

IBM Sustainability

AI, GenAI and Watsonx in Sustainability Software

Klaus Roder
Data, AI & Sustainability Enthusiast
WW Technical Sales Leader
IBM Sustainability Software



Surfing the AI wave




Surfing the **AI** wave
Surfing the **AI *hype cycle*** wave





Surfing the **AI** wave
Surfing the **AI *hype cycle*** wave
Surfing the **AI ~~hype~~** wave

Agenda

- 
- How Foundation Models Work
 - GenAI in Sustainability Software (Maximo)
 - AI in Sustainability Software

AI & Sustainability Goals
are best accomplished
when they are in sync
with key
Business Imperatives

My personal question ...

How can I,
How can we,
working for, in, with
companies
implement
AI & sustainability
in action?



Where is the opportunity for AI in sustainable operations?



Companies with a network of Physical Assets

- Reduce emissions
- Energy efficient facilities and assets
- Longer asset life
- Climate risk mitigation
- Regulatory compliance

20%–30%

Increased productivity

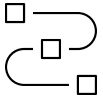


Companies with a large IT Infrastructure

- Reduce emissions
- Lower cost of IT
- Improved power utilization
- Regulatory compliance

30% in 6mo.

Reduced cloud and data center consumption



Companies with complex Supply Chains


- Enable track and trace
- Lower product carbon footprints
- Reduce fuel consumption
- Deliver sustainable sourcing
- Scope 3 emissions
- Regulatory compliance

52%

Reduction in expedited freight costs

Agenda

- How Foundation Models Work
- GenAI in Sustainability Software (Maximo)
- AI in Sustainability Software

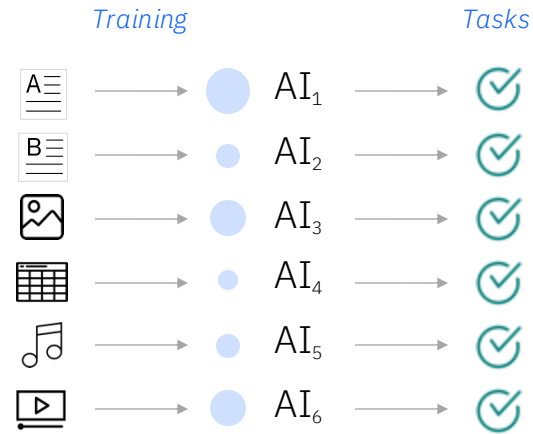


How do
foundation
models
work?

Foundation models establish a new paradigm for AI capabilities

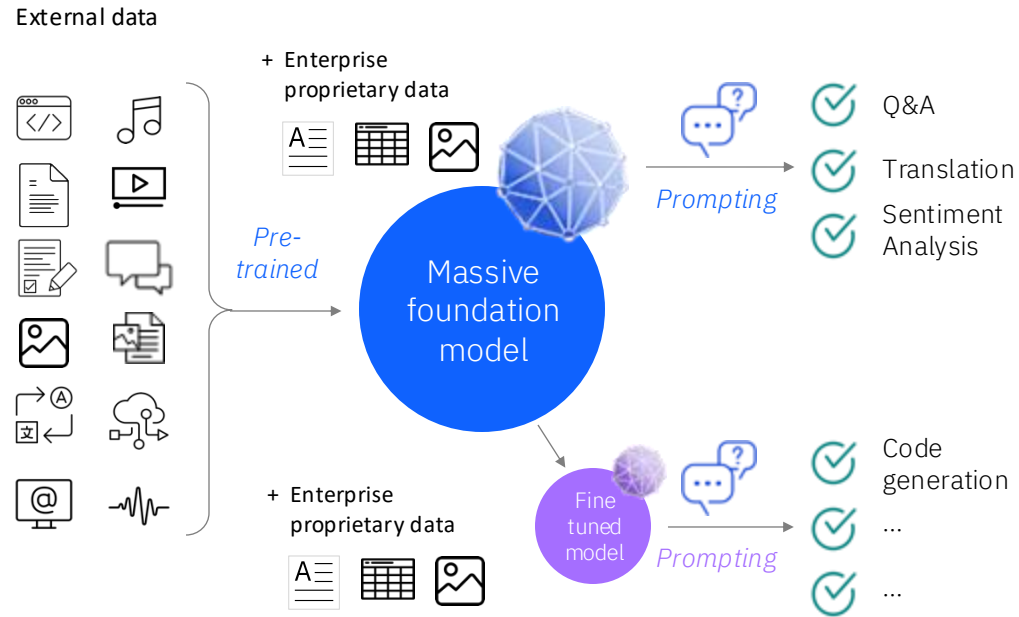
The impact of generative AI

Traditional AI models



- Individual siloed models
- Require task specific training
- Lots of human supervised training

Foundation models



- Massive multi-tasking model
- Adaptable with minimized training
- Pre-trained unsupervised learning
- Massive unlabeled data
- Self-supervision at scale

Enhanced capabilities

- Summarization
- Conversational Knowledge
- Content Creation
- Code Co-Creation

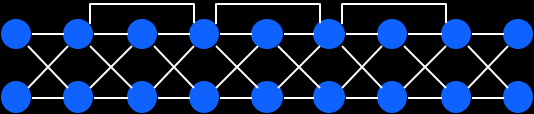
Key advantages

- Lower upfront costs through less labeling
- Faster deployment through fine tuning and inferencing
- Equal or better accuracy for multiple use cases
- Incremental revenue through better performance

