AI Foundation Models for Earth Sciences
The Biodiversity Use Case

Athanasios Trantas

Dutch Organization for Applied Scientific Research – TNO

Advanced Computing Engineering

SURF Advance Compute User Days – 12/12/2024

BIODT
biodiversitydigitaltwin

TNO innovation for life

Agenda

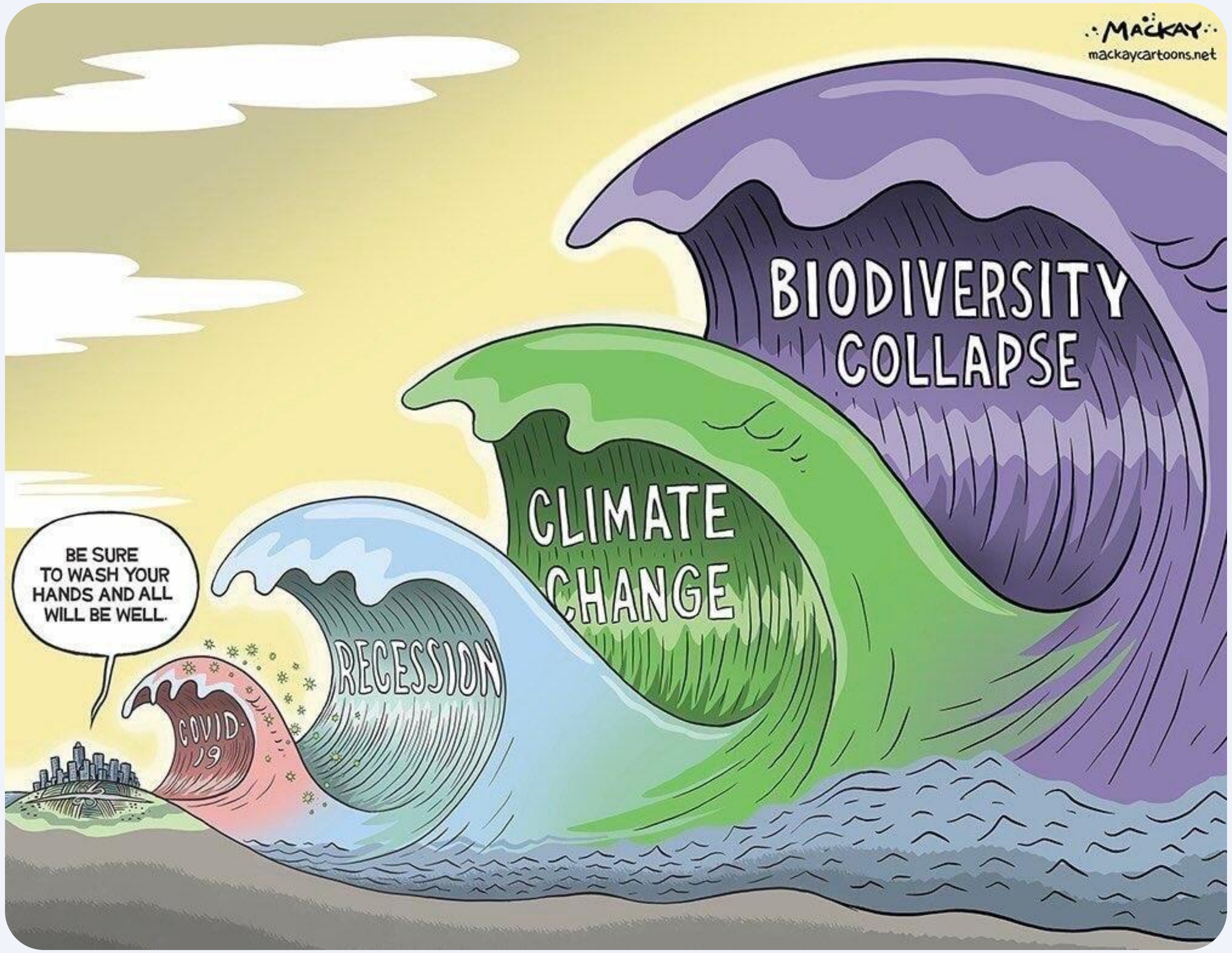
- Introduction
- Data
- Model
- Next steps
- Live Discussion

INTRODUCTION

Introduction - Biodiversity

The **resilience** of our **ecosystem** depends on the intrinsic interdependence between all the different **species** and their interactions

- Is vital for the prosperity of all life on earth
- Under threat due to broad anthropogenic changes
- Significant impact on Earth’s ability to sustain life

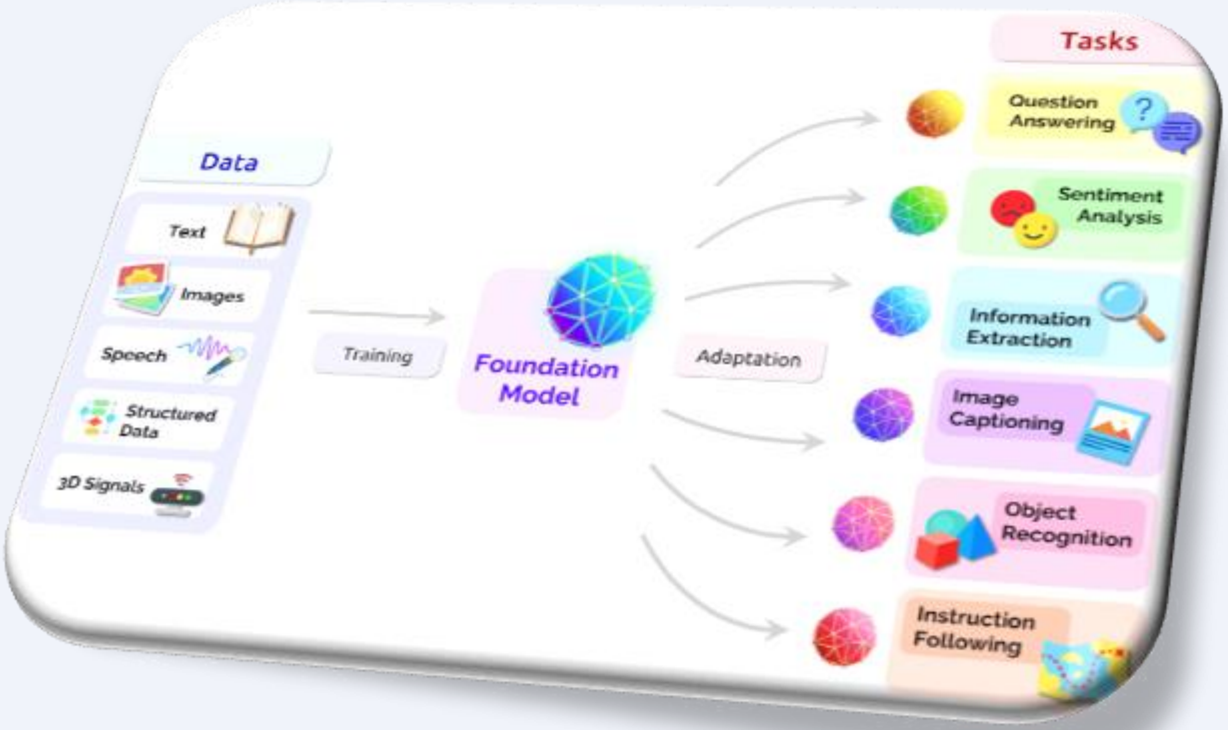


Trantas, Athanasios, et al. "Digital twin challenges in biodiversity modelling." *Ecological Informatics* (2023): 102357.

Introduction – Foundation Models

Foundation Models describe a broad paradigm shift in AI that encompass models like:

- pre-trained models
- self-supervised models
- large language models
- language vision models
- general purpose models
- multi-purpose models
- and task-agnostic models



Bommasani, Rishi, et al. "On the opportunities and risks of foundation models." *arXiv preprint arXiv:2108.07258* (2021).

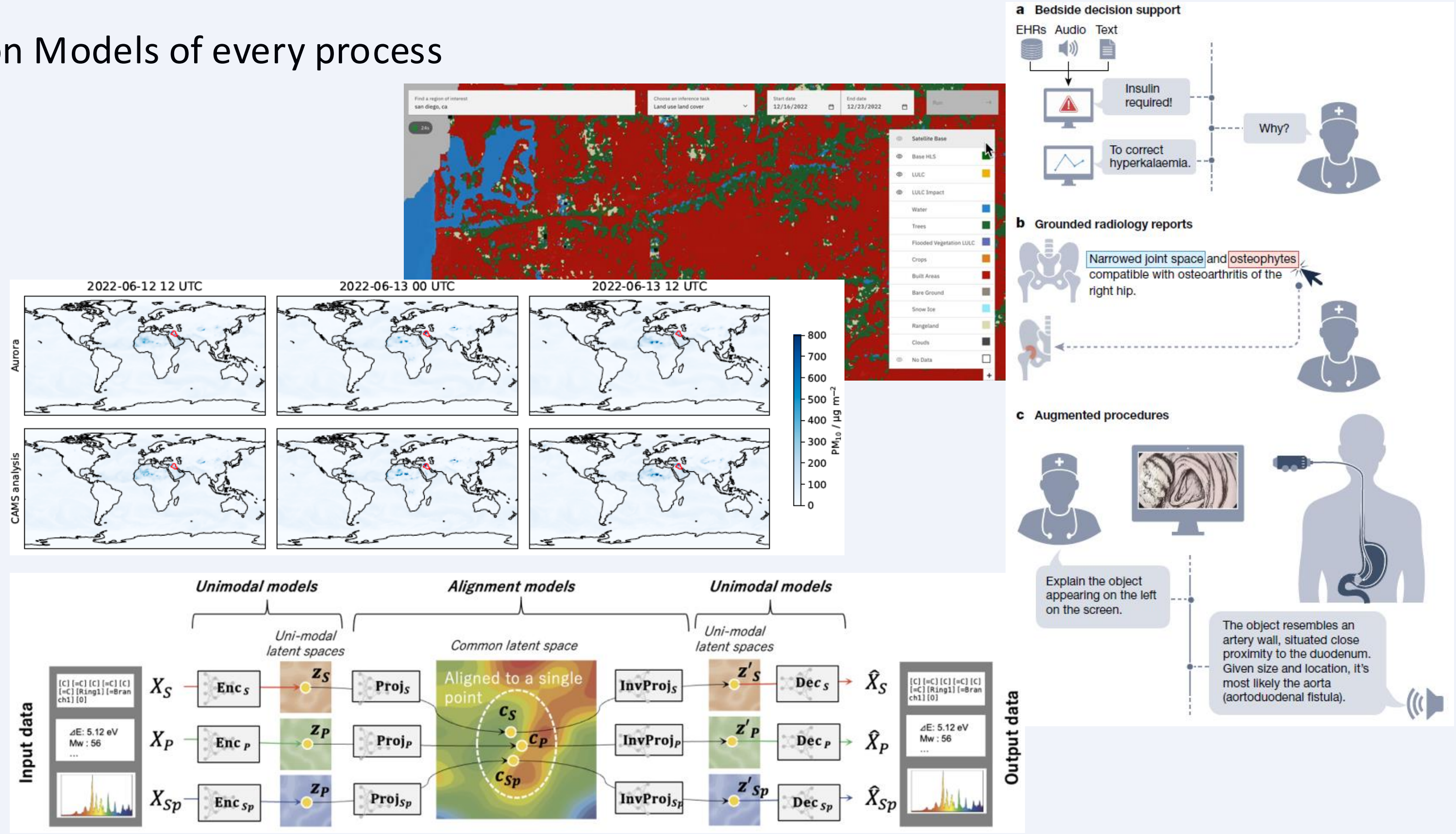
A Foundation Model is an AI Neural Network — trained on mountains of raw data, generally with unsupervised learning — that can be adapted to accomplish a broad range of tasks. The term "foundation" connotes the significance of architectural **stability**, **safety** and **security.**

Foundation Model examples

All the AI community is striving to create Foundation Models of every process

Examples

- Foundation Model of Earth [1]
- Foundation Model of Atmosphere [2]
- Foundation Model for Medicine [3]
- Foundation Model for Material Science [4]



Also:

There are big advancements on <u>how</u> to **train** and do **inference** on Foundation Models, while the hardware is continuously improving to handle **larger** and larger models that encompass more and more **modalities [5]**.

Why we need a Foundation Model for Biodiversity?

