TRILLIUM TECH

Earth Systems Predictability: How can AI advance planetary stewardship?

A ROADMAP FOR A PLANETARY NERVOUS SYSTEM

esa Φ-lab

UNIVERSITY OF OXFORD

FDL EUROPE

A QUANTIFIED PLANET.

Authors
Cormac Purcell, Anne Spalding, Lior Torenberg, Jonathan Knowles, James Parr
and the Frontier Development Lab community.

Designers
Ana Freitas, Leo Silverberg, and Jon Rushton.

Editors and Reviewers
Nicolas Longepe (ESA), Lorien Pratt (Quantellia), Alison Lowndes (Nvidia)
and Ed Gillespie (Greenpeace).

Acknowledgments
The ESP Forum was supported by the European Space Agency's ɸ-Lab and Oxford University, with additional support from Nvidia Corporation, Google Cloud, Climate-KIC Europe and Quantellia.

Icons in this document are sourced from the Noun Project archive.

The ESP Forum is an initiative of Trillium Technologies.

NAVIGATION

Earth Systems Predictability (ESP) incorporates multiple disciplines and stretches beyond technology into how we manage our everyday lives. The ideas in this report are organised into four categories:

1. **ESP Vision**
   What is ESP?
2. **Building Blocks**
   Core concepts of ESP
3. **Operations**
   Concepts and systems that enable collaboration
4. **Deployment**
   What is needed to build and deploy ESP systems?

Each category builds towards a fuller description of ESP and how it might be deployed in the real world. Finally, we present a roadmap and practical action recommendations to move ESP from concept to reality.

This report is a living document: **we invite all readers to get involved by building on the ideas of the action recommendations and championing ESP in your own work.**

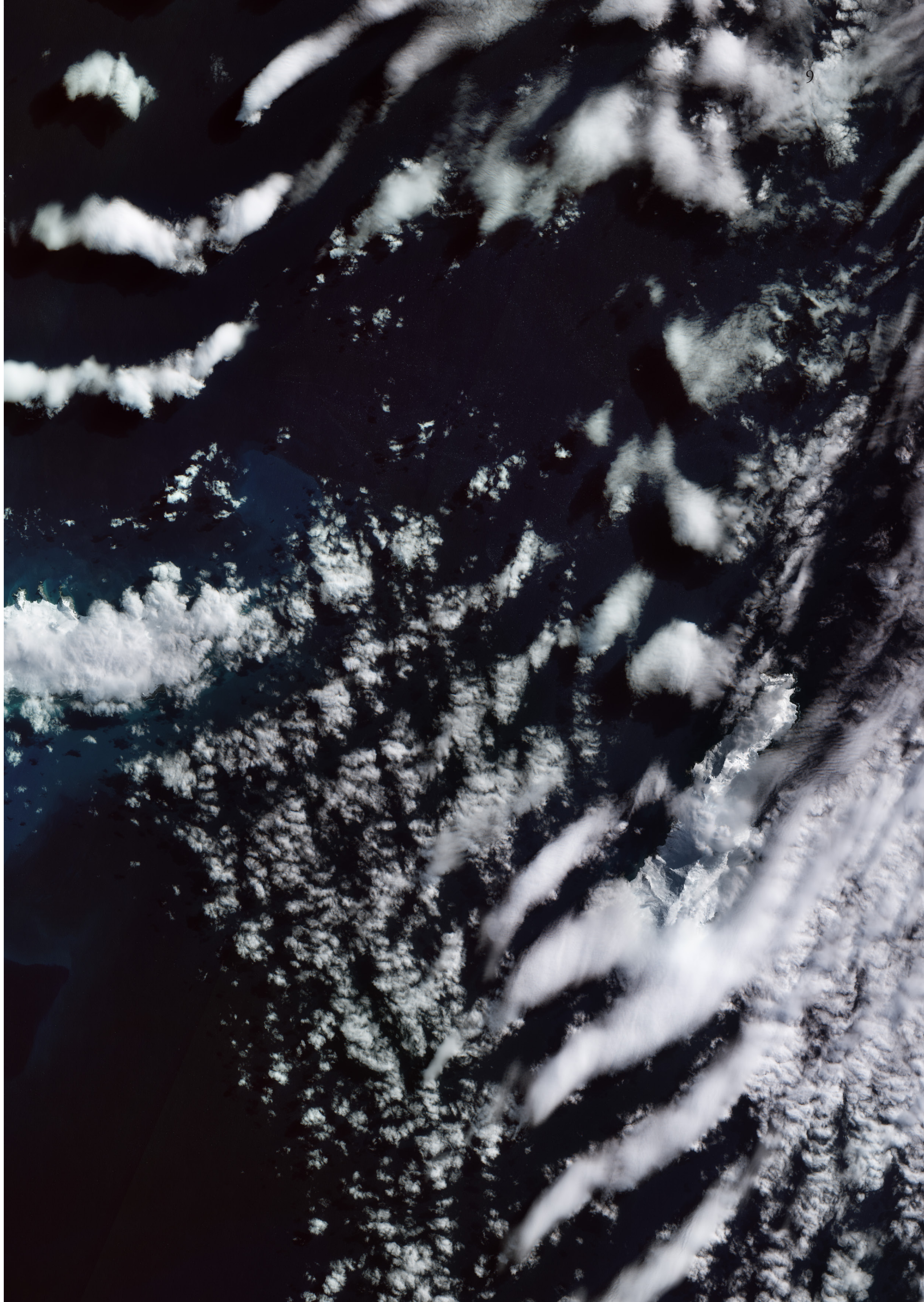
CONTENTS

KEY

In this report you will see a number of diagrams to illustrate thought processes, ideas, systems and technologies. Below are some of the most important icons used in communicating ESP visually.

- **Earth**
- **Space Data**
- **Ground Data**
- **Large Pre-trained Foundation Model/LLM**
- **Adaptor**
- **ESP Model**
- **Reporting**
- **Outcome/Task**
- **Government**
- **Person/Stakeholder/User**
- **ESP Cloud**
- **Fuel**
- **Dependency Link**

- **External Input as an Assumption (scalar) or Prediction (complex).**
- **Choice (scalar) or Plan of Action (complex).**
- **Performance Measurement or metric.**
- **Final or Intermediate outcome.**
- **Outcome.**
- **Executive Function (EF)**
- **Asset**
- **Technology**



A VIABLE FUTURE FOR ALL OF US

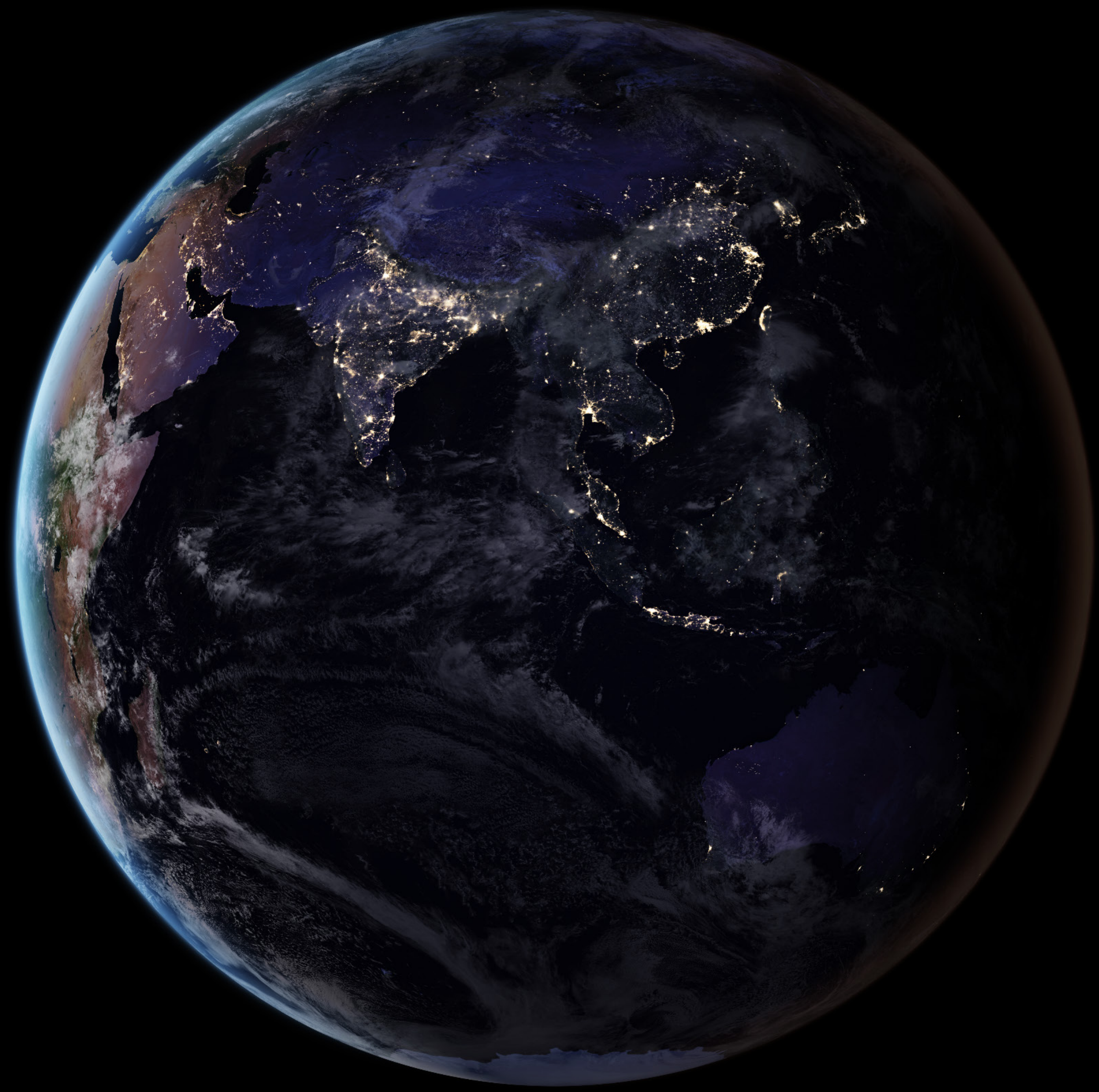
This report presents a synthesis and expansion of the outcomes of the Earth Systems Predictability (ESP) Forum - a three-day transdisciplinary workshop held online in May 2023, in partnership with the ESA ɸ-Lab, the University of Oxford and Trillium Technologies.

The effort is directly motivated by the Space for Net Zero report by the World Economic Forum, which called for '*a sort of distributed operations centre to help manage our spaceship Earth*' for the good of all humankind.

In response, the ESP Forum brought together over 150 experts in Earth science, computer modelling, machine learning and decision intelligence to accelerate our understanding of how artificial intelligence (AI) and other emerging technologies can help us make better decisions for the future of our planet.

In this document we present the combined vision of leading experts to define Earth Systems Predictability and how it can be deployed.

We imagine a brighter future for us all and provide here concrete actions to support a sustainable and thriving world.



ESP I FORUM
Earth Systems Predictability

IN NUMBERS

153 PARTICIPANTS

60 PARTICIPANTS PER SESSION

18 PROVOCATION TALKS

+33 HOURS OF IDEAS AND INSIGHTS

4 PARTNERS

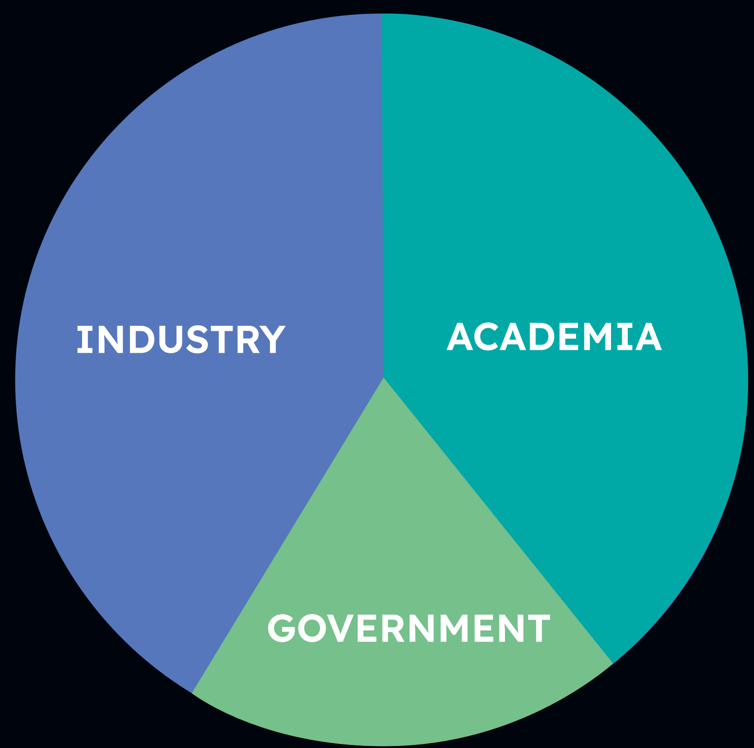
esa Φ-lab

UNIVERSITY OF OXFORD

FDL EUROPE

TRILLIUM TECH

PARTICIPANT AFFILIATION



35 ORGANISATIONS

ESA
NASA
Oxford
Google
NVIDIA
AAICO
Quantellia
Climate-KIC
UNOOSA
MIT Media Lab
World Economic Forum
Rainforest Alliance
DS3Lab at ETH Zürich
SVT Group
Terra Genesis
KP Labs
Symbiotek Systems
PopClimate
Planet Labs PBC
Mexican Space Agency
Sistematica Engineering
AI Sweden, WAI Labs
Fidelity Investments
German Research Center for AI
Fermata Energy
SparkGrid AI
Foundation for Planetary Intelligence
WRMS
Helyx SIS Ltd
Solutions for a Small Planet
SKEMA Business School
AppliedRoots
University of Nebraska-Lincoln
Insightful

SETTING THE SCENE

ESA ϕ-Lab

AI plays a pivotal role in maximizing the potential of the space assets currently in development. Our unique overview from space has taught us that we are all passengers on spaceship Earth, but we soon might have a capable "co-pilot" thanks to the combined power of Earth observation and AI.

The vantage point that space offers allows us to glean valuable insights about our planet and AI is becoming instrumental in translating huge data from Earth observations into informed decisions and actions. With AI as our co-pilot, we can better assess, simulate, predict and anticipate how to manage our planetary environment, with our destination being a more sustainable future.

We live in a very interesting time where Earth observation capabilities (and data size) are growing rapidly. Simultaneously, new tools are emerging to help derive rapid insight from the data, showing our impact on the Earth in real-time. The rate of progress in AI is unprecedented, in particular with the new generation of generative AI models capitalising on the Transformer architecture, which serve as a unifying force across diverse AI domains.

As we engage with these unified frameworks, our progress accelerates in ways previously unimaginable.

With these powerful capabilities comes great responsibility. The recurring issue of trust arises due to the complexity of working with 'black box' AI models involving trillions of parameters simultaneously. The 'why' of an outcome is often far from obvious, necessitating detailed forensic analysis. We have an imperative to routinely confront our models with the realities of the world - as scientists do in weather forecasting - so as to address concerns related to trust and safety.

Yet, amidst this acceleration, it is crucial to maintain a human-centric approach and a focus on sustainability for our planet. Our trajectory should be guided by the most objective information on the state of the environment and its evolution, keeping in mind that our powerful AI co-pilot remains a tool to assist our development.

Pierre-Philippe Mathieu
Head of ESA ϕ-Lab Explore Office
ESA ESRIN

THE CHOICE TO THRIVE WITH EVERY DECISION.

EXECUTIVE SUMMARY

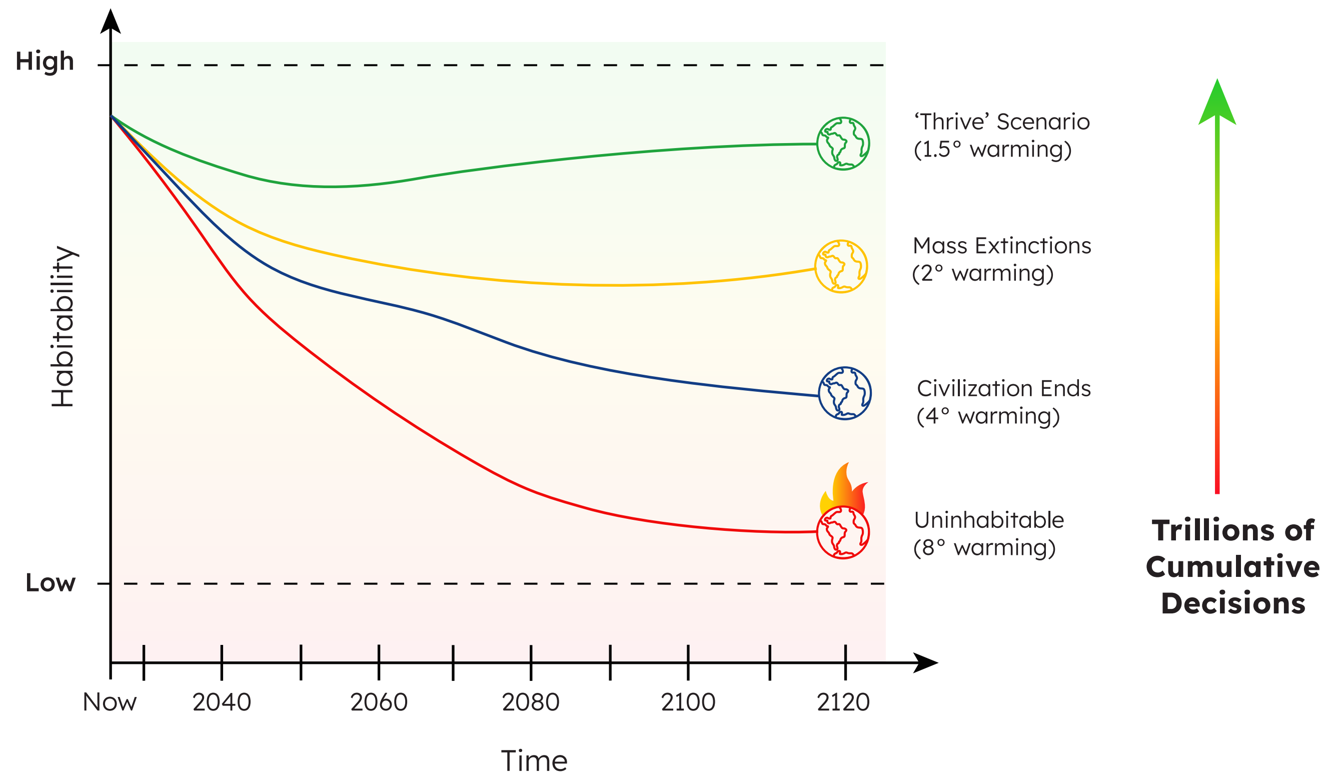
Earth Systems Predictability (ESP) enables humanity to embed considerations of long-term planetary habitability and the choice to thrive into every decision, however large or small.

Earth observation (EO) and artificial intelligence (AI) technology is making rapid assessment of climate and biosphere impacts possible for decisions ranging from international and government policy (e.g., resilience planning, crisis response), through corporate strategy (e.g., supply chains, business cases), to personal choices (e.g., travel, construction, shopping).

ESP couples Earth system models with domain-specialist simulations in economics, health, transport and the full extent of human civilisation.

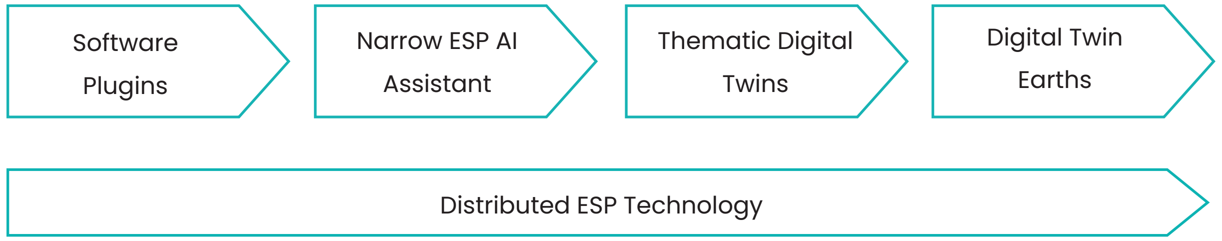
ESP technology can support decisions that are good for the long-term future of our planet, but also enable human wellbeing and enterprise to thrive across multiple linked domains

It is a vision of trillions of cumulative decisions, rather than one big idea.



ESP is envisioned as a ubiquitous decision-support technology, spanning orders-of-magnitude in scale, from lightweight software modules running on mobile devices, to sophisticated cloud-native Digital Twin Earths.

Getting to a position where human society is thriving in concert with Earth's planetary systems will require us to make **trillions of climate-, biosphere- and people-friendly decisions over the next few decades - starting now.** ESP technology must be agile and efficient enough to integrate with all of our existing software systems, wherever they are deployed. This means supporting a distributed architecture from the start, informed by the maxim of ***'good enough decisions'*** - implementing throttling so that the upper limits of fidelity and accuracy are driven by user needs. The concept of 'joint modelling' through decision intelligence is fundamental to ESP and is a research area that should be prioritised in the future.

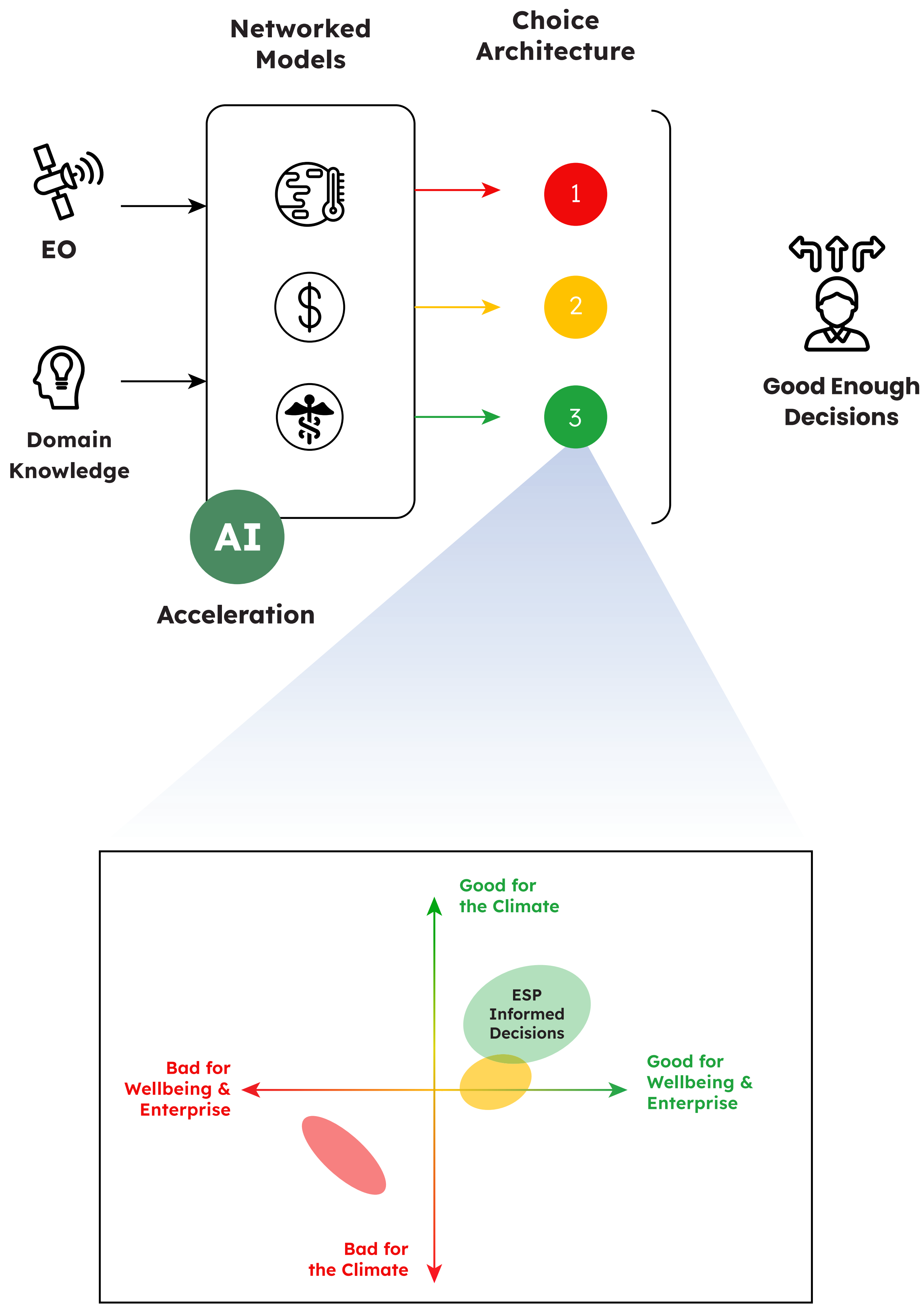


Compelling visualisations and interactive interfaces are essential components of ESP technology, helping to craft narratives around planetary stewardship and support confident mission-critical decisions.

EPS visualisations linked to interactive controls will bring ESP to the classroom, incident management centre and mobile device, embedding ESP concepts in society. Simultaneously, Large Language Models interfaces will make ESP-led decision-making and expert knowledge accessible to almost everyone.

ESP must be developed from the ground-up in close cooperation with the ordinary people who will use the technology to make everyday decisions.

Machine learning and simulation codes are traditionally developed within research communities, before being adapted to real world use-cases. This paradigm must be inverted: first building deep knowledge of practical applications to inform fundamental research. Deployment stakeholders must work with developers to create continuous validation workflows for ESP technology, inspiring trust and promoting knowledge exchange.



12 KEY CONCEPTS

1. Earth Systems Predictability (ESP) is a strategy to integrate Earth observations and a broad range of predictive models into decision-making tools.

Weather and climate models are at the heart of ESP and the goal is to always make biosphere-friendly decisions, even when the primary objective is related to economics, manufacturing, health, sales, energy, education or wellbeing. ***There needs be no trade-off.***

2. The formal discipline of Decision Intelligence provides a framework for integrating data and simulations into a larger decision model.

Decision models are expressed through causal decision diagrams (CDDs), which can be thought of as blueprints for integrated 'systems-of-systems'. CDDs implemented as software could support complex multimodal decision making that optimises outcomes for a range of stakeholders. ***ESP technology enables joint modelling for better decisions.***

3. Observations from Earth-orbiting satellites are a core data-source for ESP and a wide range of disciplines.

Open and accessible Earth observation (EO) data from ESA, NASA and the other space agencies provide frequent measurements of over half the essential climate variables. But EO data is also highly useful for agriculture, the energy sector, smart cities, supply chain monitoring, carbon markets and much more. ***Live EO data will inform the simulations in ESP systems and provide situational awareness for decisions.***

4. Emerging artificial intelligence (AI) technology is introducing paradigm shifts in modelling capabilities: faster, more detailed and better informed by data.

AI-driven speed-ups of 100 - 10,000 times for weather and climate forecasting are facilitating scenario modelling, including predictions of extreme events. As the technology evolves, we will be able to predict further into the future, at higher fidelities to model impact on regional scales. ***These capabilities are fundamental to ESP technology for making confident decisions with a far time-horizon***

5. AI is making advanced modelling and specialist knowledge more accessible and understandable to non-experts.

AI-powered simulations can be run on modest computing hardware - laptops and phones instead of supercomputers. They can be integrated cheaply into decision systems, educational tools, or research toolkits. ***Large language models (LLMs) are encoding deep knowledge for easy retrieval, while foundation models are poised to make deploying AI tools accessible to anyone.***

6. Responsive, interactive visualisations linked directly to ESP models and EO data will facilitate a deep understanding of how decisions affect Earth systems into the future.

Rich interactions with ESP technology in the form of Digital Twin Earths, or more focused thematic interfaces, will promote understanding of how decisions affect the long-term climate. Widespread behaviour change should be supported by interactive narratives and conversations. However, ***ESP systems must be accessible on mobile devices with small screens, which requires a distributed software architecture.***

7. ESP technology must be developed, tested and deployed in cooperation with deployment stakeholders and end-user groups.

A wide range of stakeholders must be involved in the co-creation of ESP technology from an early phase of development. Stakeholders include: groups of people who will be affected by decisions made with ESP technology, users of ESP tools, data custodians, human-rights specialists and representatives of first peoples. ***Transparent development will result in tools that are better fit-for-purpose, in addition to promoting two-way understanding between technologists and end-users.***

8. Independent benchmarks of ESP systems must be developed, with realistic data, robust models and domain appropriate metrics.

Predictions from ESP systems should be validated by independent entities, operating in their specialist area of expertise. Validation should be an ongoing process, as AI models can exhibit drift over time.

Explainability, auditability, privacy and data-integrity should be paramount. ***Standards for ESP technology could be formalised in a certification process, which would ideally be recognised internationally.***

9. An Earth Operation Centre (EOC) would be a physical space to deploy a Digital Twin Earth containing ESP technology for decision making in support of governments and large organisations.

A 'EOC' centre could support day-to-day decision-making, long-term planning for climate change mitigation, crisis management, research collaboration, and technology development. EOCs are best supported by a cloud-native suite of ESP technologies ***that offer joint-modelling capabilities and advanced visualisation interfaces, such as augmented reality.***

10. An Earth Operation Network (EON) takes the concept of an EOC to much larger scales, but also makes it accessible to smaller organisations.

EOCs must scale to large and small decisions, meaning that a network of EOCs is needed. We imagine that such an Earth Operation Network would prioritise lightweight computing, distributing processing and visualisation tasks across the nodes. EON client software must be light enough to run on low-power computing devices (e.g., smartphones) ***to allow developing nations to participate.***

11. ESP software plugins have the potential to make climate friendly decisions accessible to a very wide range of organisations.

Robust calculations of how decisions affect our future climate could be offered to a broad range of stakeholders by developing ESP plugins. Such software modules could estimate the climate cost of business decisions, government policy, infrastructure creation and technology development by integrating seamlessly into software already in use. ***However, the application of ESP calculations should be incentivised by regulation or law.***

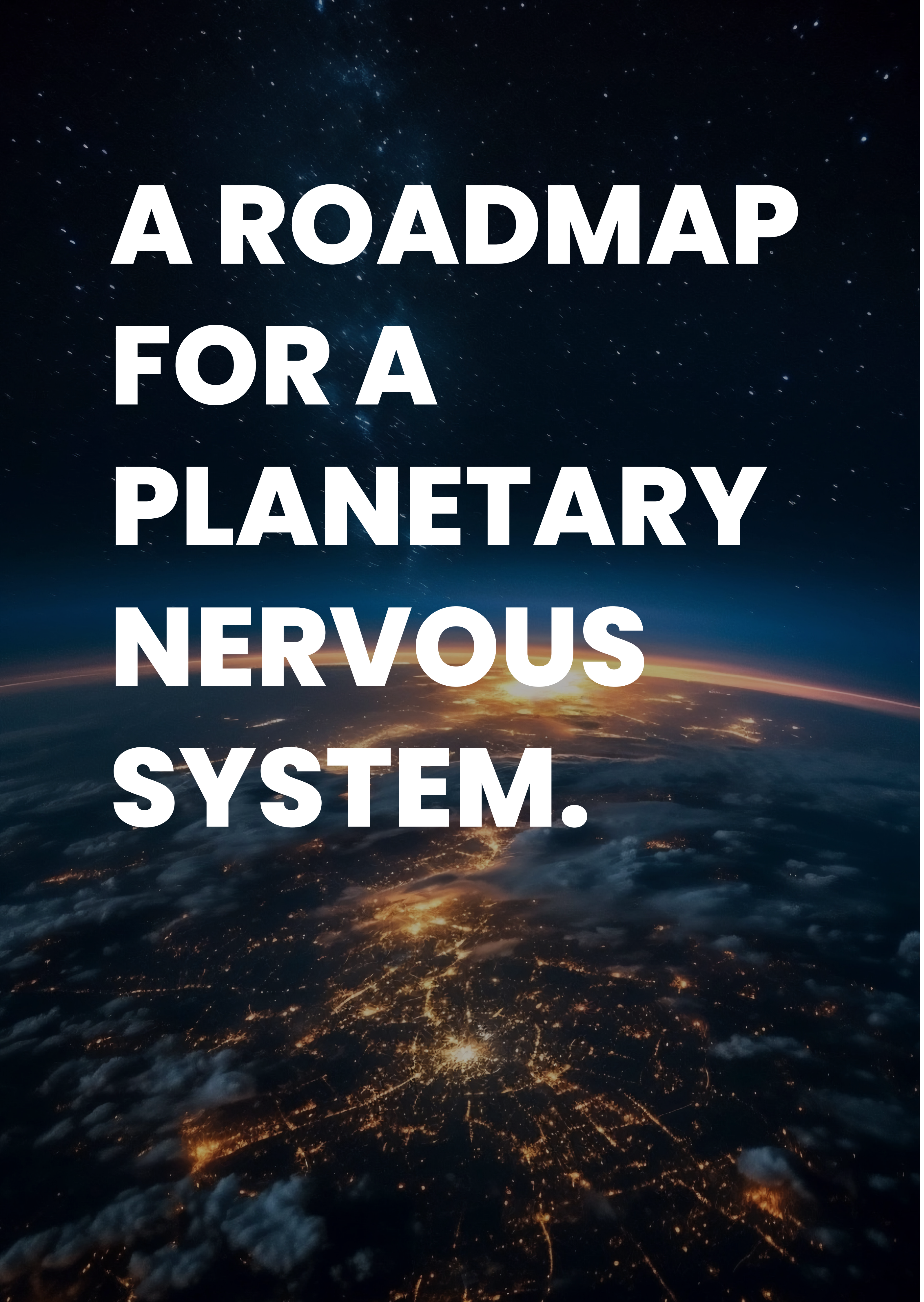
12. Our financial system has a significant role to play in driving sustainable practices, with ESP as a fundamental decision-making tool.

We need well-designed financial incentives that reference ESP modelling as a gold-standard benchmark. However, continual monitoring of ground-level impact is critical to support financial decisions. ***Inspiration and examples can be derived from the concept of social return on investment (SROI), which has a well-developed methodology to assess impact.***

"Today, ESP is a collection of emerging technologies that are lacking effectiveness because they are not well-integrated (imagine car parts strewn on the ground).

Tomorrow, these technologies will be integrated into an 'engine' that drives human decisions and their resulting actions towards achieving our shared climate goals."

Lorien Pratt
Quantellia
Inventor of Transfer Learning and Decision Intelligence.



The **ESP Technical Roadmap** described later in this report is a blueprint of a distributed *ESP Cloud Computing Software Stack:* a unifying architecture to guide the development of ESP technologies.

The ESP Technical Roadmap is complemented by **ESP Executive Functions** that provide crucial human oversight to manage ethical, cultural and social aspects of operational systems.

ESP VISION

ESP: AN EMERGING DISCIPLINE

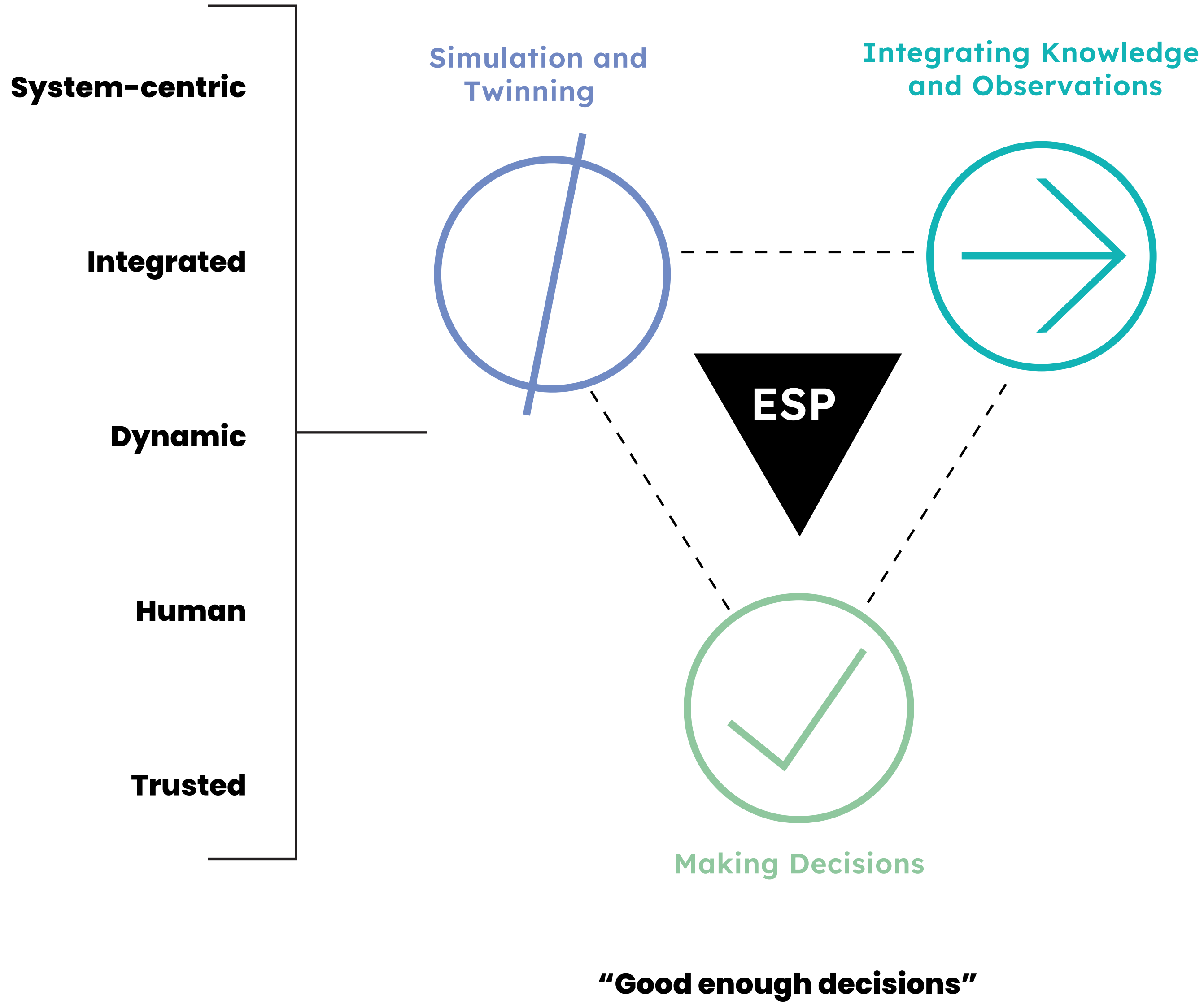
Earth Systems Predictability (ESP) is a term that describes an ecosystem and integration strategy for emerging technologies that offer the capability to measure and visualise planetary health, predict how it will change over time, understand human actions and support enhanced decisions for long-term planetary habitability and human wellbeing.

