IBM & NASA Open Source Largest Geospatial AI Foundation Model on Hugging Face

Effort aims to widen access to NASA earth science data for geospatial intelligence and accelerate climate-related discoveries – Aug 3, 2023

https://newsroom.ibm.com/2023-08-03-IBM-and-NASA-Open-Source-Largest-Geospatial-AI-Foundation-Model-on-Hugging-Face

IBM & NASA Open Source Largest Geospatial AI Foundation Model on Hugging Face

Effort aims to widen access to NASA earth science data for geospatial intelligence and accelerate climate-related discoveries – Aug 3, 2023

https://newsroom.ibm.com/2023-08-03-IBM-and-NASA-Open-Source-Largest-Geospatial-AI-Foundation-Model-on-Hugging-Face

Brian Hamlin

Piedmont, California, United States

Safety First Means Climate

- Satellite and other Remote Sensing
  expanding at an accelerating pace
- Artificial Intelligence advances
  are real, measurable, being deployed
- **FoundationModels** develop and are
  applied to geospatial data environments

Parts & Pieces to Start

- **NASA** Earthscience Satellite Imagery for time-series analysis

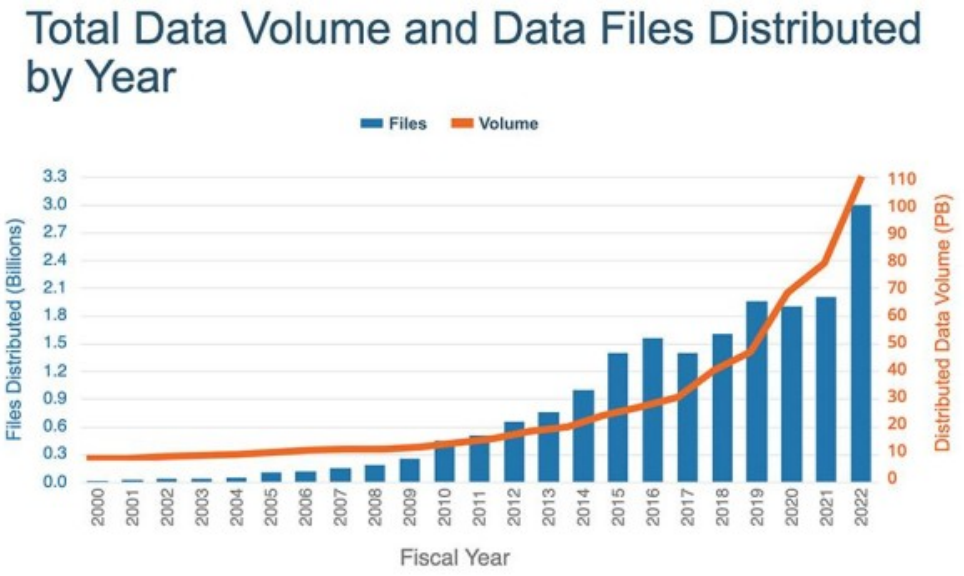
- Artificial Intelligence – **BERT** Natural Language Processing (NLP)

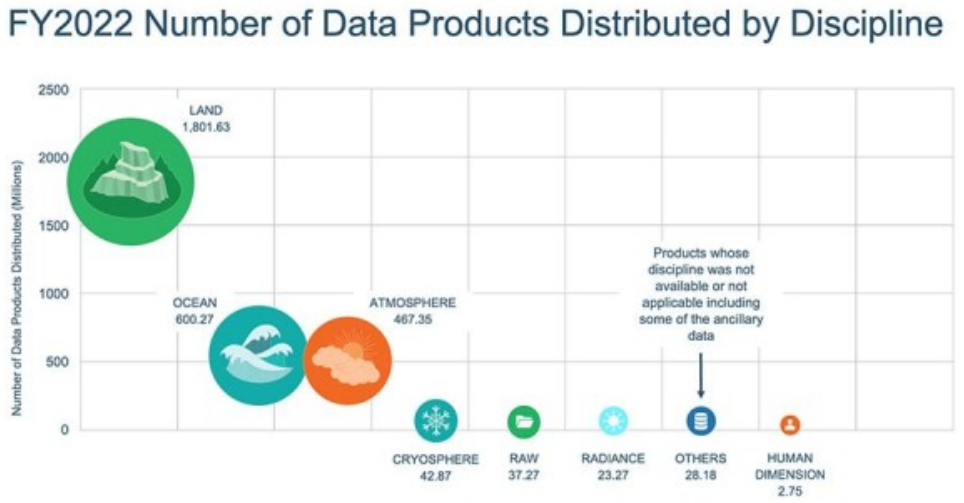
- What are **FoundationModels**

Bulk Data Metrics
2022
NASA IMPACT Labs

12 new missions to be added

Current archive holdings continue to grow (~95 missions)