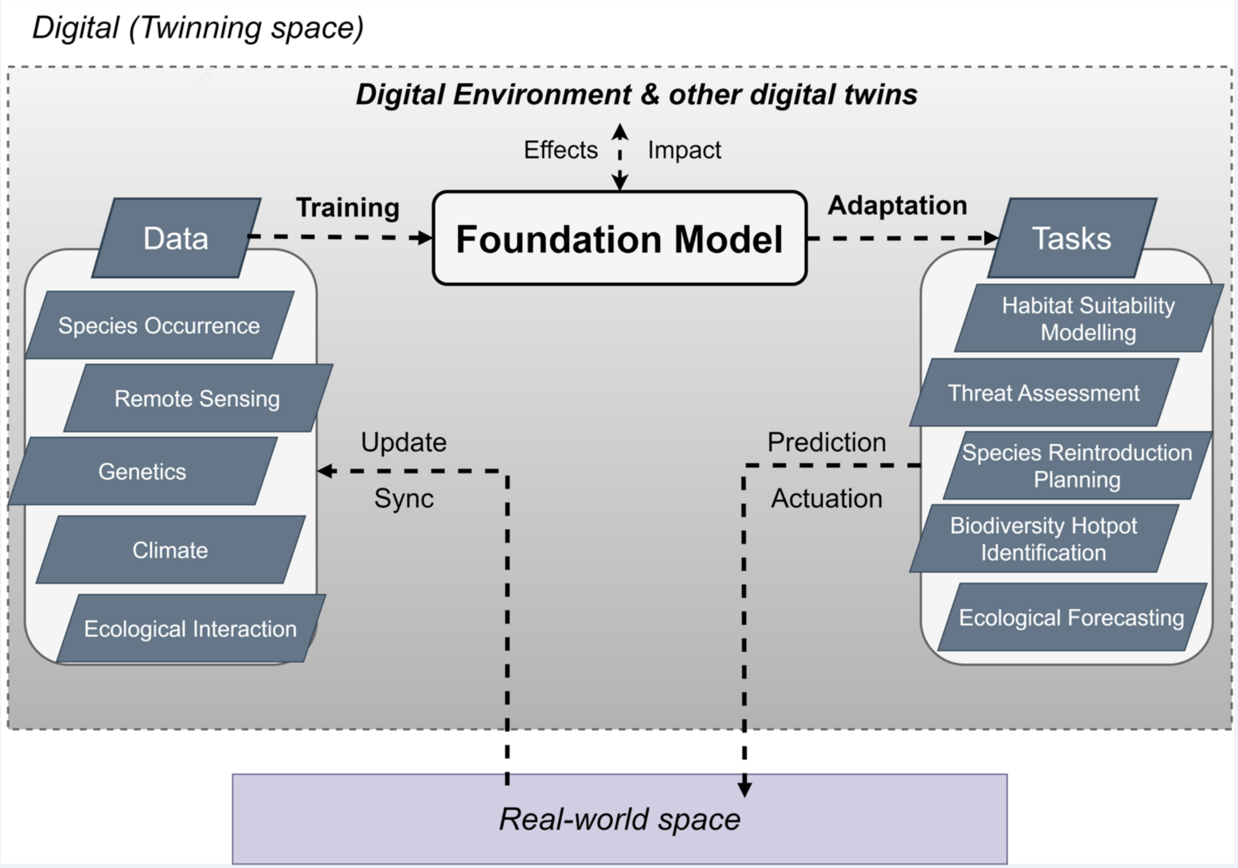
Challenges

- Diversity
- Multiple-modalities
- Complex Ecological Systems
- Scalable inter-model coordination
- Real time processing, monitoring and adaptation
- Uncertainties

Credits ESA 2023

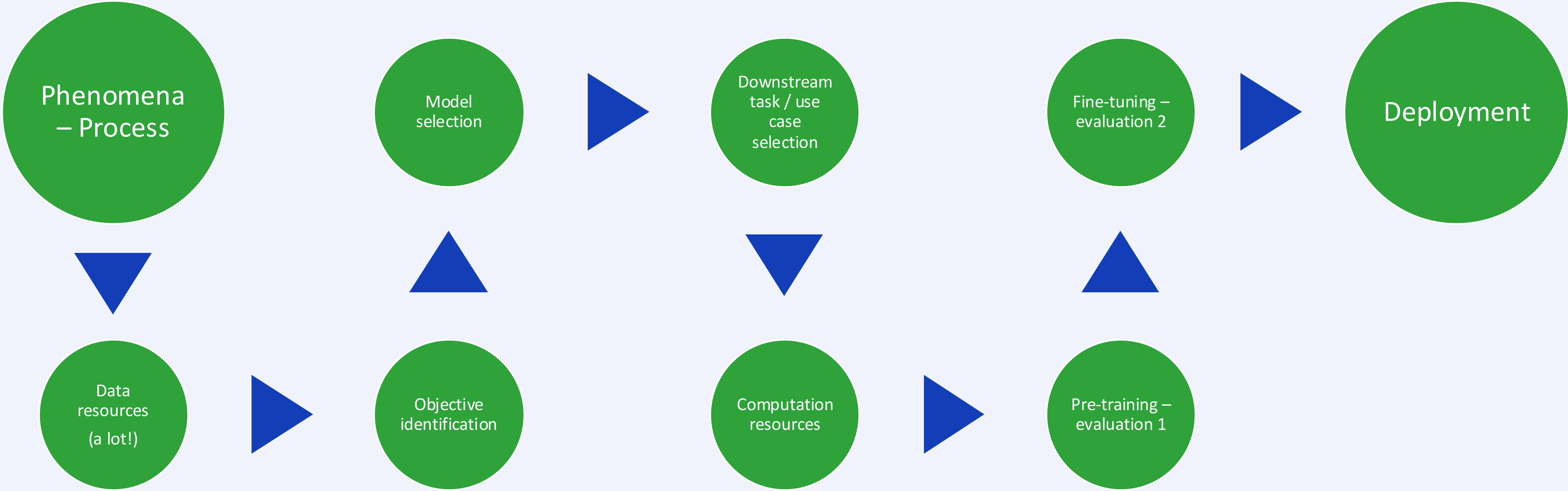
Advantages of a Foundation Model for Biodiversity

- Enhance Data Harmonisation and Analysis
- Sophisticated Modelling of Complex Ecological Systems
- New ways for User Interaction



Trantas, Athanasios, and Paolo Pileggi. "Digital Twin and Foundation Models: A New Frontier." *ICAART (3)*. 2024.

Foundation Model Development 101



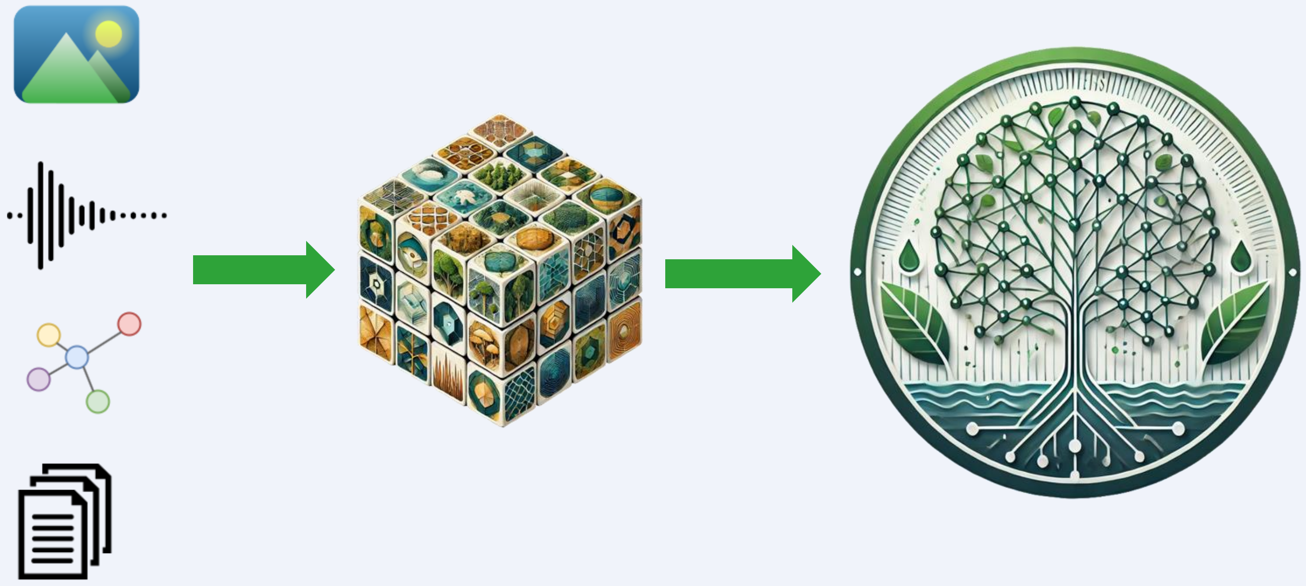
Biodiversity Foundation Model (BFM) Abstract View

Data Modalities

Data Fusion

Foundation Model Pre-training

Foundation Model Task Adaptation / Fine-tuning



- Species Distribution Modelling
- Pollination
- Grassland Dynamics
- Ecological Services
- Invasive Species
- Crop Wild Relatives

...