ESP encompasses fundamental Earth systems like the climate, weather, the geosphere and biosphere, and connects these to human systems like economics, finance and health.

ESP is enabled by revolutions in Earth observation (EO), skillful and **rapid AI-enabled predictions**, robust **uncertainty quantification, compelling visualisations** and **intuitive controls**, allowing everyone - from shop owners to policy makers - to **experiment, learn and make decisions** while factoring in **long-term impacts on the Earth.**

Although the technology is here, the vision of ESP will need to be enabled by broad partnerships between diverse end users, data custodians, technical experts and governments at all levels, and will need to be brokered by international agreements. Success will require transparent governance and a deeply ethical code of practice.

ESP is more than applied AI for climate, or "AI4EO", it is an **emerging discipline** that combines elements from Earth observation (EO), artificial intelligence (AI), simulation, visualisation and decision intelligence (DI).



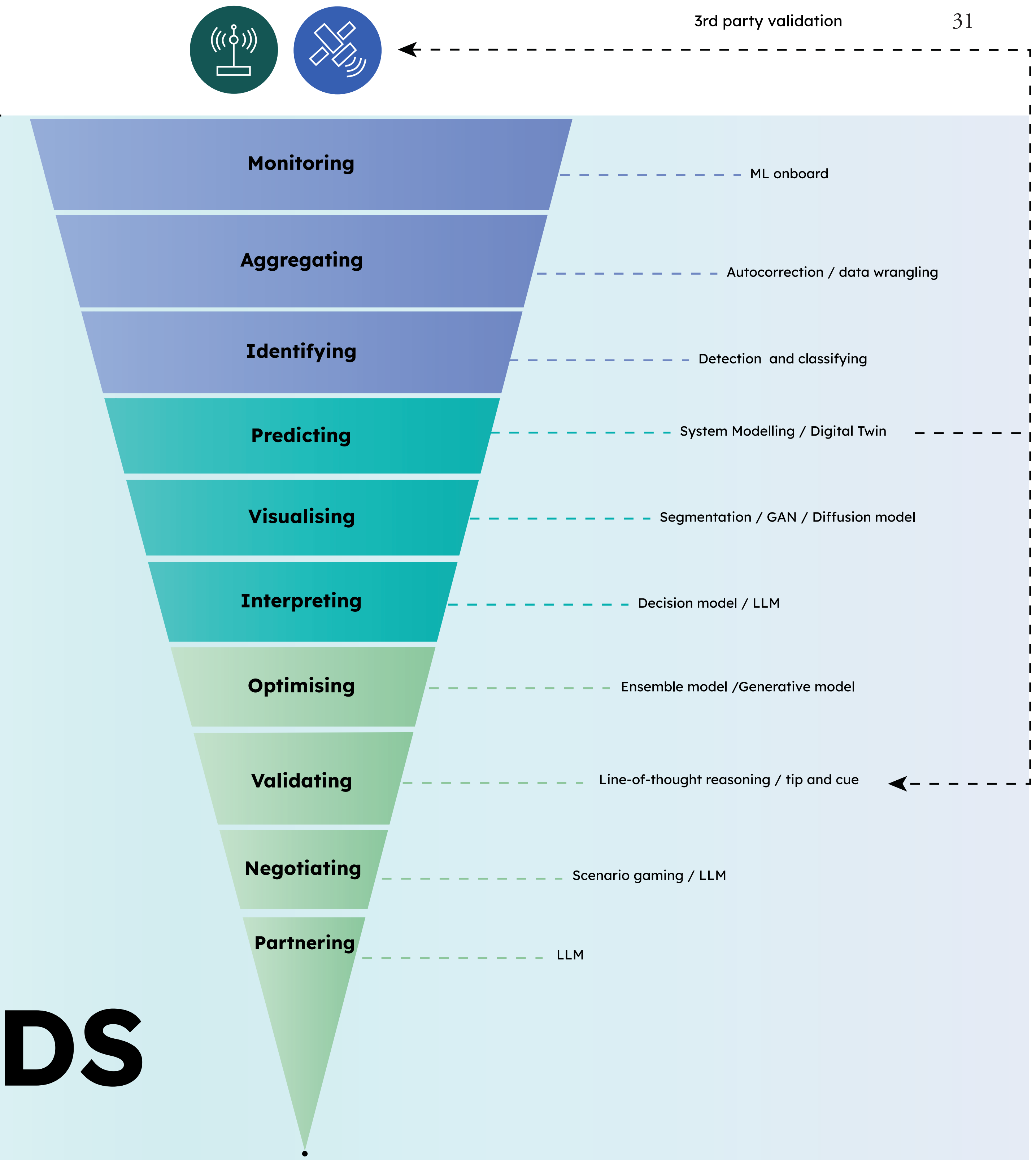
ESP merges artificial intelligence (AI), Earth observation (EO) and decision intelligence (DI) for planetary stewardship at the point of decision making.

Data

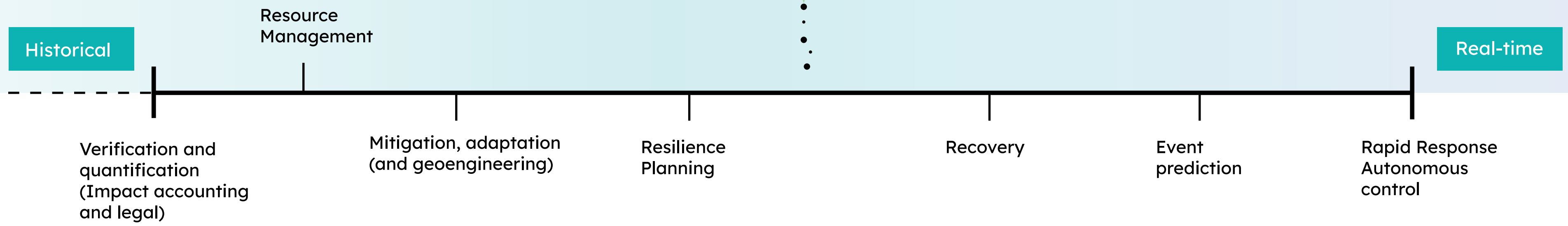
FROM DAYS

Integrated intelligence services along decision chain

TO MILLISECONDS



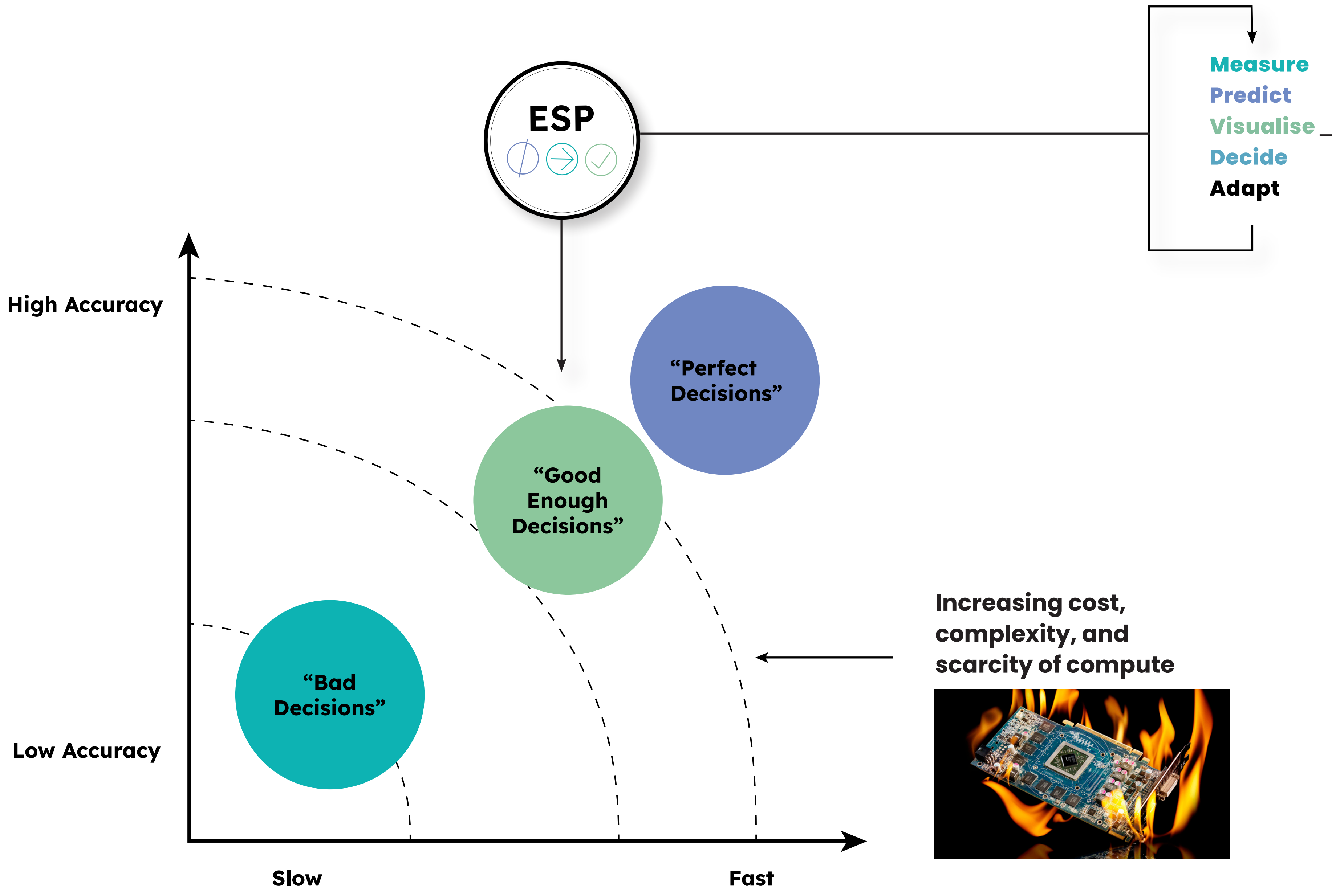
Decisions



"Good Enough Decisions"

SPEED, ECONOMY AND TRUST: THE 10 KEY PROPERTIES OF EARTH SYSTEMS PREDICTABILITY

These ten properties describe ESP systems, deployed in an open and accessible way, guided by stakeholder-centric impact, valuation and management.



Rapid Awareness
Gain insight from space

Integrated Systems
Joint modelling for better outcomes

Ultra-speed Predictions
Achieve predictions in seconds

Built-in Expertise
Leverage knowledge stores

Transparent Simulations
Explain reasons why

Trusted Operations
Validate with stakeholders

Compelling Visualisations
Facilitate understanding and insight

Informed Decisions
Explore decision scenarios and impact

Equitable and low cost
Prioritise access and ease-of-use

Adaptive Learning
Continuously integrate and improve

BUILDING ON EXISTING FRAMEWORKS

The concept of ESP, and the vision for how it might be deployed, builds upon existing work in Earth digital twinning, large-scale simulations, finance, ESG (environmental, social and corporate governance), SROI (social return on investment) and DI (decision intelligence), amongst many other disciplines.

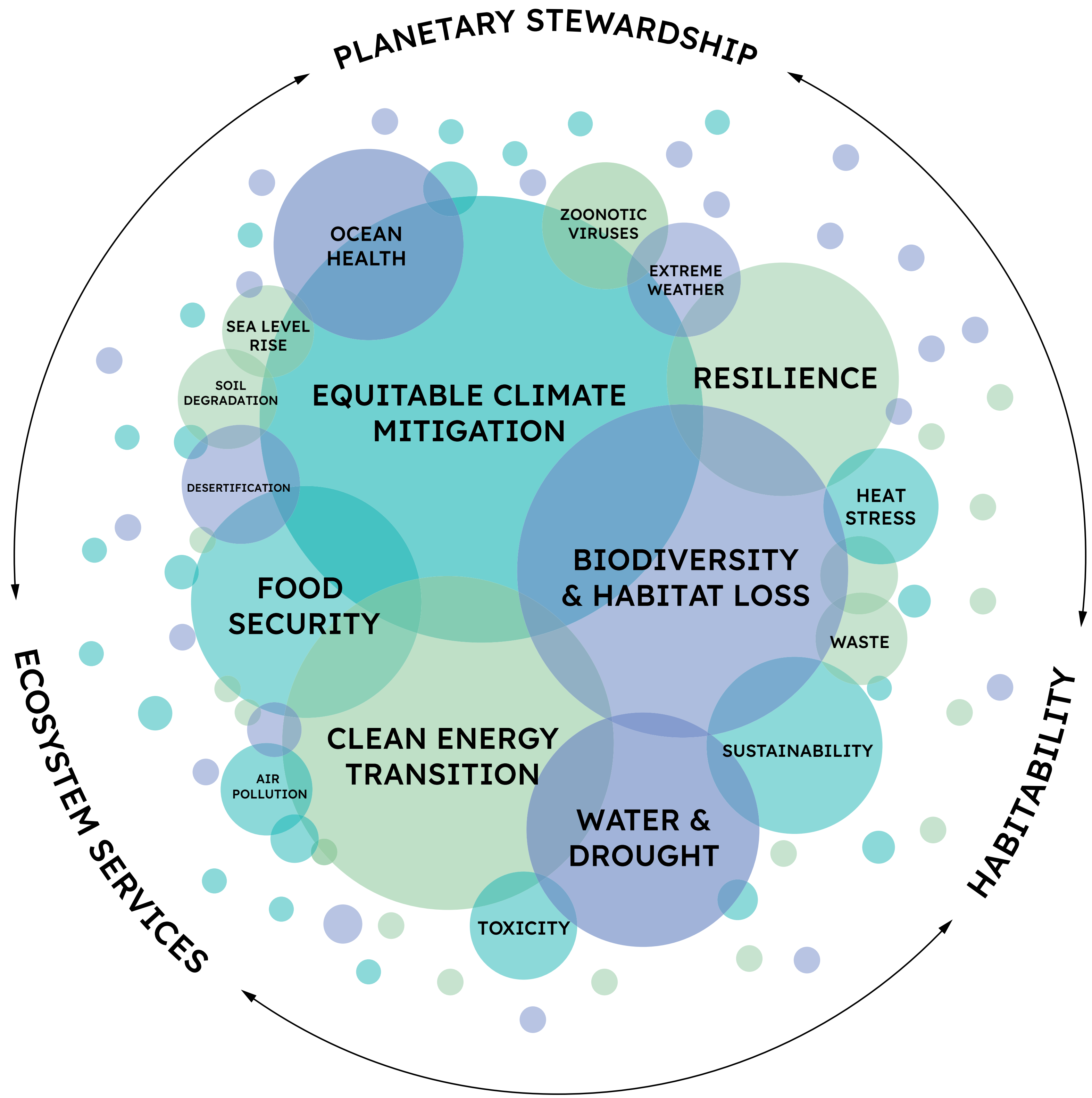
We delve deeper into many of these subjects as we explain the vision concepts of ESP in the body of this report.

For each vision concept, we also present **actions that have been recommended** by the scientific, technological and domain-specialisation communities through participation in the ESP Forum.

The first and highest level set of actions shown overleaf involves the definition of ESP itself and how it interfaces with the world.

To achieve the future we want, we must embrace a fundamental truth: **our economies and wellbeing are intertwined with Earth's systems, not apart from it.** While our sociological systems of the past have emphasised the need for technological or economic progress in order to overcome nature's limits, we now understand that we are embedded within nature.

Our activities as individuals and as nations affect the delicate balance of ecosystems in ways we now experience every day. ***Unless we measure our positive and negative impact on nature and human wellbeing, and integrate this knowledge in our decisions and actions, we are destined to destroy the very thing that nurtures us.***



ACTIONS

Establishing ESP as an emerging discipline.

The participants in the ESP Forum recommended the following actions:

→ Consult with the wider community to build and constantly validate a shared vision of Earth Systems Predictability.

→ Define both near- and long-term actions that will enable the ESP vision.

→ Create a technical and human semantic ESP roadmap that stakeholders can begin to fund and implement.

The remaining actions are gathered in the ESP Roadmap section of this report, placing them in the context of ESP technology and governance. If you are interested in a particular concept, turn to the ESP Roadmap section to find out how you could get involved.

SUSTAINABILITY

“Let’s make AI as central to the sustainability agenda for business and finance, as it is to the word itself.”

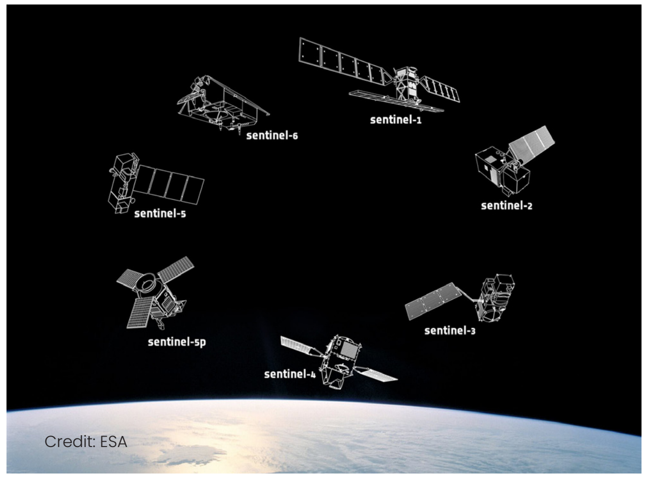
John Elkington
Chairman Volans
Creator of the Triple Bottom Line framework

ESP BUILDING BLOCKS

BUILDING BLOCK 1:
DATA FROM EARTH AND SPACE TO FEED AI

Satellites take the pulse of our planet.

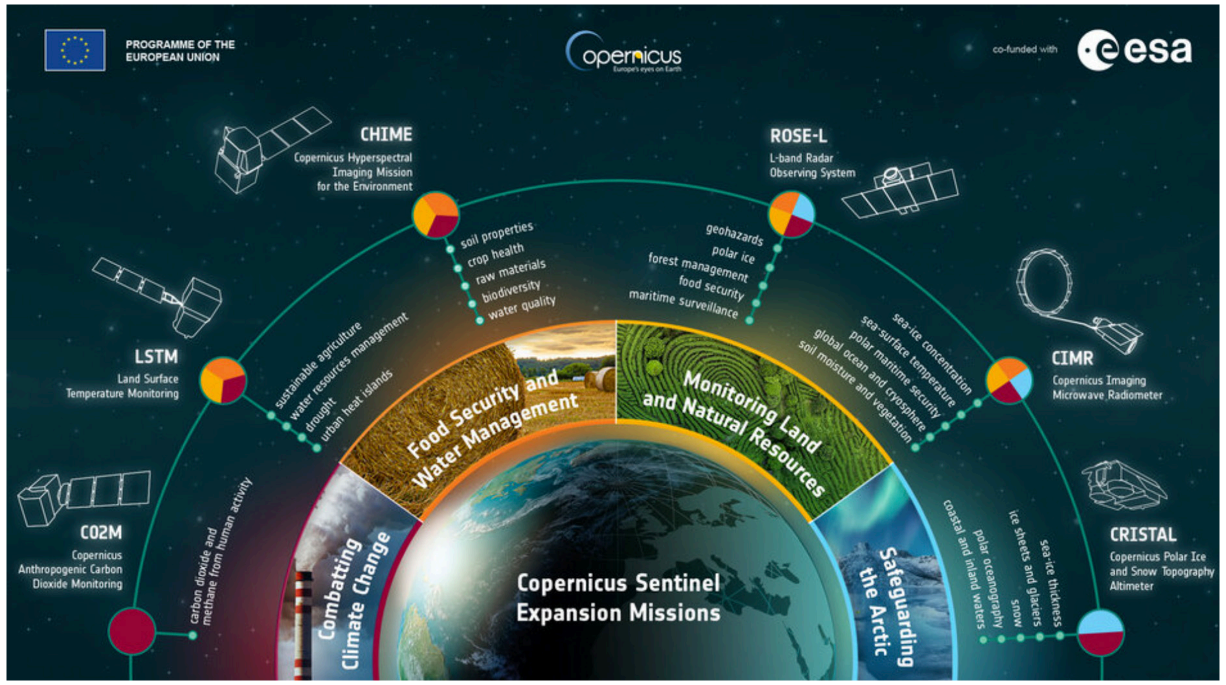
Cameras and sensors deployed in the real world are essential for feeding predictive models with data and keeping them grounded in reality. These include weather stations, landscape cameras or even social media feeds. But EO data is especially important for measuring over half of the key climate variables that feed ESP models.



Credit: ESA

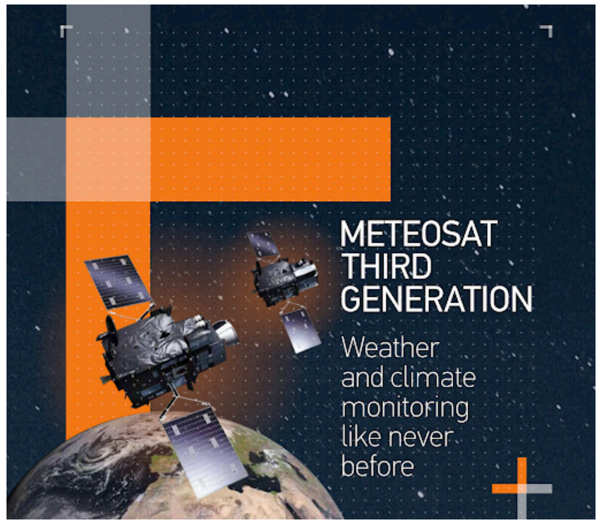
The European Space Agency's Sentinel missions monitor the surface of the planet - both land and water - to provide open data for a wide variety of uses. The Sentinel-2 mission has been particularly successful, imaging swathes of the Earth at 10-m resolution, every five days, at optical and infrared wavelengths. The NASA Landsat program fulfills a similar purpose, with a revisit time of eight days.

The Sentinel-1 mission complements Sentinel-2 by providing active radar images at a similar resolution of 10-m. Radar imagery can be acquired through clouds and at night, and so can provide an uninterrupted view of land use or situational awareness for emergency response. The two Sentinel-3 satellites are designed to measure sea-surface topography, sea- and land-surface temperature, ocean colour and land colour in support of weather and climate forecasting. Sentinel-5P monitors aerosols and trace gasses to build a picture of the atmospheric air quality, which is used in climate modelling. Similarly, Sentinel-4 and Sentinel-5 are upcoming atmospheric monitoring instruments that will be deployed by ESA partners on meteorological satellites. Finally, Sentinel-6 is an active radar mission that delivers measurements of global sea surface height in support of oceanography and climate science.



Six further Sentinel Expansion missions are being considered by ESA to address data gaps and respond to changing user needs. These include advanced instrumentation like hyperspectral cameras and novel microwave radar sensors, and specialised systems for monitoring carbon dioxide, land surface temperature and sea ice.

The ESA Earth Explorer satellite missions are designed to help answer key scientific questions on how Earth systems operate and how humanity is affecting natural processes. Proposed by the scientific community, they incorporate experimental technology and advanced design principles that are often used to develop operational missions like the Sentinels. There are ten missions in the current program, targeting scientific measurements of the atmosphere, ice, water, gravity, thermal radiation, forests, vegetation and changes due to motion. Ideas for Earth Explorer 11 are being considered for launch during 2031 - 2032.



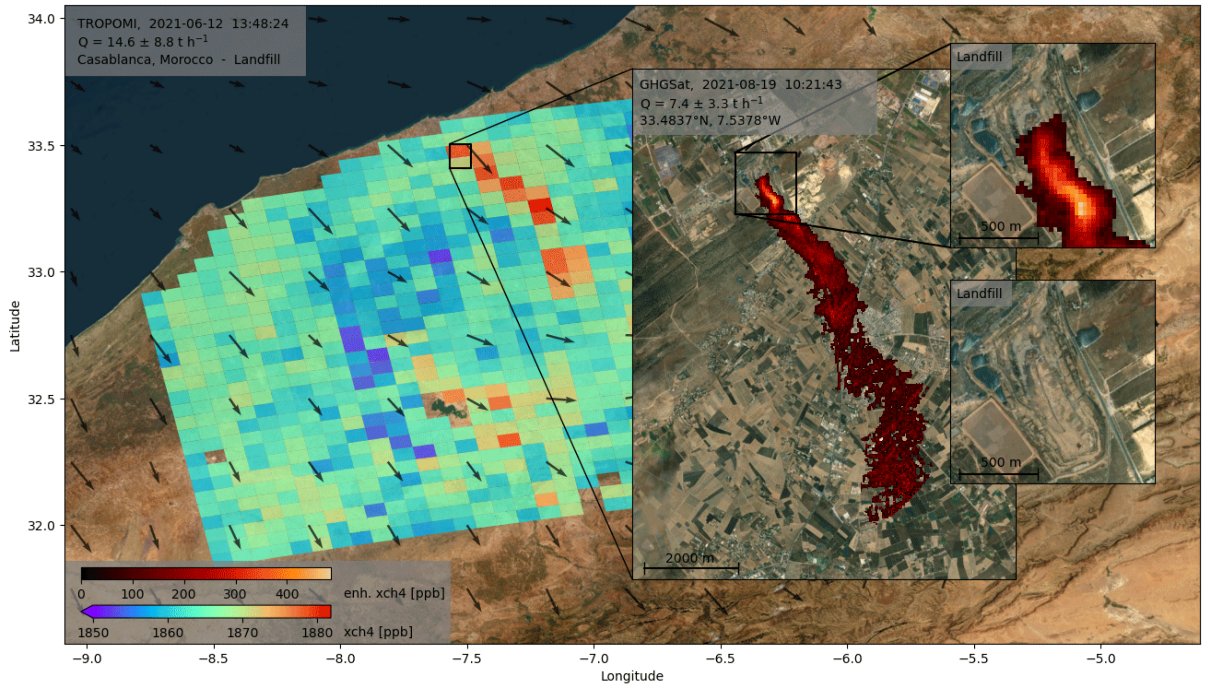
The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) is an intergovernmental European organisation that operates the geostationary satellites Meteosat-10 and Meteosat-11 over Europe and Africa, and Meteosat-9 over the Indian Ocean. It also operates two Metop polar-orbiting satellites in collaboration with the US National Oceanic and Atmospheric Administration (NOAA), and runs the Jason sea-level monitoring missions with the USA.

Earth observation data informs ESP-enabled decisions.

Earth observations data feeds into a host of use-cases.

Some examples are:

- Monitoring energy infrastructure.
- Predicting solar power generation from observed cloud coverage.
- Predicting wind power generation via weather modelling that is informed by EO and ground-sensor data.
- Measuring crop growth using EO data and crop yields models.
- Mapping land use and tracking how it changes; this includes urban expansion, renewable power capacity, carbon budget assessment, carbon sequestration, biodiversity and habitat change.
- Measuring GHG gas emission, especially methane, which has four times the heating effect compared to carbon dioxide.



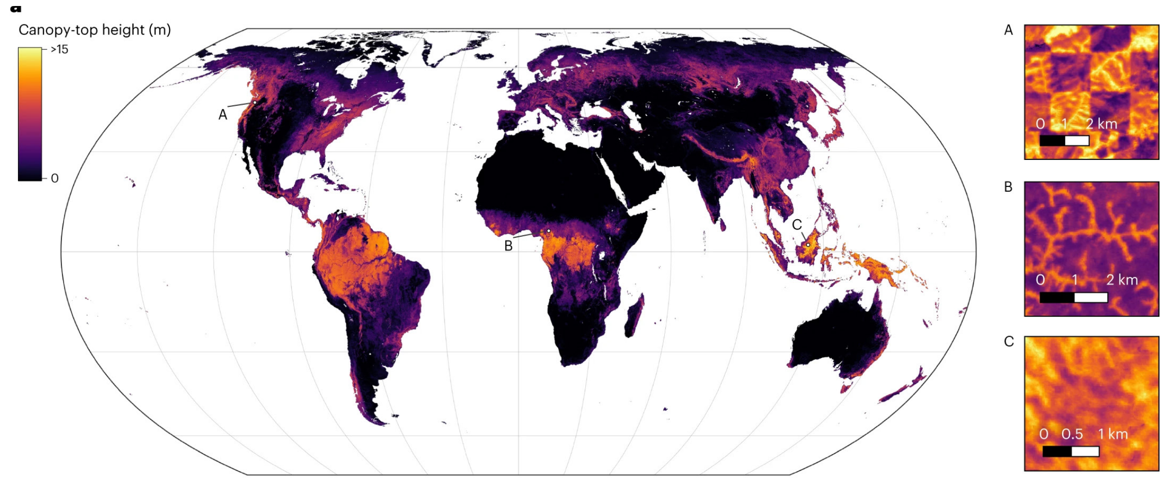
The image above shows methane plumes detected from Casablanca, Morocco on two different days during 2022 using TROPOMI and GHGSat, as presented in Schuit et al. 2023. Earth orbiting satellites are powerful tools for detecting emissions of methane and other greenhouse gases.

The ESA Climate Change Initiative (CCI) program

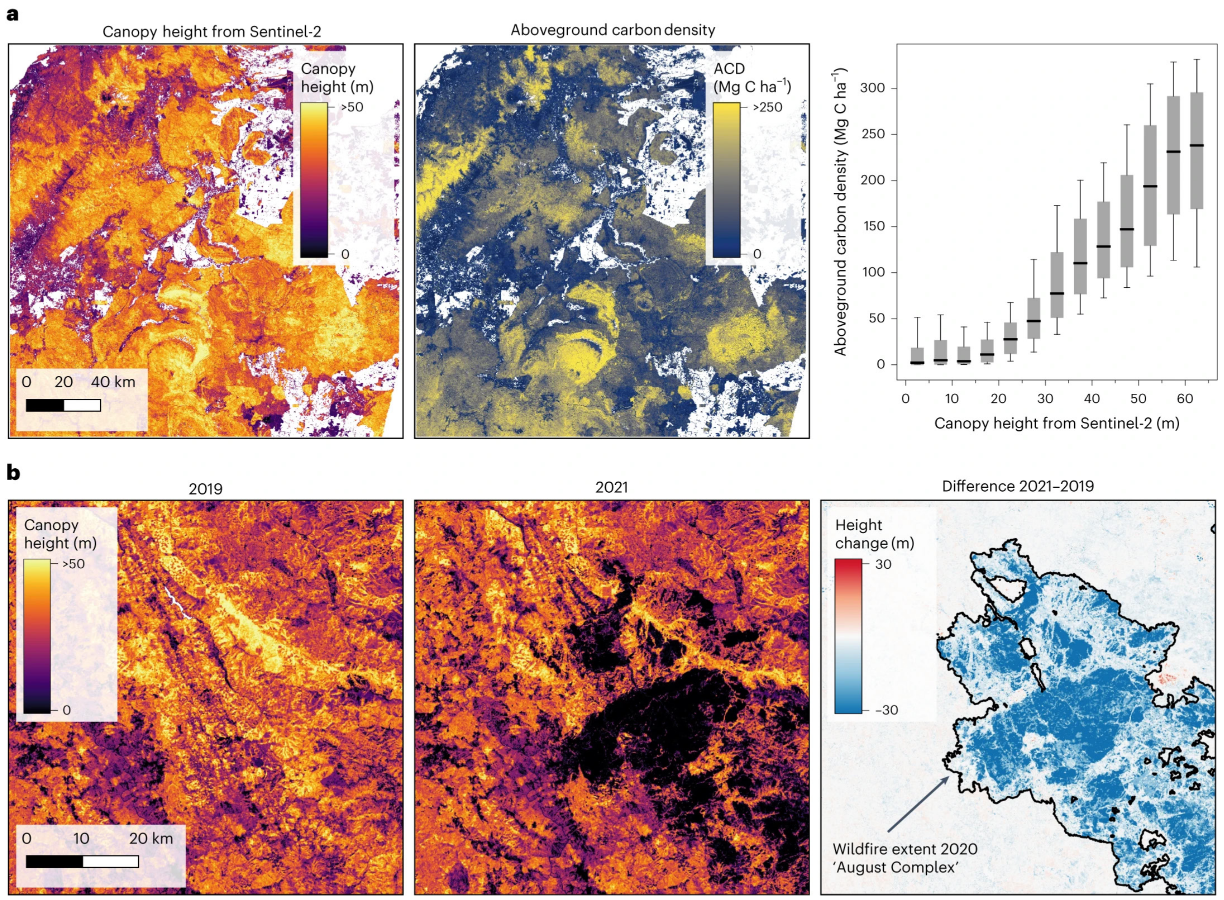
ESA runs the CCI program to support the measurement and publication of essential climate variables, many of which can be derived from Earth observation data. The CCI program contains 27 parallel projects focused on different climate variables and a dedicated climate modelling user group (CMUG) to link the observation and modelling communities. ESA provides data access through an open data portal, facilitated by a python toolbox.

NASA Advanced Information Systems Technology group

NASA's Earth Science Technology Office (ESTO) Advanced Information Systems Technology (AIST) group is working on multiple use cases: for Earth observations in digital twin Earths: wildfires, ocean carbon, water cycles, Central Africa carbon corridors, atmospheric boundary layers, coastal zone digital twins and national waterways.



Global canopy height map for the year 2020, estimated from ESA Sentinel-2 imagery fused with the GEDI spaceborne LiDAR instrument (from Lang et al 2023).



Examples for potential applications of the Global Canopy Height map. a: Biomass and carbon stock mapping in Sabah, northern Borneo. b: Monitoring damage due to wildfire in northern California (from Lang et al 2023).

Resource:
'A high-resolution canopy height model of the Earth'(2023)

nature ecology & evolution

Satellite-based global monitoring

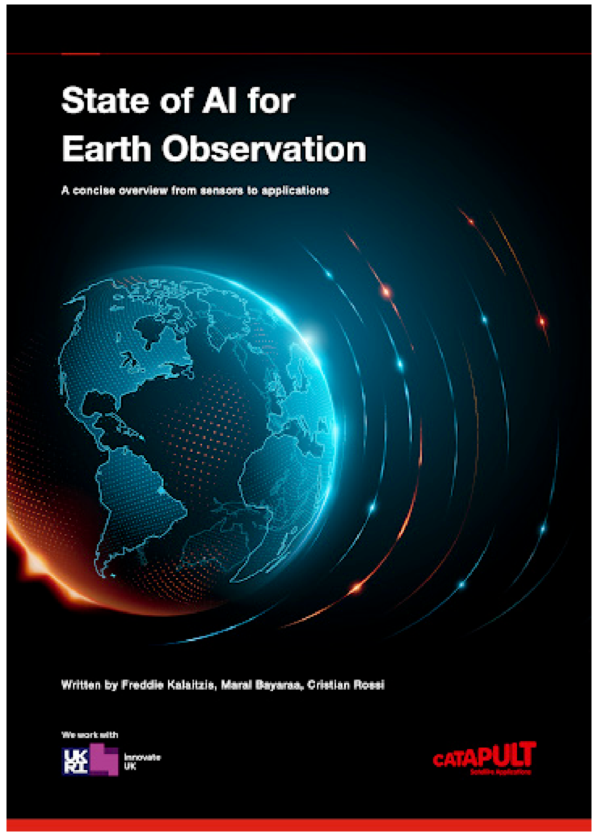
Monitoring Earth system variables and ecological health has become much easier since 2010 with the advent of cloud-based computing platforms for processing satellite imagery. Commercial services like Google Earth Engine offer access to historical and recent imagery from NASA and ESA, in addition to maps created using machine learning methods (e.g., near-real-time Dynamic World land cover maps). Microsoft Planetary Computer and Amazon Geospatial offer similar services.