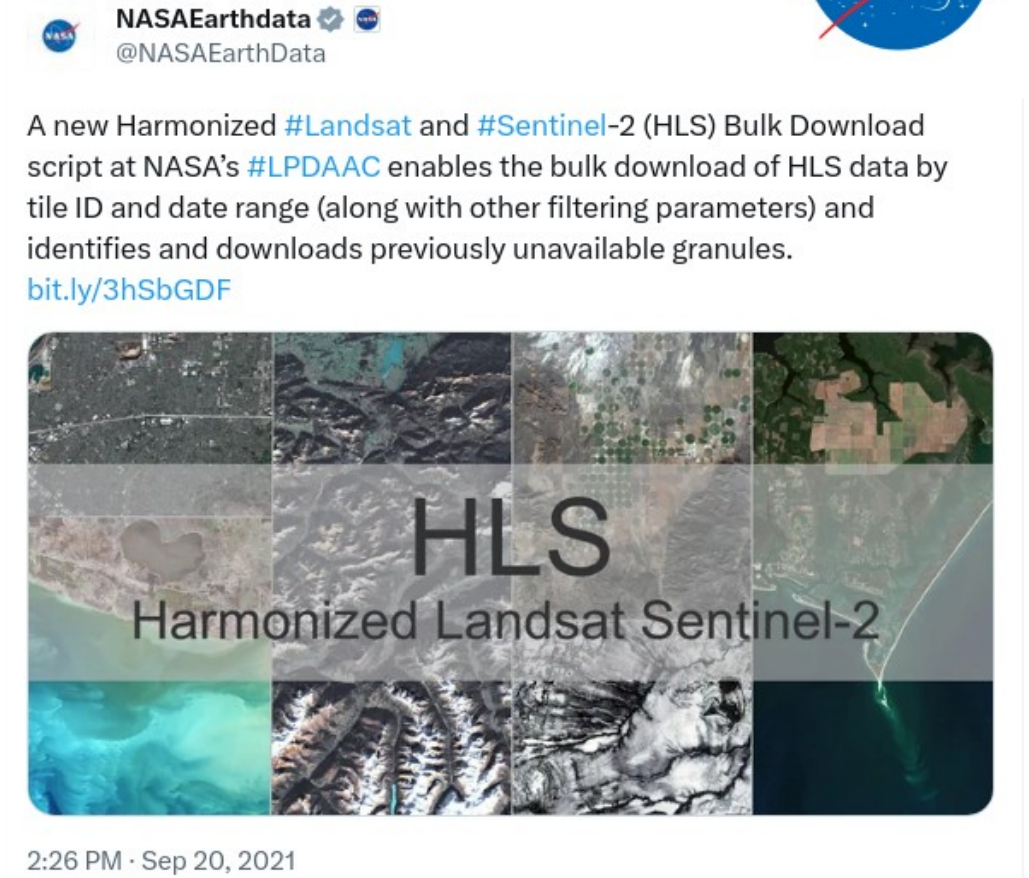
Dr. Rahul Ramachandran*, NASA/MSFC IMPACT

Contact:
rahul.ramachandran@nasa.gov

NASA SOFTWARE

IBM's new geospatial foundation model was uniquely built using Earth surface reflectance data from **NASA**'s Harmonized Landsat Sentinel-2 (**HLS**) data products. HLS data is used routinely for environmental monitoring and data production is managed, in part, by the IMPACT HLS team. By priming the foundation model with labeled images of features like fire burn scars and flooding boundaries, the model is then poised to significantly accelerate geospatial analysis of future environmental events. Early testing has already indicated significant improvement (15%) in mapping floods and fire burn scars compared to deep learning models.

`https://impactunofficial.medium.com/ibm-geospatial-foundation-model-trained-with-hls-data-acbc5a8b0a49`



In **January 2023**, under a NASA Space Act Agreement, IBM began training a foundation model on a sliver of NASA's Harmonized Landsat Sentinel-2 (HLS) dataset, which provides a full view of Earth every two to three days. At a resolution of **30-meters per pixel, HLS** images are close enough for detecting changes in land-use but not quite detailed enough for identifying individual trees.

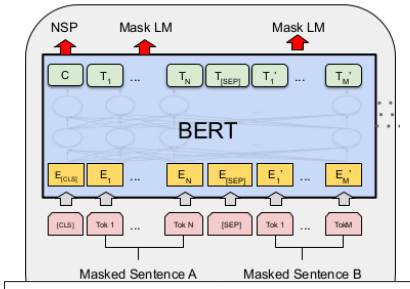
Built on a vision transformer (ViT) and a masked autoencoder (MAE) architecture, the model has been adapted to process satellite images by **expanding its spatial attention mechanism to include time**. IBM trained the model on its AI supercomputer, Vela, and leveraged PyTorch and ecosystem libraries for training and tuning on labeled images of floods and burn-scars from wildfires. In tests, researchers saw a 15% accuracy boost compared to state-of-the-art deep learning models for mapping floods and fires.

`https://research.ibm.com/blog/nasa-hugging-face-ibm?mhsrc=ibmsearch_a&mhq=geospatial%20AND%20NASA`

[Submitted on 11 Oct 2018 (v1), last revised 24 May 2019 (this version, v2)]

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova



Google AI Language
ArXiv:1810.04805v2

Bidirectional Encoder Representations from Transformers (**BERT**) is a family of language models introduced in 2018 by researchers at Google.[1][2] A 2020 literature survey concluded that "in a little over a year, BERT has become a ubiquitous baseline in Natural Language Processing (NLP) experiments counting over 150 research publications analyzing and improving the model."[3]
The reasons for BERT's state-of-the-art performance on these natural language understanding tasks are not yet well understood.[9][10] Current research has focused on investigating the relationship behind BERT's output as a result of carefully chosen input sequences,[11][12] analysis of internal vector representations through probing classifiers,[13][14] and the relationships represented by attention weights.[9][10] The high performance of the BERT model could also be attributed to the fact that it is bidirectionally trained. This means that BERT, based on the Transformer model architecture, applies its self-attention mechanism to learn information from a text from the left and right side during training, and consequently gains a deep understanding of the context. For example, the word *fine* can have two different meanings depending on the context (**I feel** fine **today**, **She has** fine **blond hair**). BERT considers the words surrounding the target word *fine* from the left and right side.

A Google AI Language team introduced BERT around 2018; by 2020 BERT was almost ubiquitous in Natural Language Processing (NLP) AI, using a new mechanism called

Transformers

WIKIPEDIA
The Free Encyclopedia

A **transformer** is a deep learning architecture that relies on the parallel multi-head attention mechanism.[1] The modern transformer was proposed in the 2017 paper titled 'Attention Is All You Need' by Ashish Vaswani et al., Google Brain team. It is

Transformers Library

Hugging Face

The Transformers library is a Python package that contains open-source implementations of transformer models for text, image, and audio tasks. It is compatible with the PyTorch, TensorFlow and JAX deep learning libraries and includes implementations of notable models like BERT and GPT-2.[16] The library was originally called "pytorch-pretrained-bert"[17] which was then renamed to "pytorch-transformers" and finally "transformers."