BIODIVERSITY DATA

SURF

TNO innovation for life

BIODT biodiversitydigitaltwin

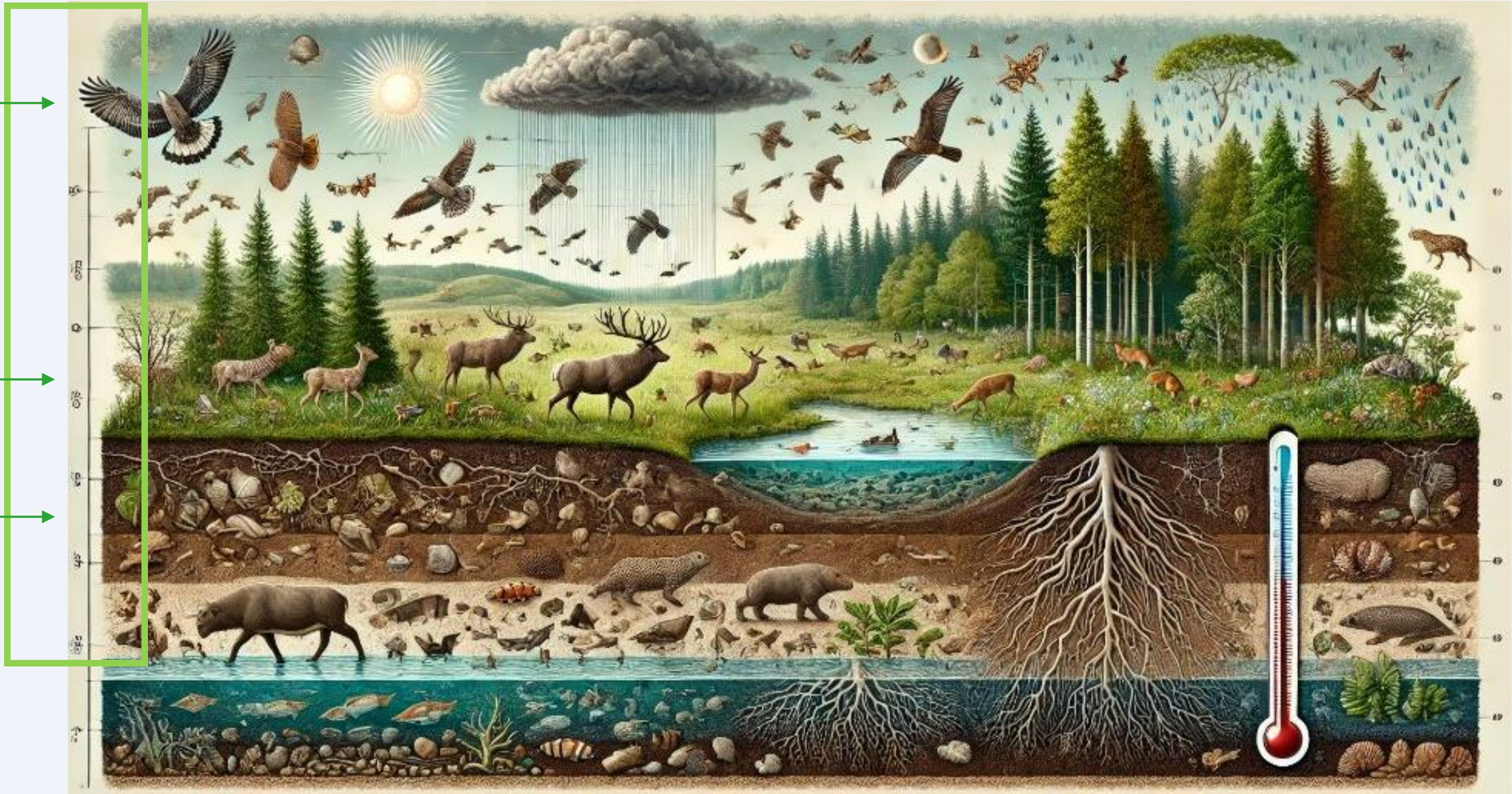
Data Introduction

Data Layers

1) Above surface

2) Surface

3) Below surface



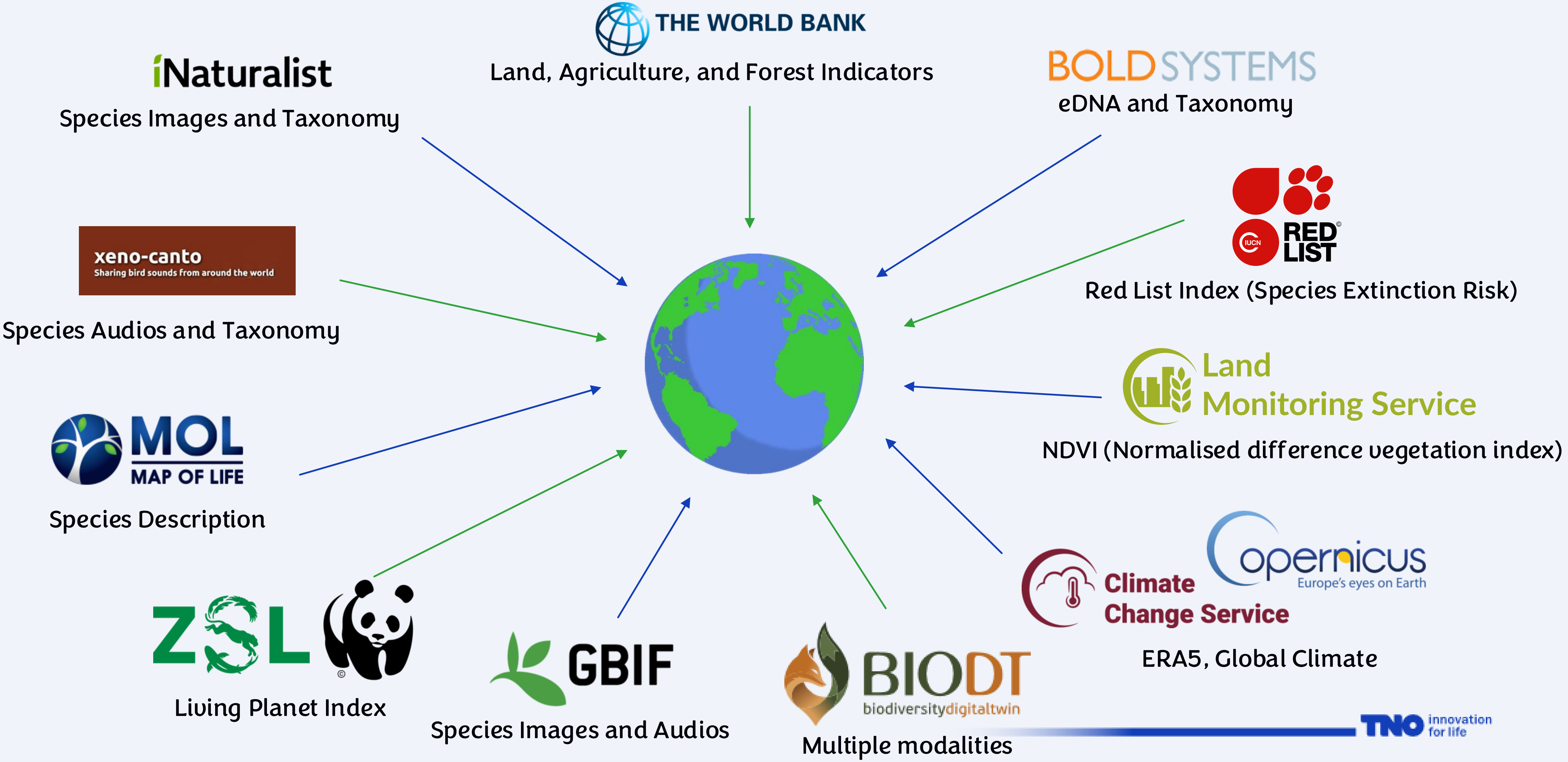
(Data) Modality

A modality, in the context of data, refers to a specific type or form of data that is captured, processed and analysed. Each modality represents a different way of encoding information, often utilising distinct sensors or data collection methods.

Most common Modalities for Biodiversity

- Imagery (photographs and videos)
- Audio recordings
- Genomic Data, Taxonomy
- Environmental DNA (eDNA)
- Satellite and Remote Sensing Data
- Geospatial Data (GPS coordinates, GIS layers)
- Climate Data (Temperature, Precipitation, Atmosphere)
- Textual Data (Scientific reports, Citizen science observations)
- Sensor Data (Environmental sensors for humidity, pH, etc.)

Data Sources



Species Data

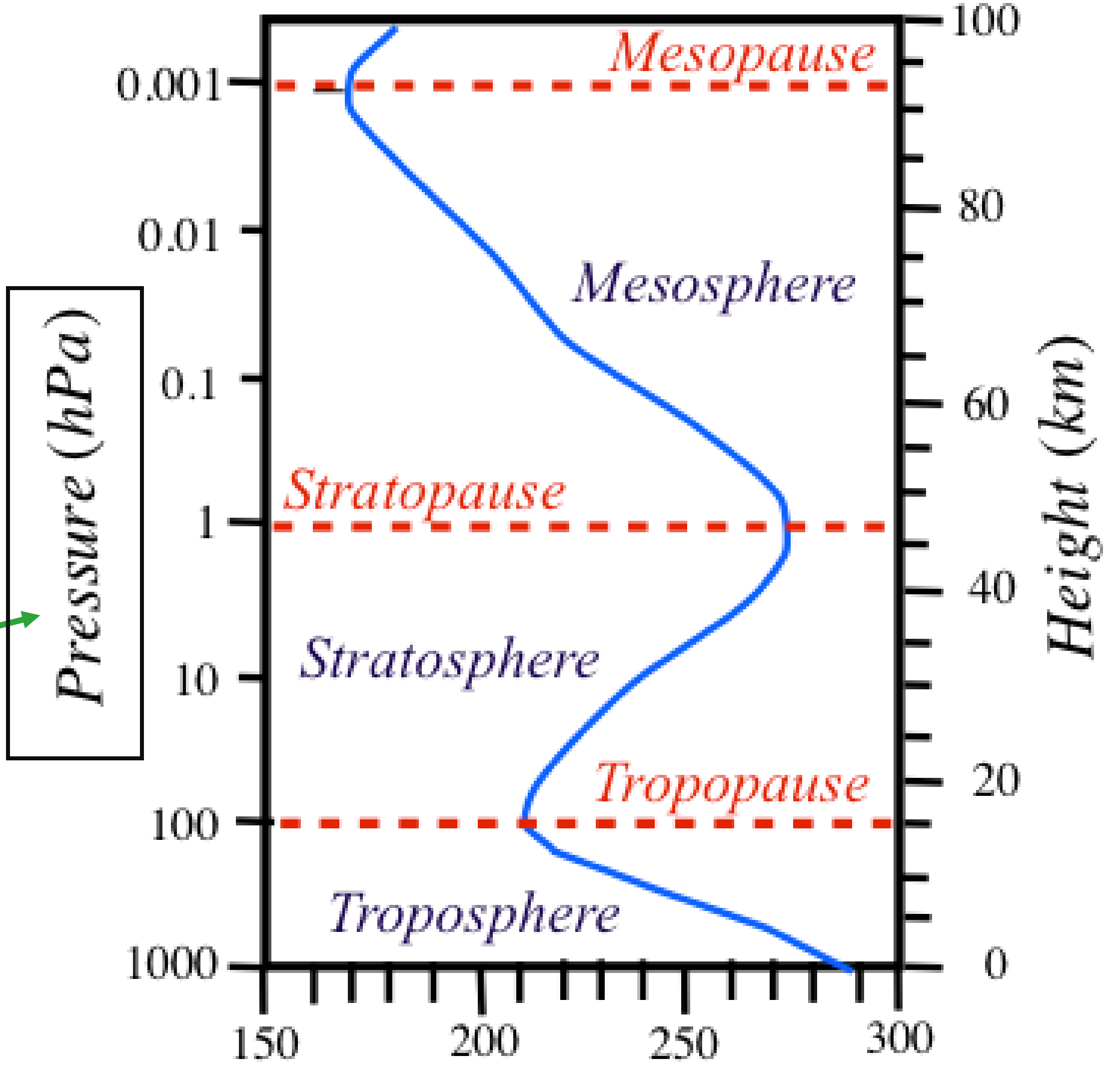
- Images (jpg)
- Audios (wav)
- Description (text)
- eDNA (ex. AACCCTATATCTAGTATTTGGC…)
- Distribution (Living Planet Index)
- Red List (ex. CE, EV, etc.)
- Taxonomy (text)
  - Phylum, Class, Order, Family, Genus
- Timestamp, Latitude, Longitude

ERA5 Data - Surface

- Surface :
  - 10m u-component of wind
  - 10m v-component of wind
  - 2m temperature
  - Mean sea level pressure
- Single :
  - Geopotential
  - Land sea mask
  - Soil type

ERA5 Data - Atmosphere

- Atmosphere :
  - Geopotential
  - Temperature
  - u-component of wind
  - v-component of wind
  - Specific Humidity
- Pressure Levels :
  - 50, 100, 150, 200, 250, 300, 400
  - 500, 600, 700, 850, 925, 1000



World Bank - Indicators

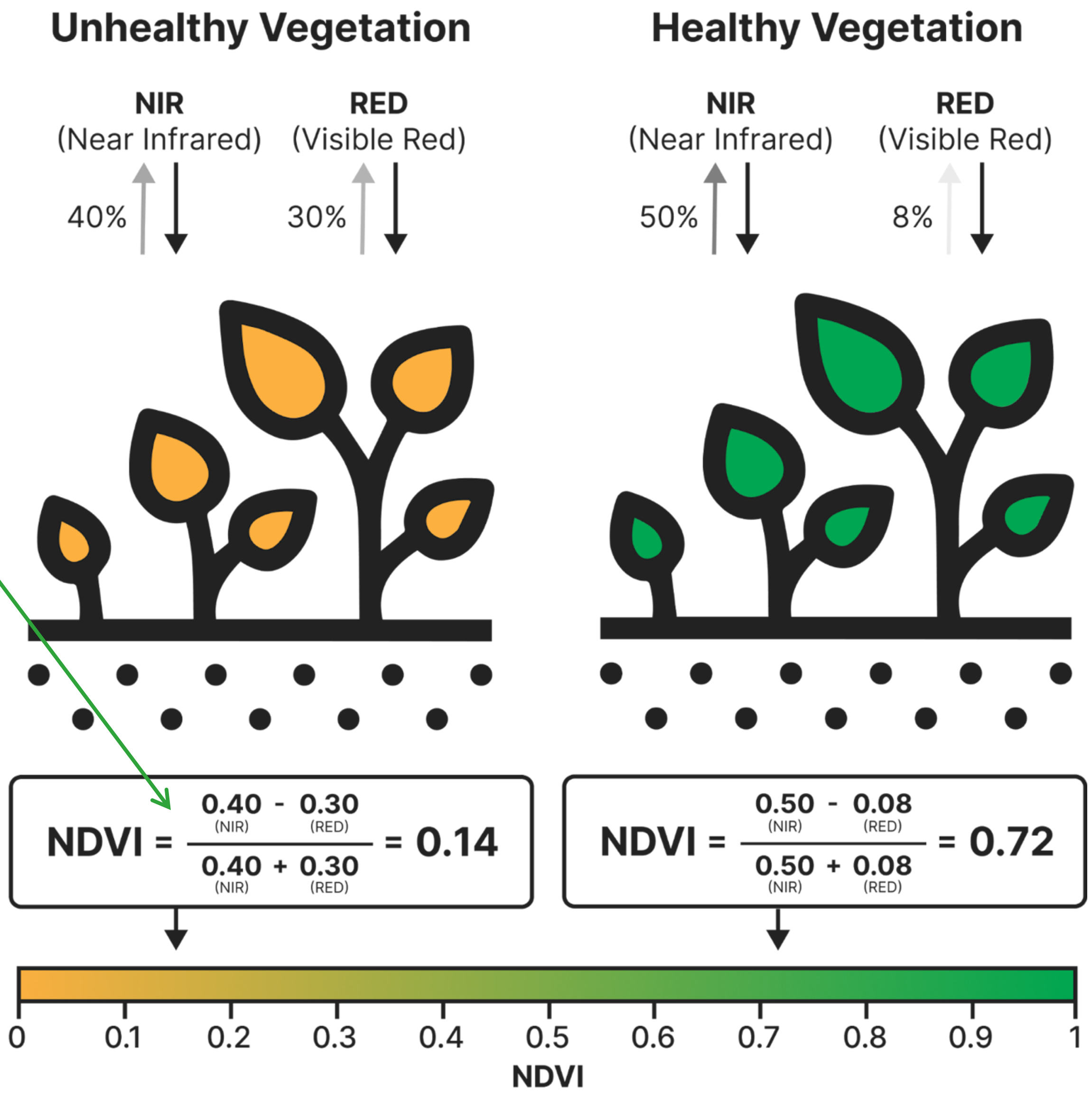
- Land :
  - Land area (% of total land area)
- Forest :
  - Forest area (% of land area)
- Agriculture :
  - Agriculture land (% of land area)
  - Agriculture irrigated (% of total agricultural land)
  - Arable land (% of land area)
  - Permanent cropland (% of land area)