ESA's new Copernicus Data Space Ecosystem is expanding to provide access to a host of Earth observation data, including from the Sentinel missions and the Copernicus Contributing Missions. It offers web-based data discovery and visualisation interfaces, APIs for querying and downloading data, and cloud-based processing tools.

Considerations for ongoing ecological monitoring

- Insight from monitoring programmes is critical for ecology and it needs to be both scalable and transparent. Data from bio-acoustic measurements, lidar and manual field surveys are all very valuable, but many of these methods are not scalable to global monitoring on realistic timescales
- Transparency and trust in monitoring programmes is especially important when thinking about financial incentives, carbon credits and the biodiversity economy.
- Successful large scale conservation requires collaboration with ecologists and regular site observation to validate ongoing progress (e.g. geobon.org).

Google Earth Engine



The recent white paper "State of AI for Earth Observation" by the Space Applications Catapult provides a concise overview of Earth observation technology and AI methods used to process data into useful information.

Faster access to Earth observation data

Some ESP deployment strategies require access to near-real-time EO data for situational awareness and feeding predictive models. Data from Europe's Copernicus Program is generally available within 3-6 hours, or even sooner (30 minutes) for dedicated services like the Copernicus Maritime Surveillance Service and CleanSeaNet, offered by the European Maritime Safety Agency.

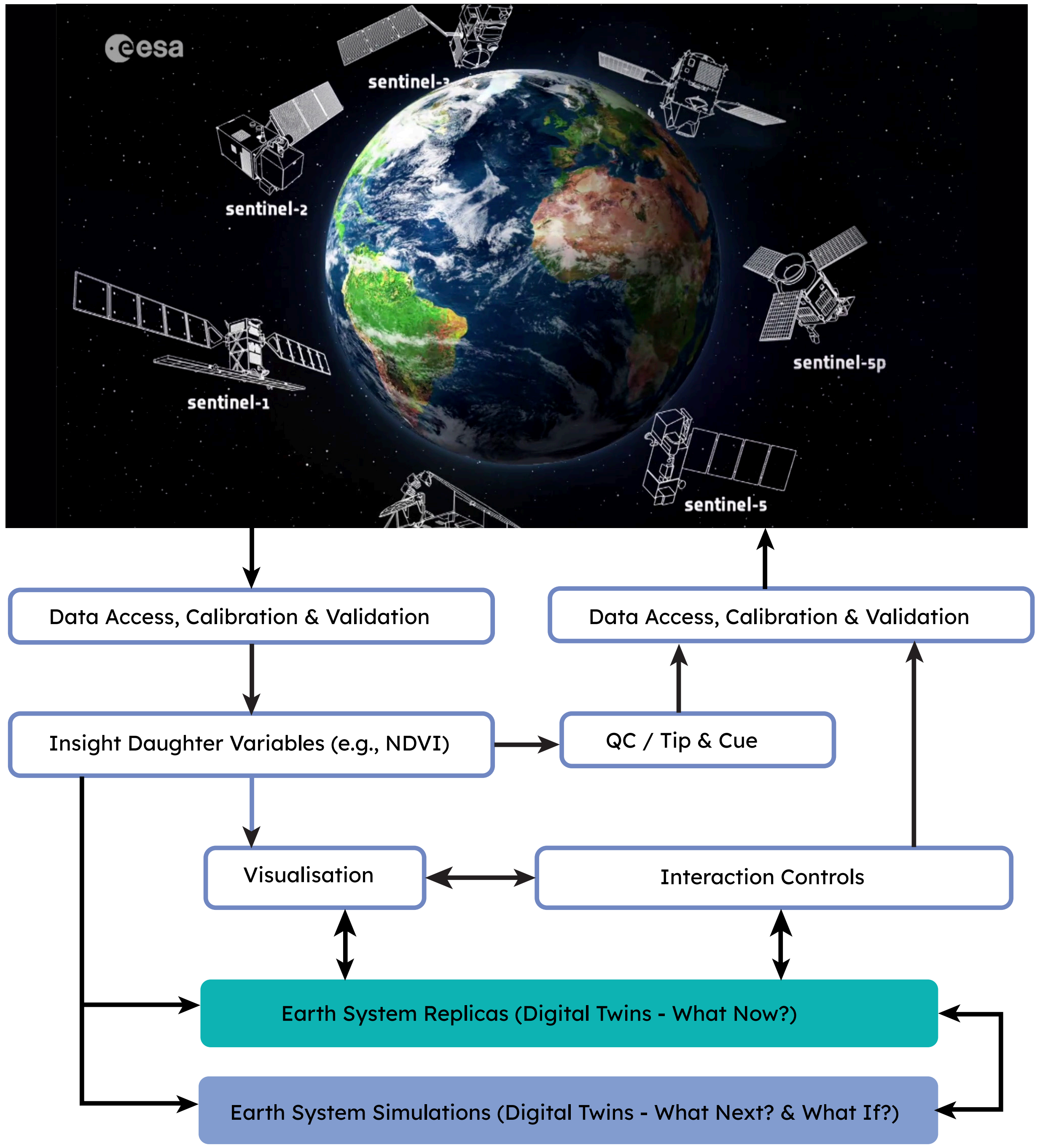
However, the lead-time for publicly available data on the Copernicus Emergency Management Service, or in cloud-based repositories (e.g., Google Earth Engine, AWS), can be up to one day. Commercial operators like Maxar, Planet Labs and Iceye offer accelerated access to high-resolution (< 1-m) daily imagery, but these services are costly, making them inaccessible to many users, particularly in the developing world.

We need faster access to EO data to support the management of dynamic situations and emergency response, which will become increasingly necessary as global heating drives extreme weather events.

The following can facilitate rapid access to insight from orbit:

- **Intelligent data compression.** Use ML onboard satellites to select only relevant data to compress and downlink.
- **Processing onboard satellites.** Use ML onboard to process raw acquired data into insight layers, which are almost always considerably smaller than the original data. This also negates the need to apply post-processing on the ground, but has the disadvantage that the transmitted data is tuned for a small number of use-cases.
- **Federated intelligence.** ML-driven federated learning is an extension of processing onboard satellites and facilitates autonomous decision-making within constellations of satellites, enabling smarter tasking and better use of limited bandwidth orbit-to-ground communication links.
- **Intelligent satellite cooperation.** Use intelligence onboard to take advantage of different types of sensor, processing capabilities, area coverage and instrument resolution on satellite constellations. Consider the 'tip and cue' strategy where a 'finder' satellite monitors a large area and tips a more specialised satellite to follow-up on an event detection.
- **Direct 5/6G dowlinks.** Transmission of kilobyte-sized data direct to end-user handsets (e.g., via Starlink, or Iridium networks).
- **Faster ground-based processing.** Deploy more high performance computing (HPC) infrastructure to shrink the time taken to process raw data into insight.

Critically, all of these systems and outputs would need to be **continuously validated** and subject to appropriate quality-control processes, to guarantee a high level of trust.



The infrastructure connecting EO satellite constellations to downstream services (like Digital Twin Earths, for example) are ripe for automation at the connection points; see the arrows in the above diagram.

- **Automated onboard calibration and validation** would facilitate automated decision-making onboard satellites, where decisions depend on quantified measurements. For instance, measurements of methane releases above a certain absolute threshold, or noise levels prompting re-observation.
- **Open-access visualisation and interaction controls** would place humans in the learning and decision loops, which is critical for AI reinforcement learning and safe operation of AI decision-making systems.

Barriers to access and use of Earth observation data in ESP

Machine learning can play a huge role in converting EO data into insights that are meaningful for stakeholders. It can also help enhance the spatial and temporal accuracy of EO imagery so that it is more broadly useful when paired with other data.

However, the greatest challenges lie with: 1) making EO data accessible, and 2) deriving useful tailored insight from EO data.

How do we enable more people to have access to EO data, interact with it and derive meaning from it?

Barriers to effective utilisation of EO data can be summarised as follows:

Lack of Awareness of EO Data

→ There is a general lack of **awareness** of EO data in most domains that stand to gain from it.

→ There is a lack of **understanding** of how EO data can translate to domain-targeted insight. Many stakeholders simply don't realise that EO data could be useful for their needs.

Lack of Availability of Suitable EO Data

→ For some stakeholders, EO data does not provide sufficient resolution or measurement **accuracy** to provide the confidence they need.

→ Some applications require a **high cadence** of return rate (sampling on short timescales) that open data programmes cannot currently provide. For example, tracking the spread of active wildfire may require better than hourly cadence under extreme weather conditions.

→ The **inhomogeneity or noise properties** of some EO data limits the overall availability of data for use in some use-cases (e.g., noise levels in some SAR data can limit the confidence of some analysis).

Technical Barriers for Using EO Data

→ Polarisation and phase information in SAR data, and spectral features in multi- and hyperspectral data require significant skill to analyse and interpret.

→ There is a lack of understanding of how AI and ML can be used to analyse EO data and what the advantages are in terms of accuracy, speed, scale and other variables.

→ Viewed through the lens of equity and accessibility, EO and AI systems are technically difficult to use, especially for those who already feel behind the curve in utilising some of these technologies because of lack of skill or curtailed access to education - something that disproportionately affects the developing world.

Business Barriers for Using EO Data

→ Example use cases often focus on data utilisation and miss the crucial point, which is the value extracted from the insights gained from EO data for stakeholders down the supply chain.

→ We need transparency on how data and insights are being used to create meaningful change that can be measured (e.g., quantitative impact maps) to convince people of the utility of EO data and AI.

→ Governance and policy barriers can also be issues, including lack of data standards and poor understanding by policy makers.

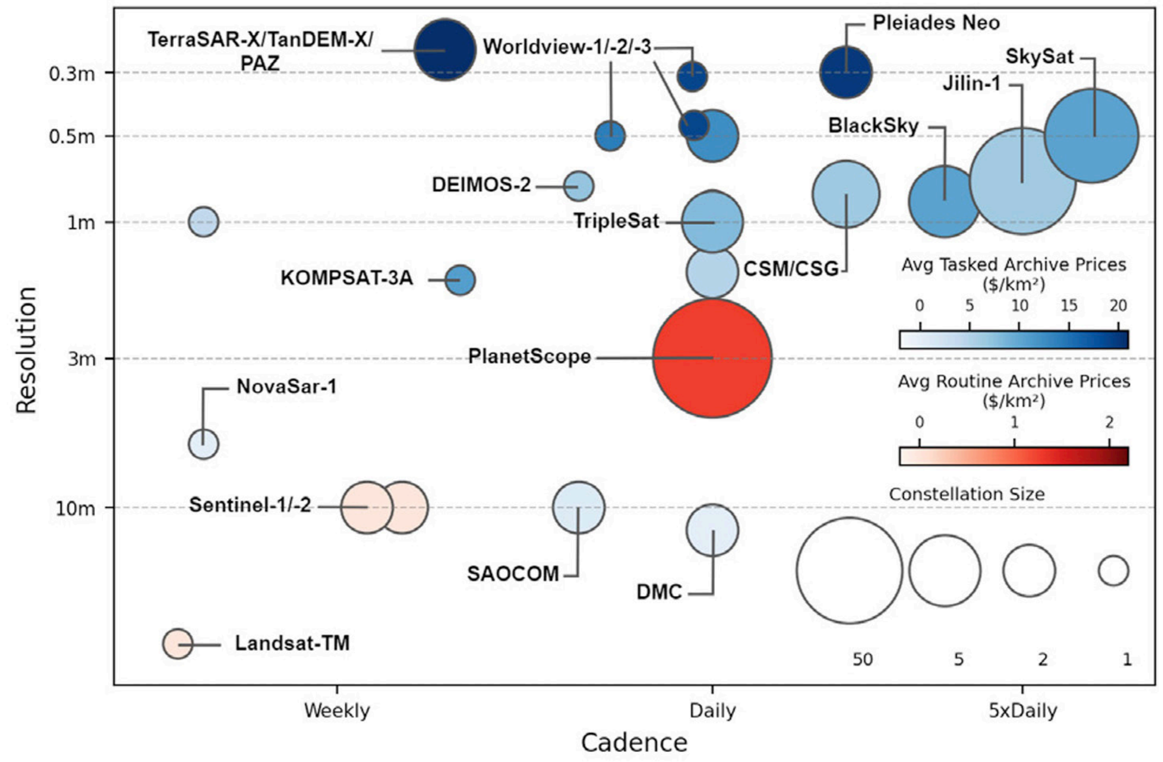
→ We need to provide ready-made incentives for using EO data:

- New economic incentives (gains, efficiencies) in different arenas.
- Proven business models for collection, dissemination and management of Earth observation data.
- Validated case studies quantifying the value creation of that economic impact.

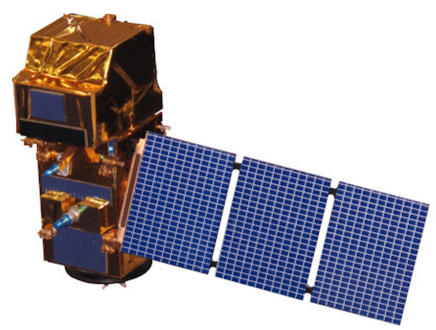
Earth observation platforms, from McCabe et al 2017.

Current EO Capabilities

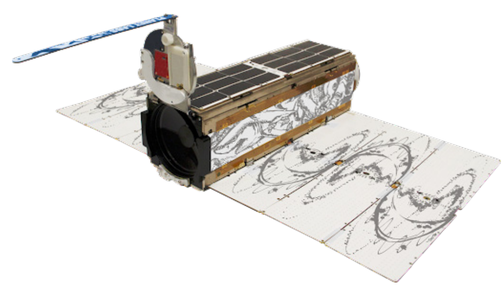
Over the last few years, there has been a significant increase in imaging satellites deployed to low Earth orbit (LEO) for scientific study as technology demonstrators and/or for commercial purposes. Most EO satellites observe optical and infrared bands but the fleet of synthetic aperture radar (SAR) instruments is also increasing. Our EO capabilities are unprecedented and improving all the time.



Revisit period vs. spatial resolution of selected EO satellite missions, classified by constellation size. Reprinted from the 2022 white paper on the [State of AI for Earth Observation] by the Satellite Applications Catapult.

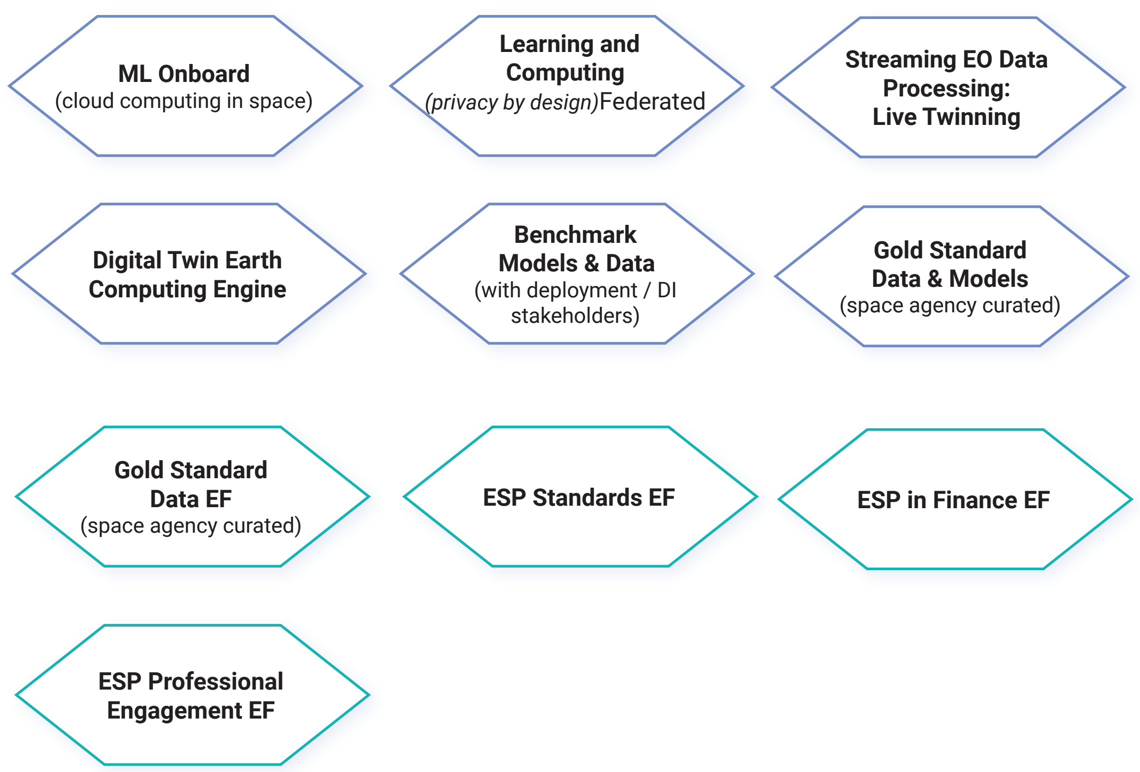


Copernicus / ESA Sentinel 2
10 metre resolution every 5 days



Planet Dove Constellation
3-5 metre resolution every day

Linked Roadmap Actions



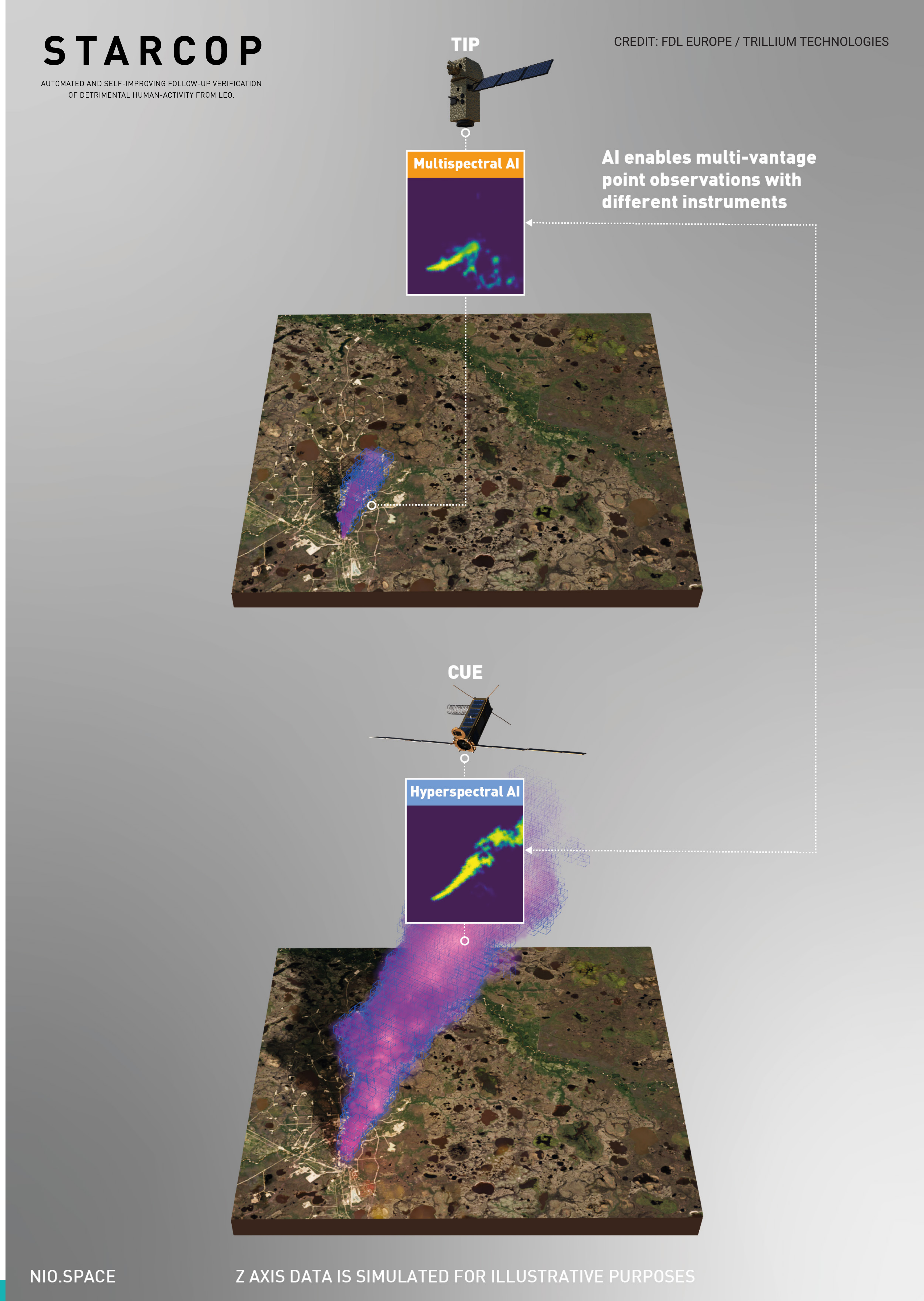
These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

The ESP Roadmap is a living document - we invite you to get involved in ESP development by reviewing it, adding actions, making suggestions and recommending resources.

Resource:
Open data tools.

Pangaea: an open, long-lived data store.

CKAN: an open-source data management system.



BUILDING BLOCK 2:
AN AI PARADIGM SHIFT

AI HAS COME OF AGE AS A TRUSTED TOOL

What's new?
Technological leaps by AI

Progress in AI and ML has accelerated in the last five years, leading to major innovations including

1. Physics-informed machine learning
2. Transformer Models
3. Foundation Models

1. Physics-informed ML

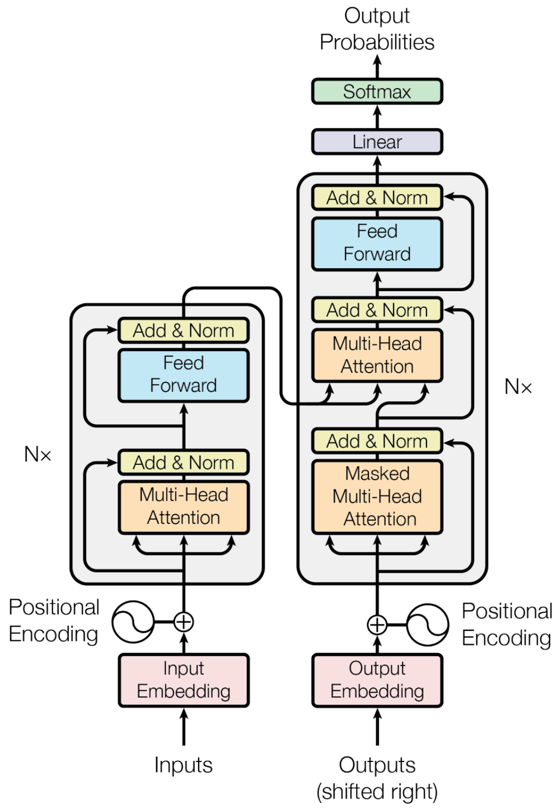
Physics informed ML promises to hugely accelerate simulations of Earth systems, while better incorporating real-world data and the modelling of systems on a wide range of scales.

2. Transformer Models

This active area of development includes Large Language Models (LLMs) and Vision Transformers (ViTs), which are revolutionising the way humans interact with machines, and are giving machines human-like capabilities.

3. Foundation Models

Foundation models are large models that generalise well and be adapted to different tasks. They are expensive to pre-train but promise to democratise AI and lower the cost of participation.



The architecture of the Transformer model is presented in the seminal paper '*Attention Is All You Need*' by researchers at Google corporation. The Transformer can be considered a 'universal architecture' and has led to a revolution in deep learning and generative artificial intelligence, starting with natural language processing.

1. Physics-informed ML

Modelling of multi-scale dynamic systems like Earth's is traditionally done by solving their governing partial differential equations (PDEs) through numerical methods. While these **physics-based models** can be accurate, they are costly to develop and run, **too slow to provide real-time predictions** and may lead to incorrect predictions because of unresolved physics questions.

In addition, physics-based models **struggle to incorporate 'messy' real-life data**, which can be noisy, have missing values or mischaracterised uncertainties and be sampled on disparate spatial and temporal scales.

Traditional machine learning, in contrast, excels at empirically modelling complex multi-scale systems if enough data is available for training an appropriate model. However, these ML systems often learn unintended observational biases, leading to poor predictive performance: they fail to generalise.

A new branch of machine learning, called **"Physics-informed machine learning"** addresses these challenges. It incorporates fundamental knowledge of physical systems into the ML algorithms, either in how they are trained (e.g. via physical loss functions and data manipulation) or into the architecture of the models themselves. These hybrid models have 'best of both worlds' advantages:

- They can be **thousands of times faster** than numerical models.
- They can **sample sub-grid physics** for more accurate predictions.
- They can **easily incorporate real-world data** of different resolutions.
- They can be more accurate and **generalise better than traditional ML models.**
- They require **less training data** compared to traditional ML models.

For example, the development of Fourier Neural Operators (FNOs), Adaptive FNOs (AFNOs), Spherical FNOs (SFNOs) and have led to giant leaps in weather forecasting capabilities and promise to revolutionise global climate modelling.

Resource:
A review of 'Physics-informed machine learning' in NATURE (2021).

nature reviews physics
Physics-informed machine learning

Resource:
'Physics-aware machine learning revolutionizes scientific paradigm for machine learning and process-based hydrology' (2023).

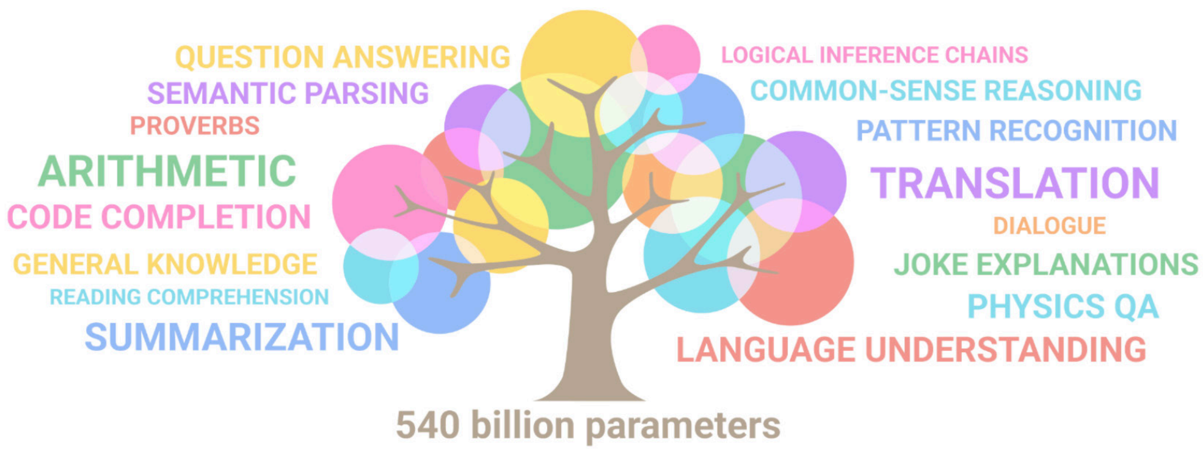
2. Transformer Models

Large language models (LLMs) based on the Transformer architecture are becoming the 'glue' of AI systems. They can perform certain complex tasks with reasonable skill, such as writing computer code, composing long-form answers to questions, composing artworks and generating lifelike photos from text descriptions. These types of complex and creative tasks can now be automated, to a degree, depending on the quality of output desired. However, we must think carefully about how to deploy AI to support of human capabilities, rather than supplant them.

Currently, LLM-based digital assistants are prone to 'hallucinations' (generating entirely plausible spurious results), but research is progressing rapidly to mitigate these issues and increase explainability (e.g., via Retrieval Augmented Generation).

The new and emergent capabilities of LLMs appear to scale directly with their size; models with more parameters are generally more skillful and capable (although algorithmic innovations are also leading to enhanced capabilities; the MISTRAL 7B model, for example). However, their large sizes also imply trade-offs:

- They require large amounts of computing power and data to train successfully, meaning that creating LLMs is a very expensive task.
- Trained LLMs still require server-class hardware to run (although much less compute power than for training), meaning that they are usually deployed on cloud-computing infrastructure, rather than on local computing devices.



Visualisation of the tasks supported by Google's Pathways Language Model (PaLM), which has 540-billion free parameters, making it one of the largest machine learning models in the world (Credit: Google AI Blog).

3. Foundation Models

Foundation models are large ML models that are poised to revolutionise science. They have five properties that make them widely useful:

- They are trained on a very broad corpus of data
- They are large enough to learn general representations of knowledge in the data
- Their architecture supports multiple inference tasks without significant re-training
- They leverage self-supervised learning on unlabeled or weakly labelled data
- They support multimodal inputs (different types and dimensions of data)

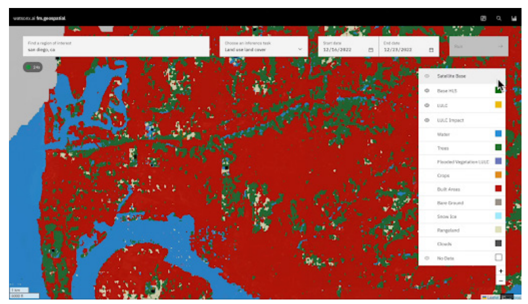
The emergence of foundation models may be as fundamental as deep learning was a decade ago, lowering adoption barriers for developing new applications and setting high standards for performance and reliability across a wide range of domains.

Most new foundation models are based on the Transformer architecture and contain billions of parameters. Their large size means that they effectively learn meaningful representations of most of the knowledge in their training data, which can be many terabytes in size. They are critical for broad and diverse applications, such as disaster response, where a powerful deep learning technique developed for one scenario may fail to work in adjacent use cases.

Foundation models are expensive to train and currently require cloud computing facilities to run. **However, the power of foundation models lies in the fact that they can be adapted for a very wide range of applications with little, or no, retraining.** This has the potential to fundamentally change how ML is deployed by removing significant barriers of cost and expertise.

i

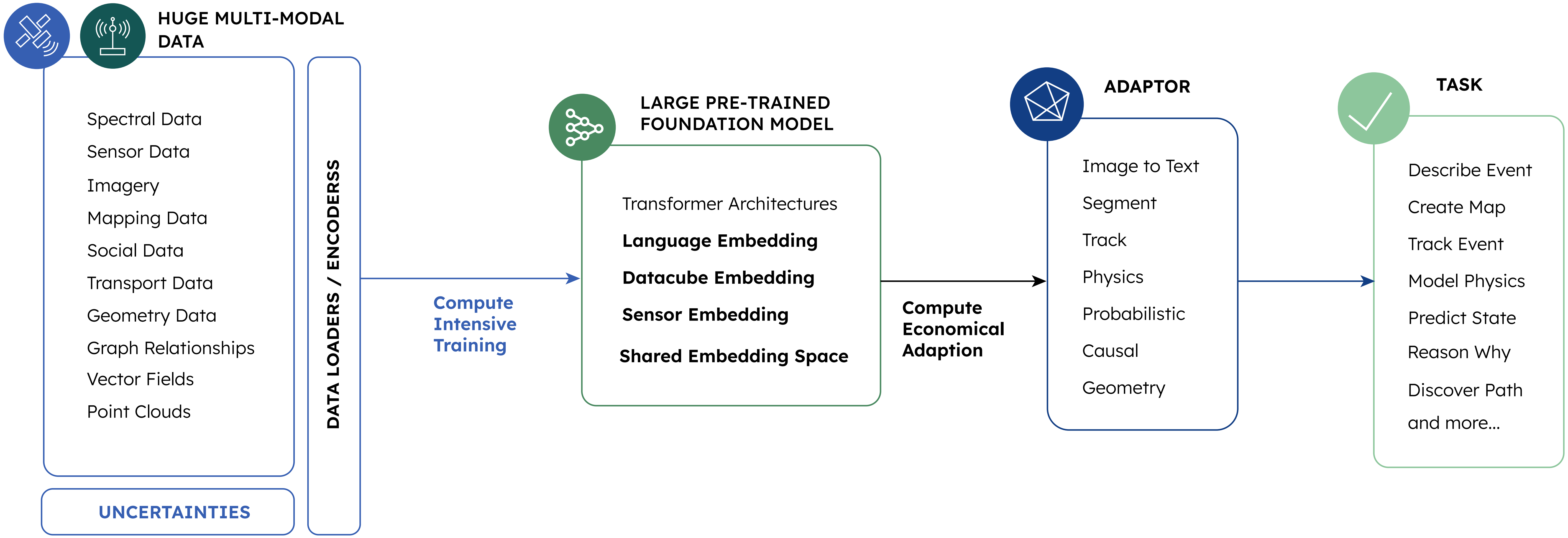
IBM and NASA have open-sourced a new large geospatial foundation model on the Hugging Face platform. The model was trained on the Harmonised Landsat Sentinel-2 dataset and can automatically generate maps of water and different types of land.



FOUNDATION MODELS

Schematic of a foundation model.

Pre-trained foundation models are large-enough to encode information supporting multiple tasks via lightweight adaptors. The process of creating adaptor modules requires far less computing power than the initial pre-training step.

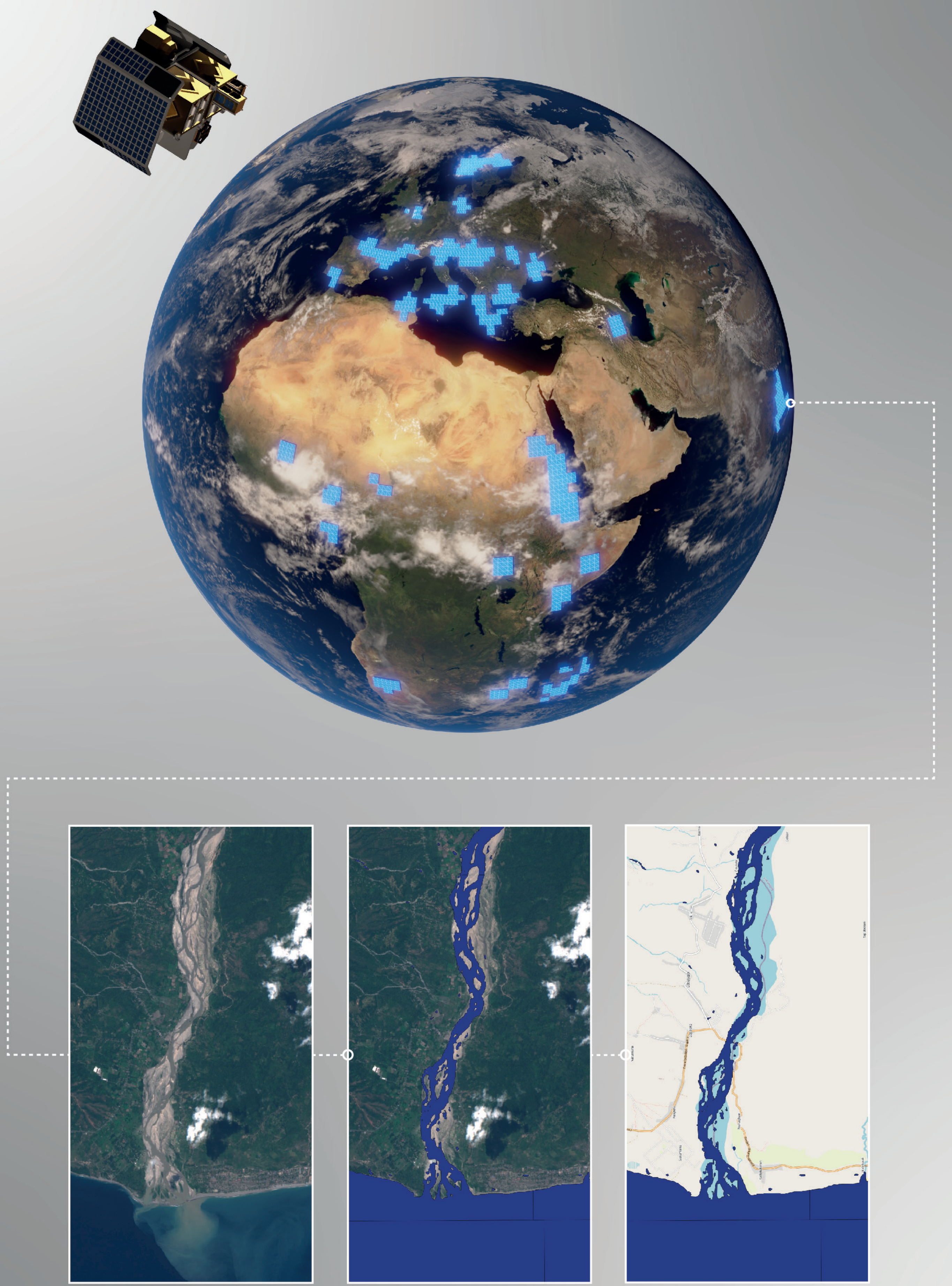


Linked Roadmap Actions

- **MLOps for ESP** (human in the loop)
- **Benchmarks** (with deployment / DI stakeholders)
- **Foundation Models** (multimodal)
- **ESP LLMs** (language + ESP knowledge)
- **Equity & Access EF** (compute, tools, data)
- **ESP Access Support Program**
- **Decision Intelligence EF** (integration)
- **ESP Challenge Program**

These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

CREDIT: FDL EUROPE / TRILLIUM TECHNOLOGIES



WORLDFLOODS ml4floods.com

BUILDING BLOCK 3: AI WEATHER AND CLIMATE MODELS

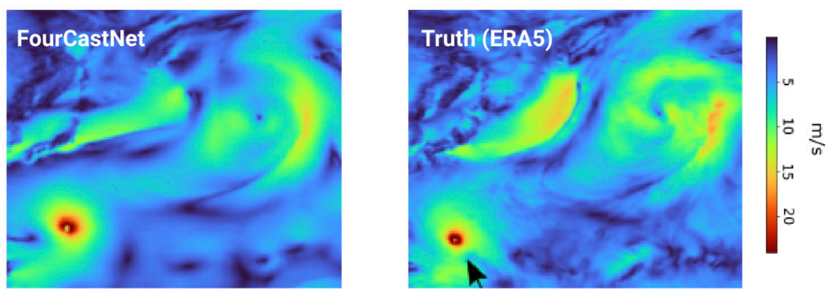
AI IS UNLOCKING RAPID AND DEMOCRATISED WEATHER PREDICTION.