- Perceiver – Machine learning algorithm for non-textual data
- BERT (language model) – Language model developed by Google
- GPT-3 – 2020 large language model
- GPT-4 – 2023 text-generating language model
- ChatGPT – AI chatbot developed by OpenAI
- Wu Dao – Chinese multimodal artificial intelligence program
- Vision transformer – Machine learning algorithm for vision processing
- BLOOM (language model) – Open-access multilingual language model

What are **FoundationModels**

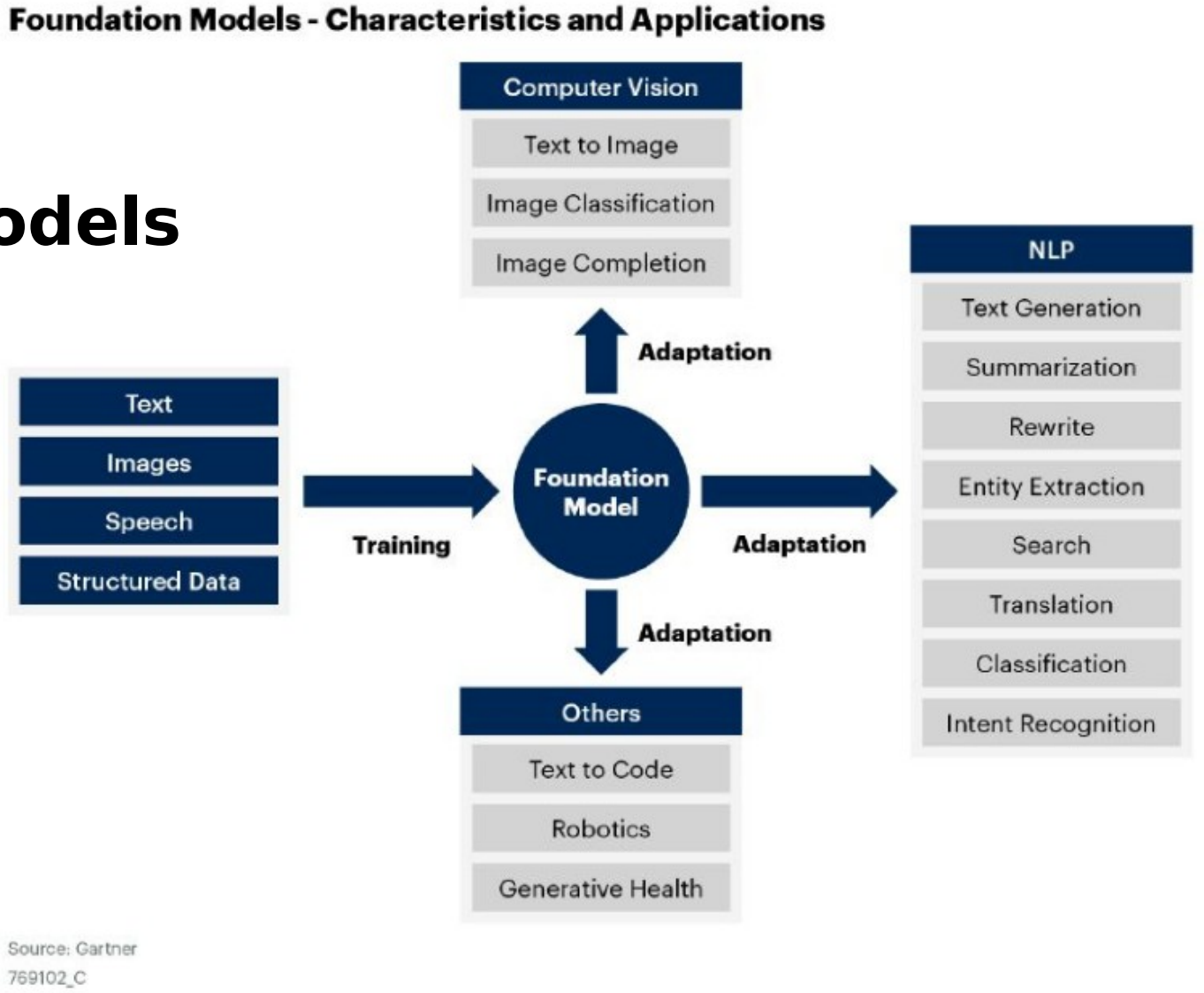
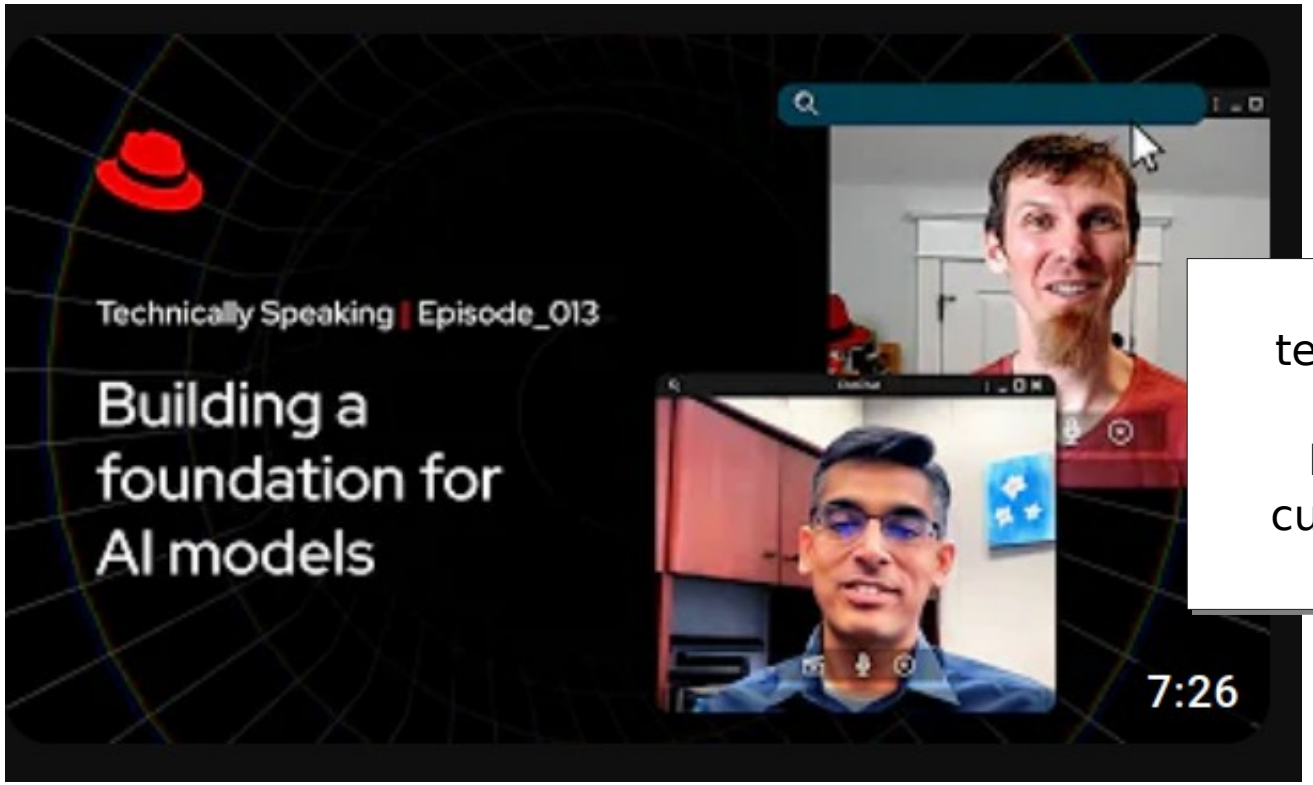