NDVI

NDVI is an indicator of the greenness of the biomes.

NDVI is calculated as a ratio between the red (R) and near infrared (NIR) values.

- NDVI will be a value between -1 and 1.
- An area with nothing growing in it will have an NDVI of zero.
- An area with dense, healthy vegetation will have an NDVI of one.
- NDVI values less than 0 suggest a lack of dry land.
- An ocean will yield an NDVI of -1



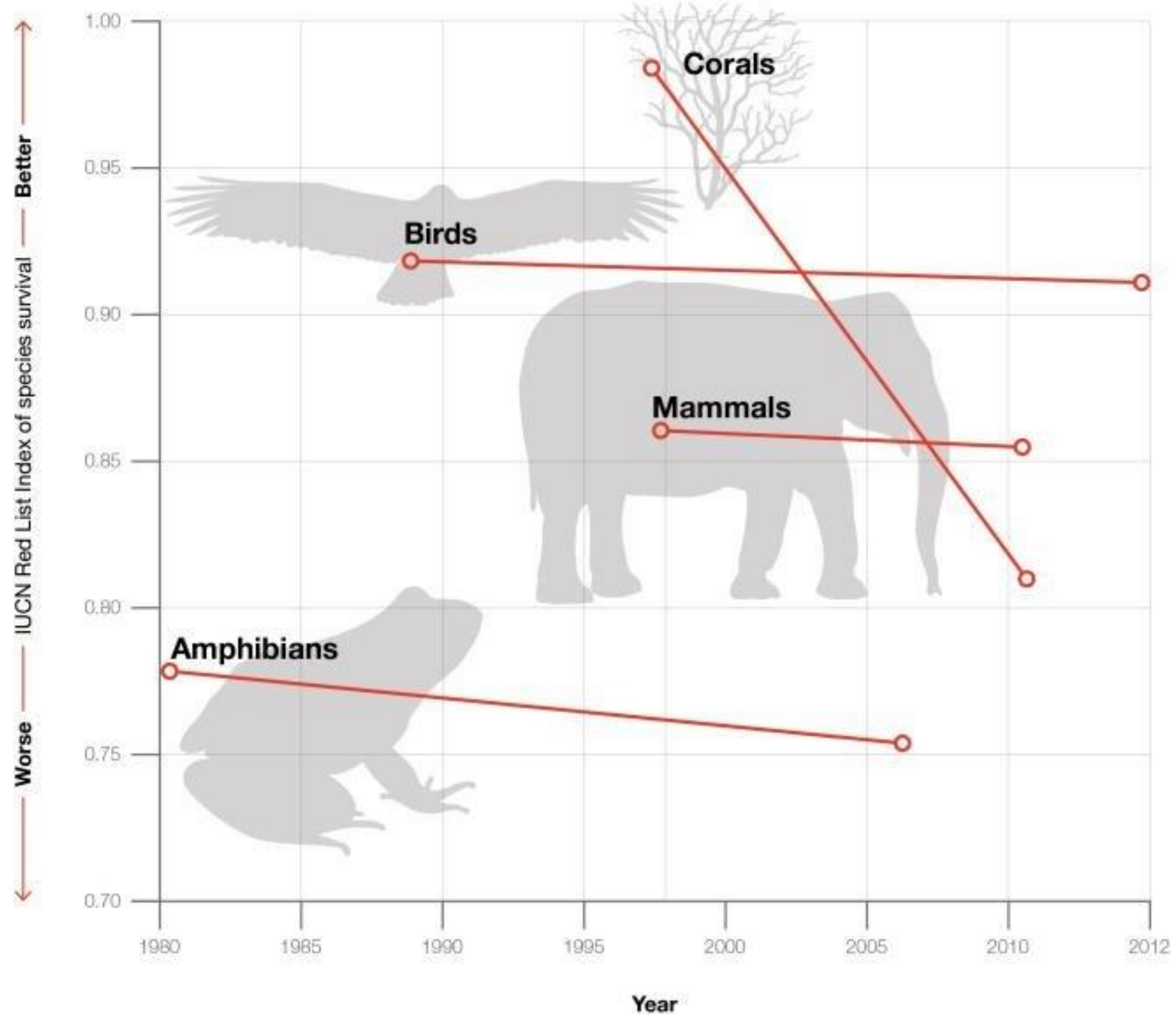
Red List Index

The Red List Index (RLI), is an indicator of the changing state of global biodiversity.

It defines the conservation status of major species groups, and measures trends in extinction risk over time.

A Value of 1.0 equates to all species being categorized as 'Least Concern'. A value of 0 indicates that all species have gone extinct.

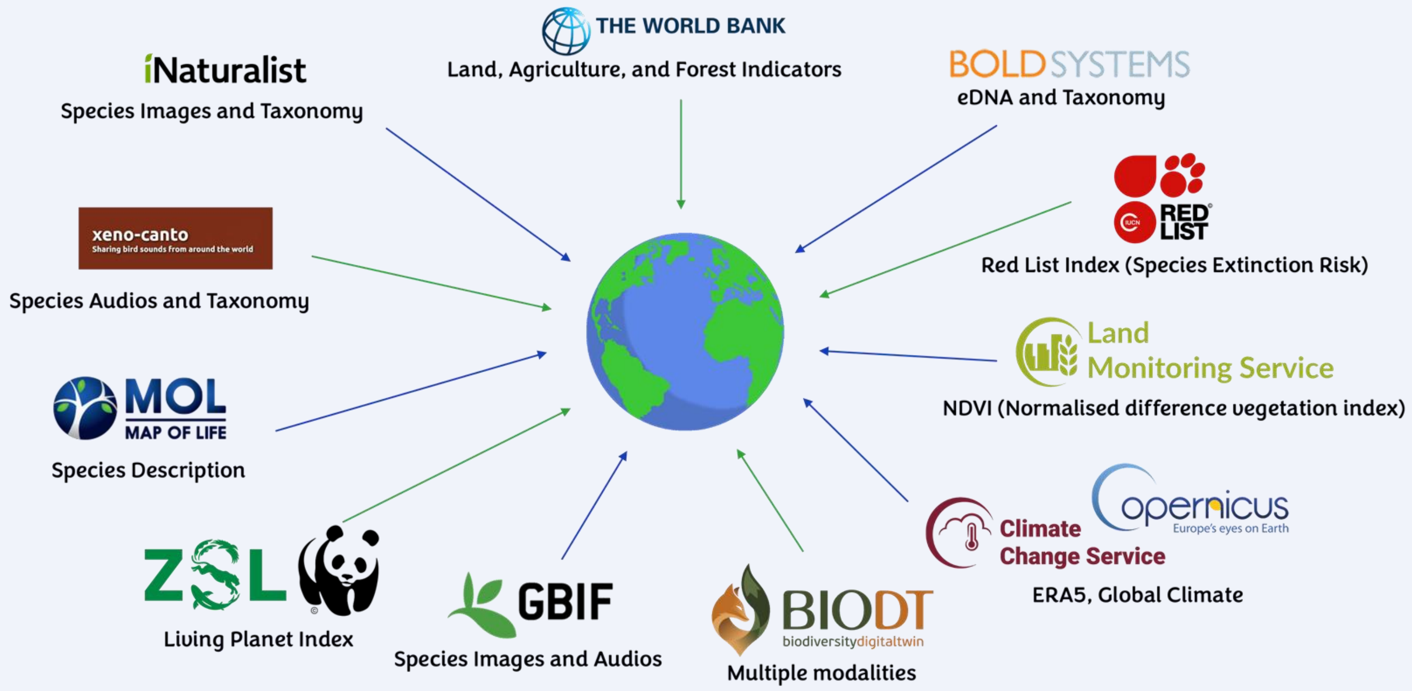
- Red List Categories :
  - EX (Extinct), EW (Extinct in the Wild), CR (Critically Endangered), EN (Endangered)
  - VU (Vulnerable), NT (Near Threatened), LC (Least Concern)



Data Challenges

Biodiversity data are not easy to work with:

1) **Fragmentation:** Dispersed among various sources
2) **Lack of protocols:** No uniform data collection and storage protocols
3) **Quality:** Varying and no validation metrics
4) **Size and storing:** From MBs to TBs
5) **Sampling frequencies:** hourly, daily, monthly, yearly
6) **Multiple modalities**
7) **Handling:** Diversified pre- and post-processing
8) **Geodesic grounding:** Different coordinate systems and representations/projections



ENGINEERING A MULTI-MODAL DATASET FOR BIODIVERSITY RESEARCH

SURF

TNO innovation for life

BIODT biodiversitydigitaltwin

Engineering a Multimodal Dataset for Biodiversity

WHY: To solve some of the previous mentioned challenges

Goal:

- Create a Unified dataset for diverse analyses
- Scalable handling of large multimodal data
- Flexible integration of spatial and temporal features

Data Acquisition

APIs Limitations: 10.000 observations (iNaturalist), random date names(Xeno-Canto), downloading time (ERA5), species names request (MOP)

Challenge : World Bank and Red List

Files only with the name of the country and values

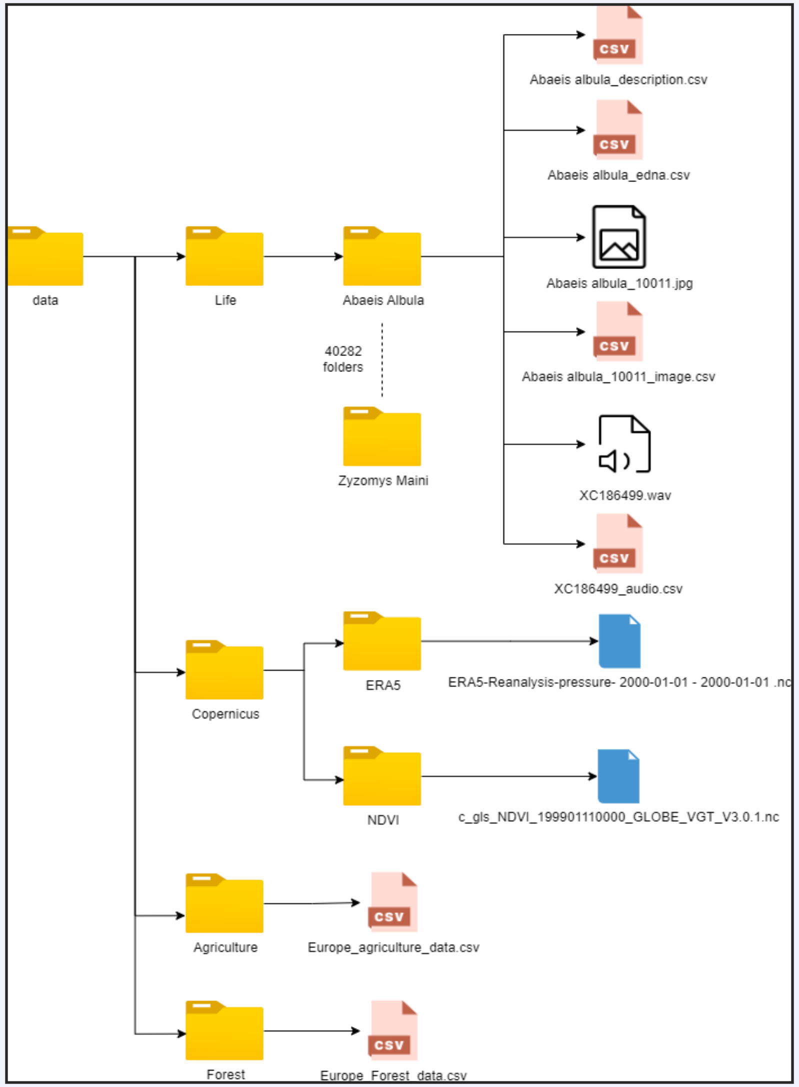
Solved by getting bounding boxes.