AI-driven emulators of weather models (such as NVIDIA's FourCastNet, Huawei's Pangu-Weather and DeepMind's GraphCast) can generate accurate predictions over 10,000 times faster than traditional physics-based models. This capability is game-changing:

- **Large ensembles** containing hundreds, or even thousands, of models can be run for low computational costs, mapping out large volumes of parameter space in a few seconds on a GPU.
- The results of thousands of AI ensemble models can provide well-calibrated and **constrained uncertainty estimates** compared to their numerical counterparts, which only have of order ~50 members.
- Large AI ensembles can sample the long tails of forecast distributions, allowing **more reliable predictions of extreme events.**
- AI weather models are much less power-hungry than their numerical counterparts, meaning that they have a **lower energy** and carbon footprint.

These advantages in speed and cost mean that AI weather modelling can be run at-will in support of hypothesis testing. Their efficiency opens the possibility of interlinked modelling of Earth and human systems to explore multiple scenarios in support of climate-friendly decision making.

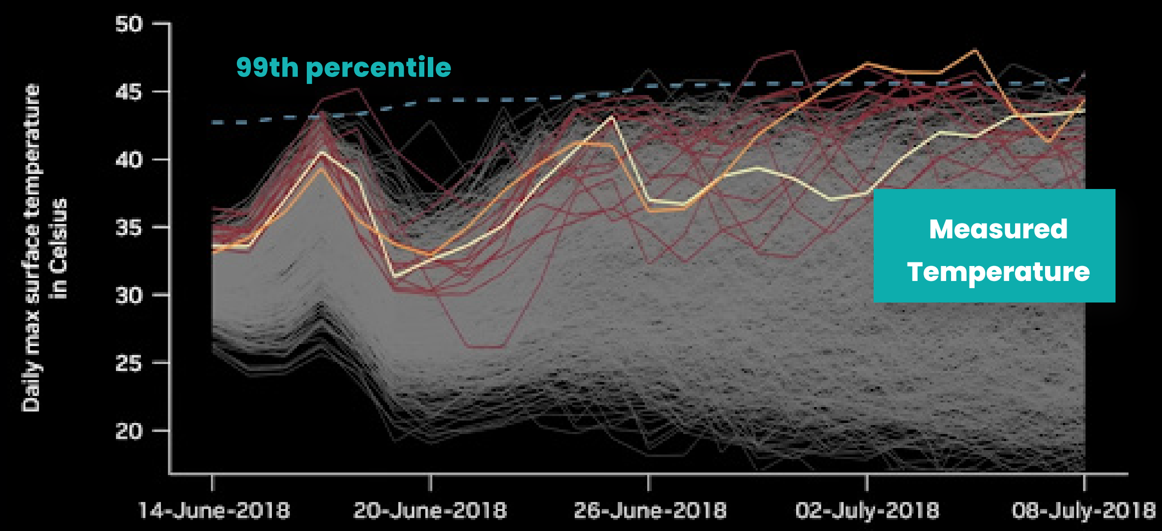
AI models (including physics-informed models and emulators) are still highly dependent on the quality of their training datasets. The ECMWF ERA5 global reanalysis dataset has been essential for training the latest weather models. We expect that future improvements in the quality of measured data will lead AI weather and climate models to encode fine-scale physics and surpass the current best numerical models.



The near-surface wind speed predicted by NVIDIA's FourCastNet AI weather model compared to the truth (ERA5 data) at a lead time of 96 hours. The arrow points to Typhoon Mangkhut forming during September 2018. The resolution of the simulation is 25° (~30 km).

2018 Heatwave in Algeria - 51.3 C

FourCastNet Ensemble Temperature Predictions



Large ensembles are critical for exploring the likelihood of extreme weather events. Here 12 FourCastNet models cross the 99th percentile line for the predicted temperature, compared to none in an ensemble of 50 numerical weather prediction (NWP) models (Credit: NVIDIA Research).

Resource:
'How AI models are transforming weather forecasting: a showcase of data-driven systems' (2023)

ECMWF

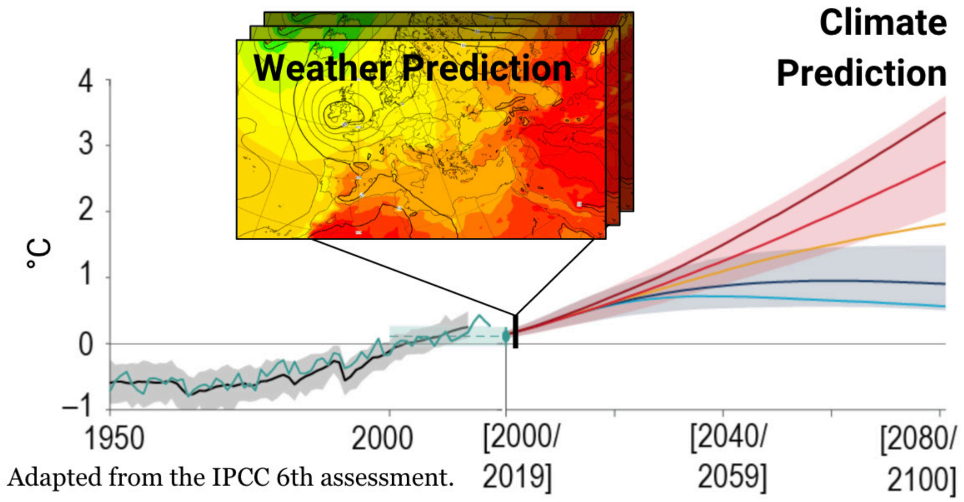
Climate vs weather models: differences in scale and aims

Climate and weather modelling are based on similar physics, but differ fundamentally in scale and operation.

Weather models aim to accurately predict high-resolution maps of atmospheric phenomena on kilometre scales and hourly intervals, given a detailed set of initial conditions. The chaotic nature of Earth systems on these scales mean that propagated uncertainties increase to unusably large values after about 10 days. However, daily weather observations can be assimilated into the latest model to continuously correct predictions.

Climate models aim to predict future climate statistics decades into the future, driven by changes in the energy balance of large-scale Earth systems. This balance is governed by forcing, an externally driven imbalance imposed on the climate system (including by human activities), and feedback, internal processes that amplify or dampen changes due to forcings (see https://doi.org/10.17226/11175).

Climate change manifests as long-term trends that are only separable from natural variability after decades of in-model time. Unlike weather modelling, there is no opportunity to assimilate new observations, so climate models are limited by their uncertainties, errors and unresolved processes (see Schneider (2023).



Adapted from the IPCC 6th assessment.

FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators

FourCastNet is a data-driven, global, medium-range weather forecast model created by a team at NVIDIA Research. It can produce a 10-day weather forecast in less than two seconds on a commercial GPU that is as skillful as the equivalent numerical weather prediction NWP model.

FourCastNet also only requires a small fraction of the full global state vector, which makes it much more robust than the NWP model.

This AFNO-based model has recently been improved to incorporate spherical harmonics, accounting for the fact the Earth is a sphere, not just a flat grid, and the skill of this AI-based technology is expected to improve.

Accurate medium-range global weather forecasting with 3D neural networks

Pangu-Weather introduced a new AI model for accurate, medium-range global weather forecasting. The developers showed that three-dimensional deep networks equipped with Earth-specific priors are effective at dealing with complex patterns in weather data and that a hierarchical temporal aggregation strategy reduces accumulation errors in medium-range forecasting.

Trained on 39 years of global data, Pangu-Weather obtains stronger deterministic forecast results on reanalysis data in all tested variables when compared with the world's best NWP system: the Integrated Forecasting System of the European Centre for Medium-Range Weather Forecasts (ECMWF). When initialised with reanalysis data, the accuracy of tracking tropical cyclones is also higher than that of the ECMWF-HRES physics-based model.

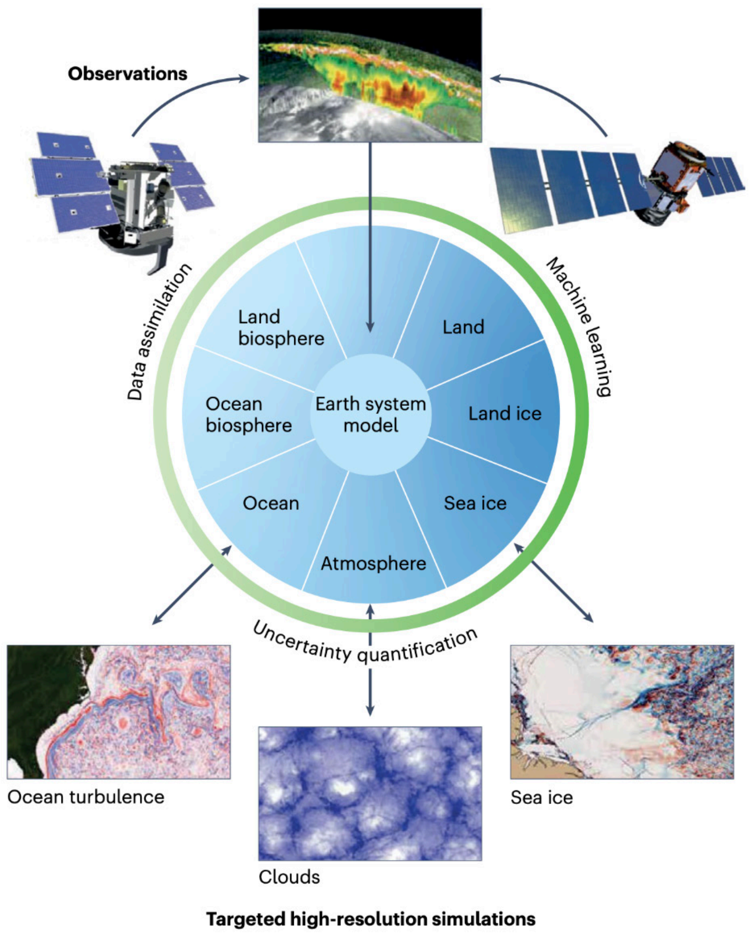
Resource:
'Inductive biases in deep learning models for weather prediction' *(2023)*

A vision for practical AI-enhanced climate predictions

Schneider et al. (2023) outlines a vision for creating the climate risk assessments needed by decision makers. Instead of pursuing kilometre-scale models, they settle on a moderate increase in spatial resolution to 10-50 km, which would still allow the generation of large (10,000) ensembles to quantify uncertainties and begin to resolve tropical cyclones and mesoscale turbulence.

They focus on:

- Calibrating the climate simulations with Earth observation data.
- Using unsupervised, or self-supervised AI, with ensemble Kalman methods to learn about unresolved processes (e.g., cloud microphysics).
- Move to high-resolution (1 km scale) regional modelling to support detailed hazard assessment.

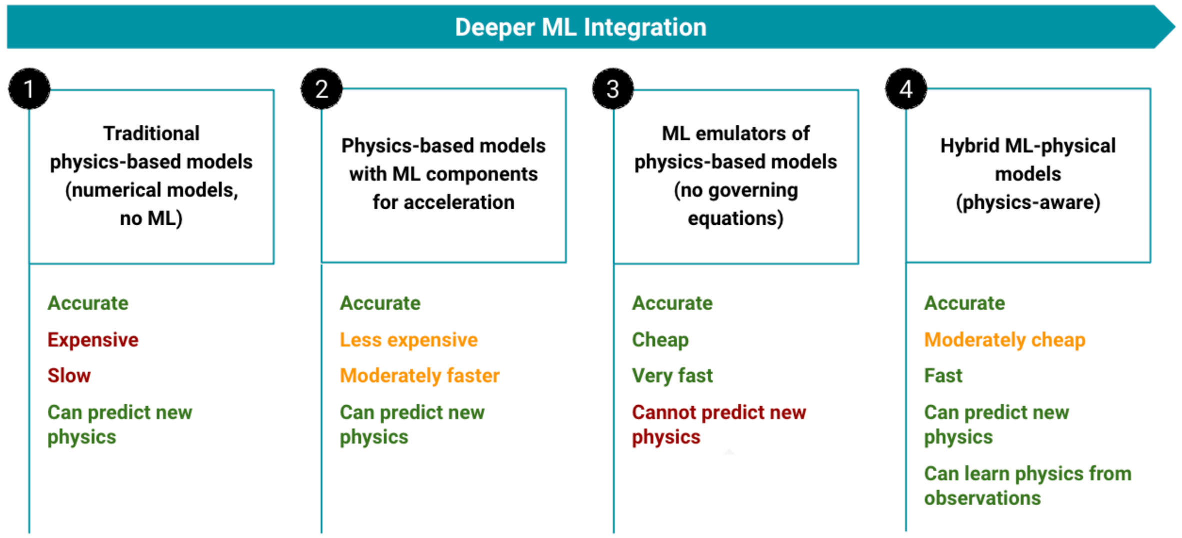


Resource:
'Harnessing AI and computing to advance climate modelling and prediction'(2023)

nature climate change

AI-driven global climate modelling efforts

The application of machine learning to global climate models can be considered along an integration scale:



Traditional physics-based numerical models (1) sit at the bottom of the scale, with the advantage of high accuracy at the cost of high compute cost and speed. These physics-based models can be accelerated by replacing components (e.g., PDE solvers) with faster ML algorithms (2), but the speed gain is relatively small. ML emulators trained on physics-based models (3) are proving to be as accurate as the physics-based models, but orders of magnitude faster. However, they cannot learn new physics that has not already been revealed by a physical model. Hybrid physics-aware ML models (4) potentially offer the best of both worlds: speed and accuracy, with the ability to predict new physics or even learn from observations.

Machine learning also has a significant part to play in downscaling climate predictions from coarse to fine resolutions, as covered later in this report.

The application of machine learning to climate models (global and regional) is a very active area of research. Bauer et al. (2022) make the case that climate and weather models should be developed between public- and private-sector institutions. For example, at team at NVIDIA research are training AI climate model emulators from CMIP-6 data, generative post-processing models for downscaling and precipitation, and models that can compress, reproduce and visualise extremely high resolution data in a process known as tethering. A comprehensive review of public research is provided by O. de Burgh-Day et al (2023).

Resource:
'Machine Learning for numerical weather and climate modelling: a review' (2023)

The need for better climate observations to support hybrid models

In an evolving global climate, future states cannot be predicted by training machine learning models on past data in the same way as for weather models. The longer time-scales involved mean that many more variables influence the evolution of the climate, like the state of the deep ocean, soil parameters or ice fractures. Hybrid ML models can likely learn relevant physics from historic extremes, but these events need to be well sampled in the observational record.

There is a pressing need for more Earth system data that:
- Cover the key variables necessary for climate modelling.
- Sample globally, to the same exacting standards.
- Are frequently refreshed and easily accessible in open data lakes.
- Are validated for quality using recognised benchmarks and tests.

Numerical models are still essential for pushing boundaries

ML-based surrogate models (AKA emulators) can only imitate the full physical models of climate and weather. **Physics-based models are currently the only sure way to generate physically accurate predictions of climate conditions that haven't yet occurred.** It is important to push these models to ever-higher resolutions if we want to obtain more accurate synthetic training data.

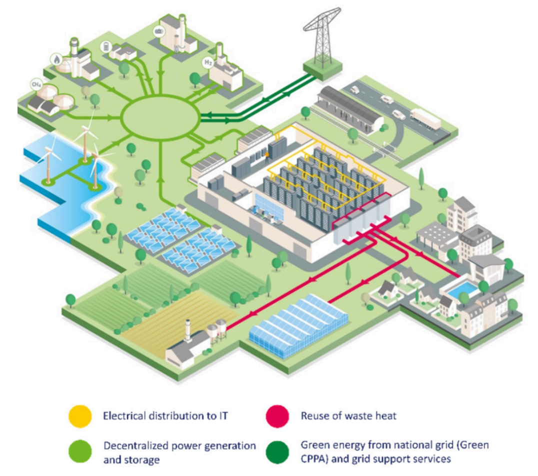
ML may also be able to accelerate numerical modelling directly, either by replacing components (such as modules to calculate subgrid physics via FORTRAN bridges), helping design more efficient architectures or by finding more efficient new algorithms. For example, DeepMind developed a ML system called AlphaTensor that discovered previously unknown optimisations in matrix multiplication.

Other approaches to solving the scaling problem

Fully differentiable Earth System models may have distinct advantages over numerical models and make hybrid-ML solutions possible. They may reduce uncertainties and sensitivities to hand-tuned initial conditions - see the recent paper 'Differentiable programming for Earth system modeling'.

Data Scaling

The data needed to fuel weather and climate models is enormous. The European Centre for Medium-range Weather Forecasts (ECMWF) is a leading organisation that produces daily forecasts using NWP and curates the ERA5 atmospheric reanalysis data of global climate, covering a period from January 1940 to the present day. They currently store 440 PB of primary data and 130 PB of secondary data, and add 230 TB of climate measurements each day. The ability of weather and climate models to ingest big data is also a bottleneck that limits spatial and temporal resolution. ML-based emulators scale well with data size compared to numerical methods.



Environmental impact of supercomputing centres

Supercomputing centres like the ECMWF's new facility in Italy consume significant electricity (~30 MW/day; equivalent to a medium-sized town) and require industrial cooling and heat management solutions. Modelling one year of the Earth's climate at 1 km resolution costs 596 MWh, which is equal to one hour of a coal-based power station's output.

Although running AI models is much more efficient, the computing cost of training large AI models is still huge. Popular public AI services like ChatGPT still use significant energy because of the number of users making queries.

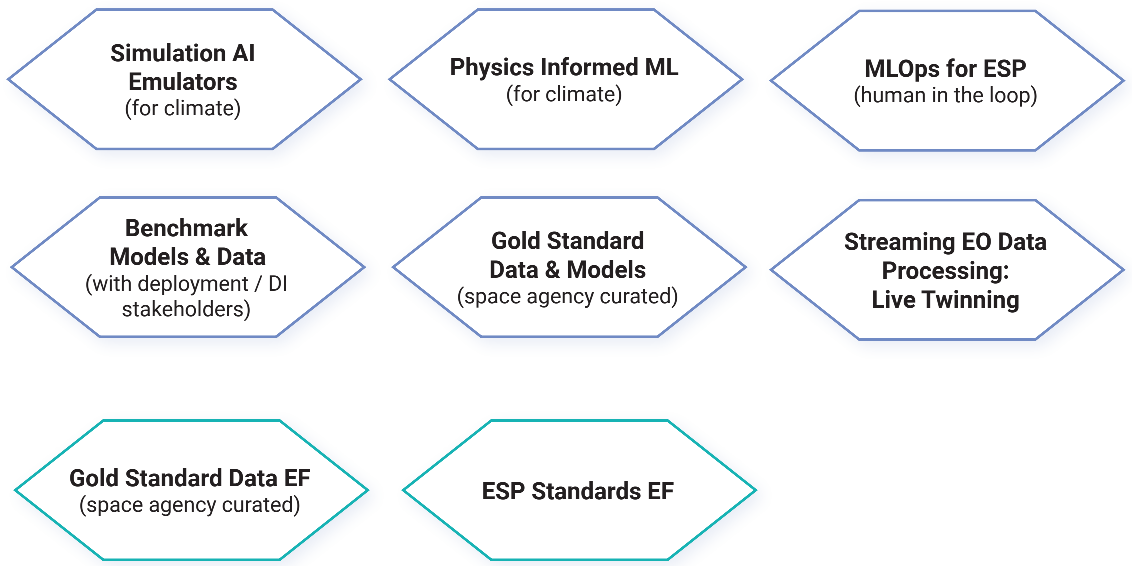
Can we be smarter about cooling solutions? For instance, by using grey water or mineral oil, locating them in colder places or using excess solar energy to provide a cold-sink?

Can we minimise the use of large-scale compute through distributed computing? When would this make sense and when would it not?

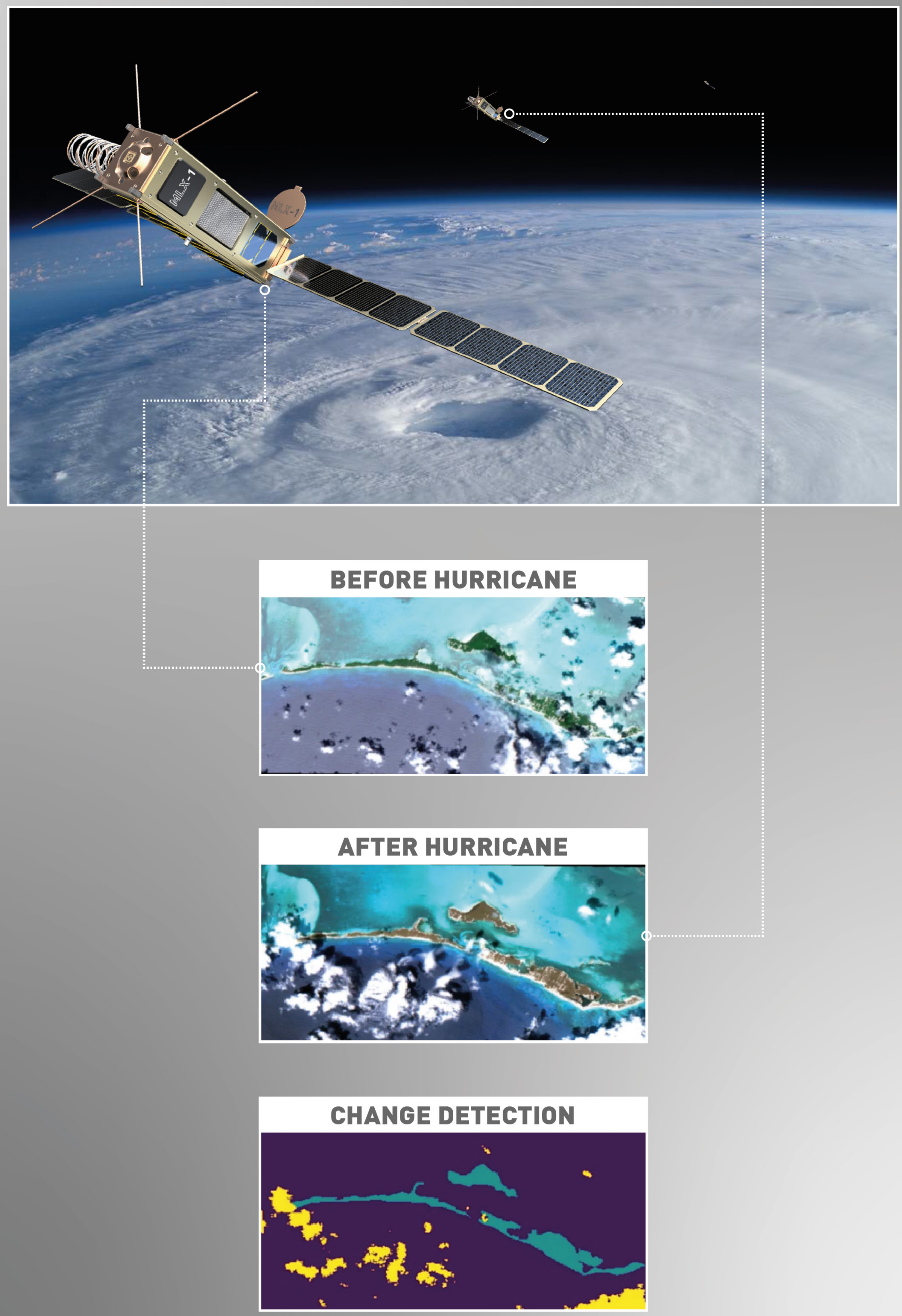
Resource:
'Deep learning and process understanding for data-driven Earth system science' (2019)

nature

Linked Roadmap Actions



These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.



RAVÆN

nio.space

BUILDING BLOCK 4: VISUALISATIONS AND INTERFACES FOR ESP

TRANSLATING DATA TO INSIGHT IS KEY.

Fast and accessible ML-driven climate emulators can empower us to make better choices for the environment in our homes, businesses, governments and on an international scale. They are a fundamental part of ESP, but like any new technology aiming to gain widespread adoption, climate emulators must have friendly and effortless user interfaces. They must:

- Present their results through compelling visualisations that are easy to understand, even for non-experts.
- Allow for easy experimentation and learning via intuitive interactive controls.
- Be accessible on a broad range of devices. Mobile devices, especially, should be first-class citizens of ESP tools.
- Speak the native languages and jargon of different domain areas (commerce, economics, governance, science and more).
- Integrate seamlessly with key software and tools that are commonly used in deep subject domains (e.g., economic or impact consulting).
- Integrate with popular platforms in use by the general public (e.g., office apps).

Interactive ESP visualisations will be especially powerful at communicating the impact of decisions on long-term climate change. End users will be able to **explore the effect of making different choices** to find which variables are important and which are not. When combined with joint modelling of other specialisations, they can build understanding of the trade-offs required to make decisions for combined good.

Interactively exploring to understand impact and change

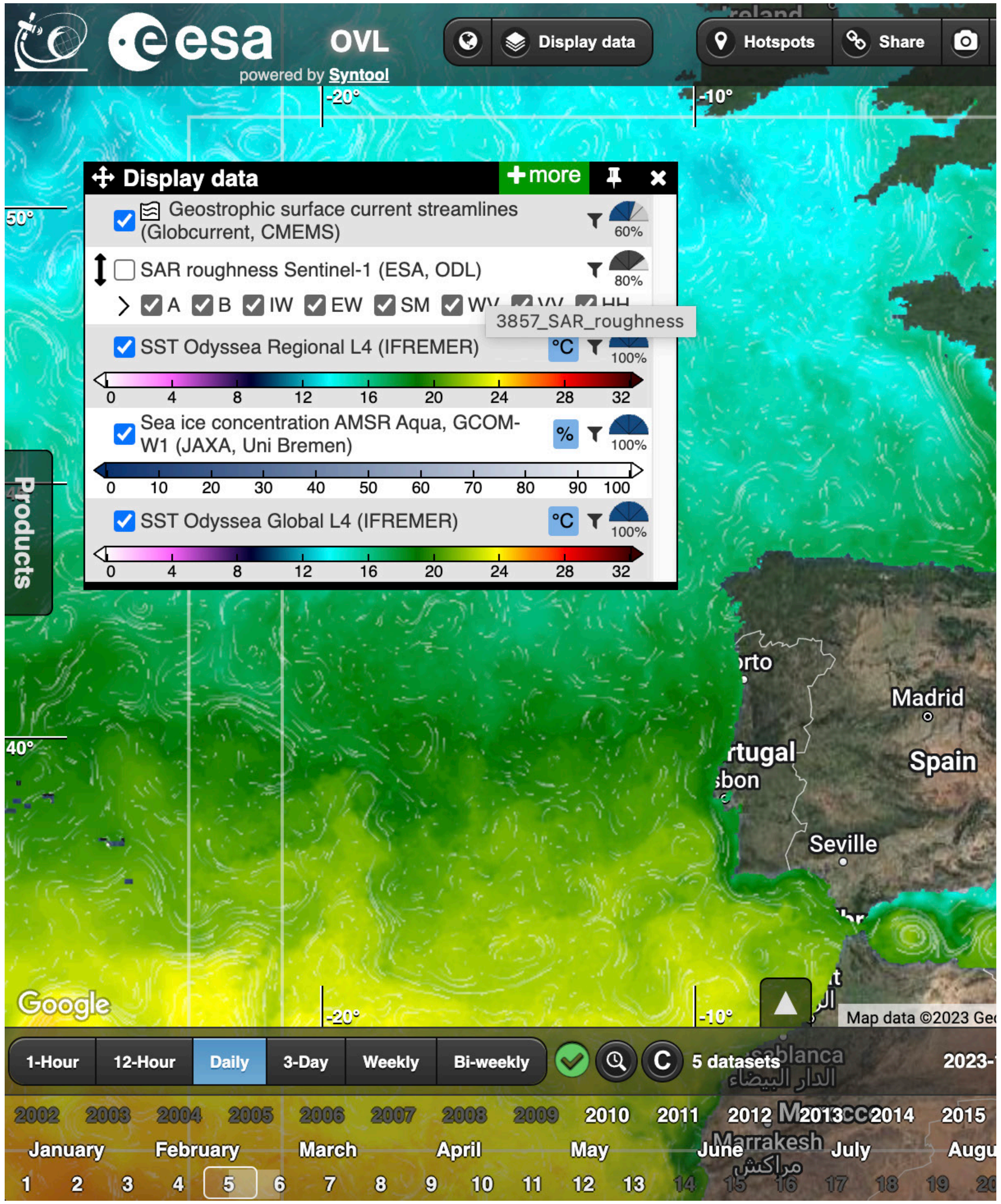
Visualisations of all sorts are a pillar of digital twins and ESP in general.

Non-Climate Domains

Diverse users in business, government and at home, will need intuitive visualisations to help understand how their choices impact long-term planetary health.

Research & Development

Researchers, developers and technical experts need informative and dynamic plots to diagnose how ESP systems are operating - including for ongoing monitoring.



ESA's Ocean Virtual Lab

Large and small organisations are already taking advantage of the power of modern web technology to build complex visualisations of planetary-scale data. This image is a screen capture of the ESA's Ocean Virtual Lab for discovering, manipulating and analysing remote sensing data in the world's oceans. Other examples include the ESA Earth System Data Lab for viewing data on multiple Earth systems, Earth NullSchool for real-time weather, and Bushfire.io for weather and fire variables.

What would a climate emulator do?

Climate emulators are especially powerful when integrated into a joint modelling framework. We cover this concept in the section on integration through decision intelligence, later in this document.

En-ROADS, created by MIT, is an excellent example of a climate solutions emulator with attractive visualisations coupled with intuitive controls. Such interactive emulators are useful for exploring the large-scale climate consequences of different policy decisions. Users can visually alter policy levels and immediately see the effect on projected greenhouse gas emissions up until the next century. A wide range of secondary graphs are available allowing for deeper exploration.

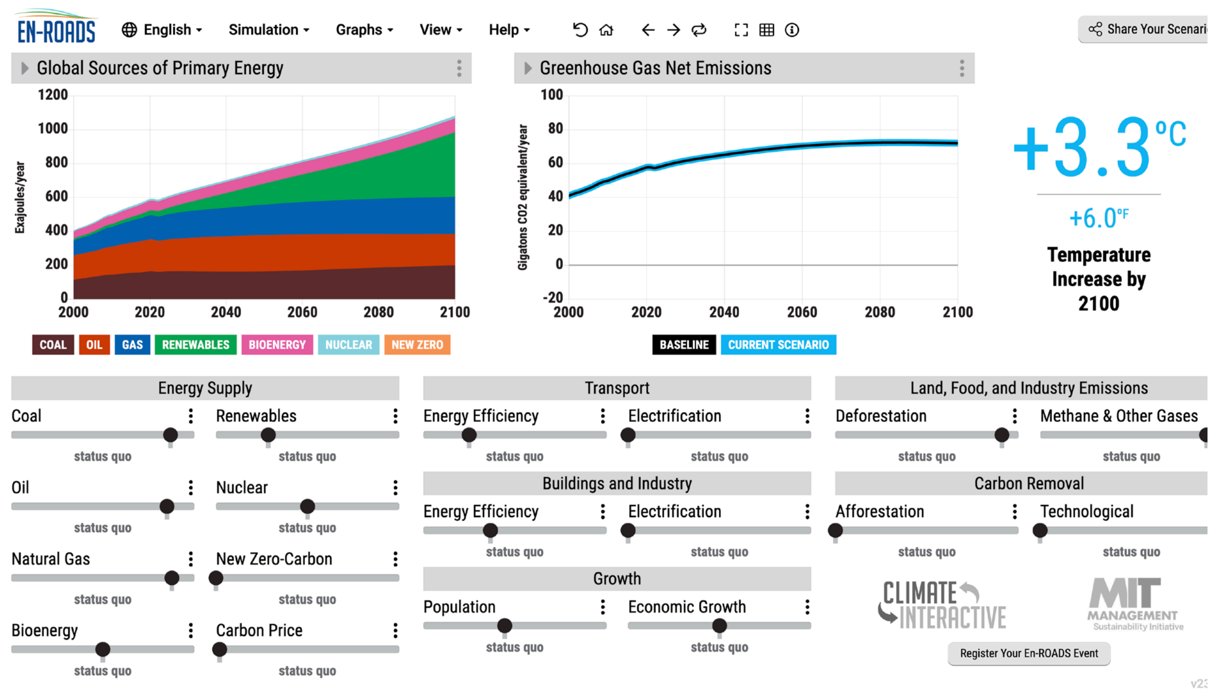
Emulator systems like En-ROADS could be built for niche industries, or even individual choices, by including a population scaling factor. Instead of policy choices, the emulator might offer the sort of choice individuals or small organisations might make and scale that to a fraction of the population. This scaling factor could itself be a choice and reflect the depth of behaviour change.

Such an emulator could:

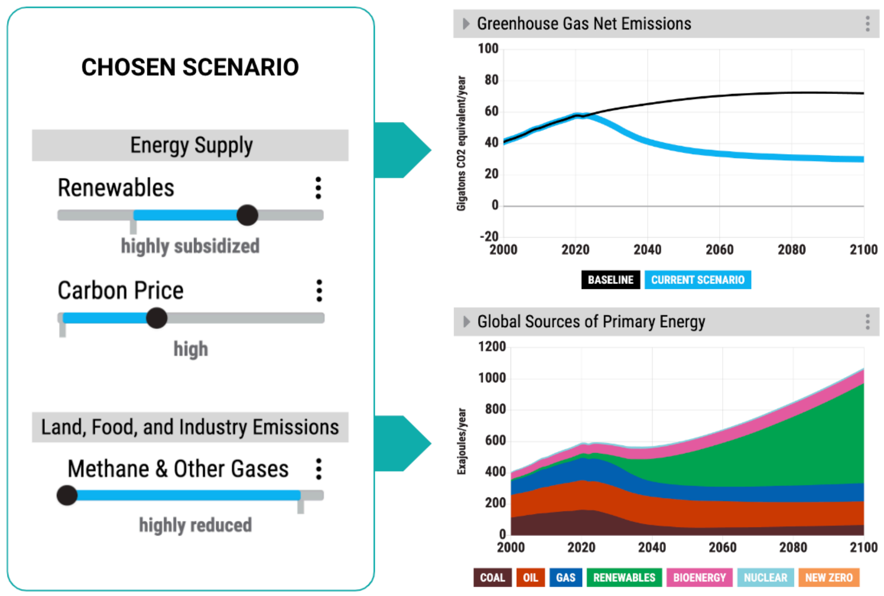
- **Help build an ESP-literate population** by translating ESP information into conversations about outcomes in the classroom and community
- **Visualise potential impacts** on a broad range of **interconnected systems,** such as rural-urban tensions, supply chains, built-environment infrastructure, mobility and logistics, energy systems, products and materials, industrial processing, food, water and more.
- **Support collective actions** for business and other sectors by helping them become aware of their interconnections.
- **Simulate cause and effect** to: a) illuminate decision options, b) address lack of imagination as to outcomes, c) detect weak signals of unintended consequences, and d) improve human capacity to integrate information.