Image source: Gartner Report - Innovation Insight for Artificial Intelligence Foundation Models, Published 27 Oct 2022



RedHat, IBM and other technology partners are productizing FoundationModels for customer deployment on premises, or cloud

Apr 27, 2022 #AI #RedHat #FoundationModels Why can't we build and reuse AI models? Mo data, mo problems? Learn how AI foundation models change the game for training AI/ML and join **Red Hat CTO Chris Wright** and **IBM Research AI VP Sriram Raghavan** to explore foundation models, an emerging approach to machine learning and data representation. Even in the age of big data when AI/ML is more prevalent, training the next generation of AI tools like NLP requires enormous data, and using AI models to new or different domains may be tricky. A foundation model can consolidate data from several sources so that one model may then be used for various activities. But how will foundation models be used for things beyond natural language processing? Don't miss this episode to explore how foundation models are a paradigm shift in how AI gets done.

Snapshot: 09/06/2023 20:14

Foundation models

A foundation model is a model trained at broad scale

that can adapted to a wide range of downstream tasks

Examples: BERT GPT-3 CLIP

The technology is not new:

Self-supervised learning with neural networks

What is new?

Scale and the ability to perform tasks beyond training

Emergence and homogenisation

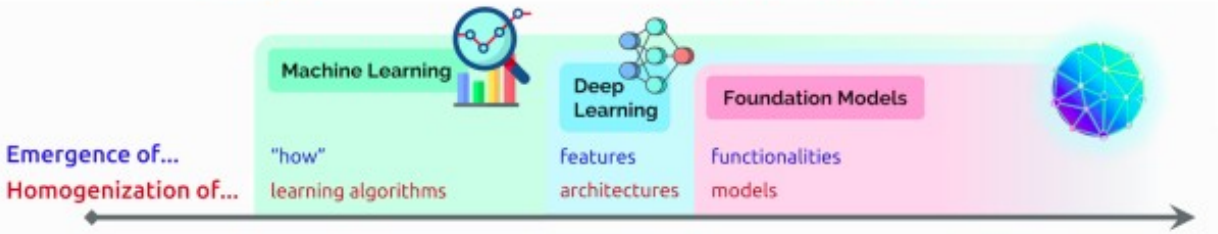
Two key ideas underpin the significance of foundation models:

Emergence

- system behaviour is implicitly induced rather than explicitly constructed
- cause of scientific excitement and anxiety of unanticipated consequences

Homogenisation

- consolidation of methodology for building machine learning system across many applications
- provides strong leverage for many tasks, but also creates single points of failure



On the Opportunities and Risks of Foundation Models

R. Bommasani, D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, M. S. Bernstein, J. Bohg, A. Bosselut, E. Brunskill, E. Brynjolfsson, S. Buch, D. Card, R. Castellon, N. Chatterji, A. Chen, K. Creel, J. Q. Davis, D. Demszky, C. Donahue, M. Doumbouya, E. Durmus, S. Ermon, J. Etchemendy, K. Ethayarajh, L. Fei-Fei, C. Finn, T. Gale, L. Gillespie, K. Goel, N. Goodman, S. Grossman, N. Guha, T. Hashimoto, P. Henderson, J. Hewitt, D. E. Ho, J. Hong, K. Hsu, J. Huang, T. Icard, S. Jain, D. Jurafsky, P. Kalluri, S. Karamcheti, G. Keeling, F. Khani, O. Khattab, P. W. Koh, M. Krass, R. Krishna, R. Kuditipudi, A. Kumar, F. Ladhak, M. Lee, T. Lee, J. Leskovec, I. Levent, X. L. Li, X. Li, T. Ma, A. Malik, C. D. Manning, S. Mirchandani, E. Mitchell, Z. Munyikwa, S. Nair, A. Narayan, D. Narayanan, B. Newman, A. Nie, J. C. Niebles, H. Nilforoshan, J. Nyarko, G. Ogut, L. Orr, I. Papadimitriou, J. S. Park, C. Piech, E. Portelance, C. Potts, A. Raghunathan, R. Reich, H. Ren, F. Rong, Y. Roohani, C. Ruiz, J. Ryan, C. Ré, D. Sadigh, S. Sagawa, K. Santhanam, A. Shih, K. Srinivasan, A. Tamkin, R. Taori, A. W. Thomas, F. Tramèr, R. E. Wang, W. Wang, B. Wu, J. Wu, Y. Wu, S. M. Xie, M. Yasunaga, J. You, M. Zaharia, M. Zhang, T. Zhang, X. Zhang, Y. Zhang, L. Zheng, K. Zhou, P. Liang, arxiv 2021

Digest (of the introduction) by Samuel Albanie, June 2022

crfm.stanford.edu

Workshop on Foundation Models

AUGUST 23-24, 2021

VIRTUAL EVENT

HAI Stanford University Human-Centered Artificial Intelligence

Center for Research on Foundation Models

Image credits/References
(BERT) J. Devlin et al., "Bert: Pre-training of deep bidirectional transformers for language unde
(GPT-3) T. Brown et al., "Language models are few-shot learners", NeurIPS (2020)
(CLIP) A. Radford et al., "Learning transferable visual models from natural language supervisi
V. R. de Sa, "Learning Classification with Unlabeled Data", NeurIPS (1993)
R. Bommasani et al., "On the opportunities and risks of foundation models", arxiv (2021)

Geospatial FoundationModels are being implemented today;

ACTION THESIS

Stakeholders must participate meaningfully in verification,
as developers build implementations.
Iterate governance, construct checks and balances;
act and develop the platforms in consequential ways.