Source	API	Files
e-DNA(BOLD)	✓	
NDVI(Copernicus Land)	✓	✓
ERA5(Copernicus)	✓	
Indicators(Word Bank)		✓
Images(iNaturalist)	✓	✓
Audios(Xeno Canto)	✓	
Description(MOF)	✓	
Living Planet Index		✓
Red List Index		✓
Images, Audios, Occurrences (GBIF)		✓

Data Acquisition - Workflow

1. **Data Retrieval**:
    - Reads pre-existing CSV, NetCDF, JSON. Filters data based on date, geolocation and content.
    - Fetches data from sources through APIs. Retrieves metadata, and data files.
2. **Data Processing**:
    - Extracts and harmonises relevant features, as data, and metadata.
    - Rounds spatial data for uniformity and organized temporal values. Same formats and ranges, in location and timestamp.
    - Processes downloaded images, audio and geospatial metadata to ensure standardisation.
3. **Storage and Organisation**:
    - Saves outputs in structured directories based on regions, species, and data modalities.
    - Creates species-specific subfolders for images, audio, and metadata.
    - Ensures seamless integration of API-based data (e.g., taxonomy and audio metadata) with file-based datasets for downstream analysis.

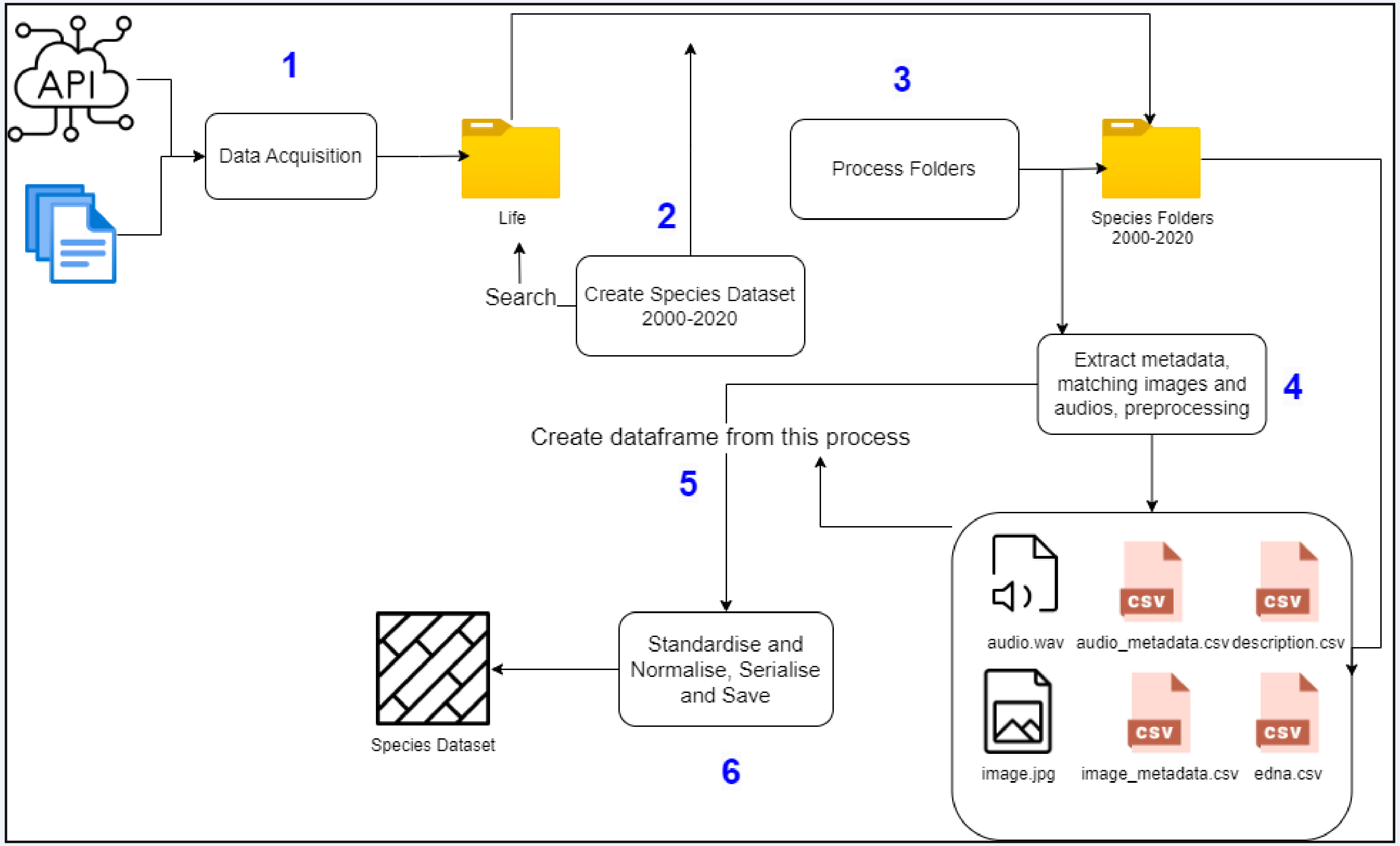
Challenge



Species Dataset - Workflow

Steps :

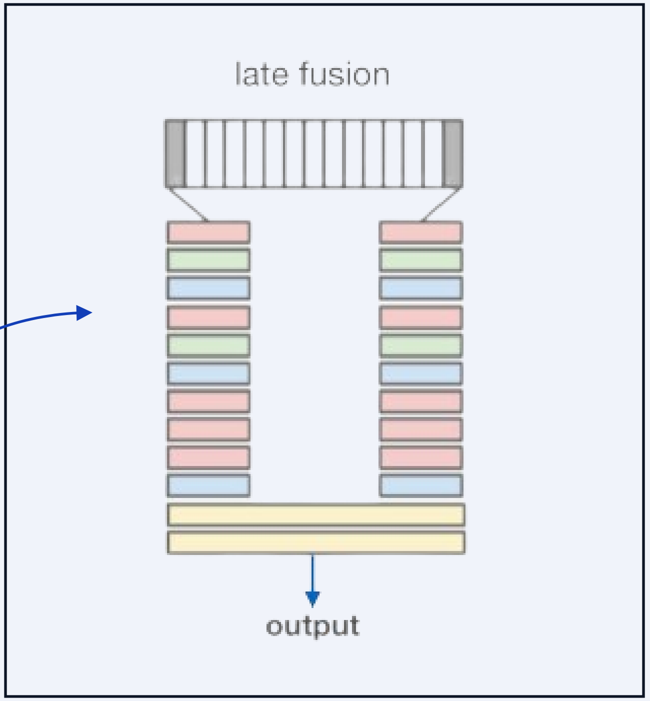
1. Data Acquisition
2. Create Species Dataset
    - Filtering to get specific folders, hash-table approach
3. Process Folders
4. Extract metadata from csv files. Matching audios and images. Get edna, description, and distribution data from csv
5. Create dataframe with these data
6. Standardise and Normalise values. Serialise data to be compatible with parquet format. Save

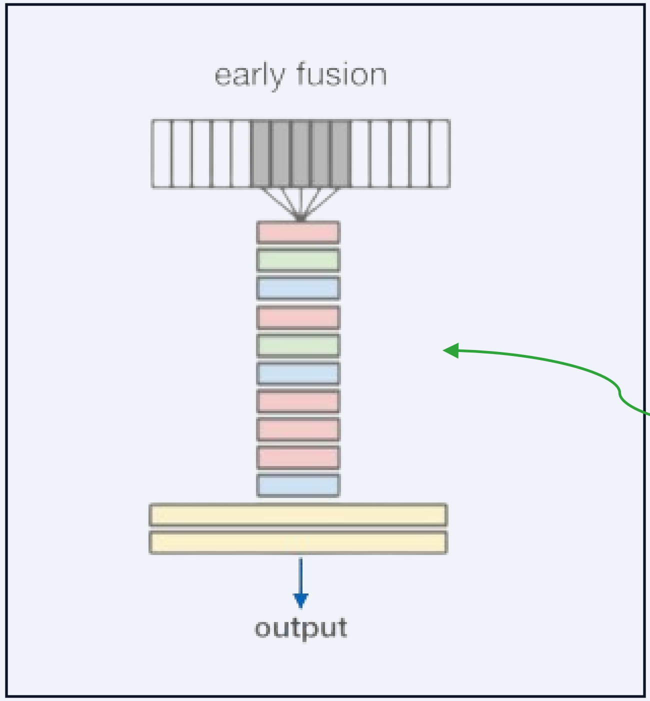


Challenge: Cannot load all the folders -> Extreme memory utilisation

Data Fusion

Data Batch Creation





Input to the Model

High Dimensionality Issues

Loss of Modality-Specific Features

Scalability and Flexibility

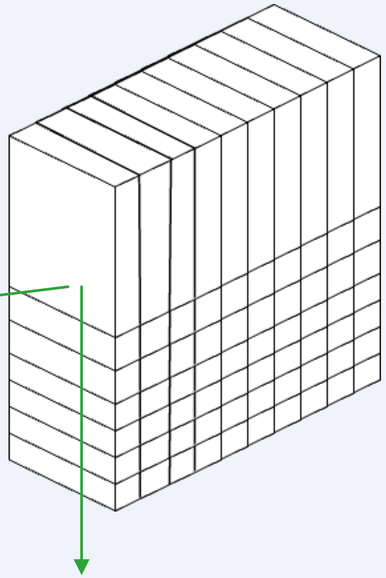
Unified Input Representation

Simplified Model Design

Data Correlation Exploitation

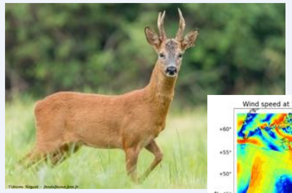
Reduced Processing Overhead

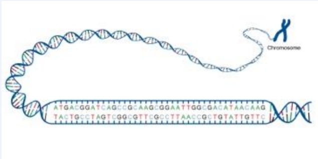
Data Batch - Concept

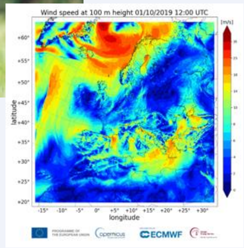


Create **a unified and structured data representation** that integrates multimodal environmental and species data for spatiotemporal analysis and machine learning.