However, climate emulators like En-ROADS are currently limited in their scope:

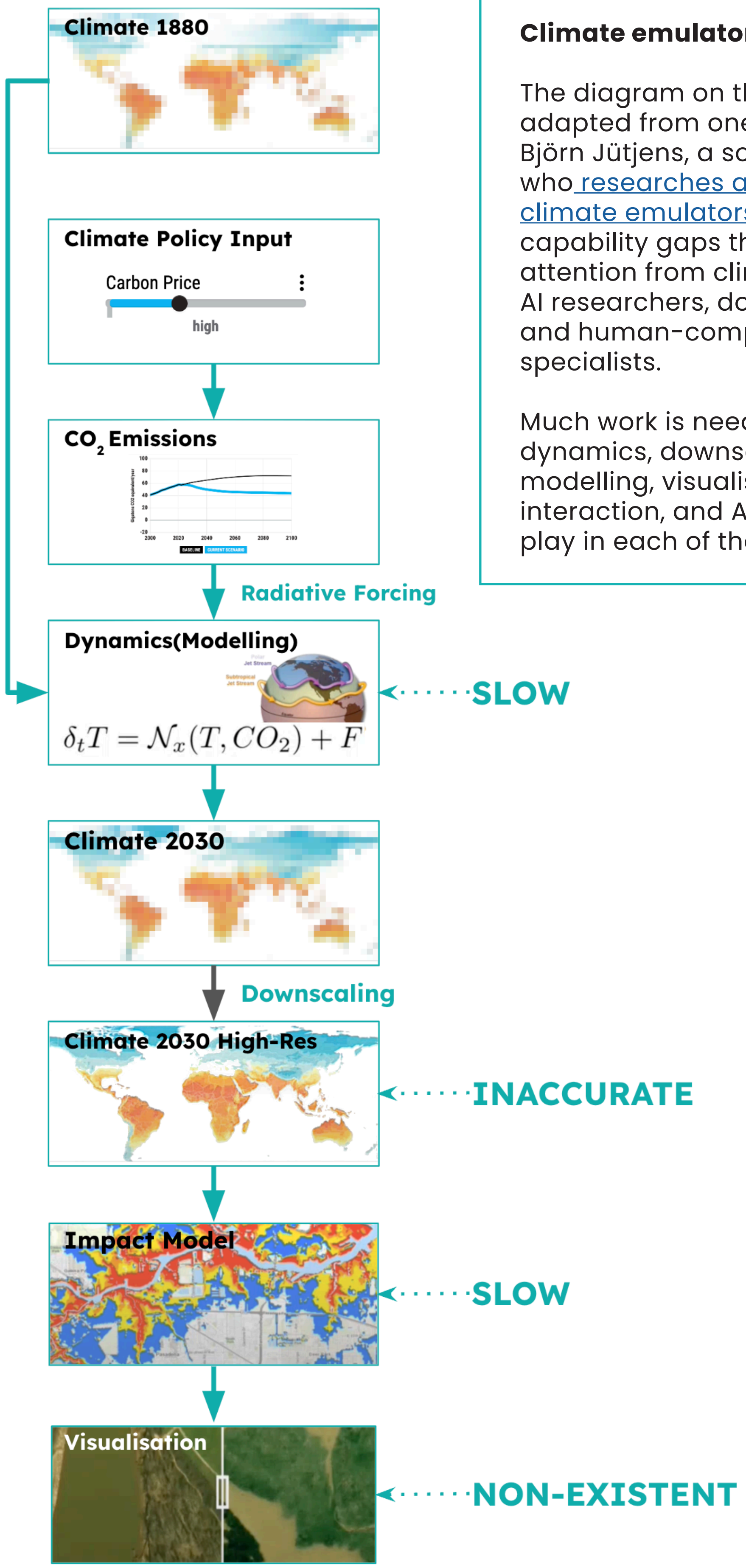
- They **can only simulate large-scale systems;** global policy leads to globally averaged outcomes.
- They **cannot currently predict outcomes for local regions,** which may be very different to the global average.
- They don't provide **accessible visualisations of model uncertainty** and tools for decision makers to understand the impact of variable initial conditions (also called 'sensitivity').
- They are **not immersive enough** to make deep connections to the public.



En-ROADS: Example of a low-resolution climate model emulator developed by MIT. It allows users to explore the relative effects of government policy on global heating.



Under default assumptions, setting a carbon price is by far the most effective policy to reduce net greenhouse gas emissions.



Climate emulator workflow

The diagram on the left is adapted from one created by Björn Jütjens, a scientist at MIT who researches and develops climate emulators. It illustrates capability gaps that need attention from climate scientists, AI researchers, domain experts and human-computer interaction specialists.

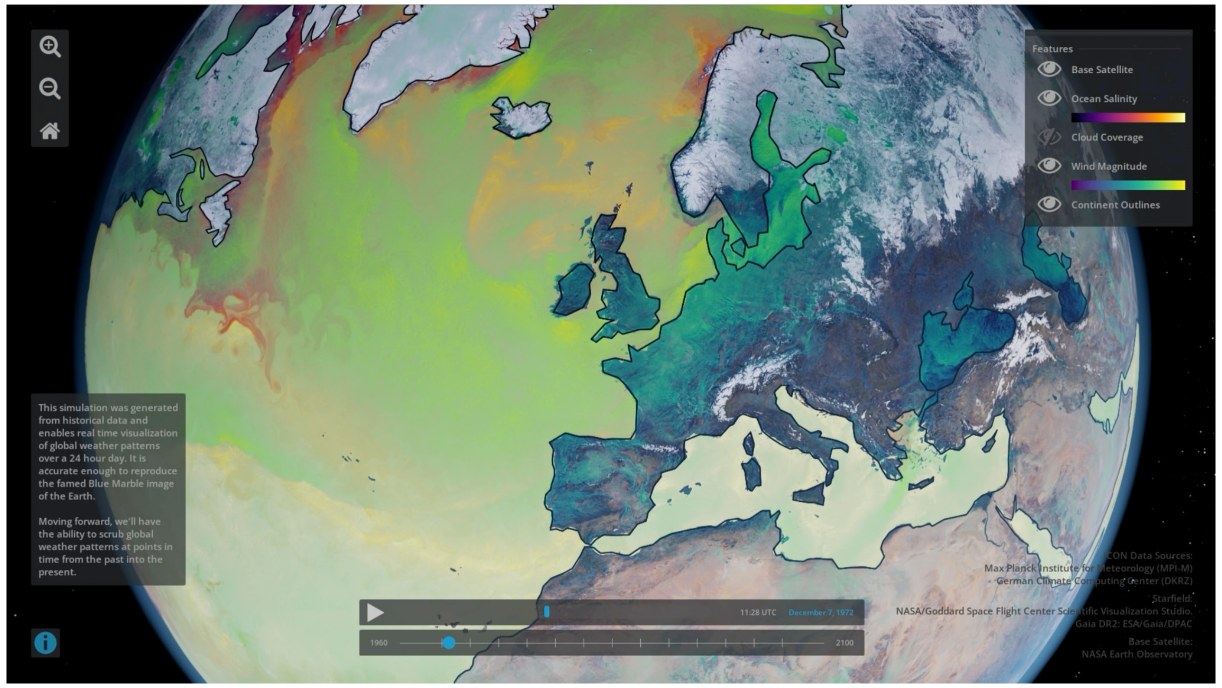
Much work is needed on dynamics, downscaling, impact modelling, visualisation and interaction, and AI has a part to play in each of these.

Insight from the gaming industry: interactive narratives needed

There are significant lessons to be learned from the gaming industry for the promotion of ESP as a concept, for creating sticky user interfaces that change behaviours and for communicating ideas in the most effective way.

Experience in the gaming industry has shown that **both conversations and interaction are key to engagement.** It is not enough to build a game where people can play with the model; it is important for the player to take part in conversations.

This is because a lot of problems, like energy decarbonisation and climate change in general, show non-linear behaviour, both in the physics and interactions of agents. Guided interactive narratives are important to help people understand the complex behaviour in these systems, and to change their thinking from linear to non-linear, inspiring and shaping positive action.

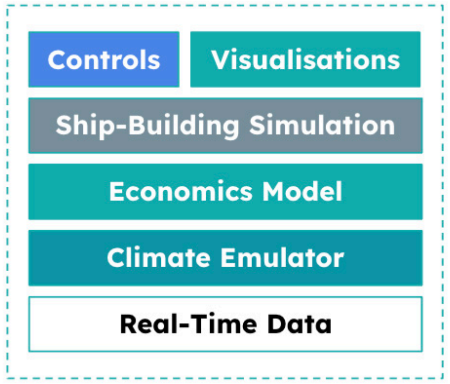


NVIDIA's Earth-2 platform enables interactive exploration of high-resolution weather and climate simulation datasets. Try the Earth-2 demo at this link.

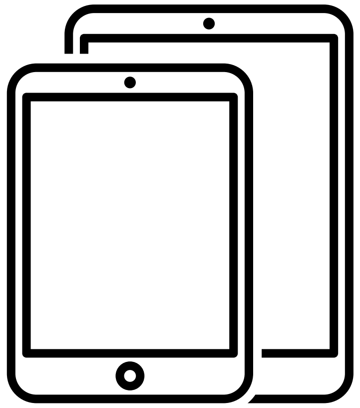
Resource:
Multi-dimensional linked data exploration implemented as a python package: https://glueviz.org/

Achieving reach via smart deployment of ESP visualisations

ESP capabilities and models should aim to be useful to a large and varied audience. Recommendations to boost reach include:

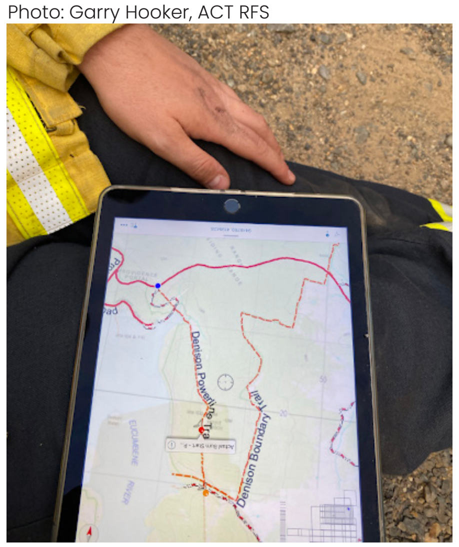


- Lightweight deployments that use fast climate emulators as a base layer, but also include thematic layers tuned to the use-cases and specialisms of the end user.
- ESP interfaces should provide access to both real-time data and future projections.



- Software capable of running on limited processing hardware (e.g., low-powered mobile devices) as well as GPU-accelerated super-computers.
- Visualisations and control interfaces tuned for small-screens. Mobile devices are the default computing device for most of the world's population, especially in the developing world.

Photo: Garry Hooker, ACT RFS



- Visualisations need to have appropriate fidelity (resolution and scale) and hide inappropriate information to guard against information overload.
- Interfaces must be tailored to the end use case. In some cases, virtual or augmented reality will offer critical capabilities.
  In others, highly interactive and rapid visualisations will be needed (i.e., fast screen refresh-rate). Different levels of interactivity will be appropriate for each use case.

Language based interfaces to ESP

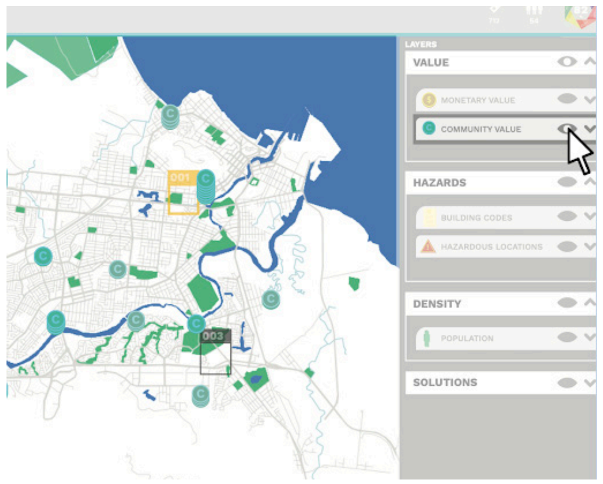
Large language models (LLMs) could provide a powerful interface for ESP capabilities - for both controlling and reporting.

- Reporting: LLMs can now translate visual data (images, maps etc.) into accessible descriptions.

> *"AI models capable of translating different modalities into natural language might be such powerful tools for making complex multimodal data accessible"*
> - Marco Zaccaria Di Fraia

- Controlling: The interfaces for interacting with LLMs can be adapted to ESP systems,to lower the barrier for accessing them and for accelerating their deployment. Imagine a ChatGPT-style interface that answers questions like *"What is going to happen in this geographic area in 20 years?".*

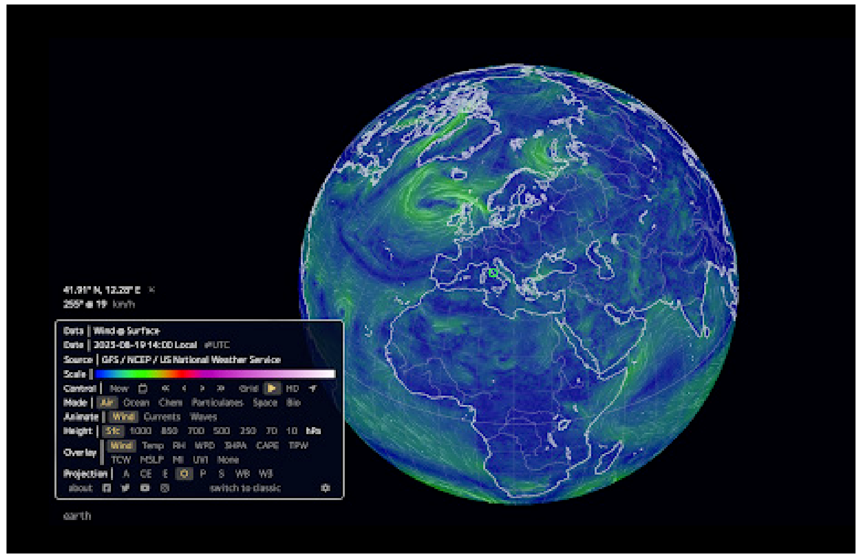
> *"You can easily build a natural language interface to most any existing software - like Omniverse, for example. By handing the documentation to the large language model, it will even be able to generate user interfaces and webpages on demand. So you don't have to create a user interface that anticipates every user's needs."*
> - David Hall, NVIDIA



What would a ChatGPT for climate look like?

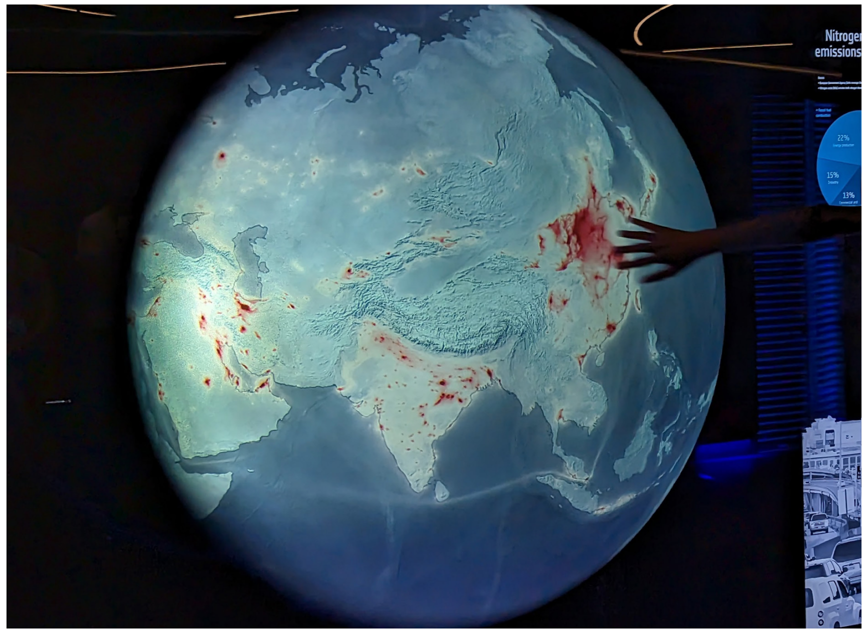
LLMs give us the ability to command complex tasks using natural language:

"Create a map of the coast of Spain in 2090 given the most likely sea-level rise, assuming greenhouse gas emissions continue to rise at the current rate. Then estimate the impact on housing and services."



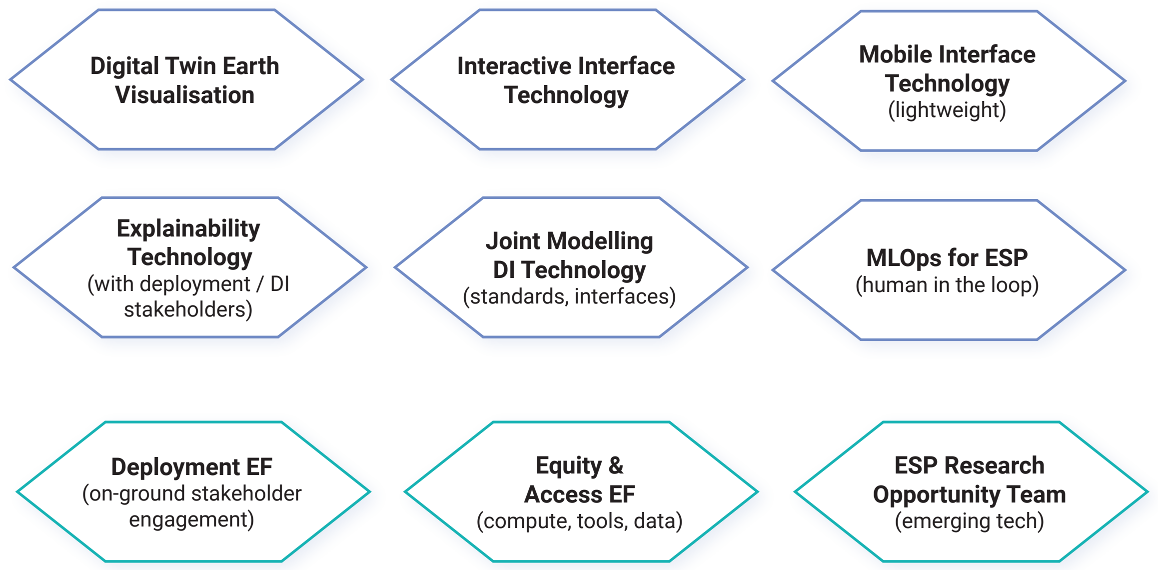
Digital Twin Earths (DTEs) are a special deployment case that leverage ESP visualisations and interfaces. We cover DTEs in more detail in the **'Deployment'** section of this document.

An interactive table-style screen at ESA ESRIN, displaying a time-series of Earth observation images. Shown here are roads cut into the Amazon rainforest as a precursor to large-scale logging operations. Users gain a better understanding through guided exploration of the data.



ESA's interactive globe screen showing a visualisation of the Earth that can be manipulated in space and time to show the progression of events. The red colour here is nitrogen dioxide emitted from fossil fuel combustion for electricity generation and transport. Ocean shipping routes are clearly visible.

Linked Roadmap Actions



These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

BUILDING BLOCK 5: GENERATIVE AI

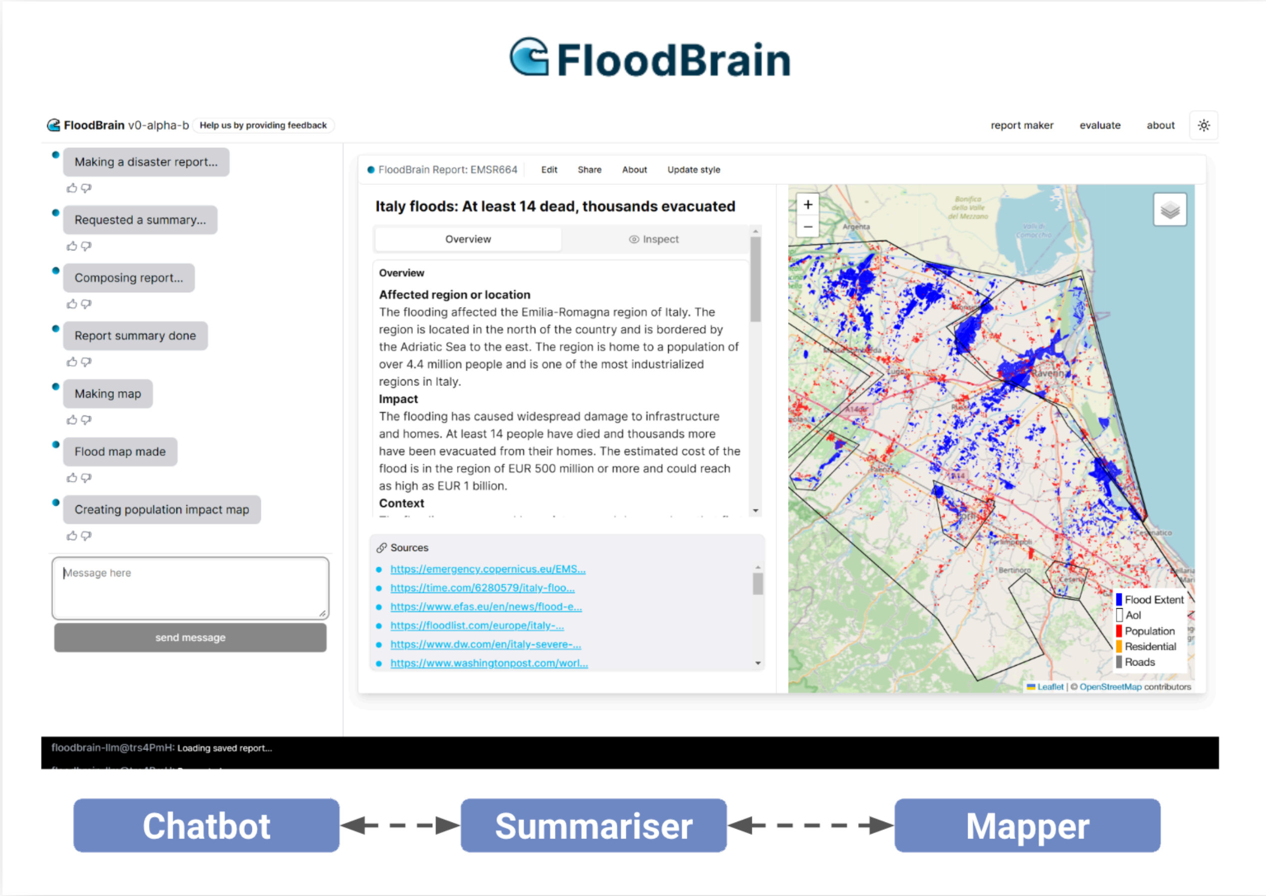
LLMs can be automators and integrators for ESP tasks

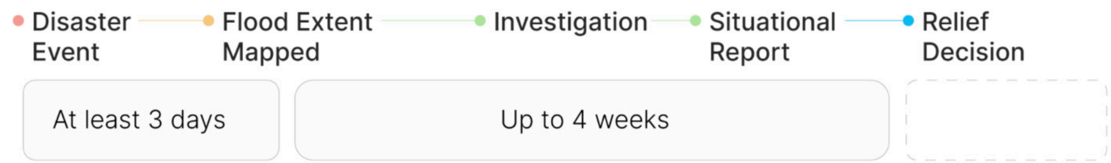
Large Language Models (LLMs) can be 'knowledge stores' of human knowledge, orchestrating many tasks:

- They can be harnessed to build a natural language interface to most existing software, like the Omniverse digital twin platform that is being used as a foundation for NVIDIA's Earth 2.0.
- Experiments have shown that LLMs can generate graphical user interfaces (GUIs) and web pages on demand after being supplied with documentation. This means that it will be easy to generate custom ESP interfaces rather than spending time creating a user interface that anticipates every user's needs.
- LLMs will be tireless assistants for everyday life (e.g., monitoring and summarising incoming communications while you rest). Recent multimodal vision-language models can even understand images; they could summarise plots and monitor satellite data streams for meaningful events.

For ESP, the capabilities of LLMs facilitate the creation of complex systems-of-systems to analyse data, connect models and make decisions in an integrated way.

> *"ChatGPT-4, for example, aces a wide variety of intelligence tests and exceeds human skill on many of them. Although it still gives silly answers to other questions, which we find obvious. They're different from us, but they don't get bored or tired or need to sleep and can perform millions of tasks in parallel, if you activate millions of copies"*
>
> \- David Hall, NVIDIA





Case Study: The FloodBrain tool (FDL Europe 2023)

One example of a digital assistant that combines language and vision processing is the FloodBrain tool developed during the FDL Europe 2023 research sprint. The FloodBrain team tackled the challenge of creating detailed and useful reports on individual flooding events and their impact. Usually, these types of reports take several weeks to complete and involve collating information from a wide range of online sources.

FloodBrain leverages LLMs (GPT 3.5, PaLM2/BISON and LLAMA2 7B Chat) to automatically answer questions about recent flood events, drawing from trusted sources, open web-searches and geolocated maps. The tool summarises the answers to pre-set questions (e.g., What caused the flooding to occur?), assesses the relevance of each piece of information, displays maps of the flood-extents and automatically analyses the impact on people, land and infrastructure. The final report is presented in an interactive web page and includes a tool that allows the inspection of all sources. ***This is an example of an LLM orchestrating and combining both quantitative spatial analysis and text-based information analysis.***

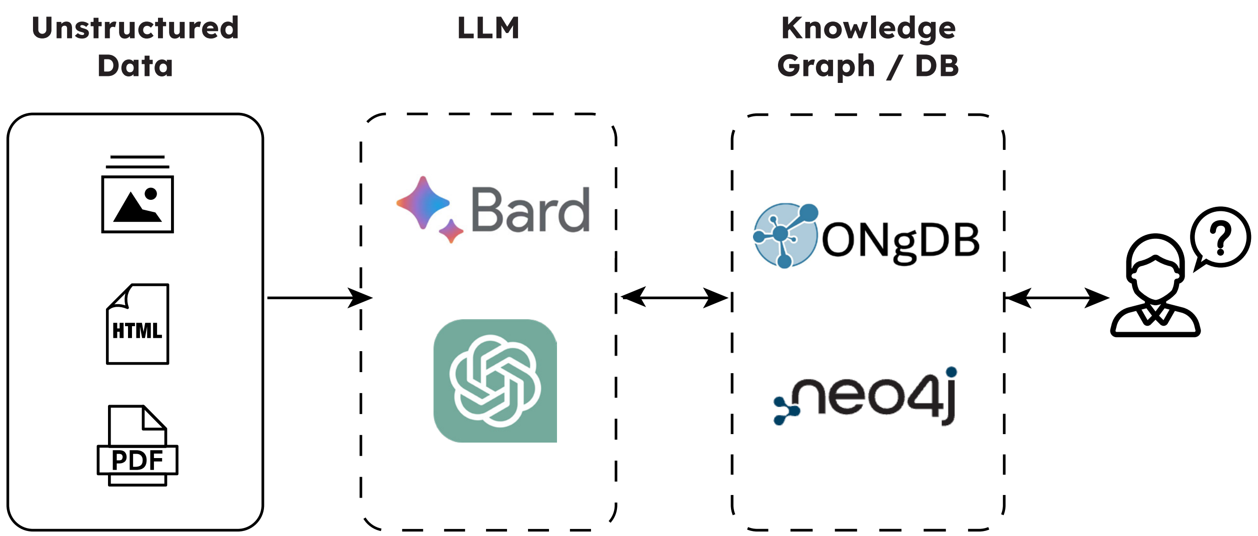
LLMs as Knowledge Stores for ESP

LLMs connected to chat interfaces are changing how we interact with computer systems. In the future, almost anyone will be able to use complex software by simply asking questions or issuing natural language instructions.

At the same time, the storage and recall of large and complex data (even entire bodies of knowledge that include contextual information and cross-links) is being enhanced by a new generation of large language models.

Research is underway into better harnessing the capabilities of LLMs by merging, or linking, them with structured knowledge stores such as relational databases and knowledge graphs in the following ways:

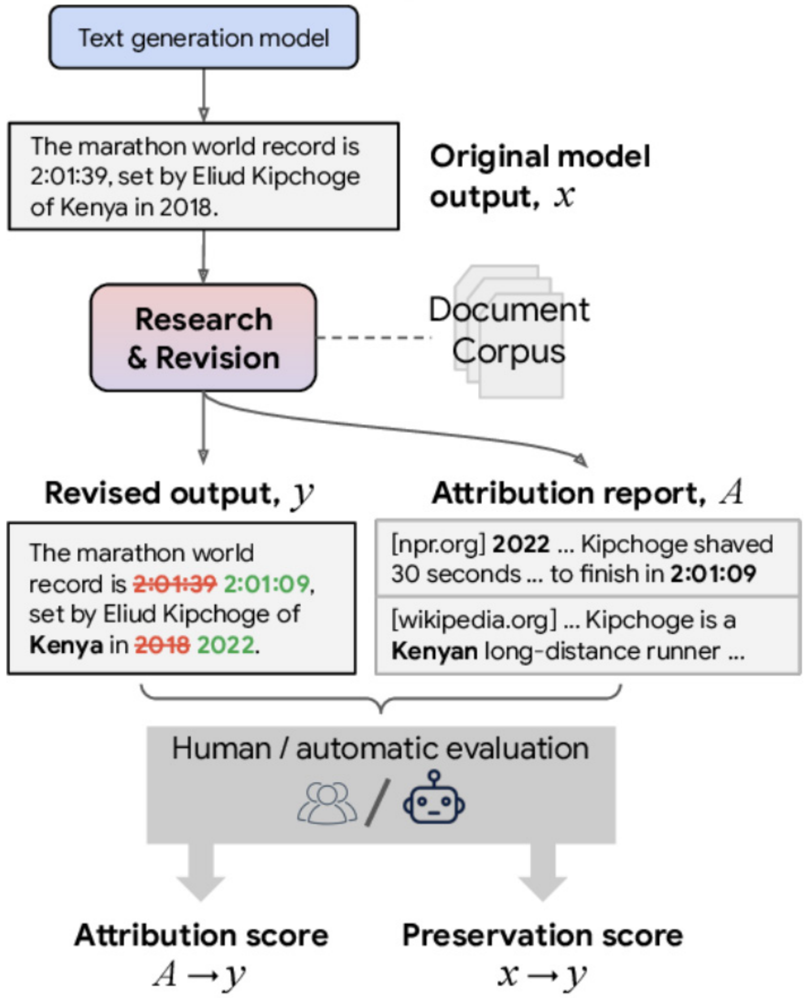
- As a natural language interface that can be used to query a database or knowledge graph, allowing complex queries to be phrased in plain language.
- As a method to create knowledge graphs by asking the LLM to decipher entities, discern relationships and eliminate redundancies by recognising duplicates.
- As a way to interact with Earth observation and climate variable databases, by applying visual question answering and semantic search.



Correcting for Hallucinations in LLMs

The tendency for LLMs to produce hallucinations is one of their biggest weaknesses, but work is underway to overcome this issue.

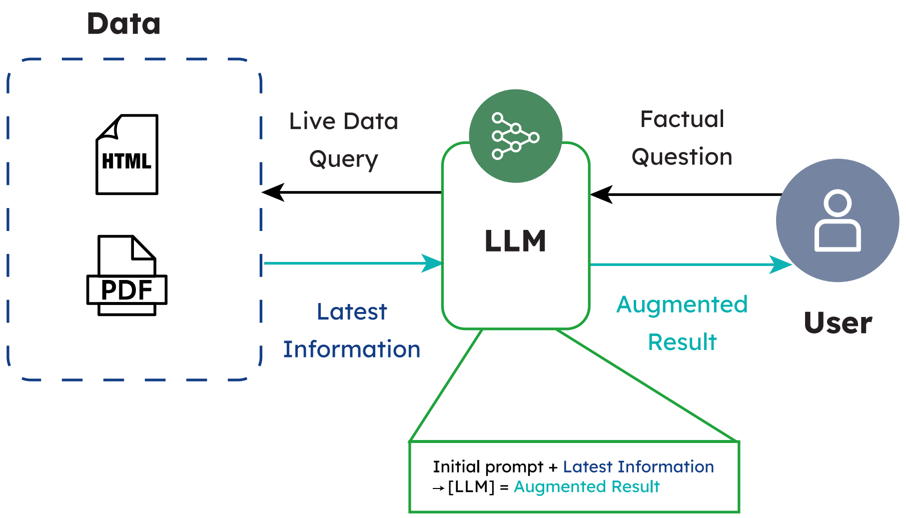
- LLMs tend to be very good at inventing facts; one way to address this is by triggering factuality modules after being prompted. In this case, the system realises that there is a factual query in the prompt, for instance: 'What is the mass of the Earth?' This creates a query to an external system. Then the output of this system is sent back to the LLM together with the prompt and the context. Now the LLM's task is *'Given this information and given what the question was, can you assemble an answer?'*
- Human interaction is required at each step in this process: to decide which queries trigger factuality processing, what external data to retrieve and to pass judgement on the outer loop to evaluate and score final responses compared to original questions.



For example, Google's RARR LLM system automatically researches and revises the output of any LLM to fix hallucinations and provide citations for each sentence. Figure credit: Gao et al 2022.

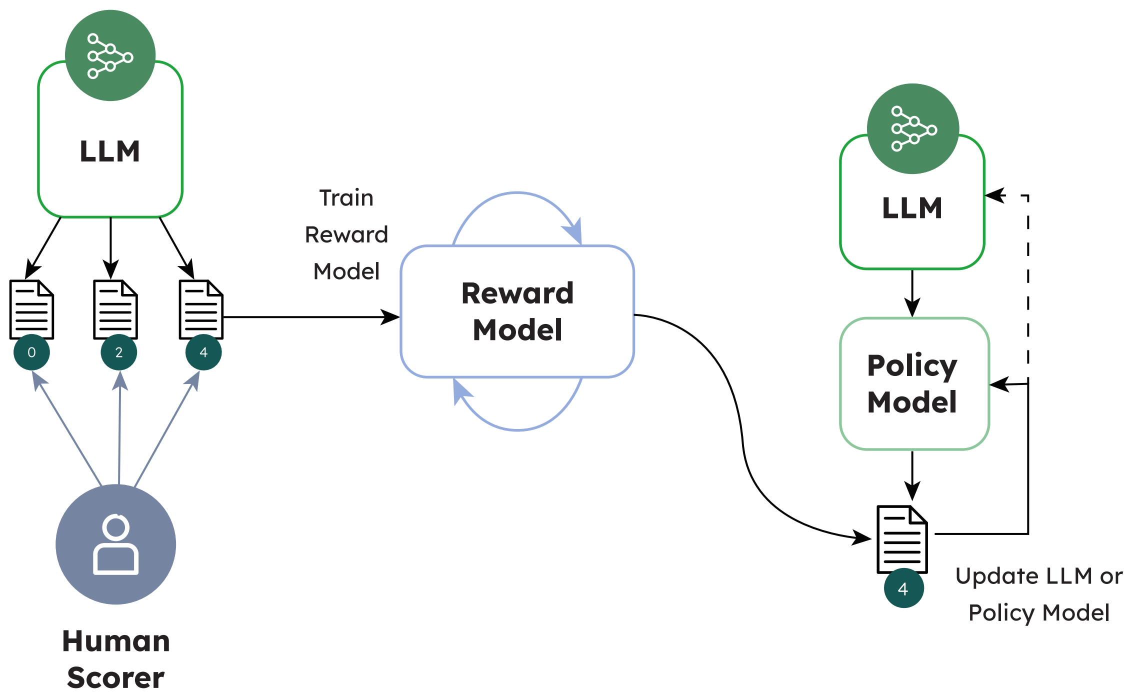
Making LLMs More Accurate With Retrieval-Augmented Generation

Retrieval-augmented generation (RAG) is a method to improve the quality of the responses that LLMs generate by anchoring them in external sources of information. It ensures models have access to up-to-date information and provides references for that information that users can validate. The LLM can also be trained to recognise when it cannot reliably answer a question and communicate that to the user.



Training LLMs with human input

Reinforcement learning with human feedback (RLHF) is key to ensuring that knowledge from domain experts is integrated correctly into LLM-based systems. This is a method to incorporate human feedback into the training and deployment of large models by using human feedback to train a secondary 'reward predictor' model, which is then used to adjust how the large model learns and behaves.



Causal Reasoning and LLMs

Advanced ESP capabilities that require complex reasoning would benefit significantly from AI-assisted orchestration. However, whether LLMs can truly determine cause-and-effect relationships is unknown and a highly active area of research.

Resource:
'Understanding Causality with Large Language Models: Feasibility and Opportunities' (2023)

arXiv

Resource:
'Causal Reasoning and Large Language Models: Opening a New Frontier for Causality' (2023)

arXiv

Even in the absence of perfect reasoning ability, LLMs are already proving useful for supporting ESP activities. This blog post demonstrates how ChatGPT can assist in building a formal decision model for a government agency looking to reduce emissions through air travel policy.

Synthetic image generation

Text-to-image generators like Stable Diffusion and Midjourney are now well known for creating photographs from text prompts. Similar technology is also useful for generating photorealistic images of future events. This image was generated by pix2pixHD, an ML model that produces imagery that is physically consistent with the output of the NOAA SLOSH storm-surge flooding model.

Such realistic visualisations of Earth systems predictions will be a vital tool for communicating with the general public and decision makers at different levels of government.

Concerns About Stacking Generative Methods

Although generative methods can produce outputs that look good to humans, often this data is not physically realistic or useful for downstream tasks, compared to data from a physics simulation. The statistics of AI-generated data can be subtly skewed in ways that are not obvious. Knowing this, it is dangerous to take ML-generated data and pipe into another ML model:

- ML-generated synthetic data that looks similar to real data to a human eye may not look similar to another deep learning model: the statistical properties may be different.
- Subtle differences in the generated data can be picked up by the downstream models and built upon. These errors can become important features, leading to the creation of realistic-looking artefacts.
- This is known as 'model collapse', where recursive training makes models forget information because of data pollution.