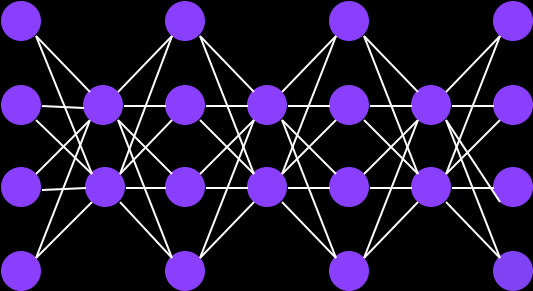
Data is the lifeblood of AI

Classical AI models: purpose-built and siloed

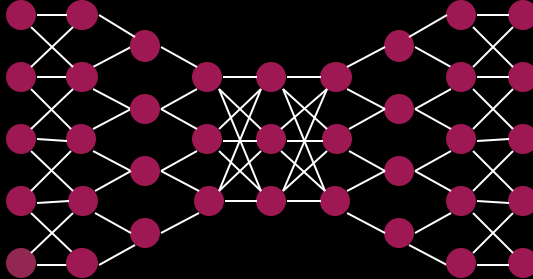
Translation



Summarization

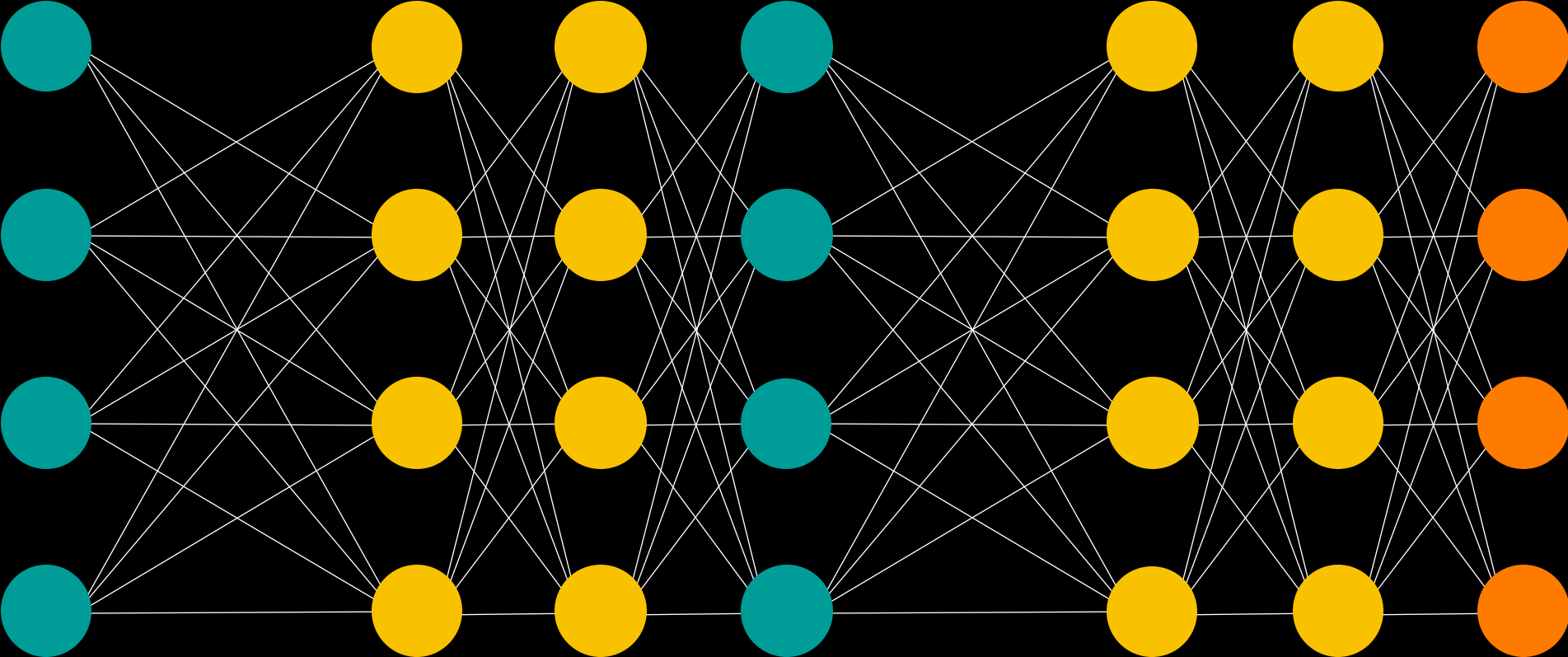


Question Answering



Each model performs a discrete task

Foundation models



Foundation model training: a lot of unlabeled data + a little labeled data

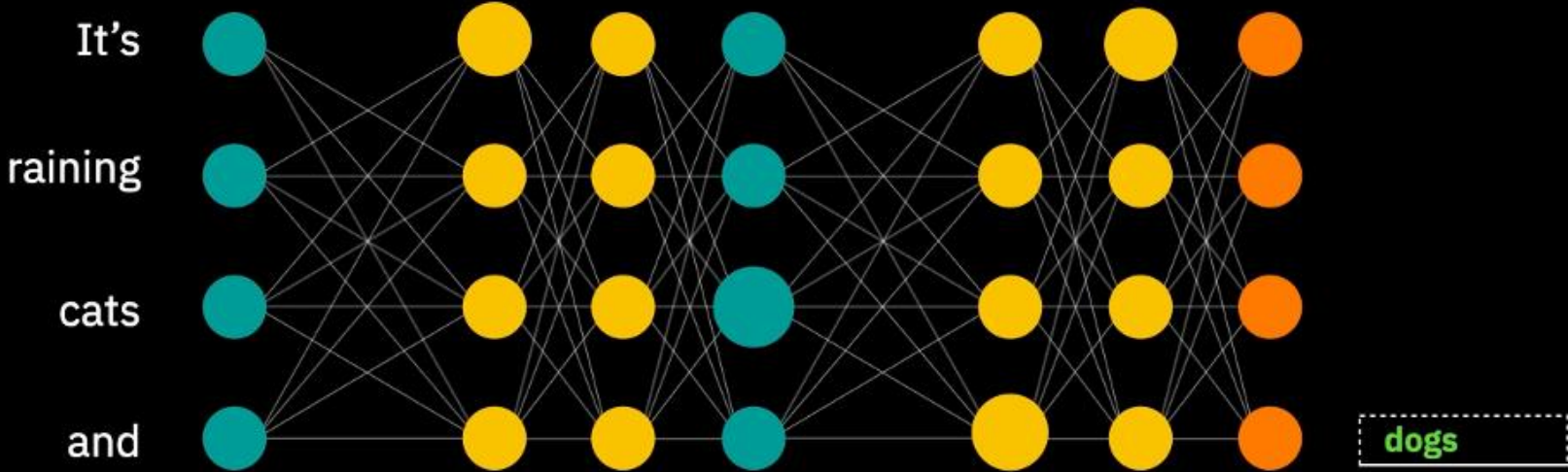


Unlabeled data for training:
100 billion sentences

Labeled data for fine tuning:
1000 examples



Training a foundation model: Self-supervision



Training a foundation model: Self-supervision

Apple announces M4 with more CPU cores and AI focus, just months after M3

Aggressive update schedule is a major departure for Apple Silicon.

by Andrew Cunningham - May 7, 2024 10:33 am



120GB/s Unified memory bandwidth

Tandem OLED display engine

Dynamic Caching

Up to

Hardware-accelerated mesh shading

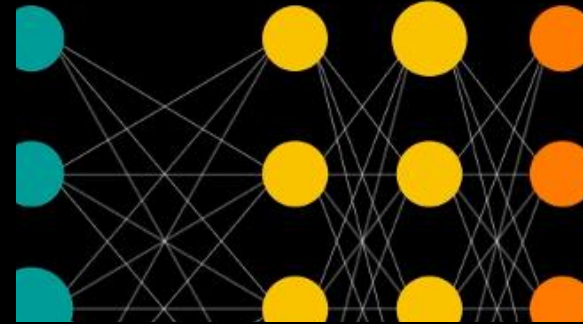
Hardware-accelerated ray tracing

Over 28 billion transistors

Second-generation 3 nm technology

Apple M4

Neural Engine with
38 trillion ops/sec



As with so much else in the tech industry right now, the M4 also has an AI focus; Apple says it's beefing up the 16-core Neural Engine (Apple's equivalent of the Neural Processing Unit that companies like Qualcomm, Intel, AMD, and Microsoft have been pushing lately). Apple says the M4 runs up to **38 trillion operations per second (TOPS)**, considerably ahead of Intel's Meteor Lake platform, though a bit short of the **45 TOPS** that Qualcomm is promising with the Snapdragon X Elite and Plus series. The M3's Neural Engine is only capable of **18 TOPS**, so that's a major step up for Apple's hardware.

Foundation model training: a lot of unlabeled data + a little labeled data



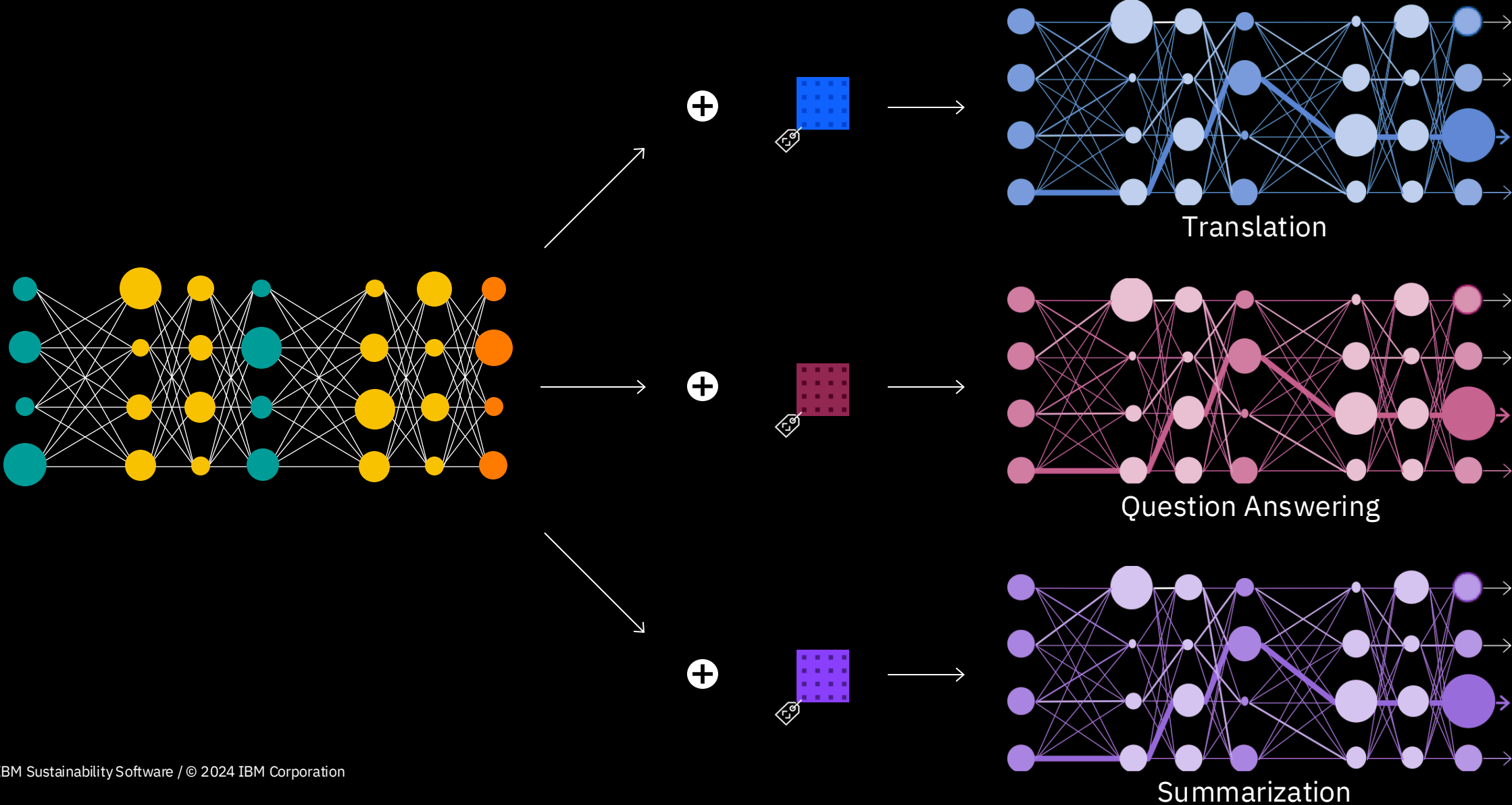
Unlabeled data for training:

100 billion sentences

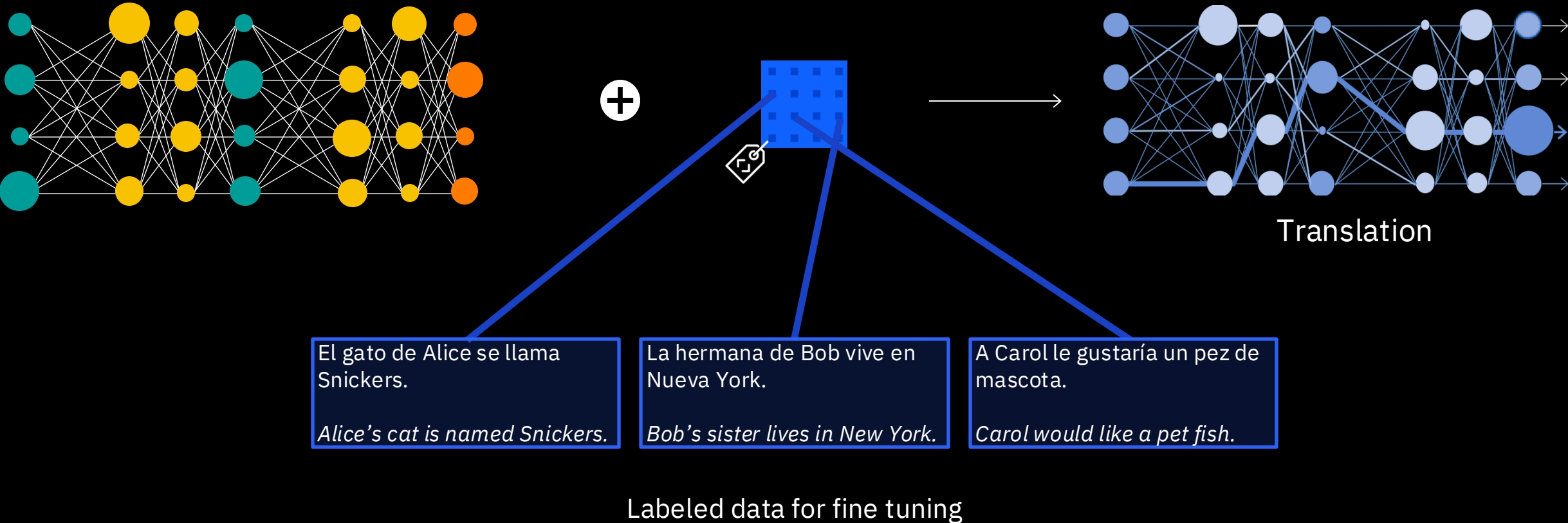
Labeled data for fine tuning:

1000 examples

Fine tuning: One trained Foundation Model for many tasks

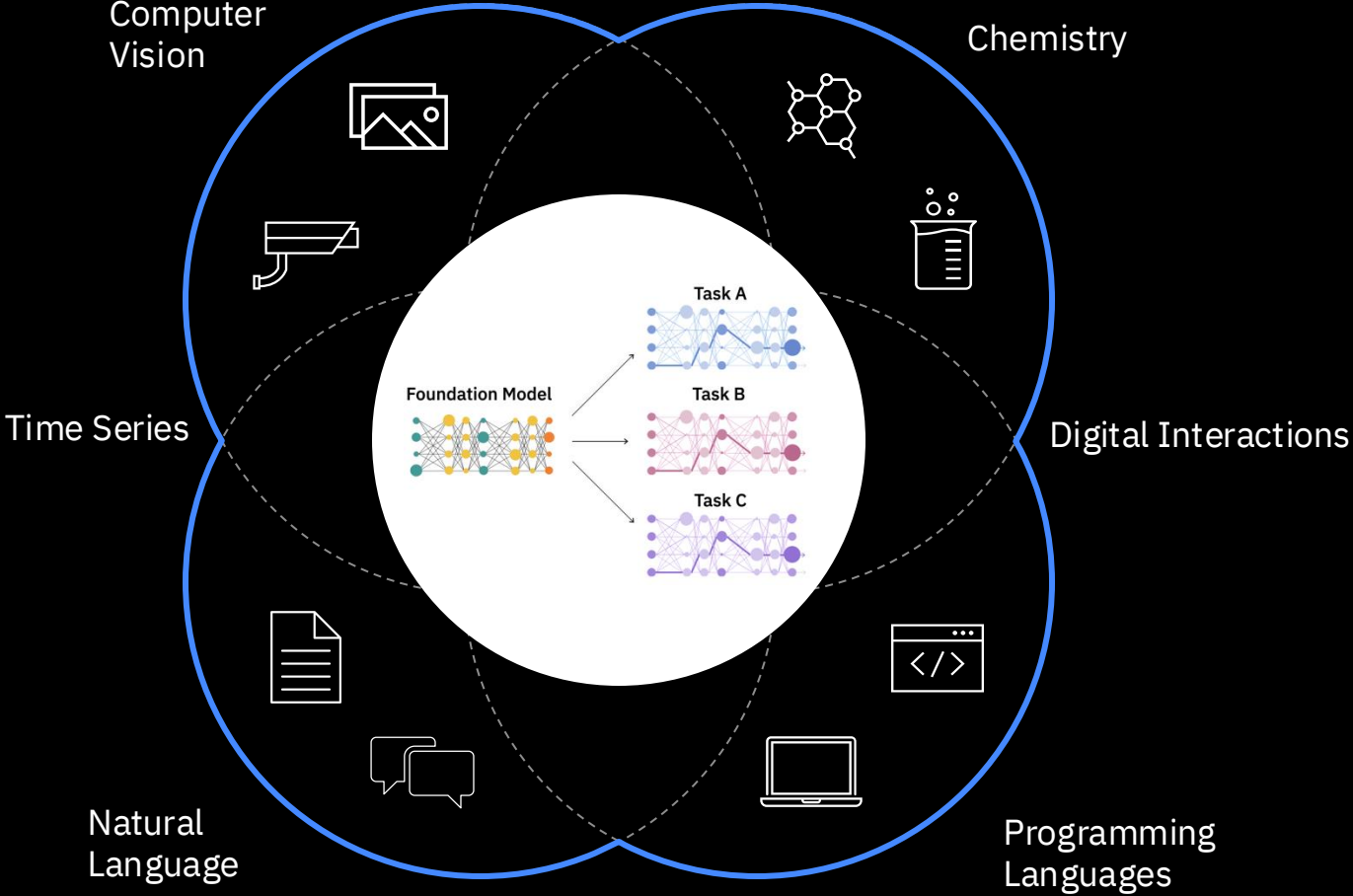


Fine tuning a trained Foundation Model



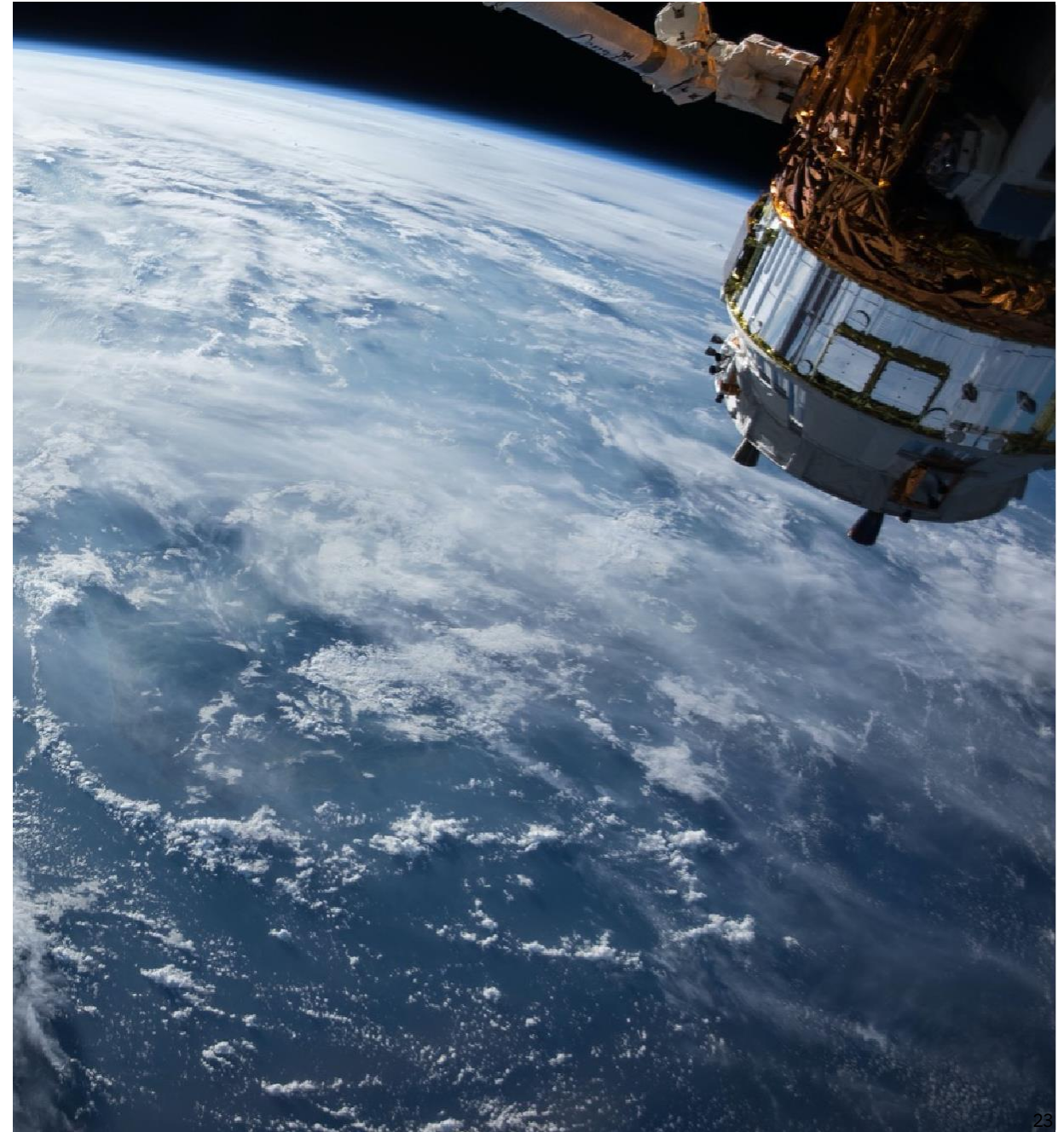
The use cases for
Foundation Models
go well beyond Large
Language Models (LLMs)

But the implications of foundation models go well beyond Large Language Models (LLMs)

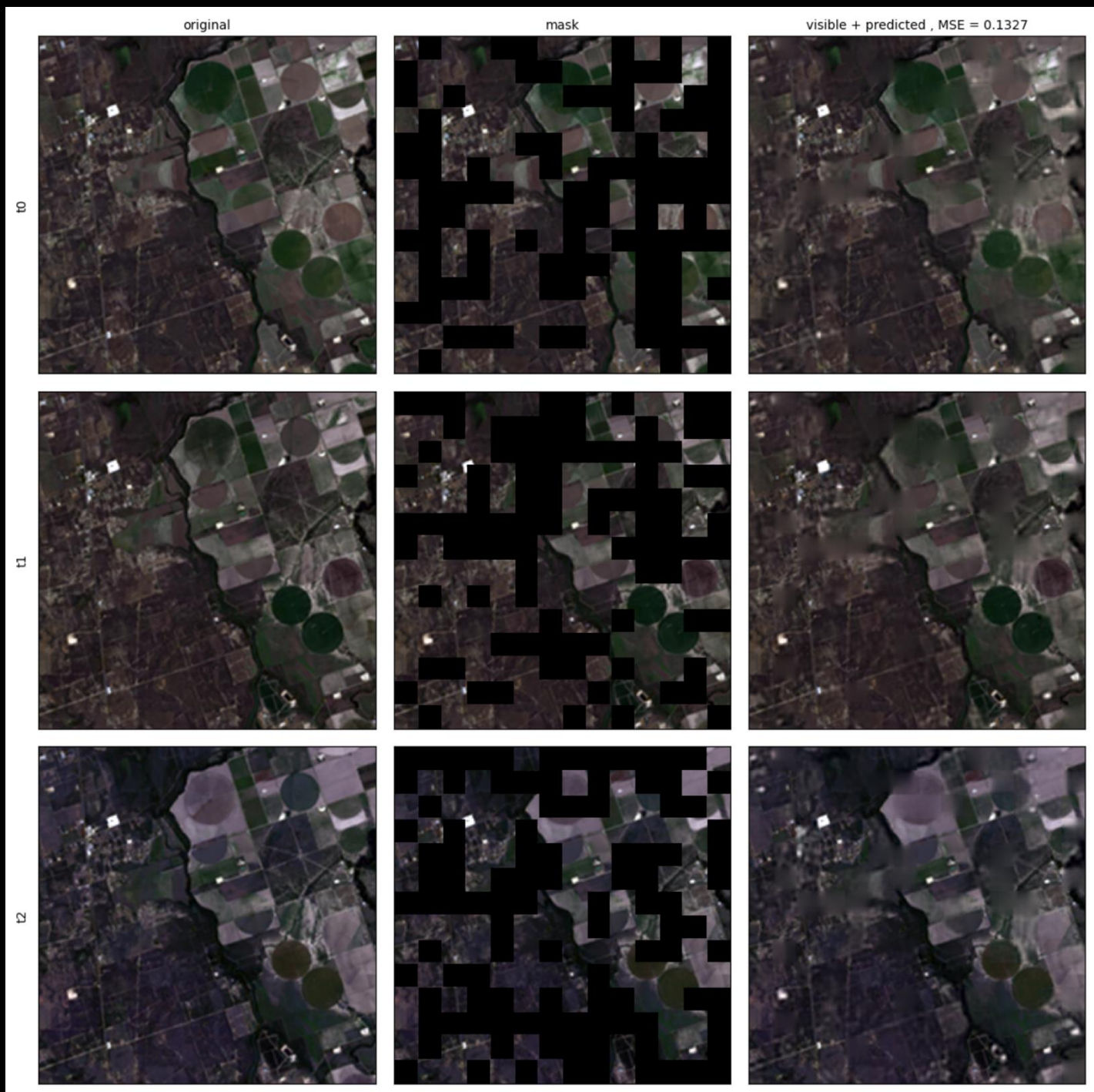


We collaborated with NASA to develop a Geospatial foundation model trained on HLS data.

The Harmonized Landsat Sentinel-2 (HLS) dataset provides global land observations every 2-3 days at 30 meter resolution.