- Defined by **latitude**, **longitude** and **timestamp**
- Ensures all modalities are aligned in a consistent spatiotemporal framework
- Independent modality initialisation, shaped by the spatiotemporal grid







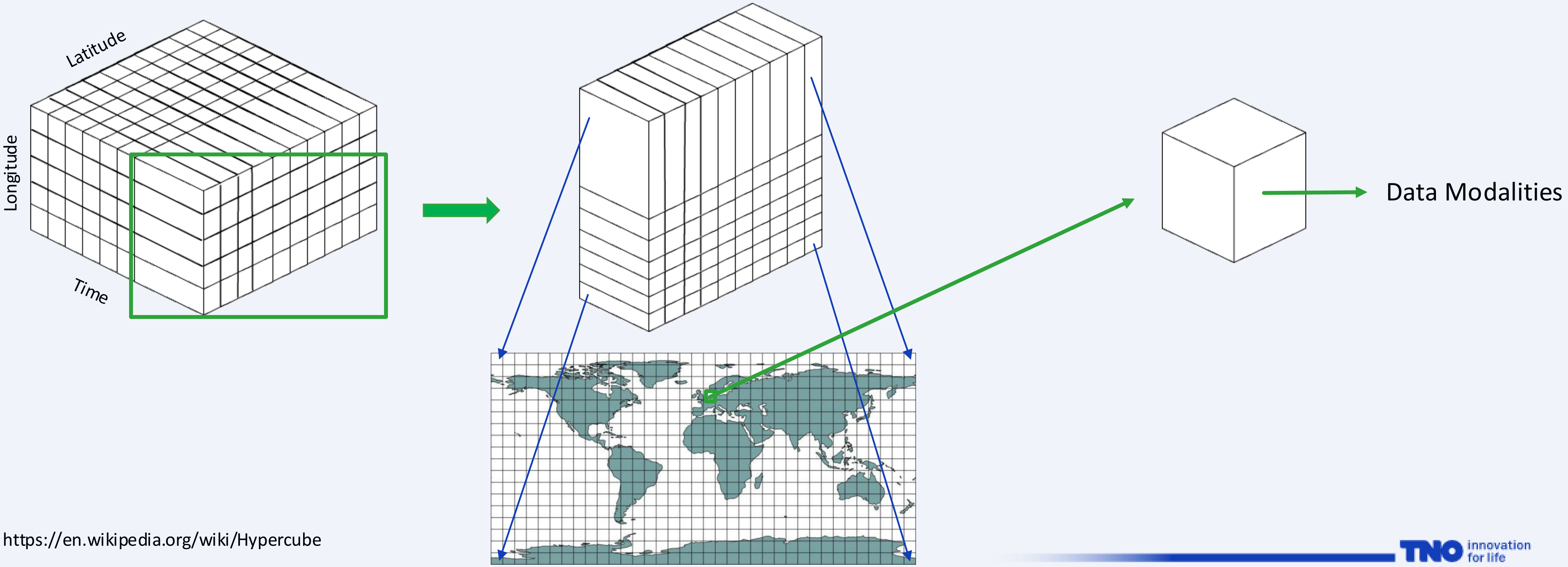
Geospatial Data Grounding

Coordinate Reference System: WGS84

Data HyperCube

Data Batch

Data Point



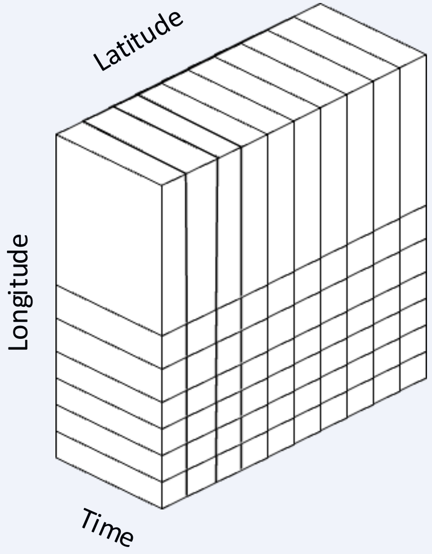
https://en.wikipedia.org/wiki/Hypercube

https://en.wikipedia.org/wiki/World_Geodetic_System

Geospatial Data Grounding

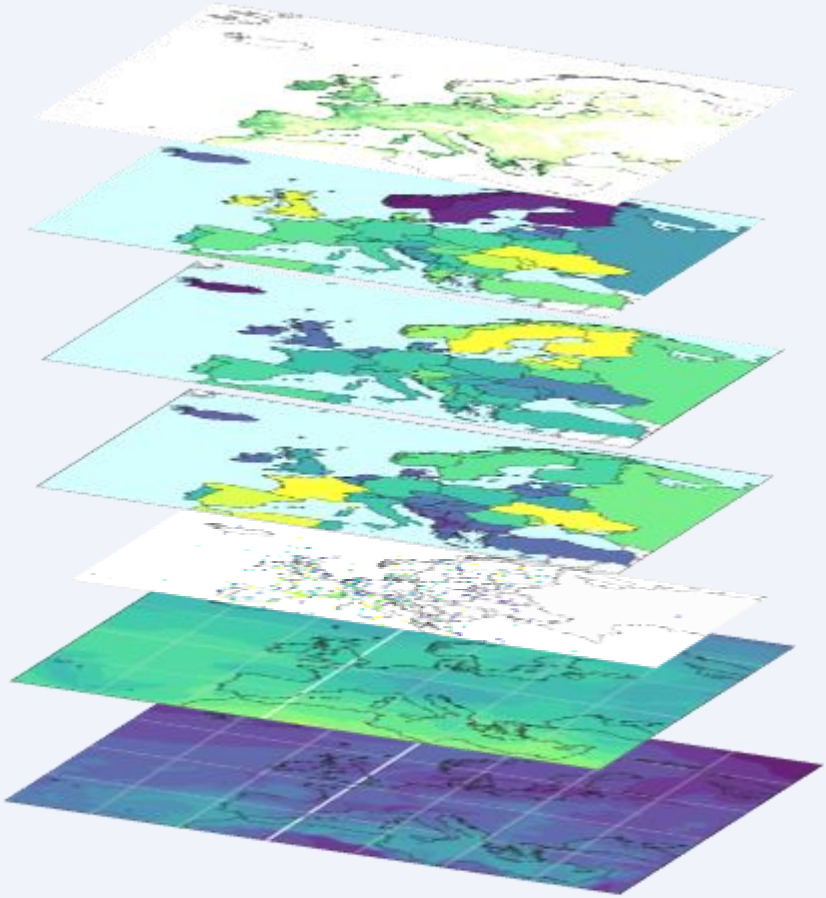
Grid Sampling: 0.25º

Data Point: [long, lat, time, modalities]



Data Modalities

- Species Taxonomy
- Species Modalities
- Climate Variables
- Weather Variables
- Vegetation Variables
- Agriculture Variables
- Cropland Variables
- Forest Variables



Data Batch - Statistics

A Full HD image has 1920x1080=2.073.600

Grid Sampling: 0.25º

Globe cells: 1.440(long) x 720 (lat) = 1.036.800

Europe cells: 321(long) x 153 (lat) = 49.113

Time Range: 01-01-2000 : 31-12-2020

Batch Size: ~70 GB

 # Batches: 13.686 = 958.020 GB => **958 TB**

Each Batch: 2 days, lat, lon, 13 pressure levels, 21 species (metadata)

Total Data Points: 14.881.239

Task	Cluster	Time
Dataset pre- and post - processing	TNO local	~ 12 days
	SURF Snellius (Rome)	~ 1.5 day
Batch creation (x1)	TNO local	~ 40 mins
	SURF Snellius (Rome)	~1 min

Dataset challenges & Limitations

❓ NaN and missing values

- Locations at sea do not contain any values as we focus on ground level biodiversity

❓ Saving and handling TB

- Each Europe batch is ~70GB (13.686 => 958 TB)

❓ Dataset pre-processing and Batch creation takes days

- Speed up to few hours with HPC scaling

❓ Limit to Europe region

- For now use the coordinates [lat, long] = [(34, 72), (-30, 50)]

❓ Data Sources Limitations

- Requests limits

TNO innovation for life

THE BIODIVERSITY FOUNDATION MODEL

Biodiversity Foundation Model - Introduction

Purpose: Model biodiversity as complete as possible, uncovering hidden relationships

Goal: Develop an operational AI Foundation Model that can be used as an emulator to:

- Accurately monitor and predict biodiversity dynamics
- Generate conservation strategies

Requirements:

- Multi-modal data representation
- Preserve spatiotemporal features
- Operate in both local and global scale
- Simulate the underlying physics in multiple scales
- Adaptable to various use-cases

Related Work

Prithvi WxC: Foundation Model for Weather and Climate

Johannes Schmude[1,†,‡], Sujit Roy[2,7,†,‡], Will Trojak[1], Johannes Jakubik[1], Daniel Salles Civitarese[1], Shraddha Singh[1], Julian Kuehnert[1], Kumar Ankur[2], Aman Gupta[3], Christopher E Phillips[2], Romeo Kienzler[1], Daniela Szwarcman[1], Vishal Gaur[2], Rajat Shinde[2], Rohit Lal[2], Arlindo Da Silva[6], Jorge Luis Guevara Diaz[1], Anne Jones[1], Simon Pfreundschuh[4], Amy Lin[2], Aditi Sheshadri[3], Udaysankar Nair[2], Valentine Anantharaj[5], Hendrik Hamann[1], Campbell Watson[1], Manil Maskey[7], Tsengdar J Lee[8], Juan Bernabe Moreno[1], Rahul Ramachandran[7]

†Equal Contribution,
‡Johannes.Schmude@ibm.com , Sujit.Roy@nasa.gov

ABSTRACT

Triggered by the realization that AI emulators can rival the performance of traditional numerical weather prediction models running on HPC systems, there is now an increasing number of large AI models that address use cases such as forecasting, downscaling, or nowcasting. While the parallel developments in the AI literature focus on foundation models – models that can be effectively tuned to address multiple, different use cases – the developments on the weather and climate side largely focus on single-use cases with particular emphasis on mid-range forecasting. We close this gap by introducing Prithvi WxC, a 2.3 billion parameter foundation model developed using 160 variables from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Prithvi WxC employs an encoder-decoder-based architecture, incorporating concepts from various recent transformer models to effectively capture both regional and global dependencies in the input data. The model has been designed to accommodate large token counts to model weather phenomena in different topologies at fine resolutions. Furthermore, it is trained with a mixed objective that combines the paradigms of masked reconstruction with forecasting. We test the model on a set of challenging downstream tasks namely: Autoregressive rollout forecasting, Downscaling, Gravity wave flux parameterization, and Extreme events estimation. The pretrained model with 2.3 billion parameters, along with the associated fine-tuning workflows, has been publicly released as an open-source contribution via Hugging Face.

AURORA: A FOUNDATION MODEL OF THE ATMOSPHERE

Cristian Bodnar[*,1], Wessel P. Bruinsma[*,1], Ana Lucic[*,1], Megan Stanley[*,1], Johannes Brandstetter[3,†], Patrick Garvan[1], Maik Riechert[1], Jonathan Weyn[2], Haiyu Dong[2], Anna Vaughan[4], Jayesh K. Gupta[5,†], Kit Tambiratnam[2], Alex Archibald[4], Elizabeth Heider[1], Max Welling[6,†], Richard E. Turner[1,4], and Paris Perdikaris[1]

[1]Microsoft Research AI for Science
[2]Microsoft Corporation [3]JKU Linz [4]University of Cambridge [5]Poly Corporation [6]University of Amsterdam

[*]Equal contribution [†]Work done while at Microsoft Research

ABSTRACT

Deep learning foundation models are revolutionizing many facets of science by leveraging vast amounts of data to learn general-purpose representations that can be adapted to tackle diverse downstream tasks. Foundation models hold the promise to also transform our ability to model our planet and its subsystems by exploiting the vast expanse of Earth system data. Here we introduce Aurora, a large-scale foundation model of the atmosphere trained on over a million hours of diverse weather and climate data. Aurora leverages the strengths of the foundation modelling approach to produce operational forecasts for a wide variety of atmospheric prediction problems, including those with limited training data, heterogeneous variables, and extreme events. In under a minute, Aurora produces 5-day global air pollution predictions and 10-day high-resolution weather forecasts that outperform state-of-the-art classical simulation tools and the best specialized deep learning models. Taken together, these results indicate that foundation models can transform environmental forecasting.

ORBIT: Oak Ridge Base Foundation Model for Earth System Predictability

Xiao Wang*, Siyan Liu*, Aristeidis Tsaris*, Jong-Youl Choi*, Ashwin M. Aji†, Ming Fan*, Wei Zhang*, Junqi Yin*, Moetasim Ashfaq*, Dan Lu*, Prasanna Balaprakash*
*Oak Ridge National Laboratory, Oak Ridge, United States
{wangx2, lius1, tsarisa, choij, fanm, zhangw3, yinj, mashfaq, lud1, pbalapra}@ornl.gov
†AMD Research and Advanced Development, Santa Clara, United States
ashwin.aji@amd.com

Abstract—**Earth system predictability is challenged by the complexity of environmental dynamics and the multitude of variables involved. Current AI foundation models, although advanced by leveraging large and heterogeneous data, are often constrained by their size and data integration, limiting their effectiveness in addressing the full range of Earth system prediction challenges. To overcome these limitations, we introduce the Oak Ridge Base Foundation Model for Earth System Predictability (ORBIT), an advanced vision transformer model that scales up to 113 billion parameters using a novel hybrid tensor-data orthogonal parallelism technique. As the largest model of its kind, ORBIT surpasses the current climate AI foundation model size by a thousandfold. Performance scaling tests conducted on the Frontier supercomputer have demonstrated that ORBIT achieves 684 petaFLOPS to 1.6 exaFLOPS sustained throughput, with scaling efficiency maintained at 41% to 85% across 49,152 AMD GPUs. These breakthroughs establish new advances in AI-driven climate modeling and demonstrate promise to significantly improve the Earth system predictability.**

forecasts are timely without compromising prediction accuracy for effective decision-making.