Rules to remember:

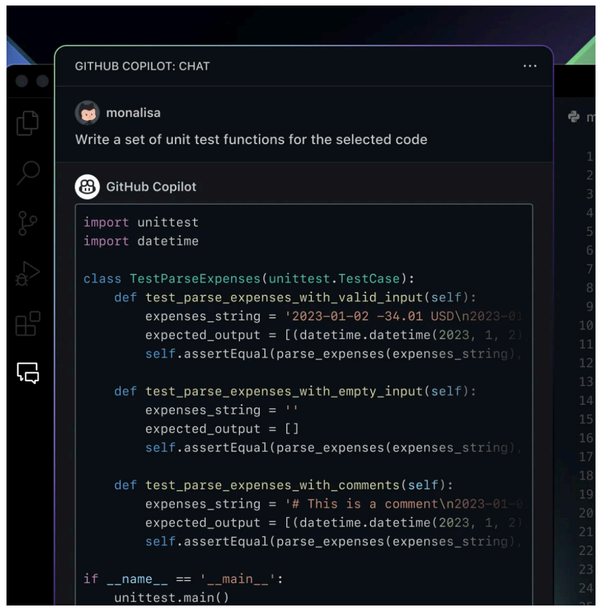
- Avoid using Generative AI to create data that is fed directly into another AI model.
- Synthetic data for use with downstream AI models should only be generated by a system that is fully understood: for example, a physics-based model. This could also be accomplished with an end-to-end system that was specifically designed to have controlled and calibrated uncertainties throughout.
- However, the output of generative AI may be suitable for human consumption, when producing qualitative visualisations, for example.

LLM Coding Interfaces ('Software 3.0')

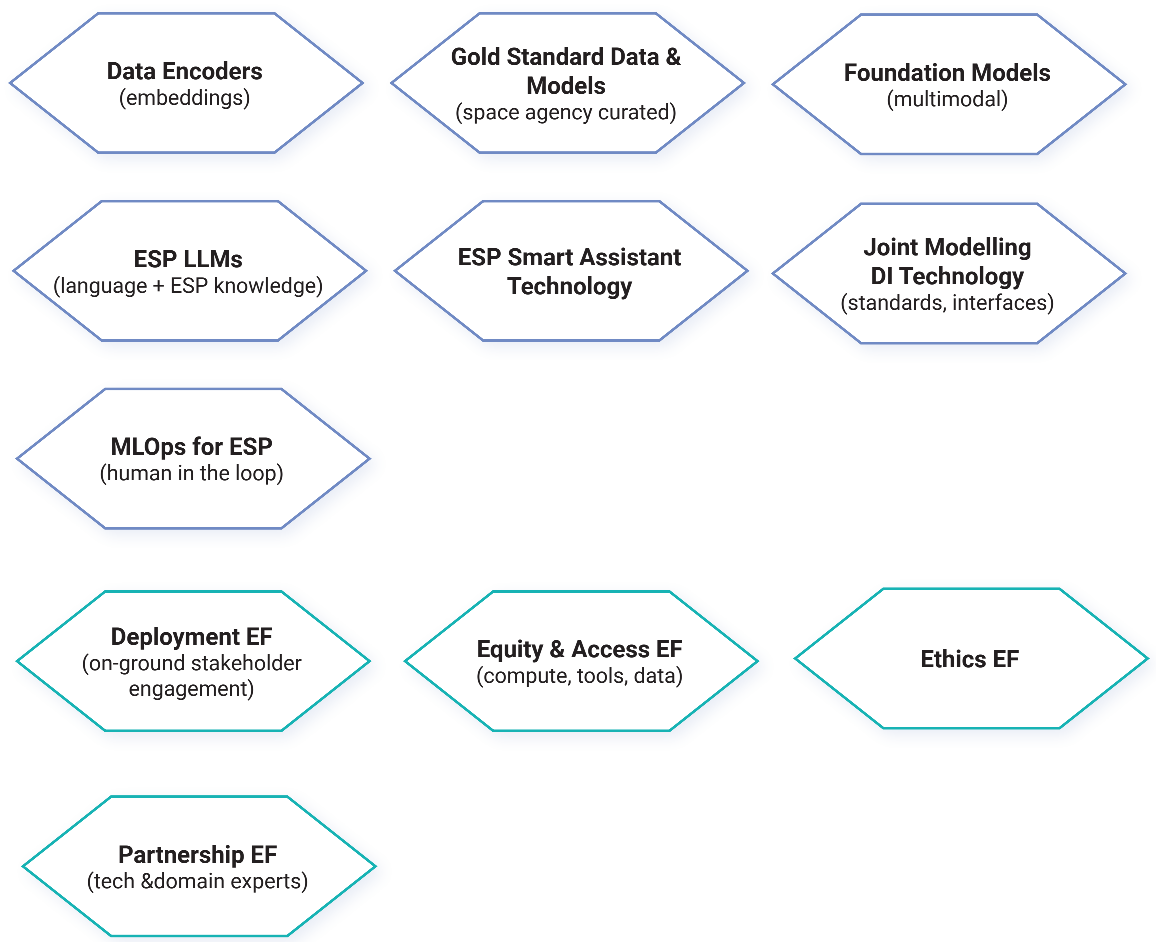
LLMs break new ground when it comes to automating tasks. One extraordinary talent that LLMs have demonstrated is a high degree of skill at generating computer code. This means that natural speech may soon become the most popular computer programming language: software 3.0.

The GitHub software development platform has introduced an AI-driven assistant called GitHub Copilot that can write functions and classes in response to language prompts. Tools like Copilot can explain the logic behind unfamiliar code, or even explain why a block of code is creating an error message and suggest a fix. LLMs trained on technical documentation can be effective at generating and updating documentation structures within the code as well as generating formatted user-guides.

Complex ESP tasks (e.g., performing analysis involving geolocated and unstructured data) can be made much easier by providing a LLM-based coding interface.



Linked Roadmap Actions



These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

Resource:
'Self-Consuming Generative Models Go MAD'(2023)

arXiv

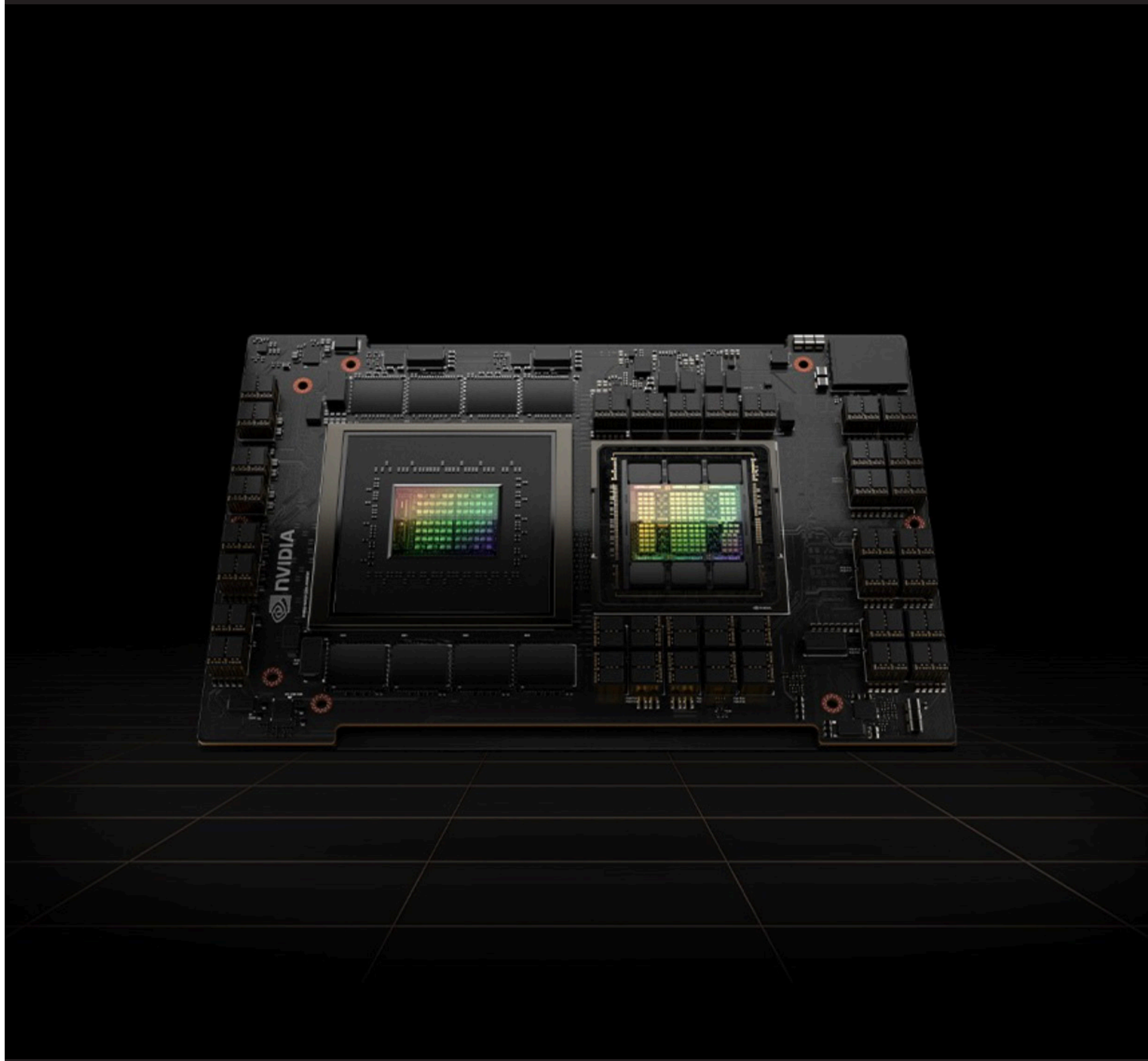
BUILDING BLOCK 6: SCALED SYSTEMS

ESP technology and systems will need to operate on large scales:

- Communicating across large geographical areas.
- Ingesting and processing data from disparate sources, including EO data, social data, networked sensors, simulated data and more.
- Ingesting and processing wide-bandwidth streaming data.
- Serving data simultaneously to client services (e.g., remote sensing data, visualisations, control packets etc.).

This will entail creating modular architectures that prioritise efficiencies at all levels, from acquisition to delivery. Machine learning will play a significant part in compressing and routing data. Some specific examples are:

- **Streaming observations.** It will be challenging to integrate continuously updating observations from multiple sources into a consistent simulation by data fusion. ESP technology will need to develop methods to integrate streams of data (e.g., from weather observations) when the available data changes over time, or different subsets are available at different locations. Academic work is currently concentrated on snapshot data (e.g. ECMWF ERA5) but continuous integration is more difficult because of the technical challenges and different timescales involved.
- **ESP software support systems.** Systems for monitoring development (DevOps) and operations (MLOps), will need to operate at scale.
- **Collaborative environments.** ESP will be a shared endeavour, distributed across the world. Development environments that allow scaled collaborations will be critical to building a developer community: interdisciplinary experts working together supported by a shared environment that allows multiple developers to collaborate in building systems, training models, visualising results.
- **Scaled computing hardware**. Massively parallel processors will be critical for implementing digital twin Earths and ESP systems that require large cloud computing infrastructure. The current crop of computing hardware (e.g., NVIDIA Grace Hopper) is only just becoming powerful enough to process streaming data. AI processing hardware on mobile devices will also be important for distributed ESP computing.
- **Accessible ML-ready data.** Both streaming and static data will need to be accessible in formats that are ready for ingestion into machine learning systems. This means they should be cleaned of missing, or poor quality, data, be characterised in terms of statistics and be stored in fast-access formats.



ESP OPERATIONS

OPERATIONS 1: INTEGRATION THROUGH DECISION INTELLIGENCE

Integration is a key concept in ESP that enables more robust predictions through joint modelling and better decision-making through decision intelligence.

Joint modelling: Climate, economics, health, finance, food-supply, energy, and ecosystem models all need to **integrate seamlessly** and **work rapidly together** to deliver consistent predictions.

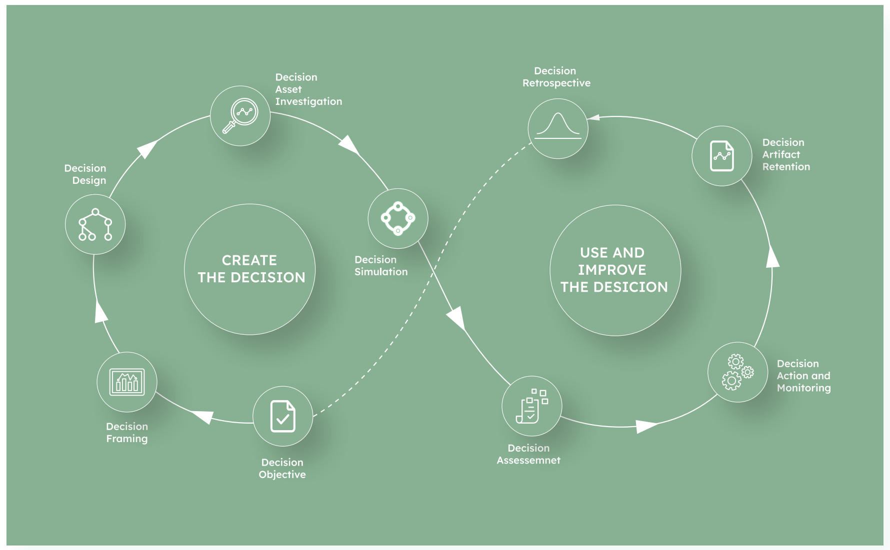
This means ESP-compatible models need to have **standardised ways of communicating,** well defined formats for inputs and outputs, and documented procedures for integration.

They need to **operate on the same rapid timescale** - which is where AI-driven models, or AI-emulators, come into their own.

All these systems have interlinked dependencies and feedback loops that influence the predicted outcomes - sometimes dramatically. Joint modelling provides more accurate predictions and allows a much fuller understanding of cause-and-effect relationships in the larger system-of-systems.

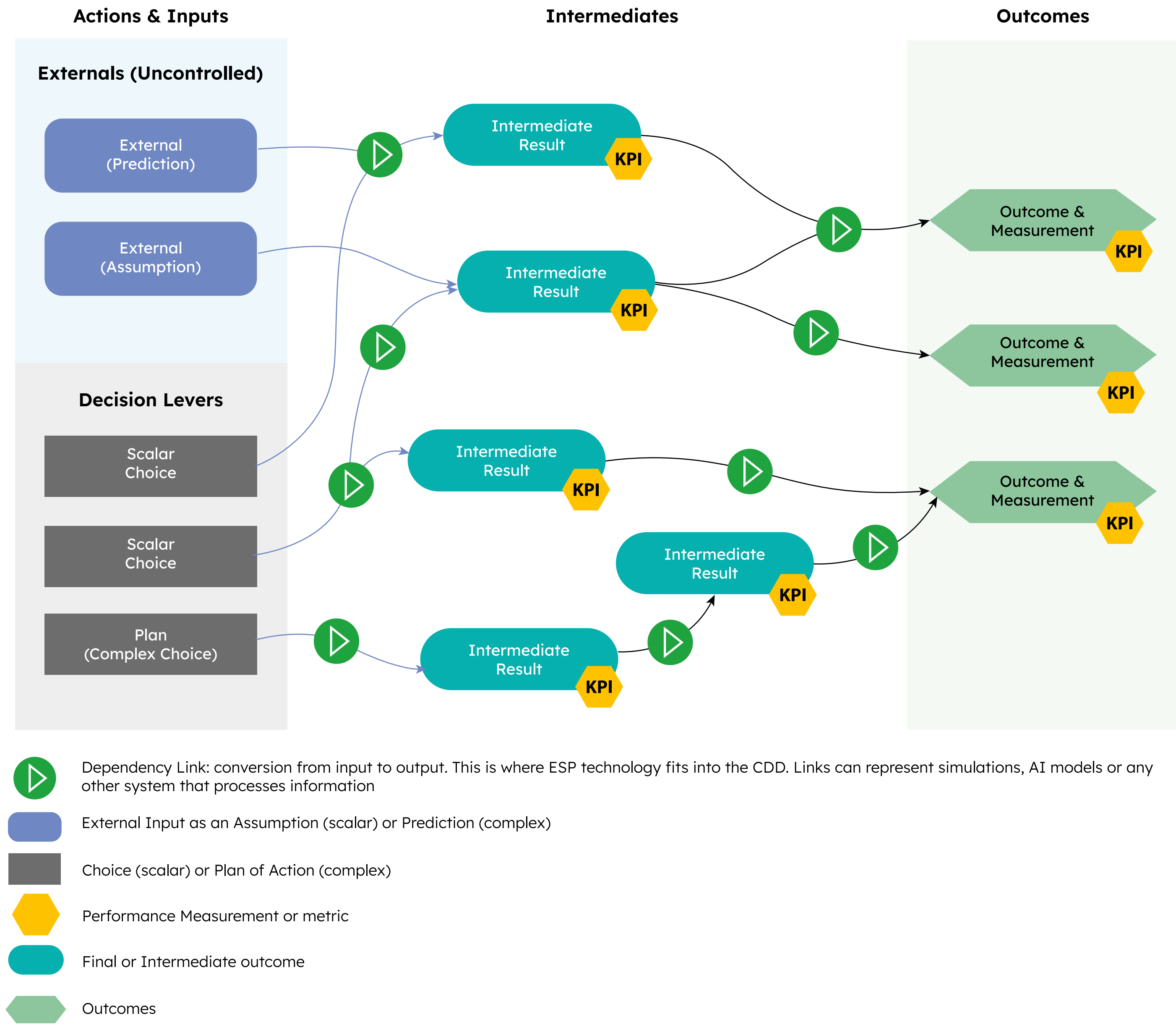
ESP - FORUM Earth Systems Predictability
ESP Certified

Certified to be compatible with ESP system standards.



Decision intelligence (DI) takes the concept of integration further by placing joint modelling at the heart of a powerful iterative **framework for designing end-to-end decisions.**

The Causal Decision Diagram (CDD)



Decision Intelligence (DI) is a method and set of associated technologies for making better, more evidence-based decisions, by mapping inputs through intermediate outcomes to desired outcomes. The central concept is that **decisions can be designed** with the aid of a diagram called a Causal Decision Diagram (CDD).

When the lens of **ESP** is added to this process, it **naturally extends joint modelling to support effective decision-making** for institutional stakeholders and overall planetary habitability.

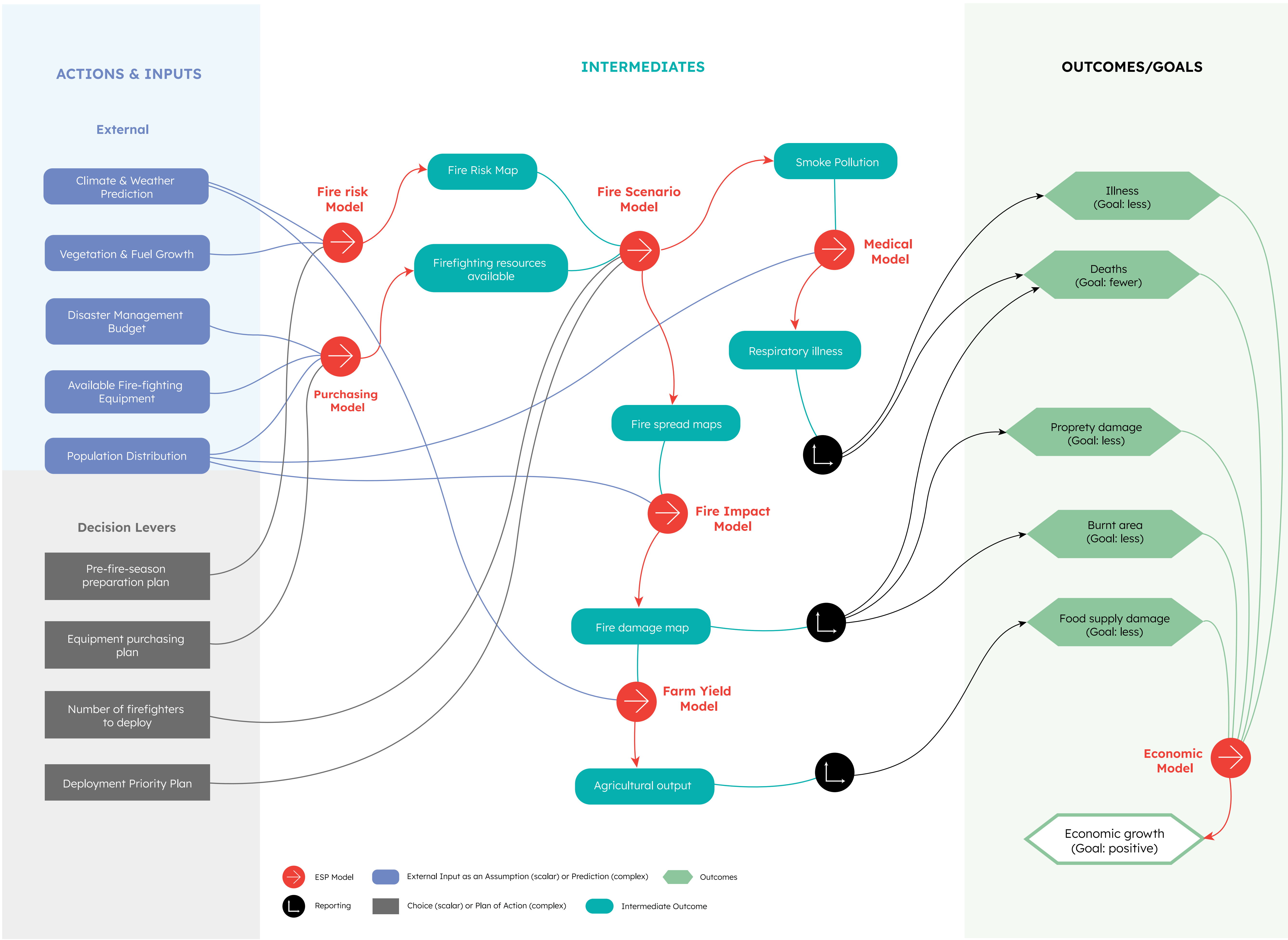
CDDs are used to map pathways from actions to desired outcomes. They incorporate simulations, feedback and performance monitoring to automate very complex decisions.

Example: A Causal Decision Diagram for Wildfire Management

Here is a simplified CDD that a government official could use when formulating a plan to manage wildfires in the future. The concept of 'integration' and model interoperability in ESP means that the entire diagram could be implemented as software, allowing the government to experiment with different inputs to arrive at an optimal solution.

What constitutes an optimal solution depends on the desired outcomes and how they are weighted against each other - a strategic and political choice.

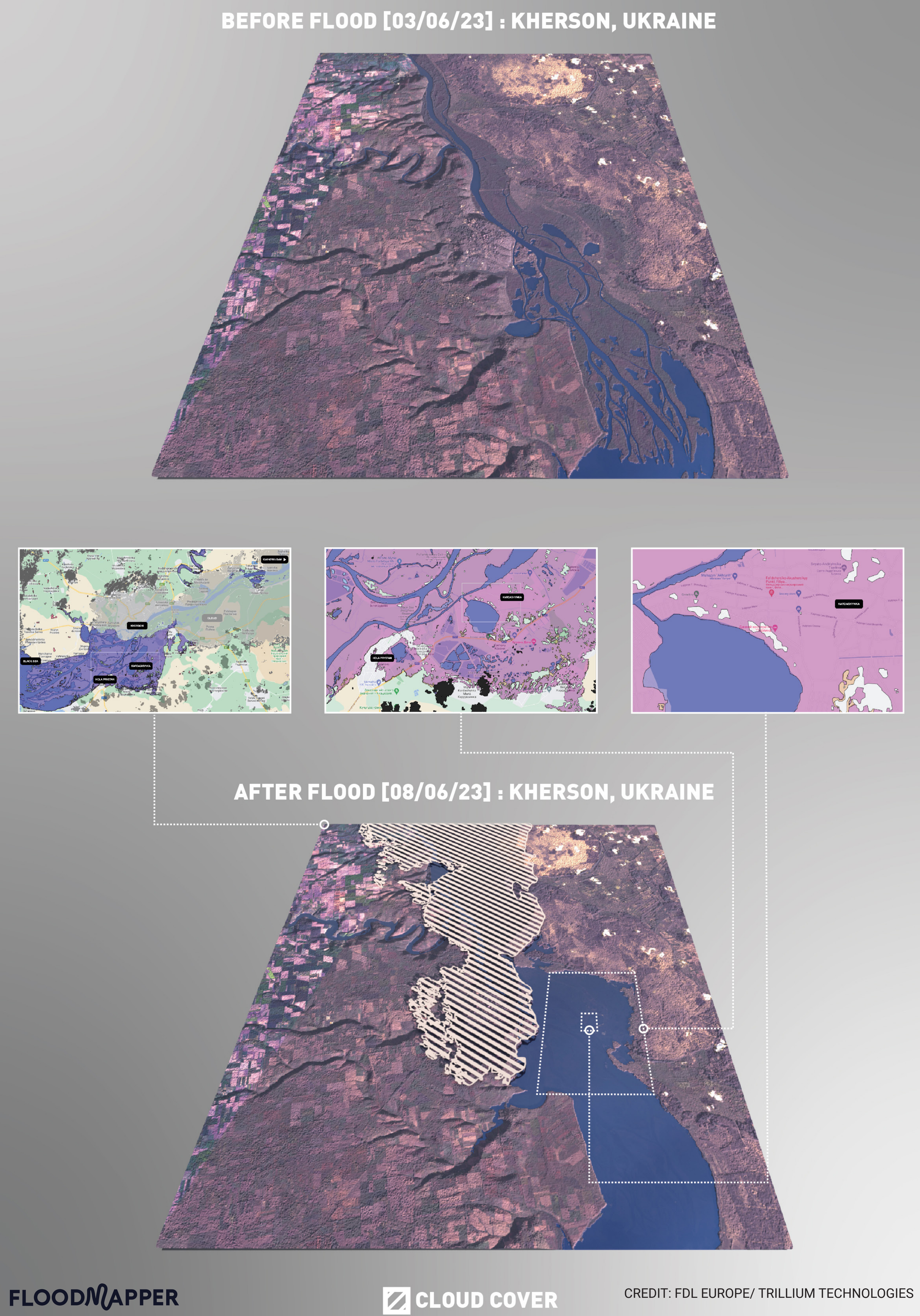
Decision intelligence is also a dynamic concept, where the process of designing the CDD is iterative, and the deployment is reviewed over time.



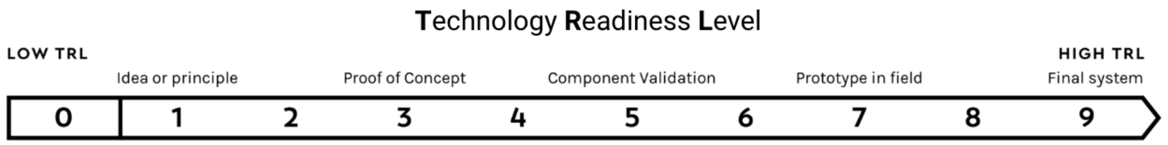
Linked Roadmap Actions

- **MLOps for ESP** (human in the loop)
- **Anthropospheric Models & Data** (Economics, Health, etc)
- **Simulation AI Emulators** (for climate)
- **Physics Informed ML** (for climate)
- **Joint Modelling DI Technology** (standards, interfaces)
- **Digital Twin Earth Computing Engine**
- **Decision Intelligence EF** (integration)
- **Application Development EF** (use-cases)
- **Ethics EF**
- **Deployment EF** (on-ground stakeholder engagement)
- **ESP Professional Engagement EF**
- **SP Conference & Workshop Series**

These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.



OPERATIONS 2: COOPERATION AND CO-CREATION



The journey from innovation to impactful deployment requires a strategic approach that includes multiple stakeholders, transparent development processes and holistic integration. Technology development must be approached from both a bottom-up and top-down perspective, working in parallel. There is a reciprocity between specialisation (zooming in on details) and integration (zooming out to an overview).

During the development of ESP systems, we must:

1. Identify levers for change.
2. Understand how EO and modelling data integrates into decision making.
3. Understand the challenges and barriers from the perspective of end users.
4. Co-design impact-measurement protocols and metrics.
5. Include human-rights, economics, biodiversity and non-human perspectives in our process.

A broad range of stakeholders should be responsible for the design and implementation of ESP systems, including space agencies, governments, industry stakeholders, academia, civil society.

> *"Collaborations between climate scientists and AI researchers are crucial to leveraging the potential of AI for improving climate model projections"*
>
> - Jonathan Bamber Earth Observation Physicist (University of Bristol)

Lowering barriers to multi-scale collaboration

Barriers put up by differences in culture and jargon can make cooperation difficult even where excellent goodwill exists. These include:

- **Barriers around different goals.** Is there a commonly understood vision? It is not clear that the goals would be the same for AI experts and climate scientists, so shared visions of the end goal are needed.
- **Language** that translates between the different domains. It is not just about the apparent direct translation between domains (e.g., EO and economics). We must consider the covert ways in which people use language to signal actions and communicate outcomes.
- **Create more opportunities** for **cross-domain collaboration.**

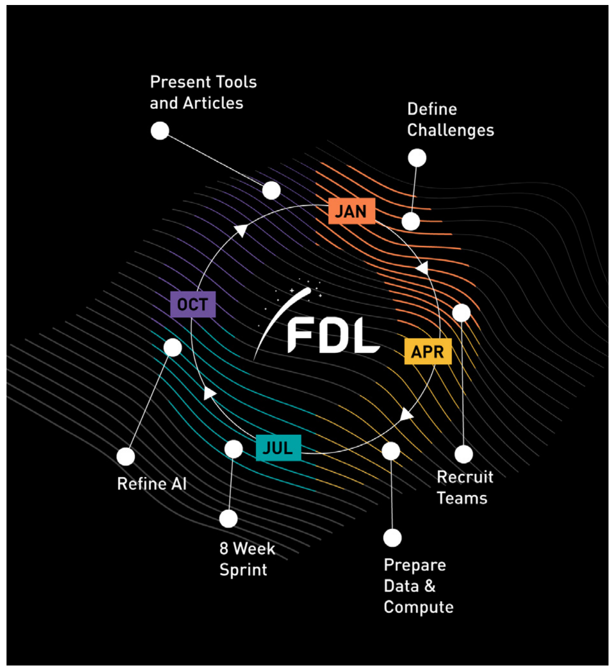
We also need to **organise challenges focused on simulations** in lock-step with data-focused research.

Cross-disciplinary collaboration

Incorporating physics into machine learning is still a cultural challenge: there is sometimes resistance in parts of ML community to collaborating directly with other fields, often motivated by a belief that the data holds all the answers.

Conversely, Earth system scientists are sometimes hesitant to trust ML methods because of a lack of understanding, or an inherent lack of explainability in the algorithms.

However, reality is that ML methods being adopted by scientists everywhere. It will always be challenging to 'embrace the chaos' when mixing fields of study, but there is much to be gained.



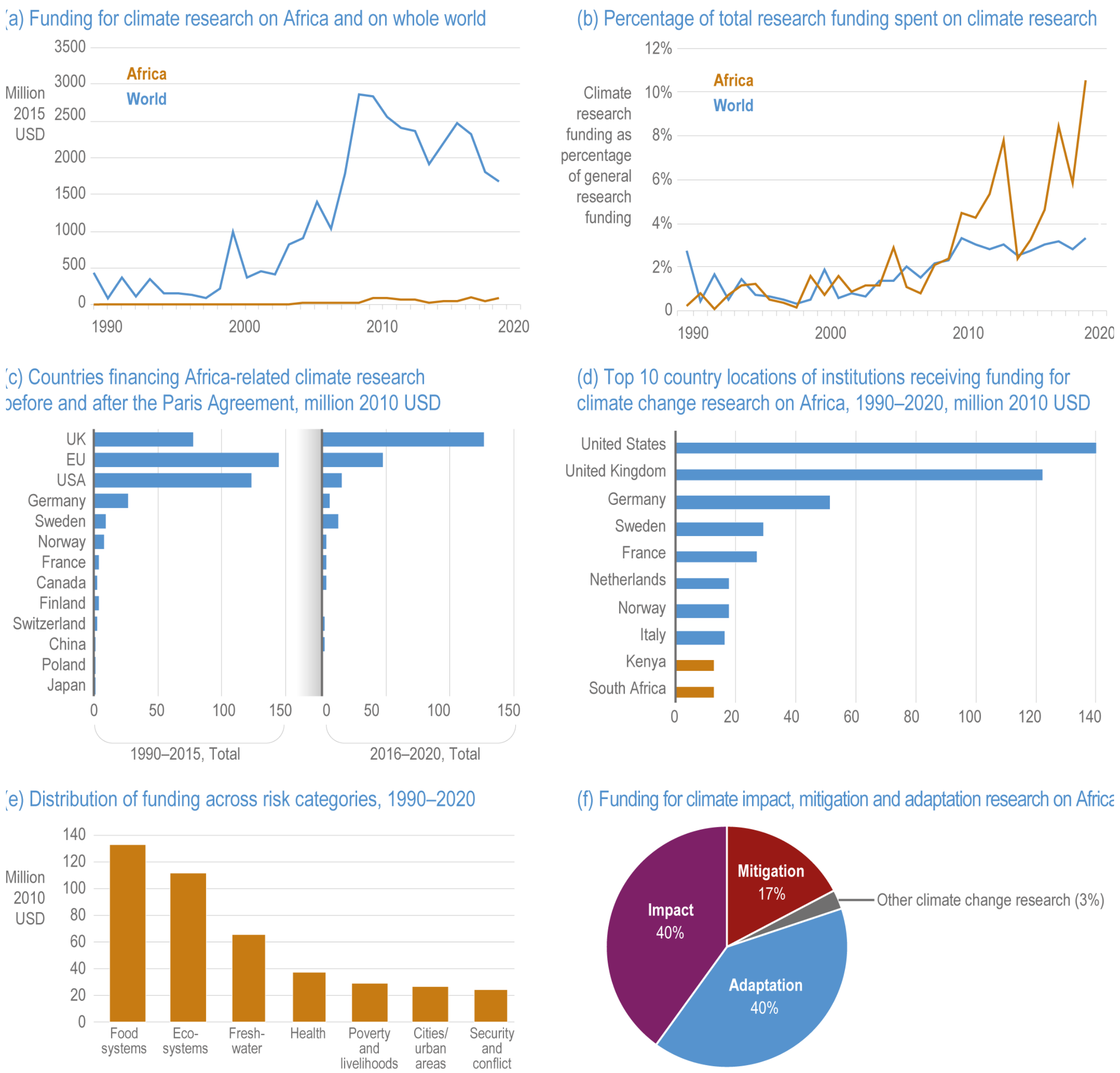
The FDL Europe program is an example of the benefits of interdisciplinarity in AI.

Co-creation with the developing world

The developing world, and especially the global south, may be worst affected by climate change, despite having contributed much less to the causes than developed countries. At the same time, developing countries have much less participation in crafting solutions to the crisis. For example, the six plots below illustrate the tiny proportion of climate research funding for Africa (IPCC 6th Assessment Report).

Distributed ESP technology should be created in full cooperation with the developing world, ideally by building research capacity in those countries. Citizens affected by climate change should be engaged in the spirit of 'listening', employing techniques such as 'concurrent design' to build ESP systems that meet end-user needs.

Funding for climate-related research on Africa is a very small proportion of global climate-related research funding



Outreach, Education and Citizen Science

For true transformation, ESP needs to be embedded in education. This is in line with ESP's core principle of co-creation through deep public consultation. ESP's potential can only be fully realised by involving on-ground stakeholders and supporting organisations in its development.

- Collaboration with government agencies and educational institutions to integrate ESP and decision intelligence into the national curriculum will nurture a generation adept at utilising these tools effectively.
- Develop classroom-ready interactive ESP visualisations and controls to support the use of EO data in educational settings.
- Create a series of professional development opportunities (e.g., training workshops) that are informed by a rich teacher-training curriculum and supported online resources. Significant numbers of educators are uncomfortable teaching science because of a lack of training and experience.
- Support personal links between scientists in ESP-related research and schools: embed scientists in the classroom to provide role-models.