Power a pipeline of skilled human participants.
Apply previously impossible capabilities.

Technology Advances

- G20 Summit, India Chair, August 2023
  Prioritize AI for Equitable Science
- State of AI in 2023

  https://hai.stanford.edu/news/2023-state-ai-14-charts
- IPCC AR6 Issues “Dire Warning”

Introduction

AI technologies are expected to bring a wide array of **economic and societal benefits** to a wide range of sectors, including environment and health, the public sector, finance, mobility, home affairs and agriculture. They are particularly useful for improving prediction, for optimising operations and resource allocation, and for personalising services.[1] However, the implications of AI systems for **fundamental rights** protected under the EU Charter of Fundamental Rights, as well as the **safety risks** for users when AI technologies are embedded in products and services, are raising concern. Most notably, AI systems may jeopardise fundamental rights such as the right to non-discrimination, freedom of expression, human dignity, personal data protection and privacy.[2]

Given the fast development of these technologies, in recent years AI regulation has become a central policy question in the European Union (EU). Policy-makers pledged to develop a **'human-centric' approach to AI** to ensure that Europeans can benefit from new technologies developed and functioning according to the EU's values and principles.[3] In its 2020 White Paper on Artificial Intelligence, the European Commission committed to **promote the uptake of AI** and **address the risks associated** with certain uses of this new technology. While the European Commission initially adopted a **soft-law approach**, with the publication of its non-binding 2019 Ethics Guidelines for Trustworthy AI and Policy and investment recommendations, it has since shifted towards a **legislative approach**, calling for the adoption of harmonised rules for the development, placing on the market and use of AI systems.[4]

Stanford University

Center for Research on Foundation Models

HAI Stanford University Human-Centered Artificial Intelligence

Response on "Foundation Models"

Author: Jitendra Malik

OPEN (FOR BUSINESS): BIG TECH, CONCENTRATED POWER, AND THE POLITICAL ECONOMY OF OPEN AI

David Gray Widder
Carnegie Mellon University
Pittsburgh, United States.

Meredith Whittaker
Signal Foundation
San Francisco, United States.

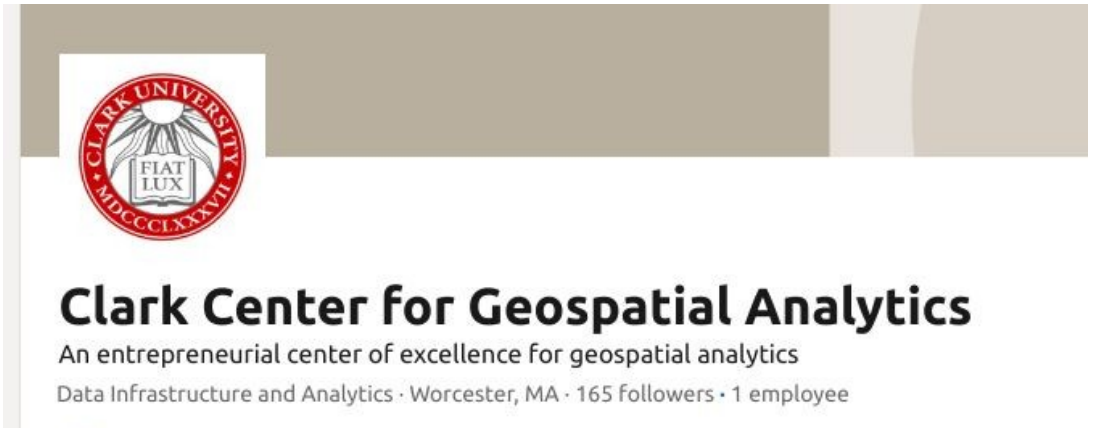
Sarah Myers West
AI Now Institute
New York, United States.

August 16, 2023

AI◦Safety is an evolving field, calling for civilian participation

Geospatial FoundationModel Verification

- **NASA** partners with researchers
  Clark Center for Geospatial Analytics
- ServiceNow Benchmarks
- Industry Expert Feedback



Clark geospatial team partners with NASA, IBM to harness AI technology; *New model to aid researchers, policymakers in addressing climate impacts*

August 29, 2023 By Meredith Woodward King
https://clarknow.clarku.edu/2023/08/29/clark-geospatial-team-partners-with-nasa-ibm-to-harness-ai-technology/

The summer of 2023 is headed for the record books, with increasing heat waves, wildfires, tropical storms, and flooding. July was declared the hottest month on earth since records began in 1880. To better understand how the earth is changing, the impact of extreme climate events, and how humans might adapt, researchers use satellite images to extrapolate data.

Hamed Alemohammad, director of Clark's new Center for Geospatial Analytics, and six graduate students — working with NASA and IBM — are hoping artificial intelligence (AI) can answer these questions. Together, they have produced the world's first geospatial AI foundation model, a milestone that will allow climate and earth scientists to access and study data more quickly and efficiently.

"In construction, you put the foundation on the ground, and then you build a customized structure on top. The foundation model is practically the same thing. But in this case, you are building a deep learning model," says Alemohammad, associate professor in Clark's Graduate School of Geography.

Using a **foundation model** to build generative AI models that can be customized for various applications, rather than building those custom models from scratch, saves researchers time and money, he explains. …

ChatGPT and Google's Bard are examples of generative models built on top of the first large language foundation model. The **Clark/NASA/IBM** project "is the first foundation model in geospatial earth science," Alemohammad says. "We want to assess the usability of foundation models in this field."