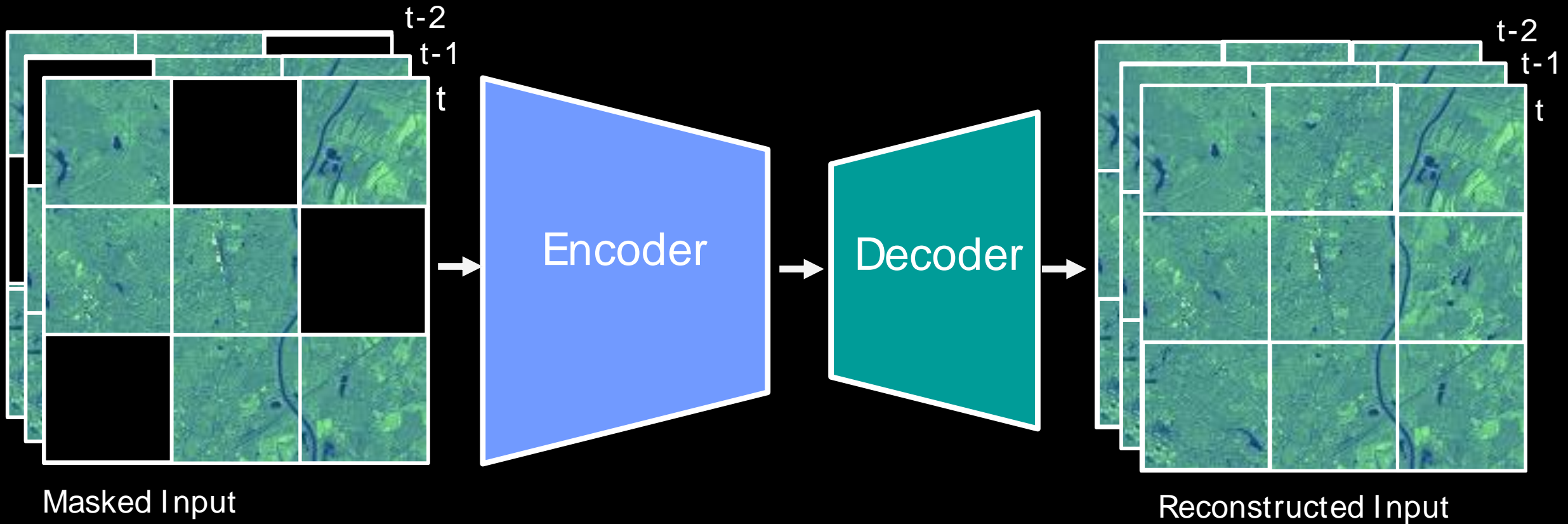


Geospatial Foundation Model Training Result

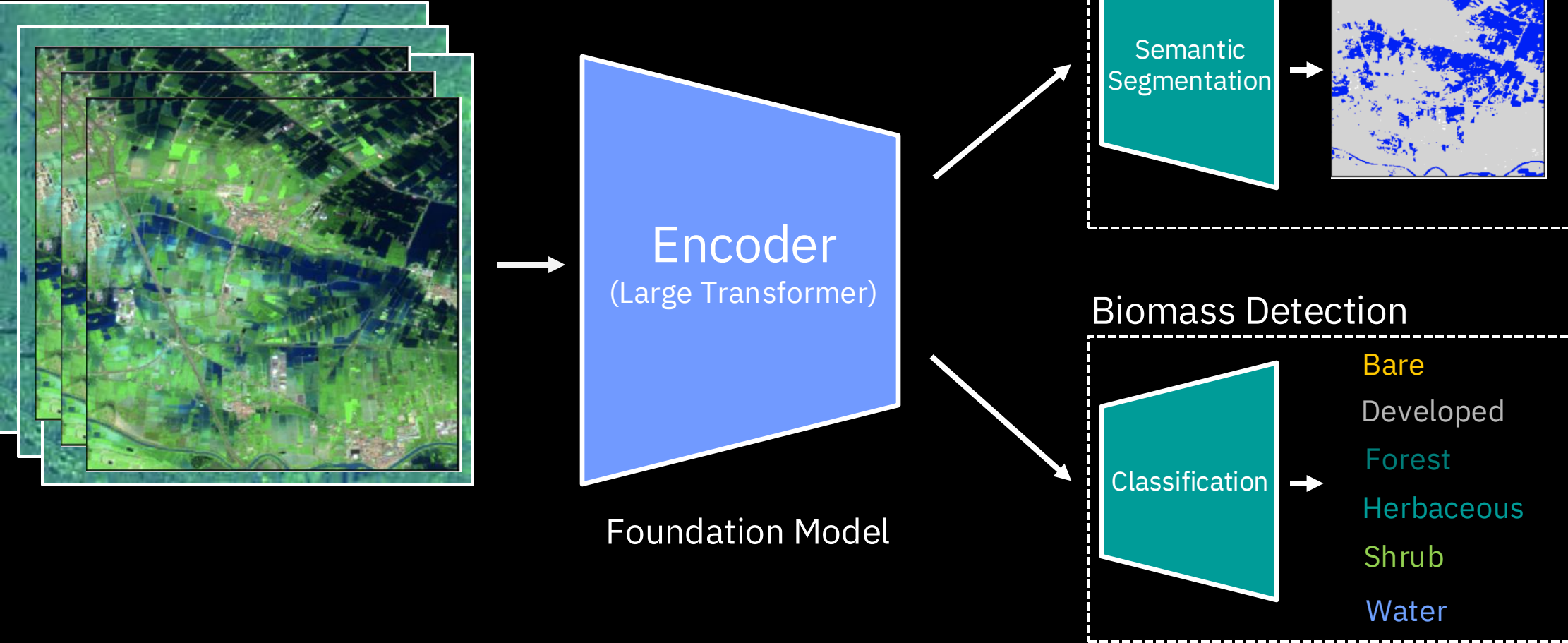


Training a foundation model: Self-supervision

Transformer Architecture consists of an Encoder & Decoder



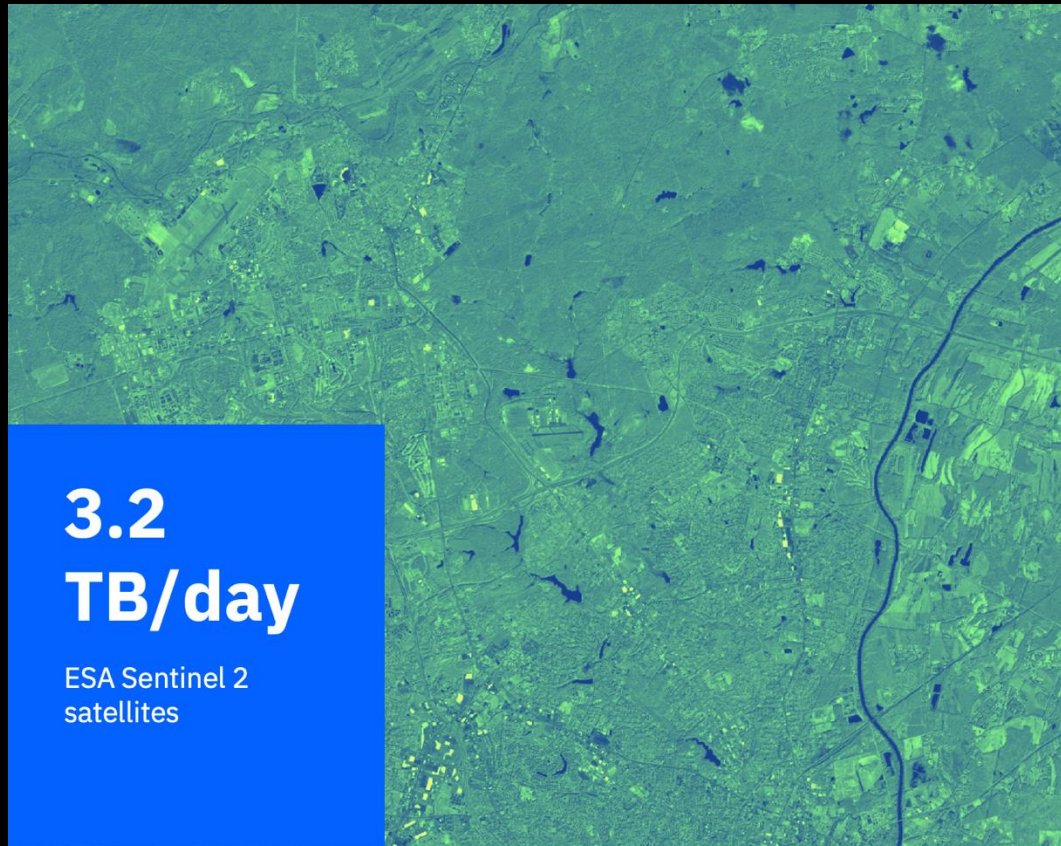
Fine tuning a trained Foundation Model



Two core types of geospatial data relevant for sustainability

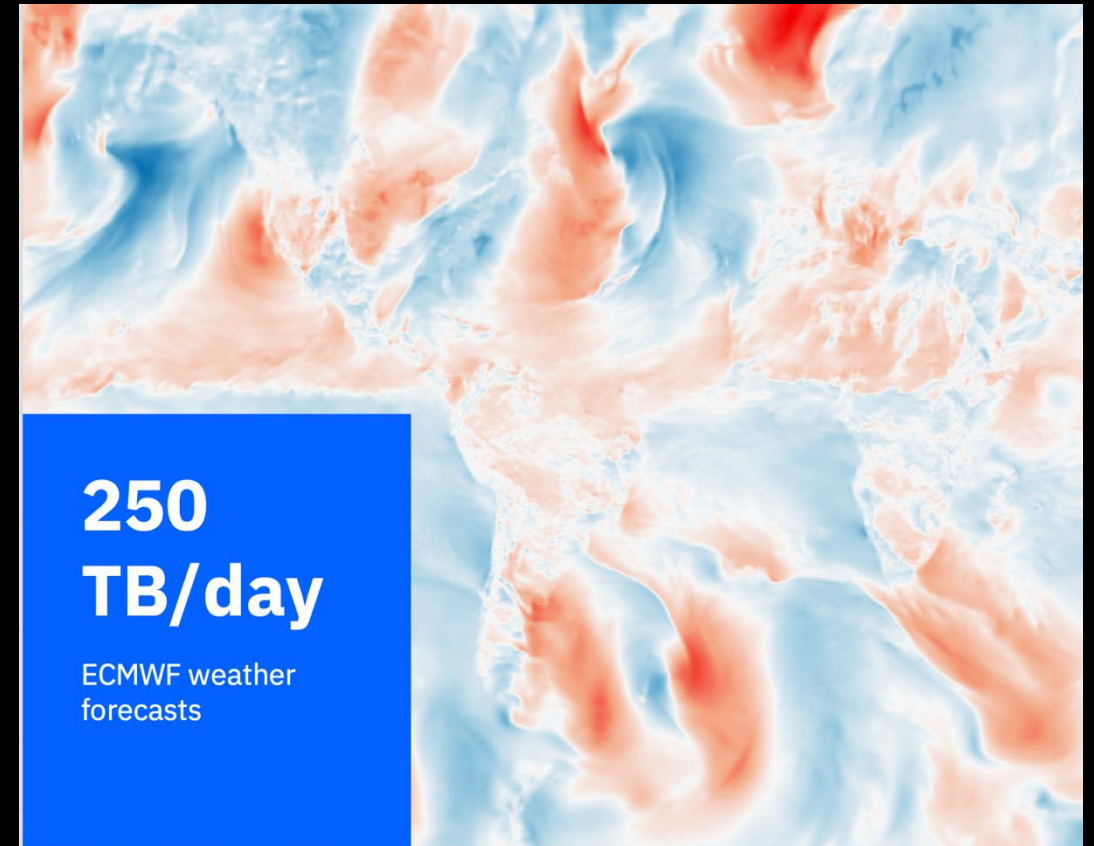
Satellite and aerial imagery

- Multimodal – images from multiple satellites representing different spectral bands