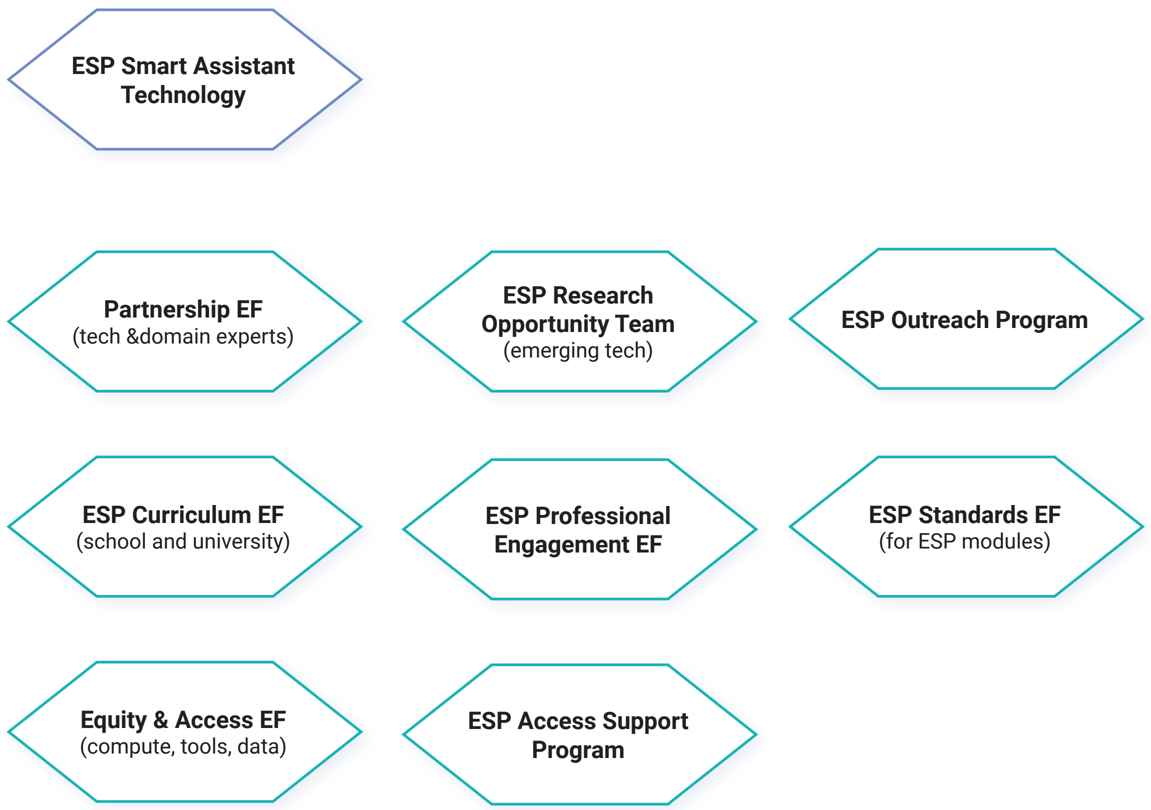
The NOBURN citizen science mobile app aims to measure the properties of fuel in the landscape to better understand the risk of wildfire and wildfire behaviour.

- Engage the public in citizen science help teach them about ESP, but also gather useful data if managed properly. Well-designed programs connect ordinary individuals to the transformative power of data-driven modelling and decision making.

Resources:
'Citizen science and the right to research: building local knowledge of climate change impacts'

nature

Linked Roadmap Actions



These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

OPERATIONS 3: ENSURING ESP IS TRUSTED AND UNDERSTANDABLE

Building trust in ESP systems requires a multi-faceted approach spanning technical, process and consultation considerations.

Trust in Outcomes: Technical Considerations

Independent Benchmarks

Predictions from ESP systems need to be validated against robust independent benchmarks that include historic events and out-of-distribution data (e.g., unseen times or places).

Domain Metrics

Assessment metrics should be driven by the use case, rather than standard AI metrics. Different use cases have different accuracy and uncertainty tolerances. For example, ESP tools that consider medical outcomes should be assessed using well-understood medical metrics.

Continuous Tracking

ESP systems should be continuously monitored and evaluated to check that they are operating well and are still fit for purpose. MLOps and deployment operations are the touchstones here, and tracking should be robust enough to detect unintended consequences.

Audit Chains

ESP systems should have transparent audit chains for data and models. Auditors should be able to back-track what models, data and software versions were used to support decisions.

Explainability

Models should provide realistic uncertainty estimates for predictions and posterior distributions for variables. They should explain how they arrived at a result and the relative importance of input variables.

Privacy & Data Integrity

ESP systems will be trusted with sensitive business and personal information, and must have a privacy-first design. This should be co-created with stakeholder groups, which includes end users.

Appropriate Transparency

Transparency is closely related to privacy and needs to be balanced against safeguarding sensitive data. Some data and models should not be released for ethical or dual use concerns.

Certified Personnel

ESP professionals should be certified in the same way that engineers and other domain experts obtain appropriate industry certifications.

DevOps

Developers of ESP systems should adopt best practices in software development, like unit-testing, module testing, integration testing, functional testing, compatibility testing, combinatorial testing, system testing, acceptance testing, performance testing and more. Machine Learning Operations (MLOps) is now also a well-established field, covering best-practices in managing the lifecycle of models and issues like model drift.

Interpretability and uncertainty in AI modelling

Interpretability in AI is a very active area of research currently, with new developments being made almost daily.

- The emerging field of 'mechanistic interpretability' aims to reverse engineer closed AI systems.
- The CERN organisation is a leader in explainable AI, publishing papers (e.g., Cunha 2020), articles and explainer videos on topics such as Bayesian deep learning.

Understanding the psychology of trusted decisions

An understanding of the psychological and cognitive mechanisms underlying trust is essential for building a trusted system. This includes trust in facts, decisions and leadership, plus the bases for trust: authority, quantitative, empirical, etc.

We need to guard against the fallacy of over-engineering model fidelity as a proxy for trust. Instead we need to situate trust within the factors considered by decision makers. We need to address questions like:

- How much data is enough for trustworthy decisions?
- What should the data quality standard be?
- What model accuracy is required for the downstream system?

These questions need to be considered in the context of a decision maker in a particular role. The academic community may often overlook these kinds of questions.

The structure of trust in AI systems follows a hierarchy: trust in data, followed by trust in models based on the data, followed by trust in decisions based on the model.

Open science, open data and open models

Scientific research is a social endeavour, greatly enhanced by the sharing of ideas and knowledge. Scientific journals, professional organisations, funding authorities and research institutions now recognise that research should always be reproducible, meaning that data, software, models and algorithms should be made open and available to the scientific community.

The AI development community is spread between academia (e.g., the OATML group at University of Oxford) and industry (e.g., NVIDIA Research, Google DeepMind and OpenAI), with "industry racing ahead of academia" according to the AI Index Report 2023.

The same report says that *"the number of incidents concerning the misuse of AI is rapidly rising"*.

Open science offers some solutions to these issues, but also poses new challenges.

- Different levels of open-access exist:
    - **Full access** to data, models (weights, biases & architecture), metadata and research methods.
    - **Sparse access** to data and models, but missing context and details of how the research was done.
    - **Limited access** to a closed model through an API (e.g., the model can be queried or run, but not inspected).
- Making data and models 'open access' supports ongoing research and helps address the reproducibility crisis (where not enough information is available to replicate scientific results).
- Open access policies also support the development of meaningful benchmarks, as models and datasets can be properly characterised (e.g., using model cards and data cards) and offered as a scaffold for follow-on research. See, for example, the 'score cards' system of the ECMWF.
- It is difficult to do research on closed AI models that are only accessible through an API endpoint. However, it is possible as this recent paper on 'Lie-Detection Black-Box LLMs' shows.
- Open access policies can lower barriers to AI research and adoption of AI tools. But the high costs of storing data and accessing compute resources will still be prohibitive to a significant population of researchers and users, especially in the developing world.

Privacy and open science

Scientists and AI developers often have privacy-related concerns about their research:

- Privacy concerns can lead to 'chilling effects' on individual scientists' willingness to make their data available under an open science policy. They may have concerns about ownership, privacy and data protection, or that their data (and models) may be taken out of context and used in an inappropriate way to train AI-driven tools.
- It seems very hard to make research available to the community, while preventing it from being scraped in an unauthorised way by web-crawlers collecting training data.
- Making all data and models accessible in the name of 'transparency' may be counterproductive and cause quality control issues. Transparency does not always mean that more data should be released as some of the data may be confusing or irrelevant. Full transparency may lead to poor decision quality without a full understanding of the context.

Risks of open data and models

Full access to data and models can introduce new risks:

- Risks of AI algorithms, models and code being adopted for purposes they are not suitable for. This can result in systematically wrong, or even dangerous, outcomes.
- Vulnerabilities and security issues within the code, or linked libraries, that could be unwittingly, and naively, adopted into mission-critical applications.

Significant responsibility lies with the organisations hosting repositories of data and models. We must ensure that:

- Data and models are not manipulated for nefarious purposes (e.g., by introducing fake, or deliberately skewed data).
- Resources will continue to be accessible an equitable basis.
- Commercial needs will not be put ahead of scientific integrity or the general good.

The Ethics of ESP

When beginning an ESP project we must consider the impacts of work during the development and after, including:

- Ethical aspects of frameworks, modes of development and outcomes.
- Social impact and the impact on the natural world.
- Legal considerations of the project.

> ***Increasingly complex actions and interventions are associated with greater risks of unintended consequences. Is there a way to measure that risk, a science of unintended consequences?***

ESP software, systems, processes and programmes should be independently certified for use in professional fields.

Certification standards and processes would need to be co-created amongst stakeholders at different levels.

There is an urgent need for trusted ground-truth data, baseline models and example predictions, and the need for associated ethical frameworks to be able to work with that data. The Ada Lovelace Institute is blazing a path on this front.

Many decisions are novel and not repeated, so we need mechanisms for evaluating the quality of decisions. ISO standards, capability maturity models (CMM) and the WELL Building Initiative are strong examples: they assess adherence to best practice standards as a proxy for ground truth when ground truth is not available.

Including the inputs of First Peoples in ESP

The First Peoples of countries like Australia hold significant cultural and practical knowledge of caring for Earth systems; for example in managing wildfire.

ESP systems should learn from this knowledge and work in tandem with different cultures.

First peoples should be consulted and involved in the design of AI systems, learning from organisations like Country Centred Design and the Atlas Biospheric Design Center (AI for decision-making by cultural leaders -see this seminar).

Combating misinformation and malign actors

The concept and use of ESP needs to thrive in the world and be deployed at scale to have an impact on planetary stewardship. This means countering misinformation, deliberate disinformation and outright attacks on ESP science.

- There is a divergence between data-driven approaches and socially driven narratives, including militancy. The vast majority of ordinary people have great difficulty understanding this, so there is bridging work to be done here.
- If misapplied, ESP technology can lead to incorrect conclusions or enable nefarious agendas. How can we guard against the accidental (through ignorance) or deliberate misuse of technology?
- We need to always brainstorm the other edge(s) of multi-use technology. AI is an amplifier, so ask "How can this be used for negative outcomes and how can that use be thwarted?"
- For knowledge systems, there may be a risk of 'non-experts' contaminating information and being counterproductive, whether by accident or design. How do we mitigate this risk? But do we also need to include 'ignorant crowds' in order to offset any 'echo chamber' effect?
- Knowledge - not least in climate action - needs to be qualified by intents, drivers, levers and contexts of those who bring or receive it.

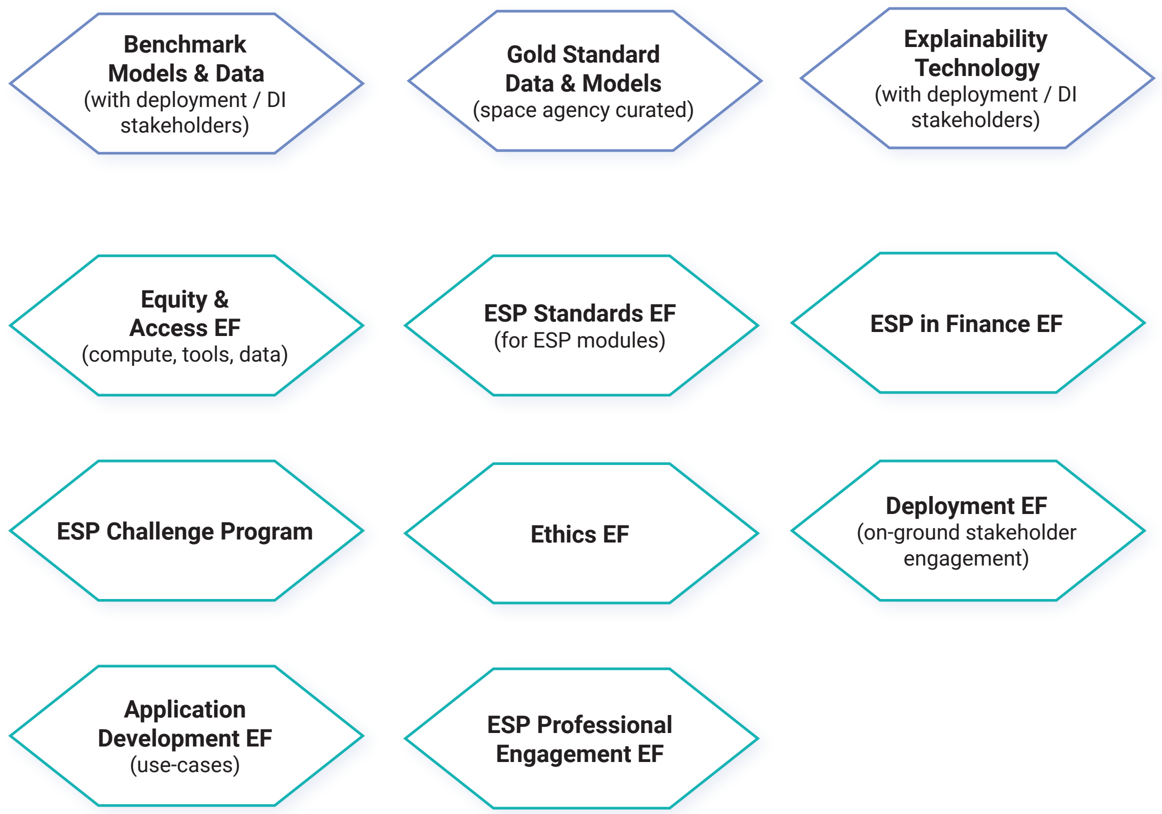
Resource on disinformation:
'The oil industry has succumbed to a dangerous new climate denialism' (2023)

Open Questions on Trust and Safety in ESP

Trust, safety and ethics in artificial intelligence are active areas of research and development, feeding into new AI-related legislation being drafted worldwide. These are just some of the open questions raised at the ESP Forum:

- → Open science also has risks. Should models/data/compute be open sourced? And to what extent? Is there a spectrum? What are the criteria?
- → How is safety complicated or simplified in a Federated Learning model?
- → Ensure equitable access to computing power for disadvantaged groups.
- → How do we guard against fake or deliberately skewed data?
- → How do we protect the information in a DTE by design?
- → How do we define privacy in global-scale digital twin models? (e.g., what's the spatial scale at which the effect of my actions become non-traceable?)
- → "If we create a digital twin of a refugee camp, how can we be sure nefarious players don't take advantage of that?"
- → It seems very hard to make research available to the community, while preventing it from being scraped in an unauthorised way by web-crawlers collecting training data. How do we tackle this issue?
- → Do we need to have some sort of constitution for AI models that have agency?
- → When is it appropriate to establish opt-in or opt-out policies when dealing with open science and privacy?

Linked Roadmap Actions



These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

OPERATIONS 4:
ESP CAPABILITY MAP

Governed by an international organisation

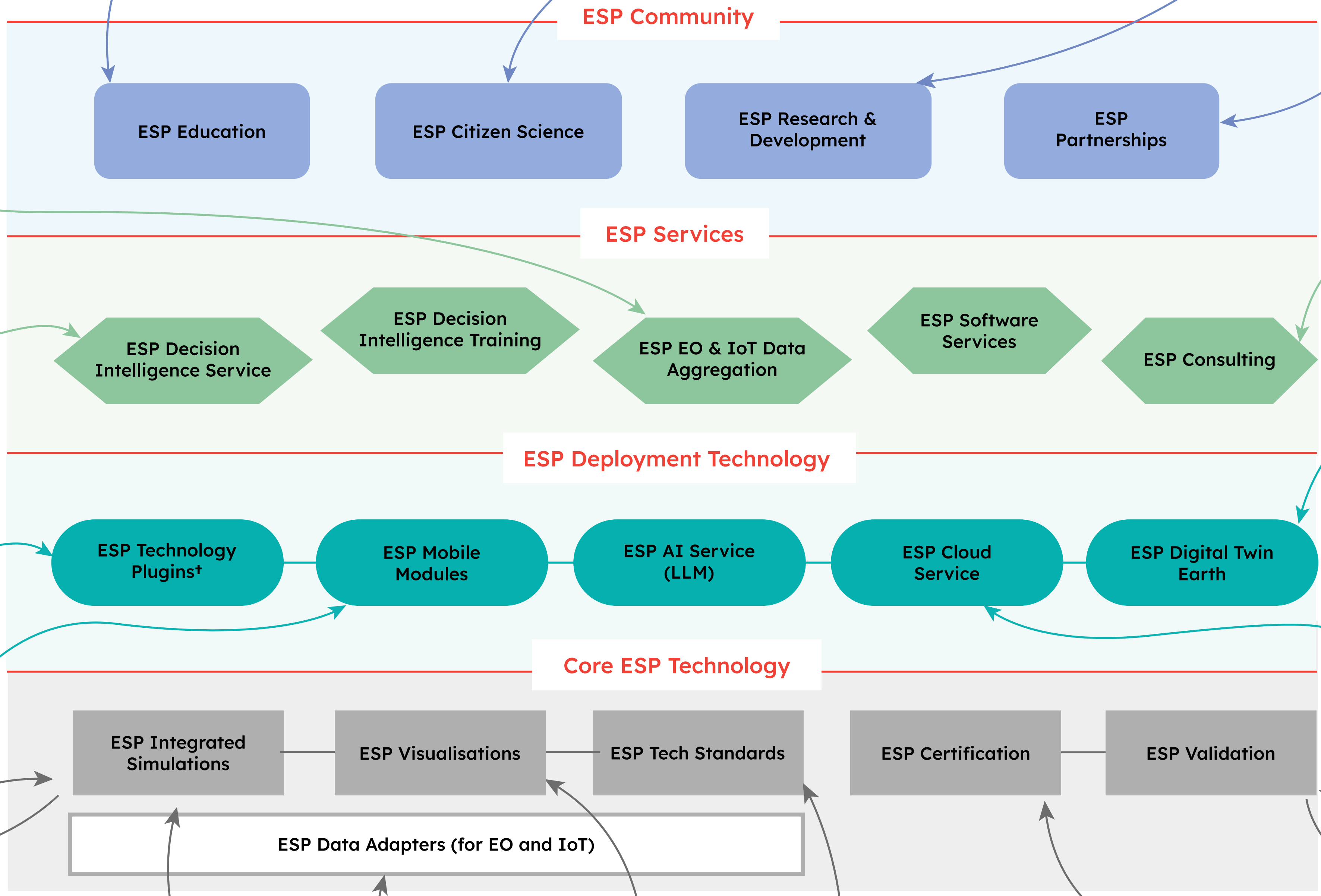
Work with government agencies and teacher organisations to embed ESP DI thinking in the national curriculum.

Citizen science, including measurements and validation, will help to broaden understanding of ESP and engage ordinary citizens.

Public-private partnerships to develop new capabilities and bring them to deployment-ready status. Development should happen in a transparent environment with open results (e.g., FDL.AI).

Provide data aggregation and quality control services for EO and Internet of Things (IoT) sensors (e.g. to support near-real-time analysis).

Deep public consultation and co-creation is at the heart of ESP. The ESP community will need to co-create capabilities by working with on-ground stakeholders and supporting organisations.



Provide deep end-to-end DI and ESP consulting services, including training on ESP tools and concepts.

Provide ESP analysis services to businesses and organisations. These should be certified to meet (or ideally exceed) regulatory needs.

ESP tools are designed to integrate with existing software systems for ease of deployment. ESP tools enhance functionality in a similar way to GitHub Copilot.

ESP technology supports regional, national and international digital twin systems, including Digital Twin Earths. Adopting ESP technology facilitates cross-collaboration and interoperability.

Deployment technology should support a wide range of use-cases and computing devices, from cloud-computing to mobile phones.

ESP technology powers centralised sites in the style of an Earth Operations Centre, but also decentralised capabilities that can be accessed or deployed remotely.

The simulations powering ESP technology are constantly evolving, so ESP developers and system integrators must work with modellers to track and respond to changes.

ESP is integration: Joint modelling of Earth and human systems.

ESP supports easy access to EO data.

Compelling visualisations are at the heart of ESP.

ESP technical standards support interoperability between models and data.

'ESP Certified' is the mark of quality.

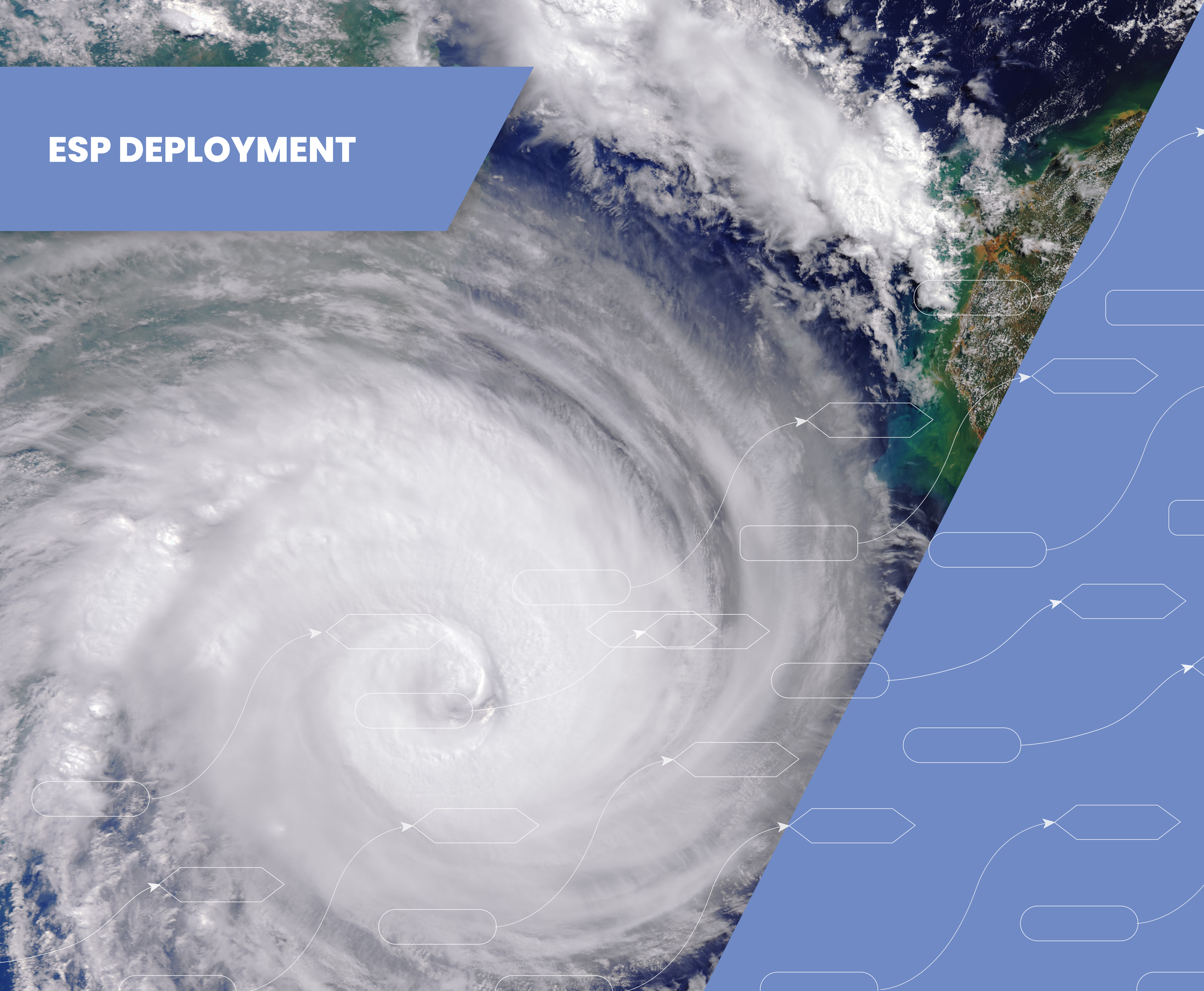
Work with professional bodies to continuously validate ESP technology for specific use cases. This should be done in an open way where possible.

† Technology plugins includes open data, open models, open submodels, open ontologies / knowledge graphs, and more:: an entire "world wide decision web" of components.

ESP technology needs to integrate into our daily lives. A map of the components*, or capabilities, in the ESP Stack might look like this.

**We don't attempt to show how each component works in this diagram*

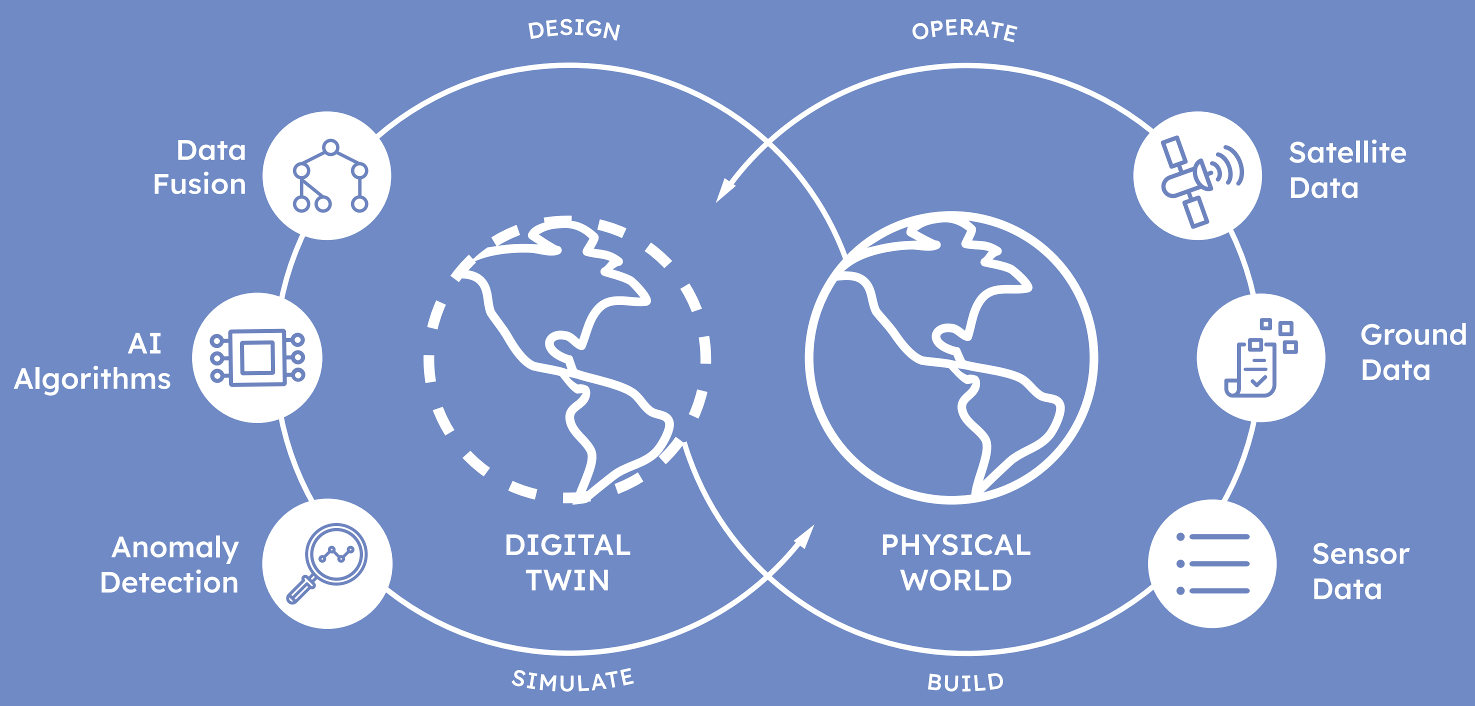
ESP DEPLOYMENT

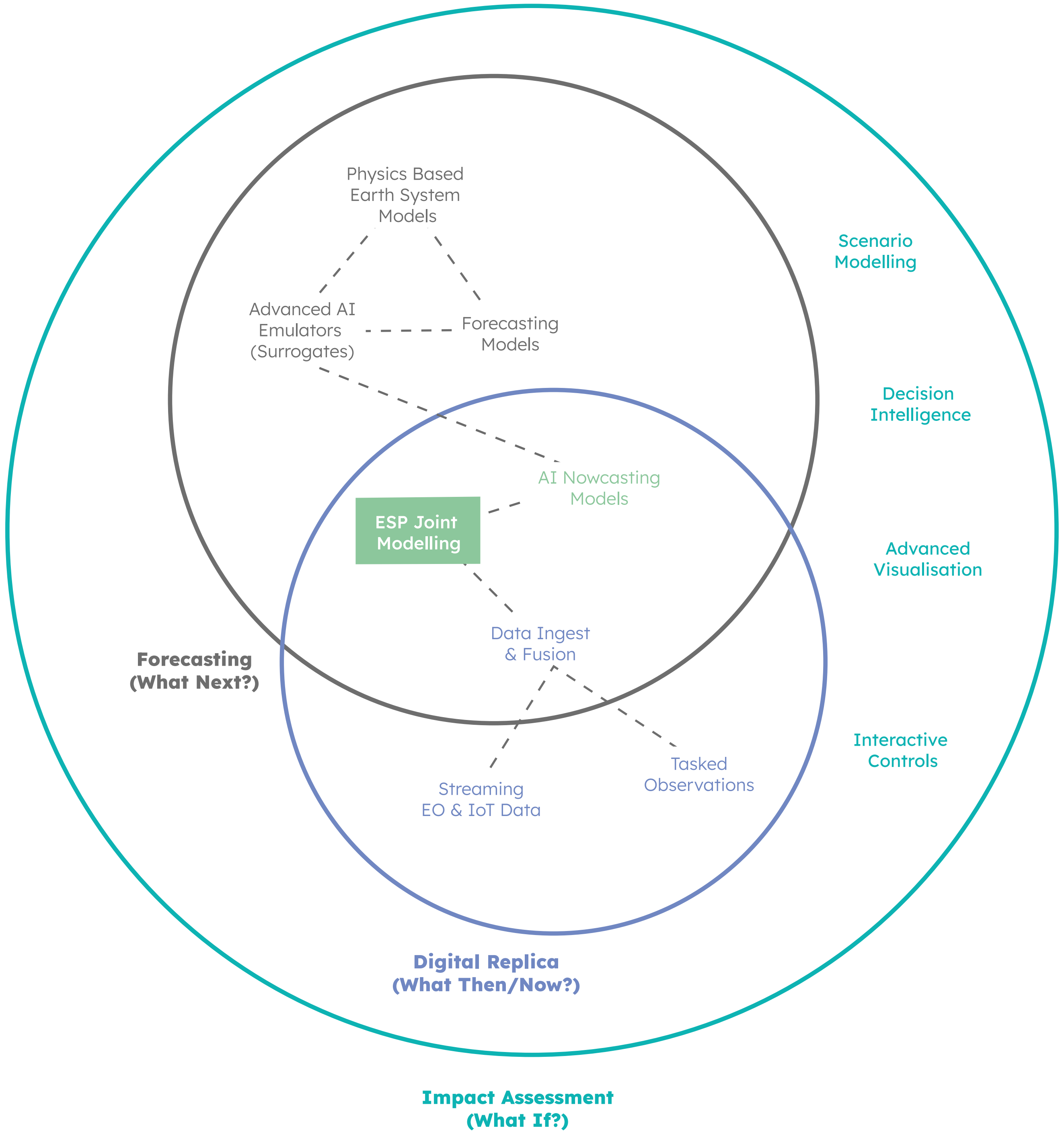


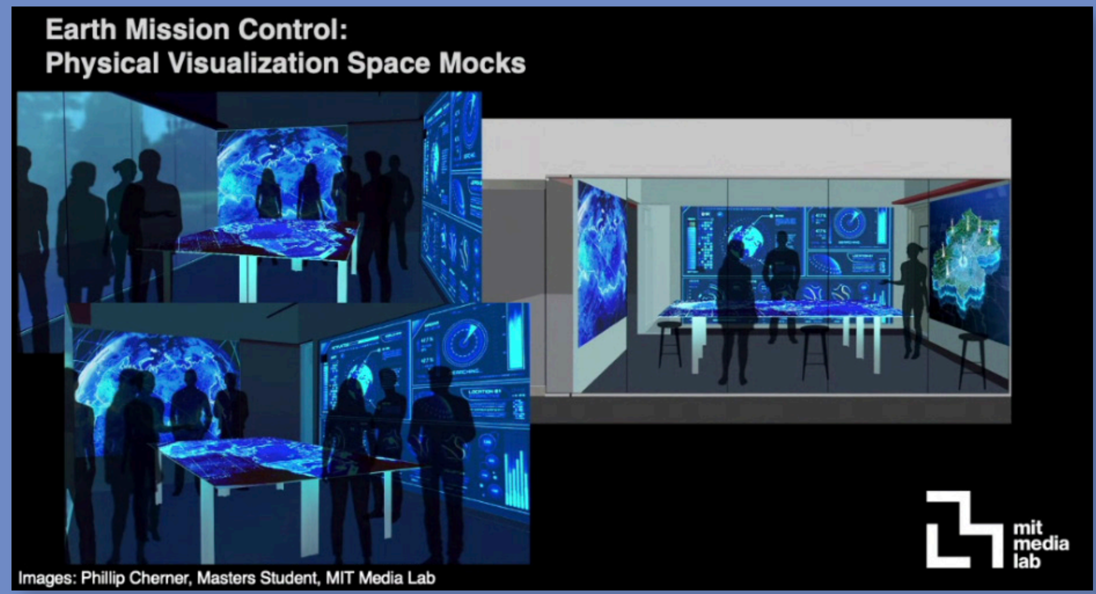
DEPLOYMENT 1: EARTH OPERATION CENTRES AND DIGITAL TWINS

The concept of a Digital Twin Earth (DTE) builds on Earth system simulations or emulators by using them as a basis for a spatial information system. A DTE consists of three main components:

- The **Digital Replica** presents images of the past and present created using continuous observations of interacting Earth and human systems (What then? / What now?)
- The **Forecasting** component presents an integrated picture of how these Earth systems are going to evolve in the future based on current information and a set of potential actions (What next?)
- The **Impact Assessment** component explores what effects will be experienced in hypothetical futures, including cross-domain modelling (What if?)
- **DTEs are interactive,** allowing models to be queried and manipulated and the outputs visualised to enhance understanding and support decision-making.
- **DTEs are connected** to near-real-time information on the status of the Earth (e.g., from EO satellites and internet-connected sensor networks).
- DTEs interconnect the simulation systems into **fully-linked systems-of-systems** that offer controls and present detailed visualisations of predicted outcomes. The core visualisation in a DTE is usually a representation of data on a virtual planet Earth, but also allowing the composition of advanced spatio-temporal queries.







The concept of Digital Twin Earths is very close to the idea of an Earth Operations Centre, or Earth Mission Control.

EOC: A case study for deploying ESP at scale

The concept of a Earth Operations Centre (EOC) was first suggested in the 'Space for Net Zero' report released by the World Economic Forum in 2021. The vision is for a mission control for Earth: **an interactive physical information space** that gathers and presents intelligence and predictions on the state of our planet. It would:

- **Prioritise** and **facilitate** the use of EO data to show the current state of our planet.
- Use ESP technology to **present insight beyond directly measured data**. This could include land-use, soil moisture, amount of carbon captured, gas leaks, hazardous events and traffic flows.
- Use ESP simulations to **predict future states,** such as water use, water availability, temperature, fire risk, flood risk, crop yield, economic metrics and more.
- Offer a plug-in system to **easily connect external models** after the models pass through a review and certification process.
- **Continuously validate and update** model performance to track real-world shifts.
- **Offer the expertise of a dedicated staff** of scientists, policy experts, negotiators and program developers.

Physical EOCs as research and collaboration nodes

An EOC could be a shared facility where politicians, civil servants and business leaders could **meet to plan decisions with the aid of world-class experts** on joint modelling and direct access to insights from EO data.

The facilities and multidisciplinary expertise at the EOC would also **support focused programmes to catalyse cooperative research and development** around the world, a gravitational attractor for researchers and decision makers facilitated by international agreements. EOCs would be:

- **Multidisciplinary by design,** bringing together scientists, economists, policy analysts and representatives of ordinary on-ground stakeholders.
- A remit and structure that encourages **thinking about ESP in a holistic way** to approach decision making; not just considering the environmental implications, but also the economic implications of what the future of our world will look like.
- **Arenas for brainstorming** and identifying key levers for change.

The services offered by Earth Operations Centres should be aligned closely with the UN Sustainable Development Goals.

EON Supported Activities:

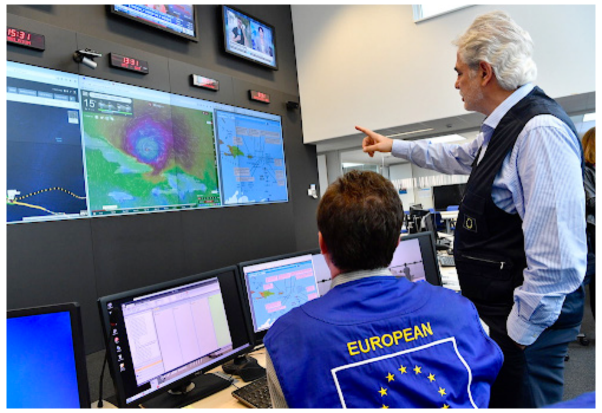
An ESP-driven Earth Operations Network (EON) would be used for multiple activities.

