI

Automated hypothesis generation

Affordable



Transforming Science with AI

AI Integration in every step of the scientific discovery and understanding

Responsible AI outputs

Efficient data utilization and model development

Collaboration and Open Science

Need for foundational science and computer science knowledge

Infrastructure and ecosystem development