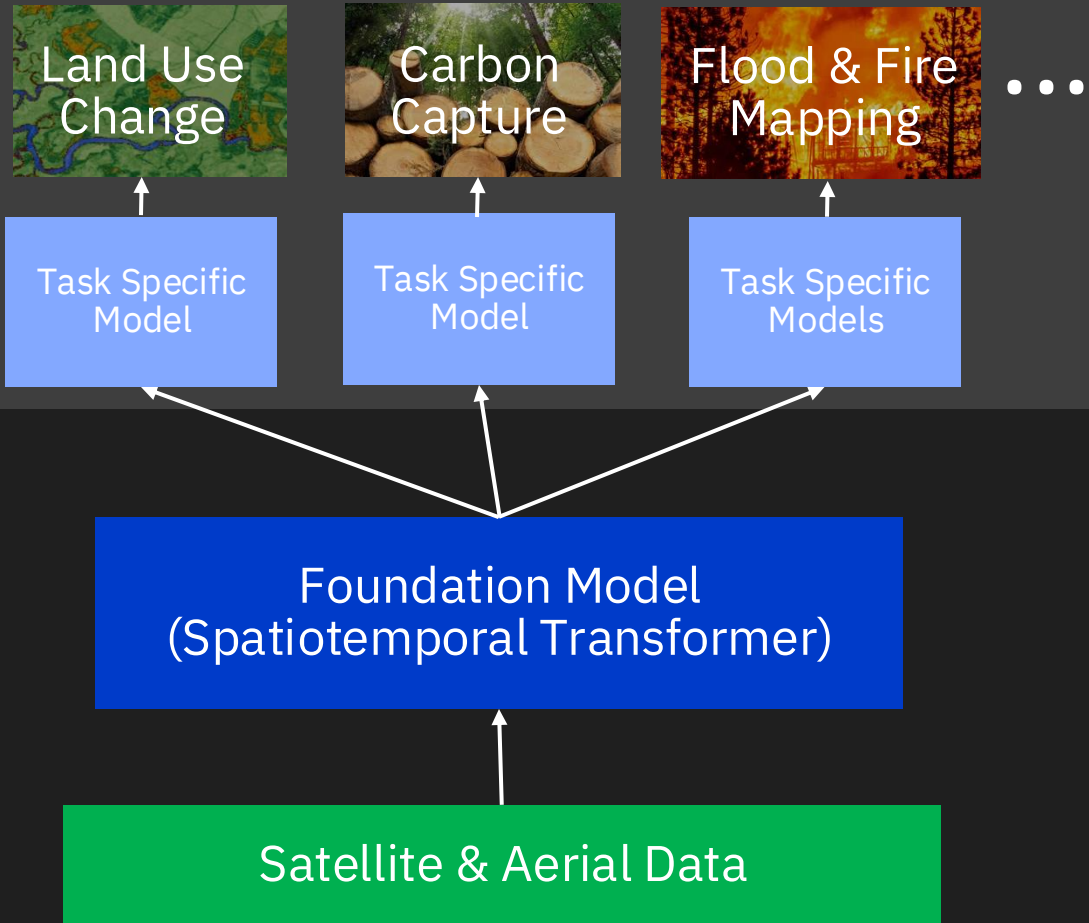
Weather measurements & forecasts

- Multimodal – time series from different processes (temperature, precipitation, wind,...)

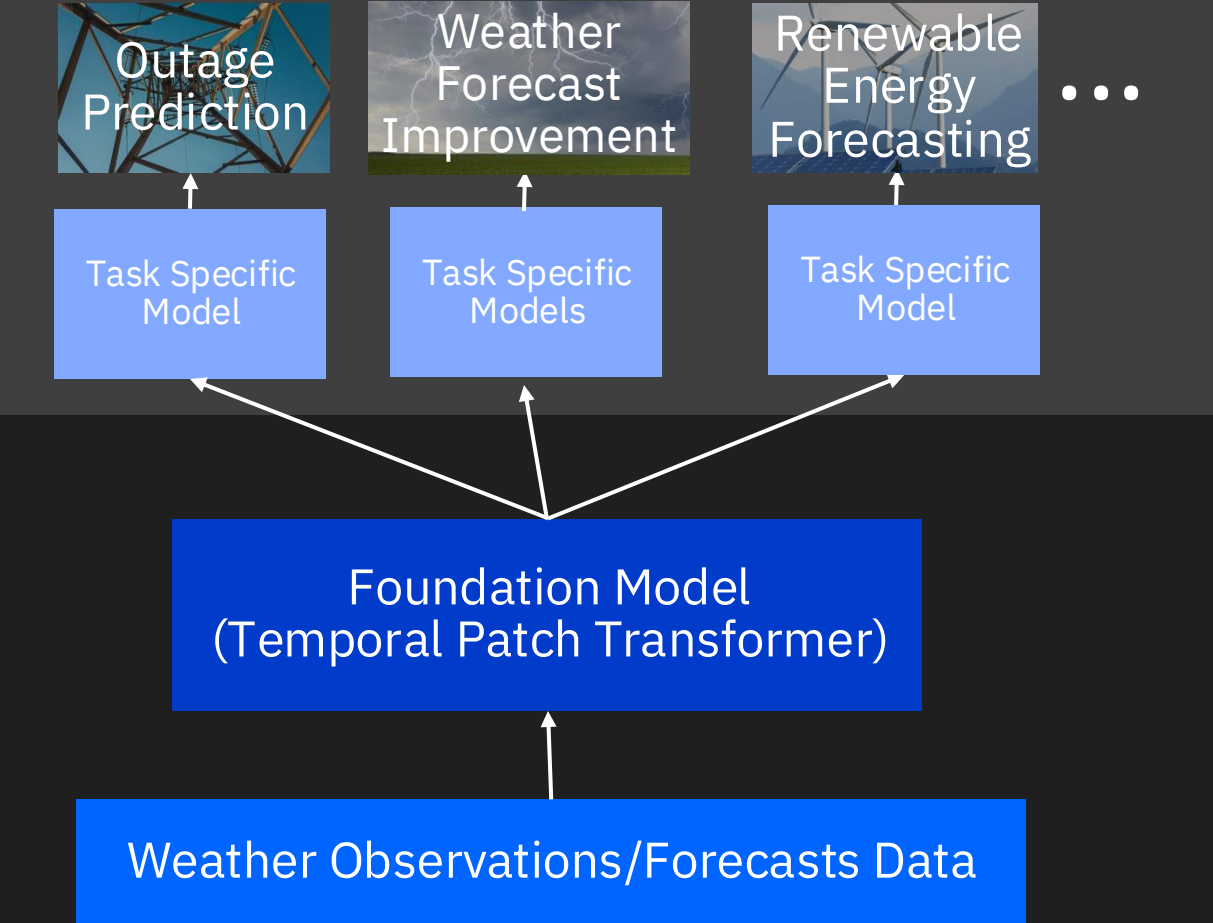


Foundation models for sustainability

Image Segmentation Use Cases



Weather Use Cases







Why NASA, IBM, And Hugging Face's Open Source Model Is A Big Deal

Ted Schadler, VP, Principal Analyst

What do you get when you combine an open source platform, a massive and critically useful dataset, and an ability to open-source an AI foundation model?