This year, IBM and NASA's Interagency Implementation and Advanced Concepts Team (**IMPACT**) deployed the geospatial AI foundation model to comb through and extract information from a year of raw, unlabeled imagery data gathered by the space agency's Harmonized Landsat Sentinel-2 (**HLS**) satellite from across the continental United States.

In July, the team released the foundation model on Hugging Face, a repository for open-source machine learning models. The foundation model — fine-tuned on human-labeled data for mapping of floods and burn scars from wildfires — so far has **demonstrated a 15 percent improvement over other state-of-the-art techniques**, according to IBM. The effort is tied to NASA's goal to make data, code, and AI models available to everyone through its Open-Source Science Initiative.

"We believe that foundation models have the potential to change the way observational data are analyzed and help us to better understand our planet," says Kevin Murphy, chief science data officer at NASA. "And by open-sourcing such models and making them available to the world, we hope to multiply their impact."

AI◦SAFETY | GEOSPATIAL FOUNDATIONMODELS | Brian M Hamlin

servicenow + ELEMENT AI

GEO-Bench: Toward Foundation Models for Earth Monitoring

Alexandre Lacoste[1] **Nils Lehmann**[2]

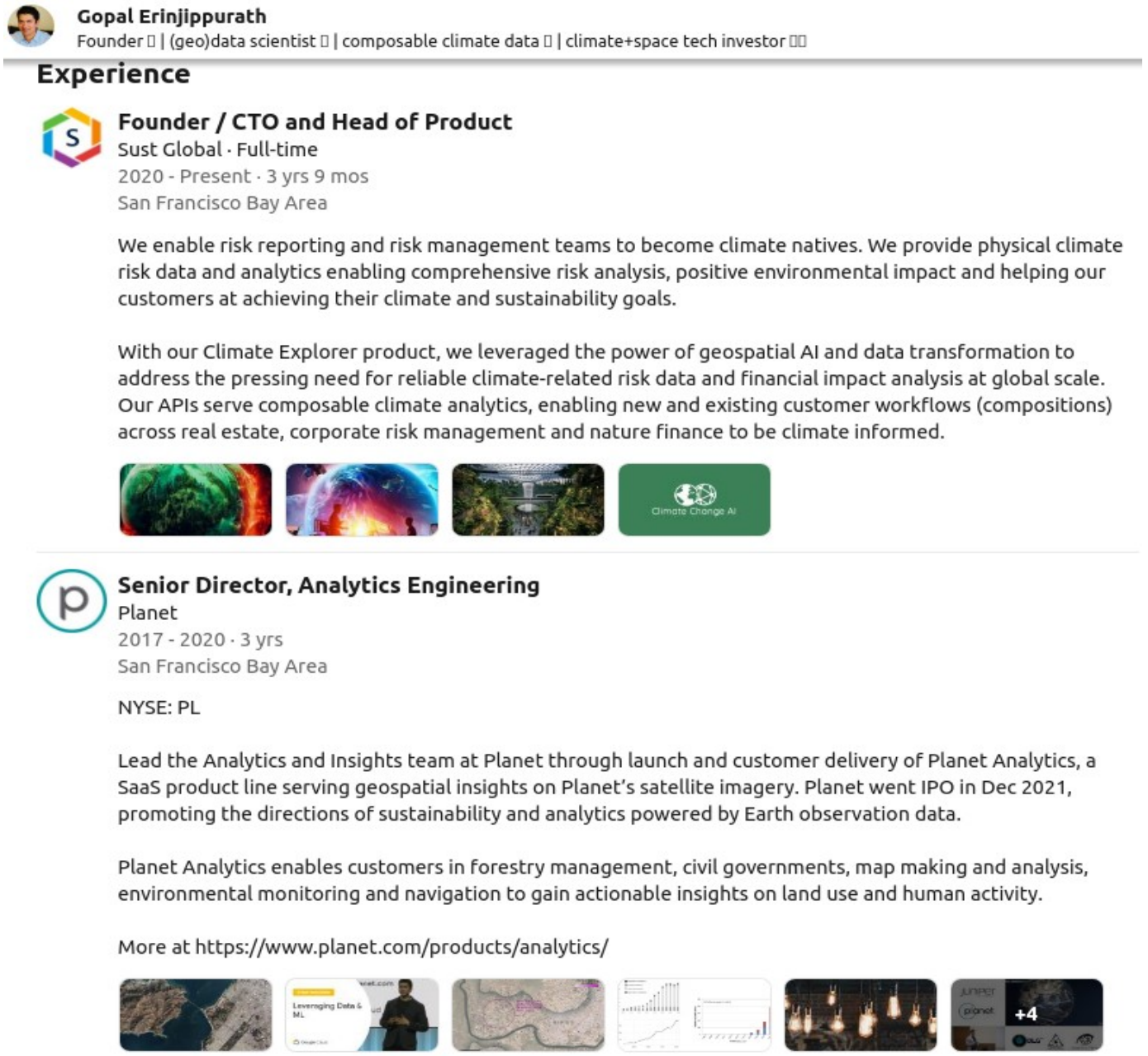
[1] ServiceNow Research [2] University of Amsterdam [3] Stanford University
[4] University of Maryland [5] MIT [6] ETH Zurich [7] Clark University
[8] Mila-Quebec [9] University of Montreal [10] Technical University of Munich