In the Earth system, what unfolds at sub-daily to daily weather scale eventually averages out and can be characterized as sub-seasonal to seasonal climate variations. Gradual shifts in weather and seasonal patterns determine the longer-term climate changes in response to internal and external forcings. Although the Earth system operates seamlessly across different scales, the current physics modeling approach involves using separate prediction systems for each scale, with one system used for weather forecasting, another for sub-seasonal to seasonal prediction, and a third for long-term decadal to centennial climate projections.

In response to the above challenges, AI-based models are emerging as a promising approach and aim to integrate prediction at different scales into a single system. These models leverage advanced machine learning techniques to handle the intricacies of the Earth system data,

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.

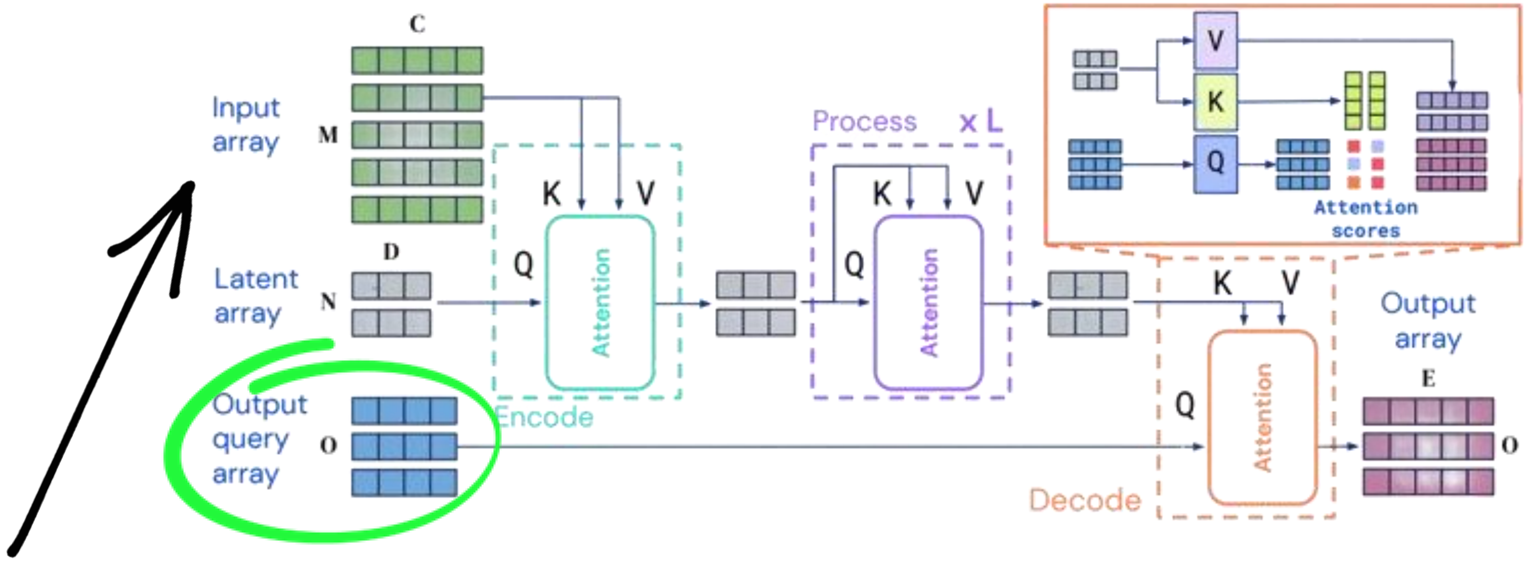
Abstract

Significant progress in the development of highly adaptable and reusable Artificial Intelligence (AI) models is expected to have a significant impact on Earth science and remote sensing. Foundation models are pre-trained on large unlabeled datasets through self-supervision, and then fine-tuned for various downstream tasks with small labeled datasets. There is an increasing interest within the scientific community to investigate whether this approach can be successfully applied to domains beyond natural language processing to effectively build generalist AI models that exploit multi-sensor data. This paper introduces a first-of-its-kind framework for the efficient pre-training and fine-tuning of foundational models on extensive geospatial data. We have utilized this framework to create Prithvi, a transformer-based geospatial foundational model pre-trained on more than 1TB of multispectral satellite imagery from the Harmonized Landsat-Sentinel 2 (HLS) dataset. Our study demonstrates the efficacy of our framework in successfully fine-tuning Prithvi to a range of Earth observation tasks that have not been tackled by previous work on foundation models involving multi-temporal cloud gap imputation, flood mapping, wildfire scar segmentation, and multi-temporal

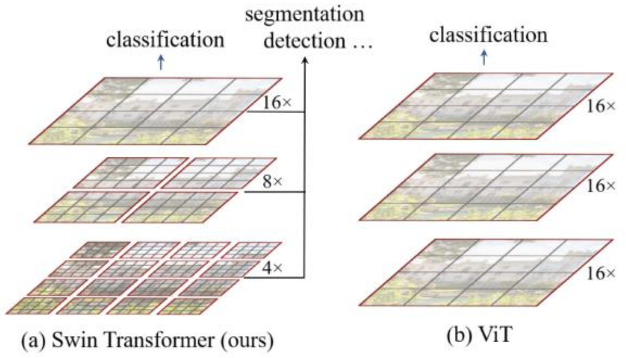
Building Blocks (Encoder & Decoder)

Perceiver IO: A General Architecture for Structured Inputs & Outputs

- **Perceiver IO** – Perceiver (thus, Transformer) –based architecture [6]

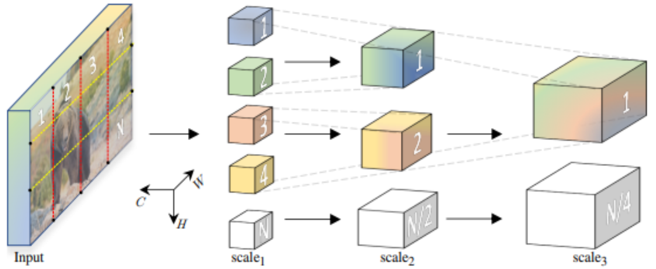


Building Blocks (Core)



Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows

- **Swin** – Visual Transformer–based architecture [7]
- Based on shifted window partitioning mechanism
  - Enables modelling of connections between different regions
- Hierarchical feature representations
  - Aligns with ecosystems structuring – from individual interactions to broader dynamics

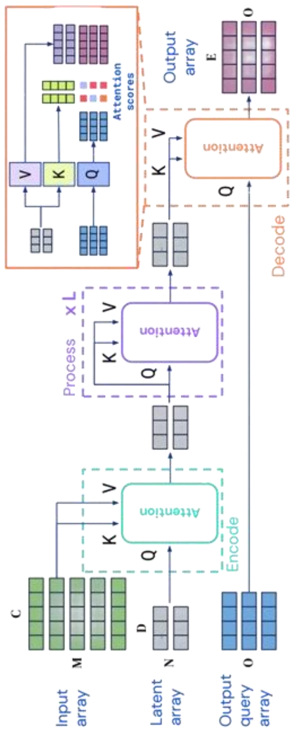


Multiscale Vision Transformer (MViT)

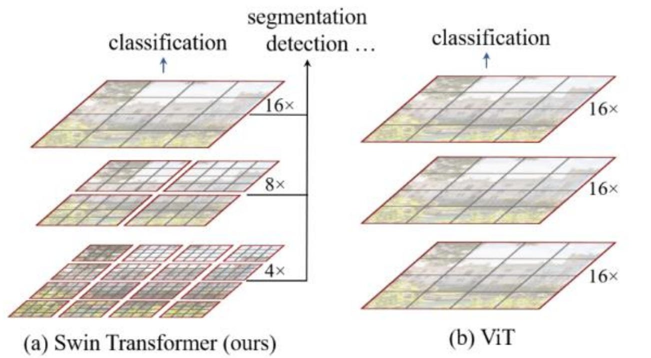
- **MViT** – Also a Visual Transformer–based architecture [8]
- Progressive spatial reduction capability
  - Understand biodiversity patterns at multiple scales
- Multi-scale feature hierarchy
  - Enable the model to understand how different spatio-temporal scales of ecological processes interact