If you're NASA, IBM, and Hugging Face, you get a massive opportunity to make geospatial data available to all through an open source geospatial AI foundation model. We like this open source geospatial intelligence resource and commitment for three reasons:




This is also a great reminder to technology executives that you will incorporate many intelligences into your genAI-fueled applications. **Do not expect or plan to rely solely on a large language model from Microsoft or Google.** Most of the specific value will lie in these domain-specific genAI intelligences.

You will create real applications by orchestrating the intelligences you need (including your own knowledge foundation models and your machine-learning models and software).

Data is the lifeblood of AI


Agenda

- 
- How Foundation Models Work
 - GenAI in Sustainability Software (Maximo)
 - AI in Sustainability Software

Generative AI Opportunities: Applying foundation models in our Sustainability Software portfolio


In flight

Failure mode context understanding




We are training a GenAI model to understand failure points. FMEA (Failure Mode and Effects Analysis) data is not available for many kinds of assets, and it can take time to acquire. This model can then be applied to other assets where data does not exist.

Work order intelligence




We are training an LLM to classify and recommend work order codes, which will assist and (if desired) enable auto-approval of work orders overcoming human error and limitations.

Health: prediction & anomaly




We are using GenAI to create sensor-level models trained on a small sample of sensor data. This greatly speeds time to value for asset health prediction.

MVI anomaly detection



MVI AI currently can lack accuracy for “Few-Shot” anomaly detection. We are building a transformer-based foundation model that will fill these gaps.

MVI prompt tuning




We are using foundation models and prompt-tuning to dramatically reduce the effort required to train the model for MVI.

Above ground biomass measurement




Use Geospatial Foundation Models to measure above ground biomass for known land cover and vegetation species.

Flood and fire detection



Use geospatial model to detect fire tracks and floods leveraging data from NASA. Would enable stakeholders to quickly understand the impacts of flood and fire.

Activity based scope 3 estimation



Leverage LLM to ingest and understand invoice level data to categorize business activities and estimate carbon emissions generated by the business at a more granular level.


In pipeline

Assist: technician assistant




We are using a GenAI assistant trained on service manuals and work orders to give technicians the latest asset service info and generate repair workflows.

MAS onboarding assistant




We are training an LLM model to answer questions that arise in implementation that can guide clients about costs, effort and risks.

Envizi SRM assistant



The sustainability disclosure environment is complex. We hope to train a GenAI virtual assistant to help guide our users through the process.

Envizi sustainability assistant



Sustainability journeys can be complex and require digesting vast amounts of data. We hope to train a Gen AI model to assist our users in this journey.

Work Order Intelligence Problem Code Recommendations using watsonx

New GenAI feature recommends problem codes to assist with existing challenges faced by customers to address poor quality (or missing) work order data, impacting effective maintenance prioritization decisions as well as time spent on review and approval.

- AI model trained using long and short descriptions from a small set of related work orders
- Work Orders displays the AI recommendations for review/accept
 - Uses new IBM Design UI elements incorporated into Graphite
 - Provides confidence score of the predictions
- Work queue surfaces the top AI recommendation to accept
- Leverages watsonx

The screenshot displays the 'Edit work order' page in the IBM Maximo Application Suite. The main description field contains the text 'Compressor pump vibrations at high operational load'. A modal window titled 'Recommendation' is open, showing a table of suggested problem codes. The table has columns for 'Problem code', 'Description', and 'Confidence'. The top recommendation is 'VIB' with a description of 'Vibration' and a confidence score of 60%. Other recommendations include 'NOI' (Noise, 55%) and 'PLU' (Plugged/chocked, 50%). A 'Regenerate' button is visible in the modal. Below the modal, the 'Problem code' field is set to 'Unspecified', and a link for 'Recommendation + 2 more' is shown, with a sub-link for 'AI Vibration 60%'. The background form includes fields for 'Asset' (983763), 'Location' (BR430), and 'Failure class' (PUMP).

Problem code	Description	Confidence
<input checked="" type="radio"/> VIB	Vibration	60%
<input type="radio"/> NOI	Noise	55%
<input type="radio"/> PLU	Plugged/chocked	50%

Hide long description

Edit Insert Format

↓ B I U ↺ sans-serif

Work type

CM

Reported by

Edward Smith

Asset and location

Asset

983763

Location

BR430

Failure class

PUMP

Specify a failure class and then select a problem code.

Recommendation

Select a recommended problem code.

Regenerate ↺

Problem code	Description	Confidence
<input checked="" type="radio"/> VIB	Vibration	60%
<input type="radio"/> NOI	Noise	55%
<input type="radio"/> PLU	Plugged/chocked	50%