Recent progress in self-supervision has shown that pre-training large neural networks on vast amounts of unsupervised data can lead to substantial increases in generalization to downstream tasks. Such models, recently coined *foundation models*, have been transformational to the field of natural language processing. Variants have also been proposed for image data, but their applicability to remote sensing tasks is limited. To stimulate the development of foundation models for Earth monitoring, we propose a benchmark comprised of six classification and six segmentation tasks, which were carefully curated and adapted to be both relevant to the field and well-suited for model evaluation. We accompany this benchmark with a robust methodology for evaluating models and reporting aggregated results to enable a reliable assessment of progress.

```
~$  dataset/geobench/*/*
11G     dataset/geobench/classification_v0.9.0/m-bigearthnet
2.8G    dataset/geobench/classification_v0.9.0/m-brick-kiln
1.7G    dataset/geobench/classification_v0.9.0/m-eurosat
2.0G    dataset/geobench/classification_v0.9.0/m-forestnet
1.4G    dataset/geobench/classification_v0.9.0/m-pv4ger
1.3G    dataset/geobench/classification_v0.9.0/m-so2sat
2.3G    dataset/geobench/segmentation_v0.9.0/m-cashew-plantation
3.4G    dataset/geobench/segmentation_v0.9.0/m-chesapeake-lancover
731M    dataset/geobench/segmentation_v0.9.0/m-NeonTree
476M    dataset/geobench/segmentation_v0.9.0/m-nz-cattle
618M    dataset/geobench/segmentation_v0.9.0/m-pv4ger-seg
3.4G    dataset/geobench/segmentation_v0.9.0/m-SA-crop-type

        Snapshot: 08/26/2023 00:51
```

Industry Expert Gopal Erinjippurath describes the Prithvi Geospatial FoundationModel



Gopal Erinjippurath
Founder □ | (geo)data scientist □ | composable climate data □ | climate+space tech investor □□

Experience

Founder / CTO and Head of Product
Sust Global · Full-time
2020 - Present · 3 yrs 9 mos
San Francisco Bay Area

We enable risk reporting and risk management teams to become climate natives. We provide physical climate risk data and analytics enabling comprehensive risk analysis, positive environmental impact and helping our customers at achieving their climate and sustainability goals.

With our Climate Explorer product, we leveraged the power of geospatial AI and data transformation to address the pressing need for reliable climate-related risk data and financial impact analysis at global scale. Our APIs serve composable climate analytics, enabling new and existing customer workflows (compositions) across real estate, corporate risk management and nature finance to be climate informed.

Senior Director, Analytics Engineering
Planet
2017 - 2020 · 3 yrs
San Francisco Bay Area

NYSE: PL

Lead the Analytics and Insights team at Planet through launch and customer delivery of Planet Analytics, a SaaS product line serving geospatial insights on Planet's satellite imagery. Planet went IPO in Dec 2021, promoting the directions of sustainability and analytics powered by Earth observation data.

Planet Analytics enables customers in forestry management, civil governments, map making and analysis, environmental monitoring and navigation to gain actionable insights on land use and human activity.

More at https://www.planet.com/products/analytics/

Gopal Erinjippurath:

https://www.sustglobal.com/

The **Prithvi-100M** model announced last week from #nasa+#ibm collab is the first large scale geospatial foundational model for #earthobservation (EO) data It's a complex model with 100M parameters trained on open #eo data with ability to work with the time dimension.

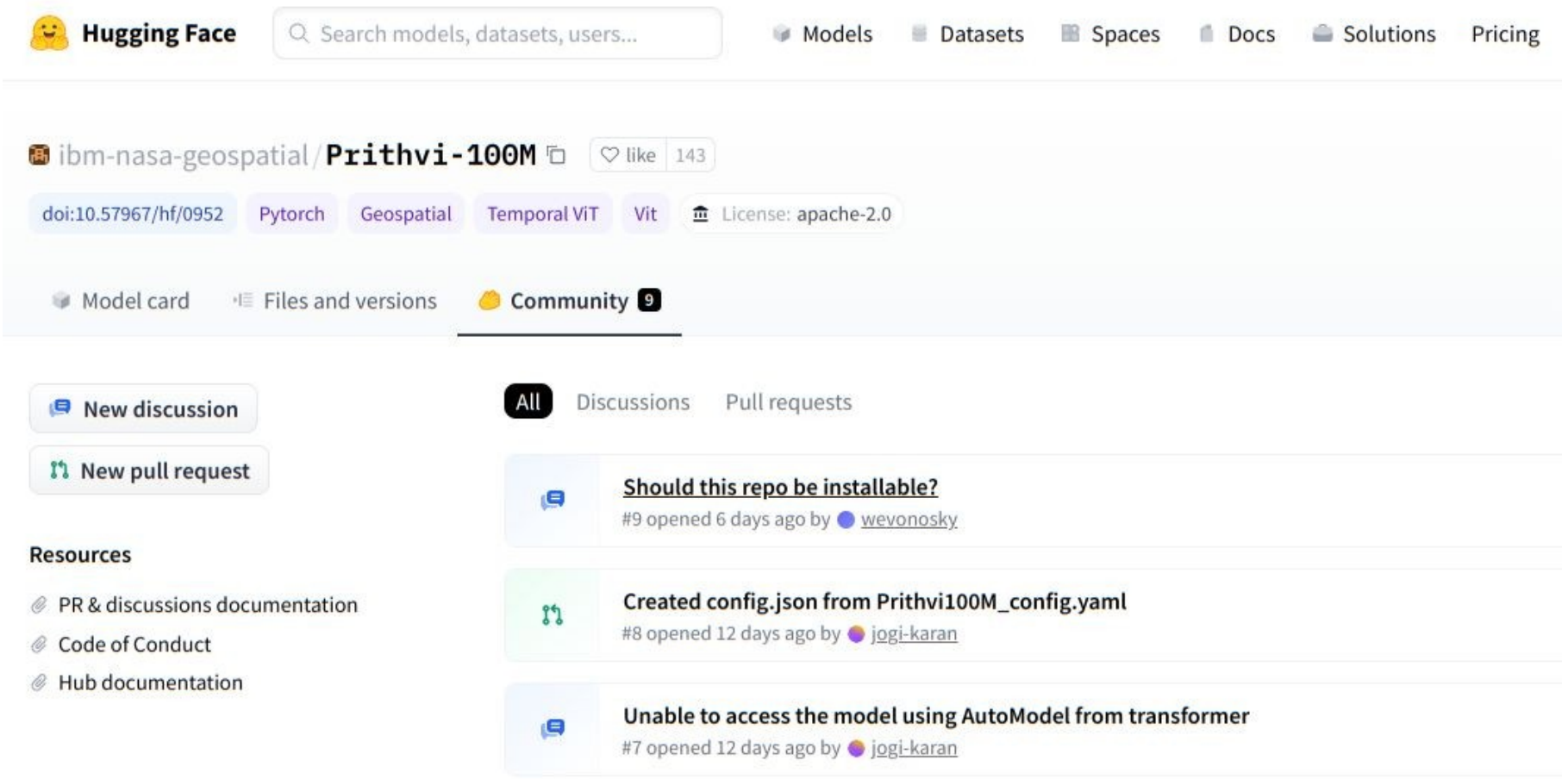
The pre-trained model is a **vision transformer (ViT)** operating as a **Masked Autoencoder (MAE)**, trained on hybrid #landsat #sentinel **(HLS) data**. What's with all those acronyms? Let's unpack these for a minute:

HLS: That's the training data used by this model. It is a data fusion product from NASA providing consistent data from the Operational Land Imager (OLI) on the Landsat 8+9 satellites and the Multi-Spectral Instrument (MSI) on the Sentinel-2A+2B satellites. IMO it’s the best of open source EO, enabling consistent global land observations at 30m spatial resolution every 3 days.

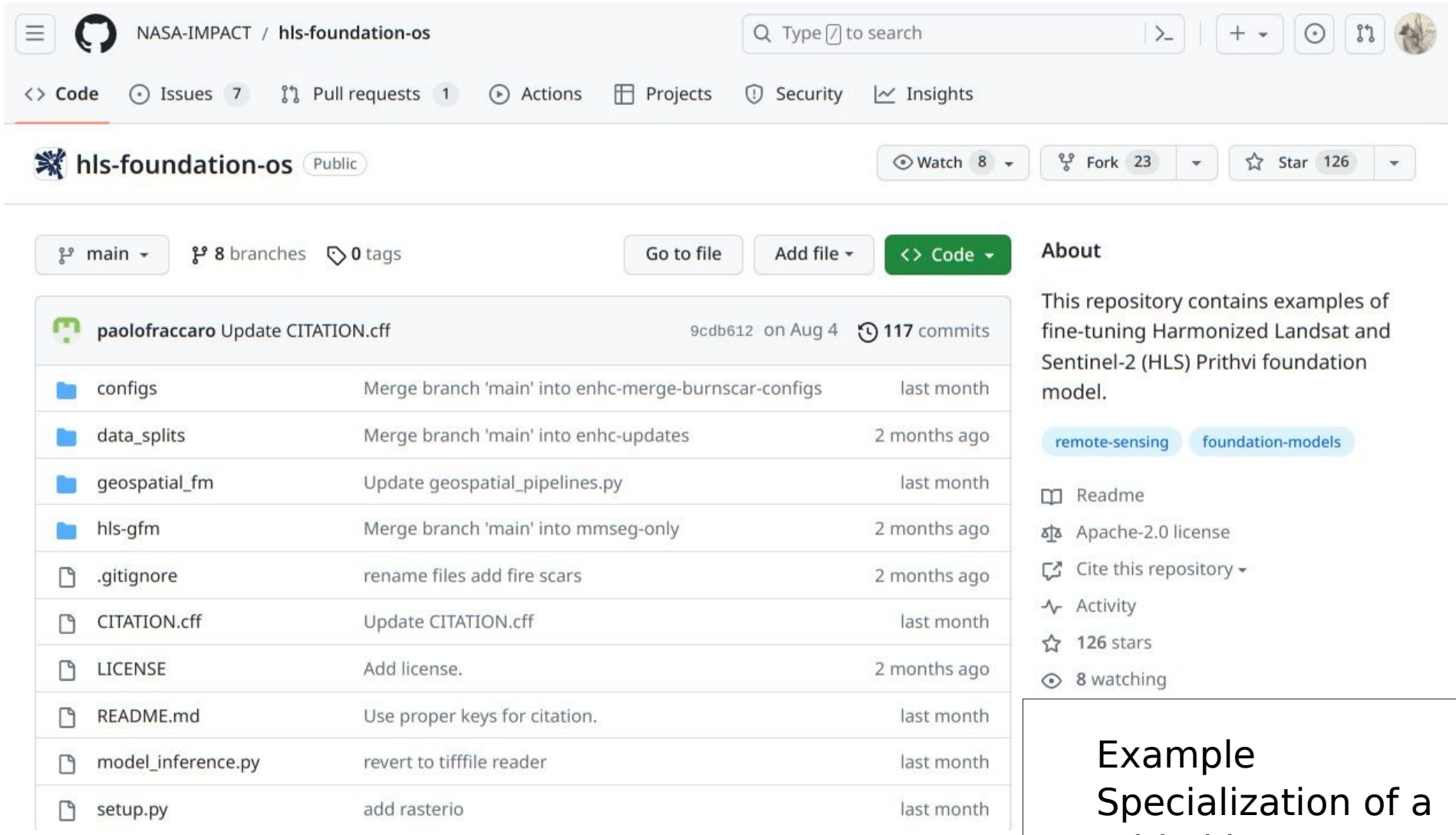
ViT: That's one aspect of the backbone network called the vision transformer. These class of models represent an input image as a series of image patches. Just like transfers in GPT 3/3.5/4 transform a series of word embeddings to text, the ViT transform these series of image patches into labels for the image. While convolutional networks (CNNs) use multidimensional pixel arrays, ViT represent the input images into "visual tokens".

MAE: Auto-encoders are powerful networks for self supervision (learning without supervision aka without labeled training data). Masked auto-encoders are trained with full images that are partially masked out and the decoder tries to predict the masked patches. Masks matter as they appear in real world EO data are clouds, smoke, haze, snow, shadow. Once trained, being a foundational model, the thesis is that the encoder can be used to create embeddings for fine tuning to broad range of tasks, examples being Flood detection using single time step 6 band imagery and weighted BCE loss and land segmentation using multi-step 6 band imagery and cross entropy loss.

HuggingFace ‘Prithvi’ Model Card



Snapshot: 09/06/2023 20:14



Snapshot: 09/06/2023 20:14

Example Specialization of a Prithvi base FoundationModel

CONTACT:

Brian M Hamlin
maplabs@light42.com
<*private-tbd*>

Berkeley, California
general inquiries
partnerships