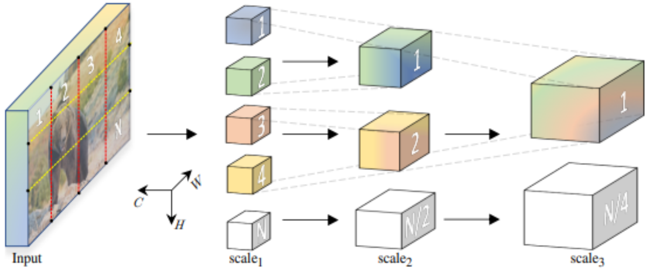
Final Architecture

Encoder

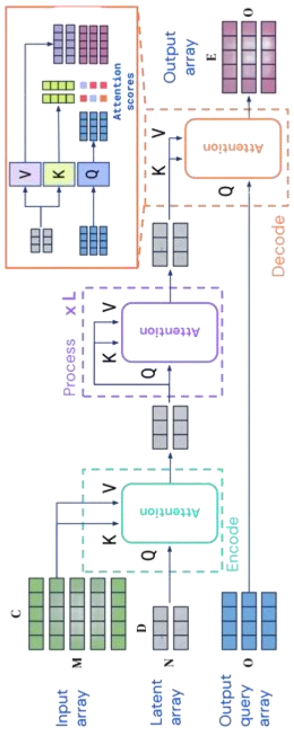


Core



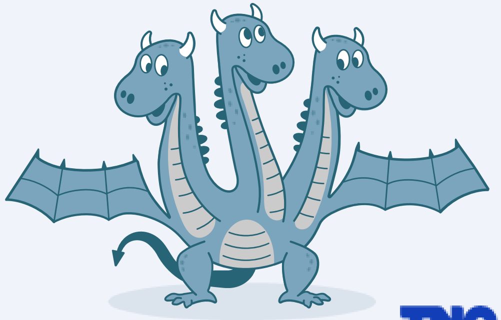


Decoder



Development status

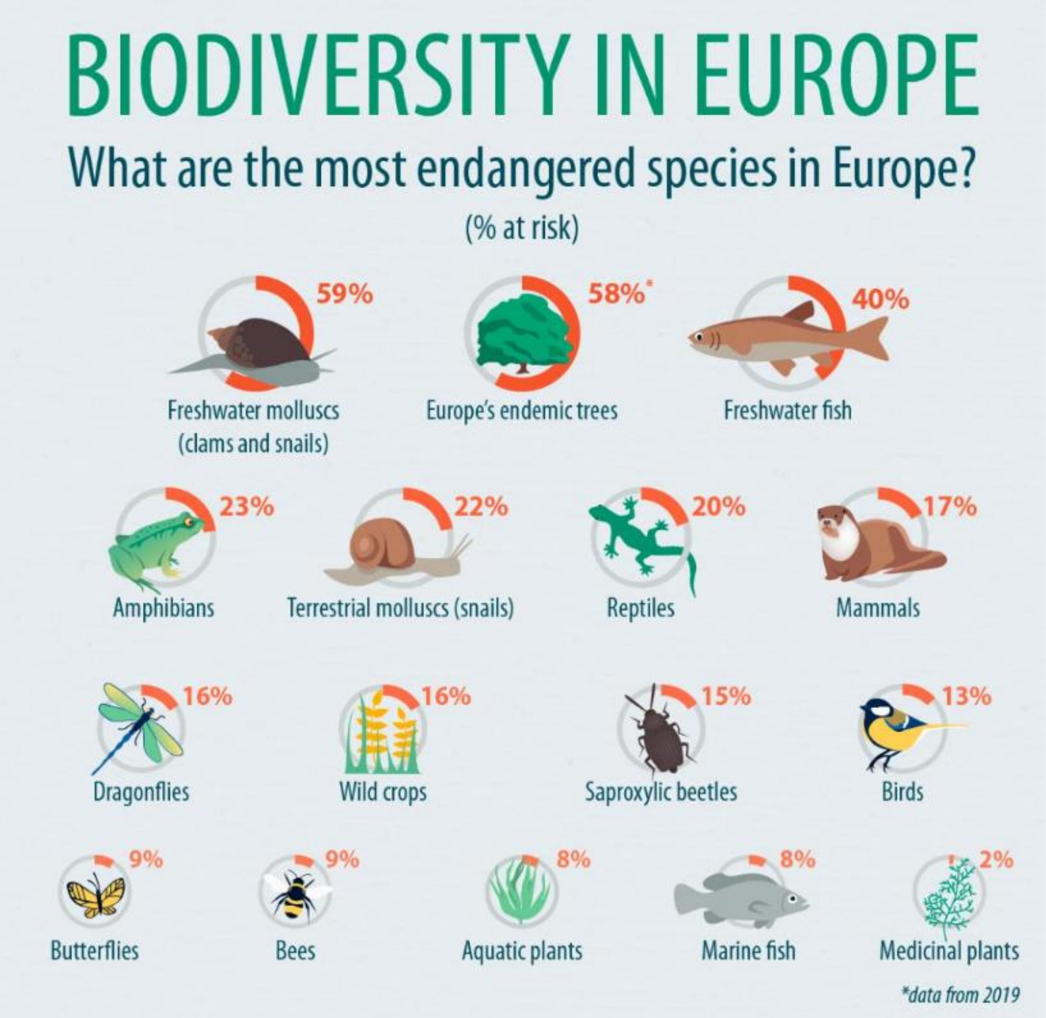
- Fully open-source frameworks
- Model sizes: 20 & 50 million parameters
- Test runs on TNO cluster
- Training runs on SURF's Snellius utilising latest NVIDIA H100



Use case(s)

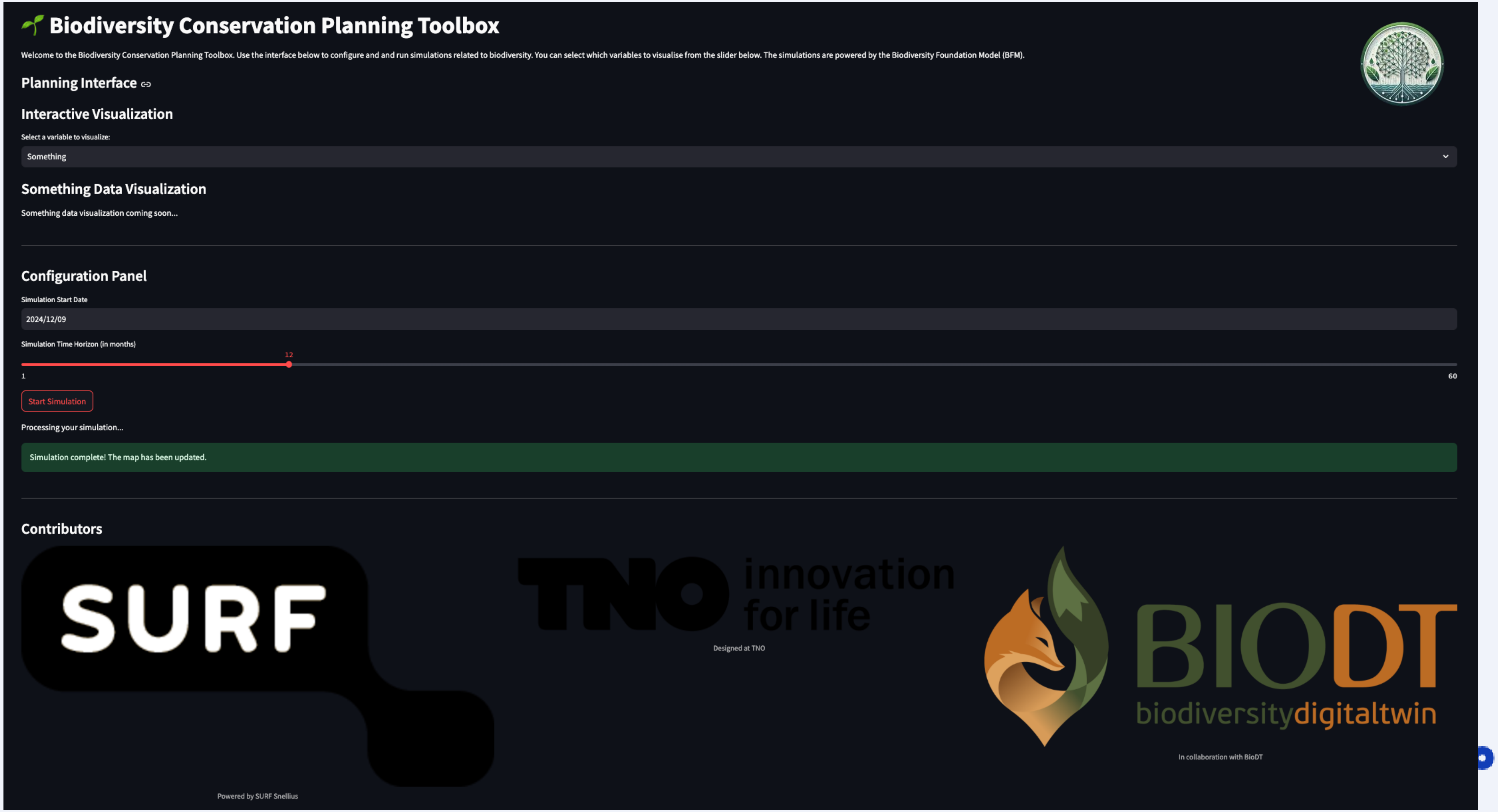
In the first iteration, Biodiversity Foundation Model will be used for:

- Species Distribution Modelling in EU
  - Focus on Endangered Species
  - 21 European-based species have been selected



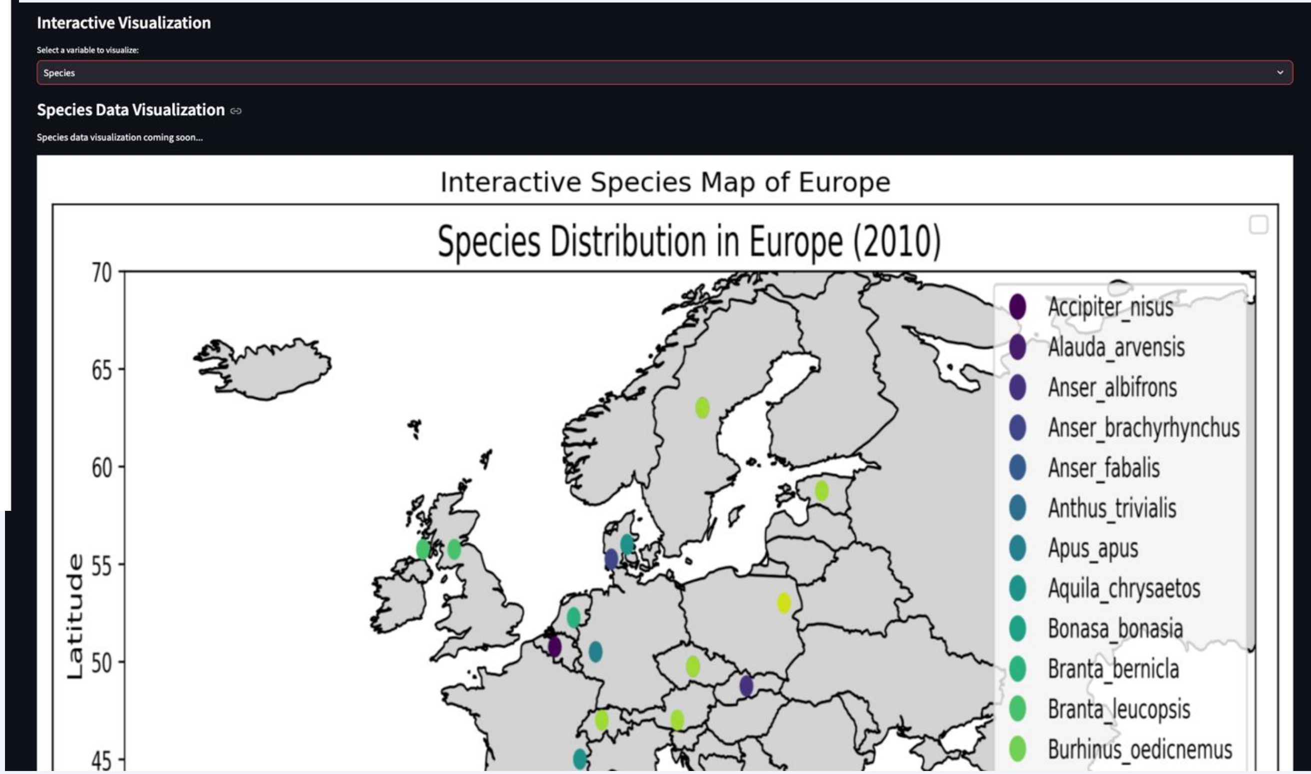
BFM Development Recap

BFM Application



innovation for life

BFM Application



Future steps

In the upcoming months

- Finalise pre-training
- Validate species distribution modelling capabilities
- Fine-tune on 2 new use cases
    - Pollination
    - Grassland Dynamics
- Add user interaction capabilities with Natural Language

BFM

Thank you for your time! Questions?

References

[1]Jakubik, J., Roy, S., Phillips, C., Fraccaro, P., Godwin, D., Zadrozny, B., Szwarcman, D., Gomes, C., Nyirjesy, G., Edwards, B., Kimura, D., Simumba, N., Chu, L., Mukkavilli, S.K., Lambhate, D., Das, K., Bangalore, R., Oliveira, D., Muszynski, M., Ankur, K., Ramasubramanian, M., Gurung, I., Khallaghi, S., Li, H., Cecil, M., Ahmadi, M., Kordi, F., Alemohammad, H., Maskey, M., Ganti, R.K., Weldemariam, K., & Ramachandran, R. (2023). Foundation Models for Generalist Geospatial Artificial Intelligence. *ArXiv, abs/2310.18660*.

[2] Bodnar, C., Bruinsma, W. P., Lucic, A., Stanley, M., Brandstetter, J., Garvan, P., ... & Perdikaris, P. (2024). Aurora: A foundation model of the atmosphere. *arXiv preprint arXiv:2405.13063*.

[3] Xu, H., Usuyama, N., Bagga, J., Zhang, S., Rao, R., Naumann, T., ... & Poon, H. (2024). A whole-slide foundation model for digital pathology from real-world data. *Nature*, 1-8.

[4] Takeda, S., Kishimoto, A., Hamada, L., Nakano, D., & Smith, J. R. (2023, June). Foundation model for material science. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 13, pp. 15376-15383).

[5] Tsaris, A., Dias, P. A., Potnis, A., Yin, J., Wang, F., & Lunga, D. (2024). Pretraining Billion-scale Geospatial Foundational Models on Frontier. *arXiv preprint arXiv:2404.11706*.

[6] Jaegle, A., Borgeaud, S., Alayrac, J. B., Doersch, C., Ionescu, C., Ding, D., ... & Carreira, J. (2021). Perceiver io: A general architecture for structured inputs & outputs. arXiv preprint arXiv:2107.14795.

[7] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).

[8] Fan, H., Xiong, B., Mangalam, K., Li, Y., Yan, Z., Malik, J., & Feichtenhofer, C. (2021). Multiscale vision transformers. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 6824-6835).

The Team

Athanasios Trantas

TNO

Martino Mension

TNO

Stylianos Stasinos

TNO

Sebastian Gribincea

TNO

Manuel Rodrigez

SURF

Marco Verdicchio

SURF

Shreshtha Sharma

TNO

Syrine Ben Aziza

TNO

Paolo Pileggi

TNO

Elena Lazovik

TNO

Aliene van der Veen

TNO

Thank you for your time!

QUESTIONS?

SURF

TNO innovation for life

BIODT biodiversitydigitaltwin

MENTI

SCAN TO JOIN

or

VISIT menti.com and use code:

26766218