Cancel

Apply

Problem code

Unspecified

Recommendation + 2 more

AI Vibration 60%

Save

Cancel

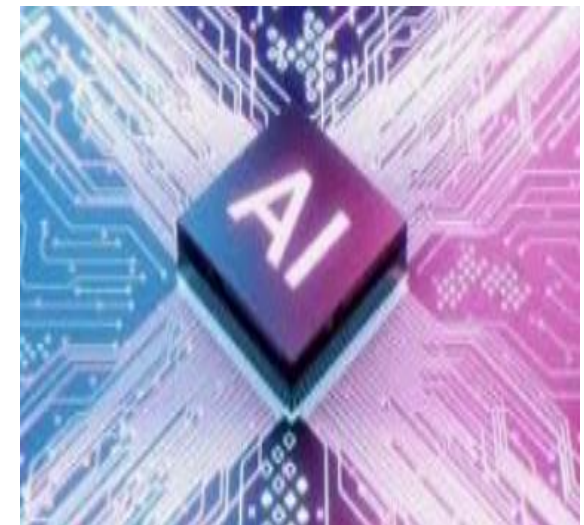
Work Order Intelligence Use Case

Problem

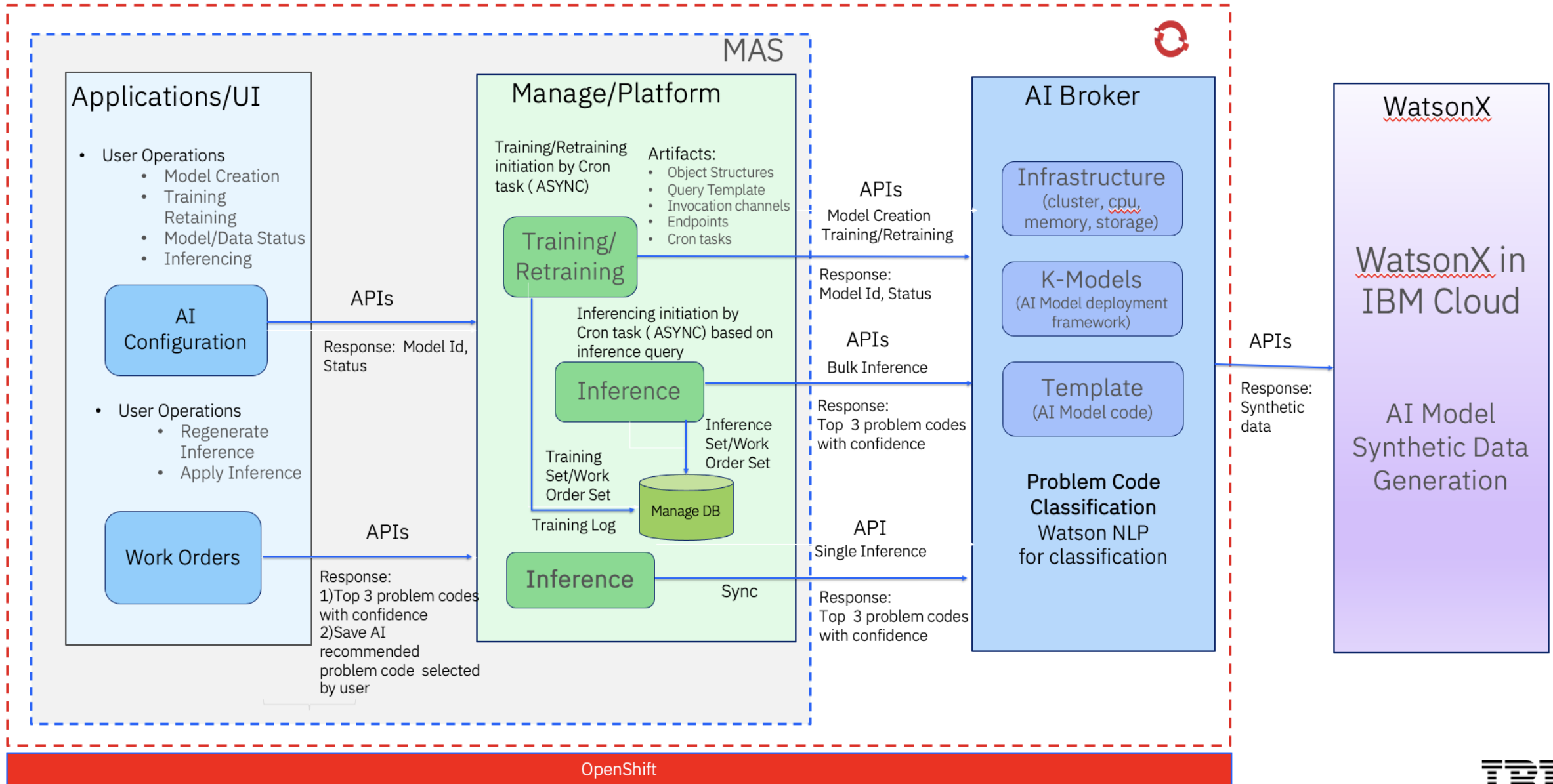
- Poor quality data impacts effective maintenance prioritization decisions as well as time spent on review and approval.

Solution


- **Automated Recommendations:**
Maintenance users will benefit from automated recommendations designed to assist in addressing issues related to poor-quality or missing work order data.
- **Maximo AI Integration:**
We will leverage new Maximo AI integration components to support use cases. These integrations will utilize Watsonx LLM Granite models to enhance data quality and streamline maintenance processes effectively.



Work Order Intelligence - Architecture



Agenda

- 
- How Foundation Models Work
 - GenAI in Sustainability Software (Maximo)
 - **AI in Sustainability Software**

Where is AI in Sustainability Software?

Route maps

Directions in Route planners

- Provide origin
- Provide destination
- Planner plots points between
- Shows route options

← from 1 Madison Ave, New York, NY 10010, USA to IBM Corporate Headquarters, 1 Orchard Rd, Armonk, NY 10504, United States

1 hr 13 min (36.2 miles) via I-87 N
Fastest route now, avoids congestion on Hutchinson Riv Pkwy N
⚠️ This route has tolls.

1 Madison Ave
New York, NY 10010, USA

- Get on FDR Dr from Madison Ave and E 42nd St
13 min (2.0 mi)
- Follow FDR Dr, I-87 N and I-287 E to NY-22 N/N Broadway in White Plains. Take exit 6 from I-287 E
34 min (27.8 mi)
- Continue on N Broadway to your destination in Armonk
14 min (6.4 mi)

IBM Corporate Headquarters
1 Orchard Rd, Armonk, NY 10504, United States

Map features: Petrol, EV charging, Things to do, Hotels, More. Live traffic, Fast/Slow indicator, Layers, Map data © 2024.

AI+ Offering Map

Offering	Business Imperative	AI Type	AI Method	AI Model	AI Capability	Application / Use Case
Above Ground Biomass	Sustainability Strategy and Roadmap ESG Data, Reporting and Risk Management Intelligent assets, facilities and infrastructure Responsible Computing and Green IT Sustainable supply chain and circularity	Decision Mgmt Interaction Deep Learning Machine Learning Generative/FM	Constraint Satisfaction Dynamic Programming Motion & Manipulation Optimization Perception Probabilistic Models Reinforcement Learning Rules Engine Self-Supervised Learning Simulation Supervised Learning Unsupervised Learning	Classification Closed Form Optimization Clustering Constraint Propagation Constraint Satisfaction Dimension Reduction Direct Policy Search Distributional Methods Dynamic Programming Linear Programming Mixed Integer Programming Nonlinear Programming Regression Search Simulation Structure Discovery Time Series Analysis Value Function Estimation	Anomaly Detection API Task Orchestration Association Rules Augmentation for Supervised Learning Computer Vision Data Augmentation Data Compression Factor analysis Failure Detection Forecasting and Prediction Fraud Detection Gradient Boosted Decision Trees Hierarchy Discovery Image Classification Image Generation Language Understanding Movement Planning Multi-objective Optimization Object Detection Pixel Segmentation Product Segmentation Recommendation Systems Relationship Discovery Sales Forecasting Scheduling optimization Spam Detection Text Generation Time Series Forecasting Time to Failure Video classification Video Generation Visualization	Action Recognition Anomaly removal from data for KPIs Anomaly detection from asset sensors Anomaly detection for doc flows Apportionment Asset Failure Probability Prediction Asset Failure Date Prediction Asset Failure Progression Prediction Business Milestone Interval Prediction Demand and Inventory Prediction Finds documents associated with a business transaction flow Fulfillment optimization Image Classification for assets Inventory Optimization Object Detection for assets Power Outage Forecasting Product Requirements Quality Analysis Product Usage forecasting Q & A Assistant on client data Scheduling optimization Scope 3 emissions estimation Vegetation proximity to Power Lines Visual Change Detection for Satellite Imagery Visual Defect Detection for assets
EIS Outage Prediction						
EIS Thematic Change Maps						
EIS Vegetation Mgmt						
ELM RQA						
Envizi						
Maximo Assist						
Maximo Monitor						
Maximo Predict						
Maximo Scheduler						
Maximo Visual Insights						
MRO IO						
SCIS Research Asst						
SCIS Watson Asst						
Sterling BTI						
Sterling BTI Doc Corr						
Sterling FO						
TRIRIGA Building Insights						

A black and white photograph of a man in a tank top looking upwards. The background is dark, and several glowing lightbulbs are suspended from the top. The text 'INSIGHTS' is written in white, bold, sans-serif font, tilted upwards, positioned above the man's head. To the left of the man, the words 'ANSWERS' and 'IDEAS' are written in white, bold, sans-serif font, tilted upwards. To the right of the man, the words 'QUESTIONS' and 'FEEDBACK' are written in white, bold, sans-serif font, tilted upwards. The overall composition suggests a cycle of thought and learning.

INSIGHTS

ANSWERS

IDEAS

QUESTIONS

FEEDBACK

謝謝
 DZIĘKUJĘ CI
 NGIYABONGA
 TEŞEKKÜR EDERİM
 DANKIE
 TERIMA KASIH
 SPASIBO
 ПАСИБО
 GRAZIE
 ПАРМЕТ СИЗГЕ
 GO RAIBH MAITH AGAT
 БЛАГОДАРЯ
 GRACIAS
 МАХАДСАНИД
 ТИ БЛАГОДАРАМ
 TAK DANKE
 RAHMAT
 HATUR NUHUN
 PAKKA PÉR
 CẢM ƠN BAN
 WAZVIITA
 FALEMINDERIT
 TAPADH LEIBH
 KEA LEBOHA
 БАЯРЛАЛАА
 MISAOTRA ANAO
 WHAKAWHETAI KOE
 DANKON TANK TAPADH LEAT
 SALAMAT
 MATUR NUWUN
 ХВАЛА ВАМ
 MULȚUMESC
 GRAZIE
 고맙습니다
 SHUKRA
 HVALA
 FAAFETAI
 ESKERRIK ASKO
 HVALA
 TEŞEKKÜR EDERİM
 OBRIGADO
 DANKJE
 EΥΧΑΡΙΣΤΩ
 GRATIAS TIBI
 AČIŪ
 SALAMAT
 MAHALO IĀ 'ŌE
 TAKK SKALDU HA
 MERCİ
 DI OU MÈSI
 ĎAKUJEM
 GRAZZI
 PAKKA PÉR
 SIPAS JI WERE
 TERIMA KASIH
 UA TSAUG RAU KOJ
 TI БЛАГОДАРАМ
 СИПОС

IBM