Long-Term Planning:

Facilitating the development of plans to manage change on national or even international scales.

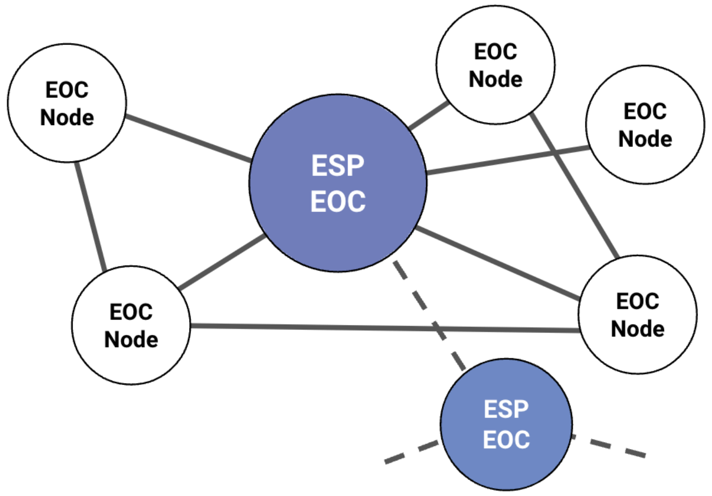
Research Collaboration & Tech Development:

EOCs would be hubs for conducting transdisciplinary research and development into new clean technologies.



Crisis Management:

The data-fusion and modelling power of EOCs could support evidence-based decision-making during natural disasters, as one example.



Earth Operation Networks (EONs)

A single EOC can only have limited influence and impact. A global network of EOCs are needed to serve different sized communities, from regional to international. This means that the systems in the EOC must be **agile, scalable** and **connected.**

Nodes in the EOC network could be operated on a franchise model with democratised tools provided to operators who meet certification requirements.

ESP Cloud

A **cloud-based ESP stack** would stretch to these scales and simplify deployment while also supporting **distributed operations.**

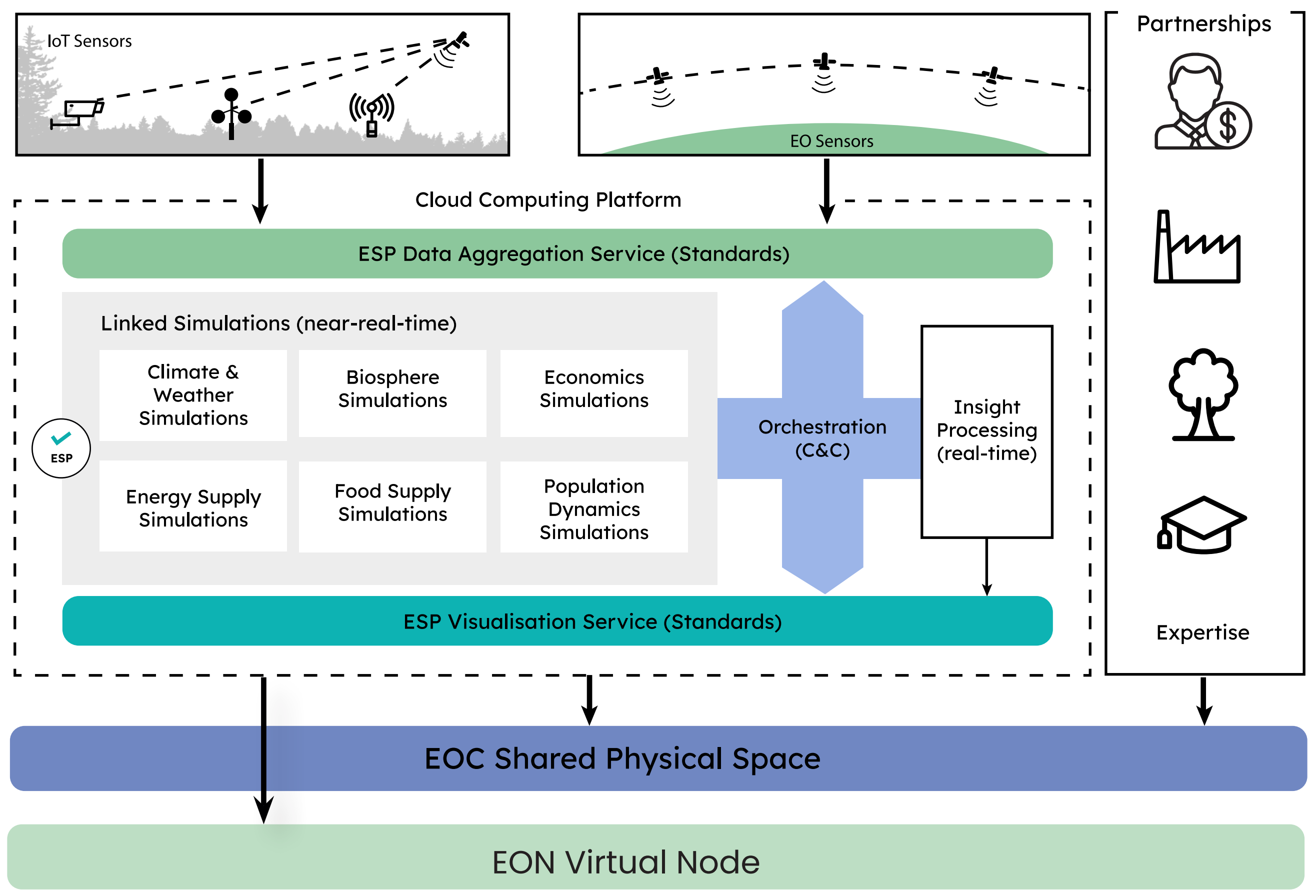
EOC front-end software must be light enough to run on **low-power computing devices** (e.g., smartphones) to allow developing nations to participate.

Some more advanced capabilities (e.g., visualisation, emulation) could be offloaded to client devices via **efficient software plugins** that are ESP Certified.

USD is the open, extensible framework and ecosystem with APIs for composing, editing, querying, rendering, collaborating and simulating within 3D virtual worlds. Alliance for OpenUSD brings together all communities working on digital twins.

AOUSD
Alliance for OpenUSD

How an EOC might work



Learning from the ALMA Project support network

The Atacama Large Millimetre Array (ALMA) is the world's largest ground-based telescope array, involving a collaboration between partners in Europe, the USA, South America and Japan.

A network of ALMA Regional Centres (ARCs) Nodes and Centres of Expertise at dedicated locations offer face-to-face and remote support on all aspects of using ALMA. The ARCs also host the developers and technical staff that build the telescope software suite and end user tools.

ALMA ARCs can serve as a useful model for a network of Earth Operations Centres.

An ecosystem of digital twins

Decision-making organisations building digital twin Earths will need to decide their scope:

- By stakeholders at different organisational levels: **International → National → Regional → Local.**
- Covering different themes, or combinations of themes: **broad and transdisciplinary → interdisciplinary → thematic.**
- Prioritising different types of input and outputs: **spatial, geographic, temporal, cultural and more.**

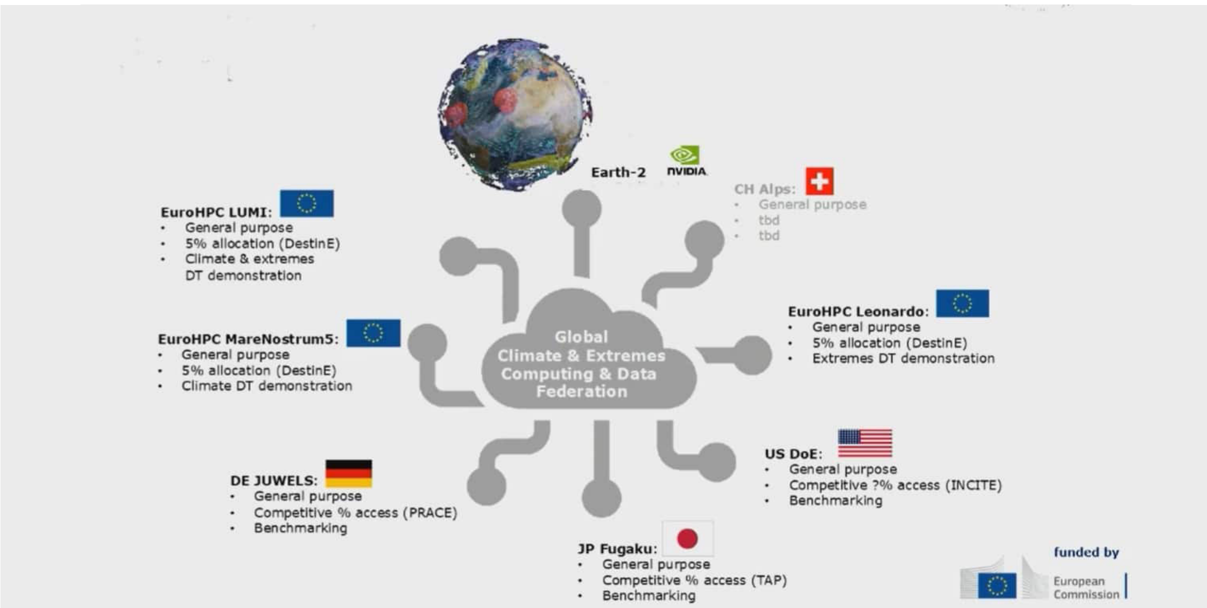
In order to avoid 'reinventing the wheel', there should be standards in the architecture of a digital twin at a high level (e.g. API hooks). For example, Destination Earth (DentinE) is now building a 'digital twin engine' that will likely be used by most of the European thematic digital twin prototypes.

Governance and financial support of future DTEs

DTEs and the underlying simulation technologies will need to be sustainable in terms of ongoing finance and management.

Because of their size and scope, DTEs must be collaborative in nature: they are unlikely to be built or sustained by one entity.

We need interoperability standards and protocols at the semantic, legal and organisational level.



Existing DTE development efforts, as outlined in a recent Nvidia blog post. Inspired by the recent concept paper on digital infrastructure for an Earth Virtualisation Engine (EVE).

Destination Earth: Europe's Digital Twin Earth

The European Commission's Destination Earth (DestinE) initiative seeks to develop a precise digital model of Earth using innovative Earth system models, cutting-edge computing, satellite data and machine learning.

The primary goal is to monitor and predict environmental changes and their impacts. Through this, users can explore the effects of climate change on various Earth components, devising adaptation and mitigation strategies.

DestinE aligns with the European Commission's Green Deal and Digital Strategy, addressing climate change, biodiversity and deforestation, while also monitoring food security and polar regions.

ESA, Eumetsat and ECMWF play integral roles, providing observation data, a multi-cloud data lake and the Digital Twin Engine.

The project is planned to gradually evolve, with milestones including the creation of the open core digital platform and initial digital twins by 2024, culminating in a comprehensive digital Earth replica by 2030 through the integration of digital twins.

ESP for Smaller Decisions: Software Plugins

A case study for deploying ESP at small scales

Many organisations have extensive and deeply embedded systems that allow them to do their jobs. These take the form of large multi-use software packages or detailed procedures designed to meet regulatory needs or cultural norms. ESP modelling software will need to work with and within such systems, perhaps in the form of software plugins or modules.

Examples of software tools large organisations use on a daily basis include:

FME for data analysis:

https://fme.safe.com/

SAS analytics:

https://www.sas.com/

Linked Roadmap Actions

- Networked Earth Operations Centre (deployment case)
- ESP Cloud Computing Software Stack (tech platform)
- Digital Twin Earth Computing Engine
- Joint Modelling DI Technology (standards, interfaces)
- MLOps for ESP (human in the loop)
- Simulation AI Emulators (for climate)
- Streaming EO Data Processing: Live Twinning
- Interactive Interface Technology
- Digital Twin Earth Visualisation
- ESP Smart Assistant Technology

- Decision Intelligence EF (integration)
- Application Development EF (use-cases)
- ESP Challenge Program
- ESP Live Lab Program (deployment trials)
- Ethics EF
- ESP in Finance EF

These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

DEPLOYMENT 2: INTEGRATING ECONOMICS AND FINANCE

Our financial system plays a crucial role in driving sustainable business and economic practices. Systems like 'triple bottom line' accounting (jointly optimising for positive economic, social and environmental impact) have been proposed, but the lack of standards, too-large array of choices and poor regulation have worked against widespread adoption.

We need:

- Well-designed financial incentives and regulation directly linked to ESP modelling.
- Recognised benchmarks and ongoing monitoring of projects.
- Adoption of the concepts of risk, trust and uncertainty.

ESP technology can support all of these goals, but we also need global coordination of policies, encoded into international treaties.

Social Investment as an ESP Trailblazer

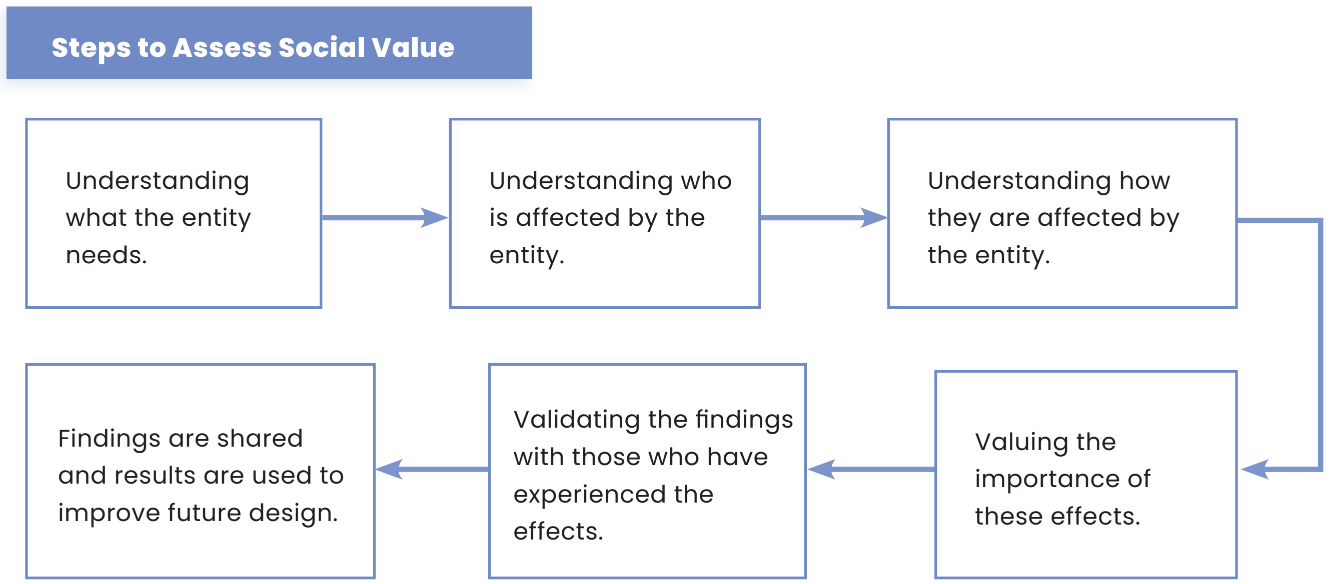
A great example for the wider financial industry is the social investment movement. The new discipline of social return on investment (SROI) analysis has grown out of entrepreneurship addressing social and environmental problems.

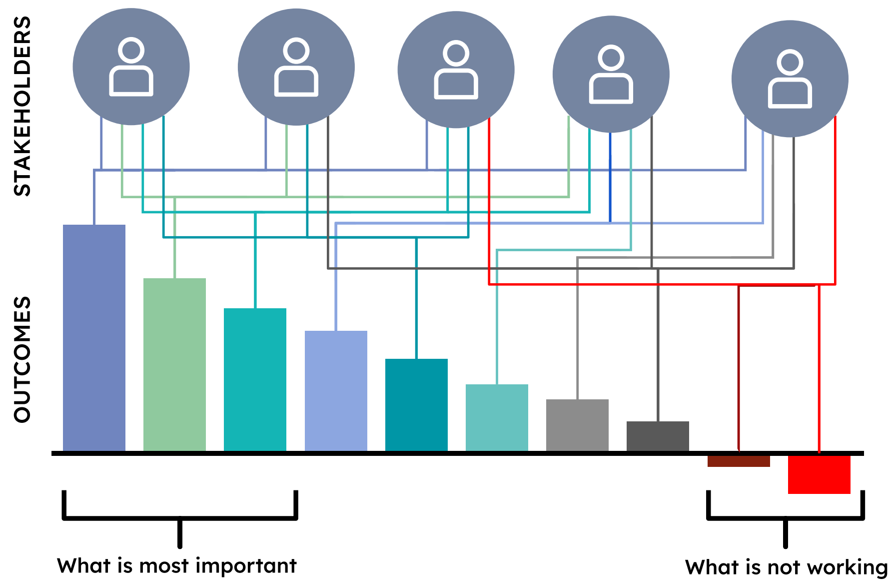
SROI has developed methods of measuring and validating the deep impacts of social investments in monetary terms, working with recipients to put a monetary value on the outcomes.

There is now a standardised auditable methodology for SROI and a global network with 26 national chapters around the world. Increasingly, governments and private equity firms are adopting SROI impact auditing, contributing to standardisation.

ESP can use SROI as a template and trailblazer for linking into the world. In SROI, assessment and validation is done by stakeholders that are directly affected by interventions. They collectively rank the outcomes and estimate a monetary value so that impact is comparable to other financial metrics.

Continual monitoring of impact is critical to SROI, especially for avoiding perverse incentives. This is the danger of creating incentives that achieve the opposite of what they were meant to do. Perverse incentives can be mitigated by continually monitoring for impact and then acting on those results.

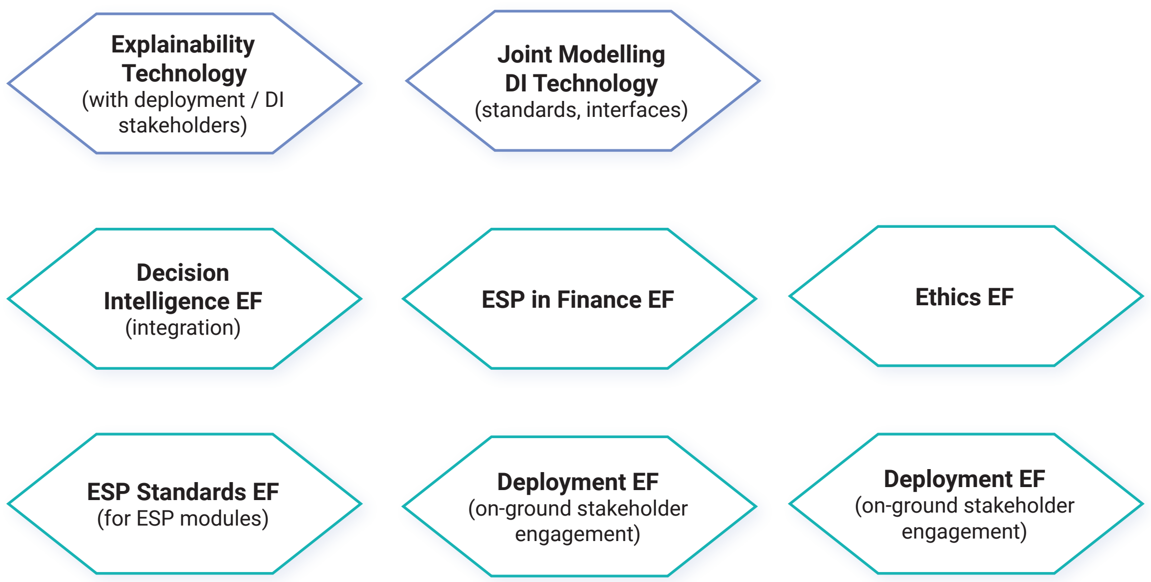




Software for SROI as an ESP pathfinder

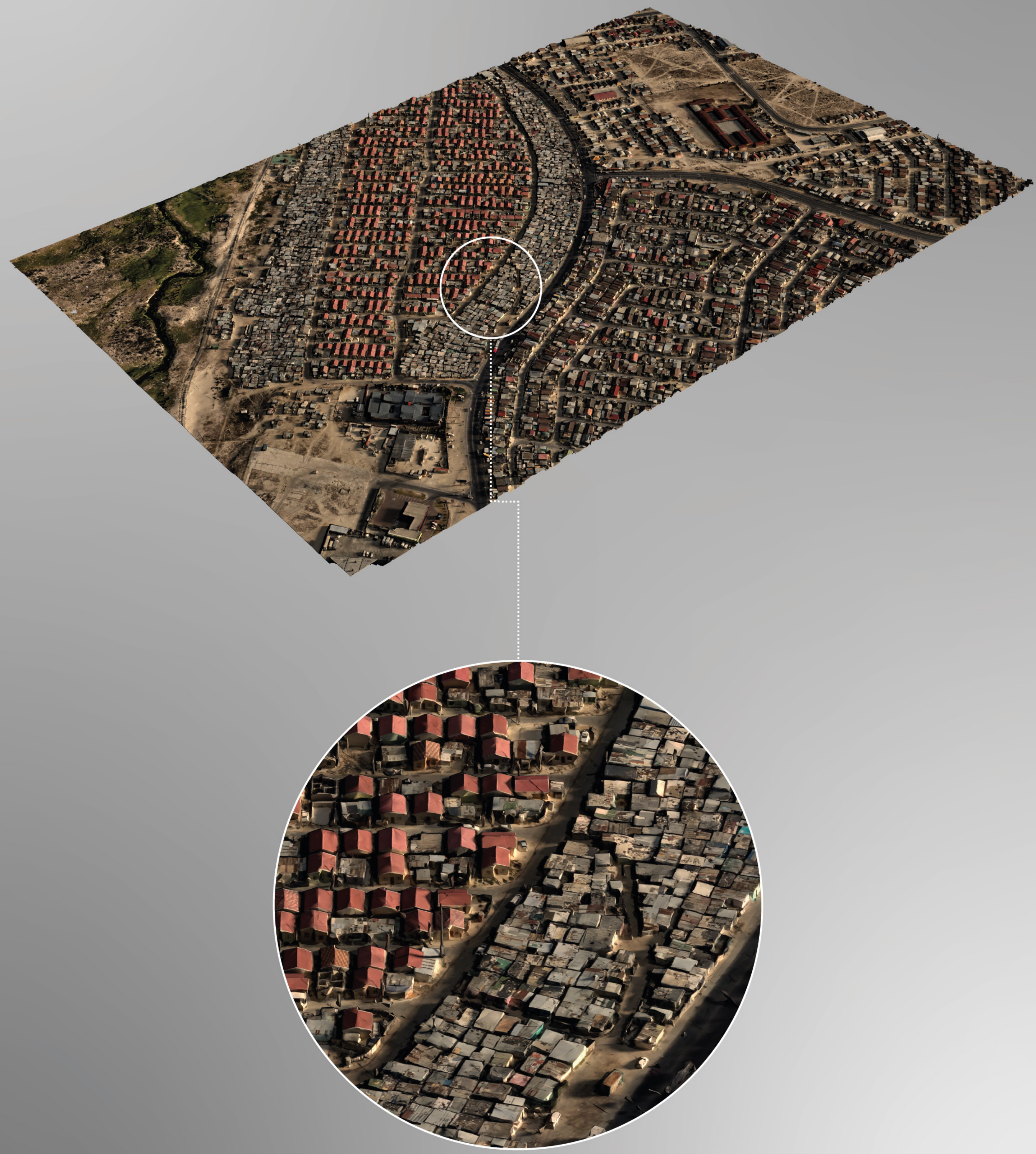
The 'Social Value Engine' software was developed in the UK in response to the Social Value Act of 2012, which requires public bodies to think about how they can secure wider social, economic and environmental benefits. It is accredited by Social Value International and provides best practice SROI measurements to include in different toolkits.

Linked Roadmap Actions



These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

IDENTIFYING INFORMAL SETTLEMENTS (E.G. THE UNBANKED) FROM ORBIT



CREDIT: FDL EUROPE / TRILLIUM TECHNOLOGIES

DEPLOYMENT 3:
SIMPLE ESP METRICS

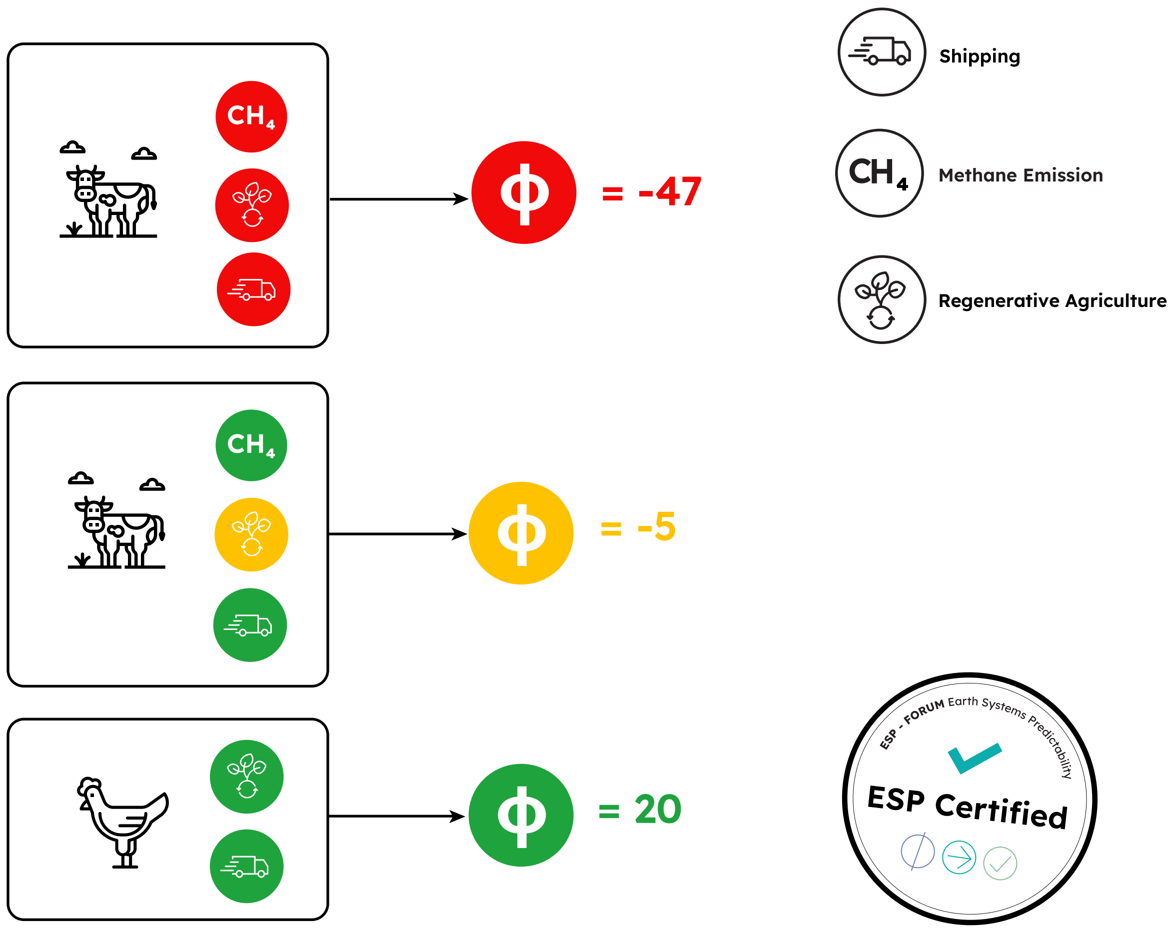
An easy way to make everyday choices for planetary health

Our vision is for ESP to be embedded in a wide array of tools used by ordinary people so that they can easily understand the impact of their daily decisions on the present and future climate. We understand there is a wealth of complexity in mapping choices to outcomes via a causal decision diagram, so we want to simplify everyday decisions.

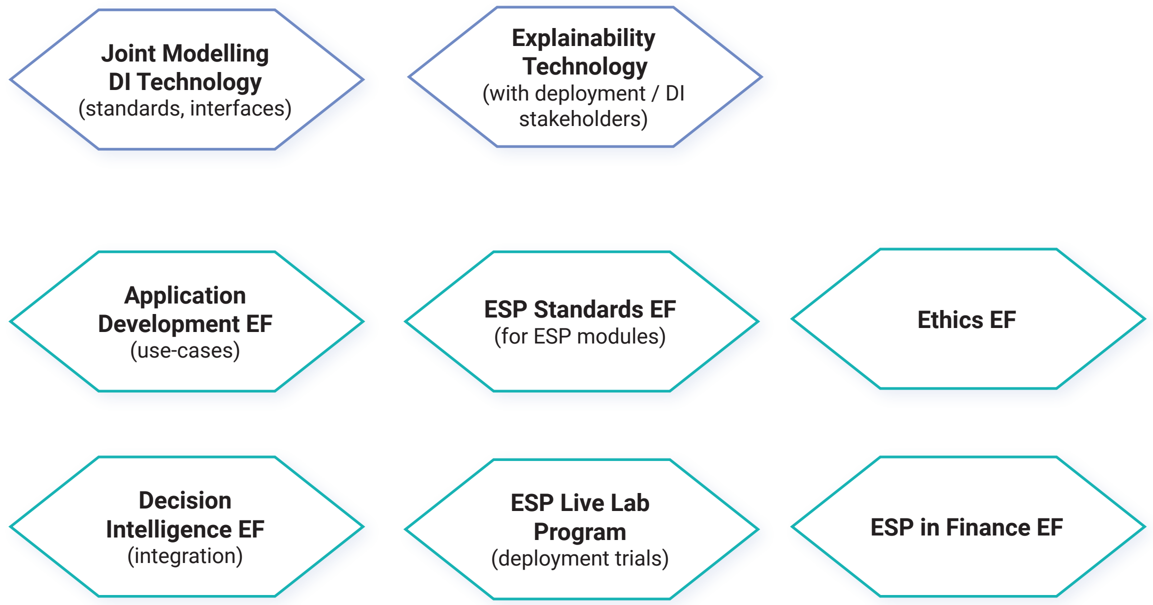
We propose encapsulating this complexity in a single metric for 'planetary health' - the Φ (Phi) symbol.

Φ would be a universal scale in the same way as energy efficiency or food health ratings, except extended to negative values to indicate net harm. Food could be labelled with a Φ score that includes production factors like methane emissions, renewable agricultural practices and shipping.

Φ scores could be pre-calculated for simple choices, but would need to be computed on-the-fly for complex decisions. This means that ESP modules or plugins would need to be part of most software packages.



Linked Roadmap Actions



These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

DEPLOYMENT 4: WHO NEEDS ESP?

Business

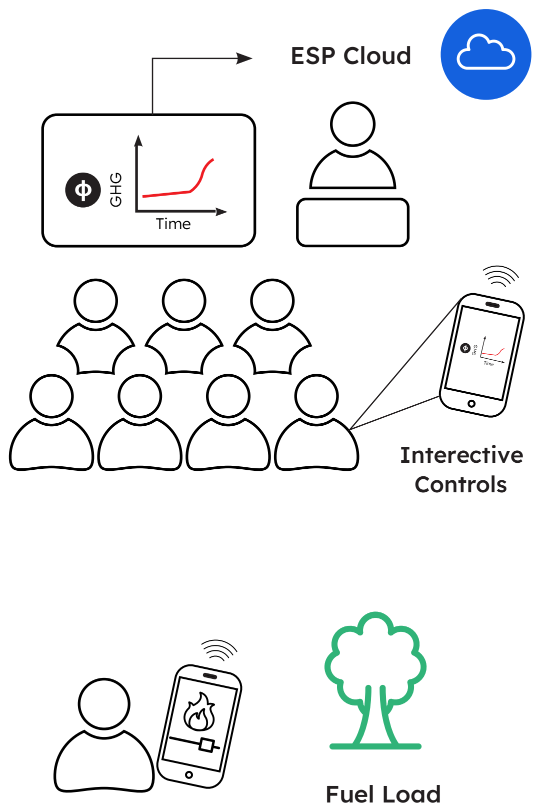
Business can use ESP technology to make choices that have a positive impact on our future climate. In this example, a restaurant manager needs to choose a coffee supplier. They use ESP-aware tools to choose a supplier that offers a good deal for the business and for the planet.

ESP Report
*Supplier **(4)** is recommended because joint modelling shows that the product is sustainably produced, yields are likely to remain stable and the climate cost of transport and production are low. (4) is als o a cost-effective choice.*

This choice needs to be properly incentivised or regulated so that the economic choice is well aligned with the climate-friendly choice and based on validated measurements.

Educators

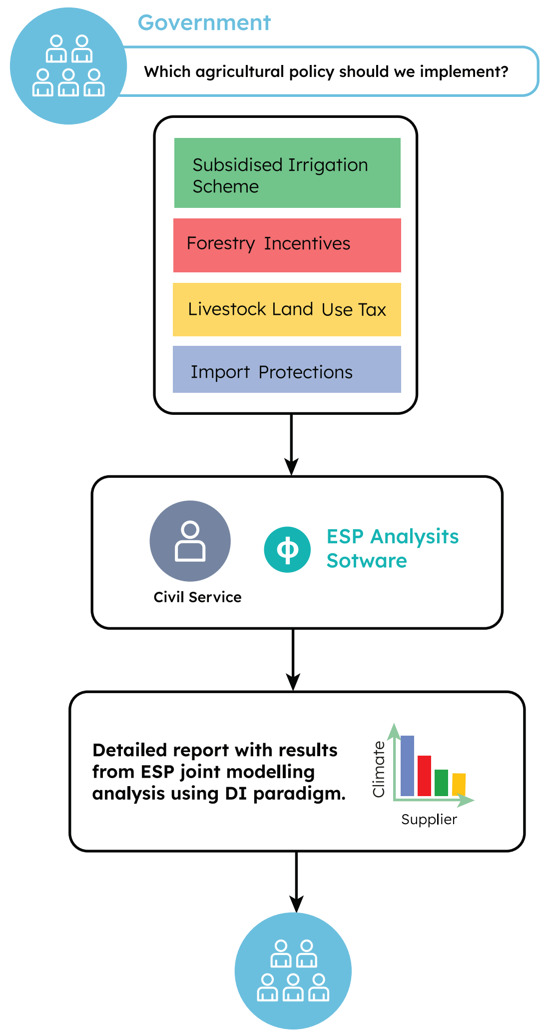
ESP technology will offer new ways for students and teachers to form a deep understanding of how large and small decisions affect the planet. The speed of AI technology combined with the latest visualisations will support student-led interactive exploration and compelling narratives.



ESP technology on mobile devices will also support citizen science through AI-assisted field measurements.

Government

Governments will find ESP technology indispensable when deciding on evidence-based policy. In this example, civil servants needs to choose between potential agricultural policies. Joint modelling that includes social aspects will help to avoid perverse outcomes.



Political decisions on large scales are complicated; binding international agreements will be needed to encourage positive behaviours.

Disaster Management

ESP technology is already critical for disaster management at all phases: mitigation, preparation, response and recovery. For example, satellite data informs predictive models of fire risk derived from long-term weather forecasts. These are used to undertake fuel reduction burns and other remedial actions. During an active fire, the incident management team uses the latest data to formulate an Incident Action Plan and direct firefighting activities.

ESP technologies will make this process more dynamic and allow planning on further time-horizons.

ESP for disaster management is closely aligned with the concepts of Earth Operations Centres and Digital Twin Earths.

Academic & Industry Research

Many new ideas and technology emerging from research groups in universities and businesses will have implications for climate change and planetary habitability. **ESP tools could be deployed in service of research:**

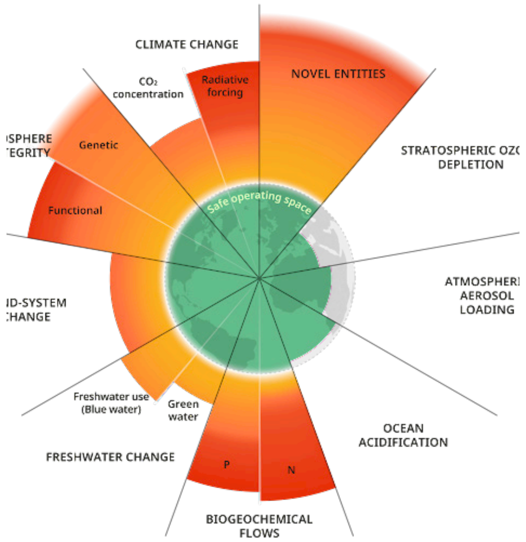
- Help quantify the range of climate and habitability impacts for emerging research.
- Assist with the development of climate-friendly, operations, business models and deployment strategies.

Ongoing fundamental and applied research on ESP technology will be needed for the foreseeable future and should be conducted in the spirit of open, trusted science. Initiatives like the Open Science Conference of the World Climate Research Programme (WCRP) illustrate the value of open collaborations in building ESP solutions..

Legal

ESP processes and technology needs to be applied on the scale of countries and continents to affect meaningful systemic change. One way to achieve widespread adoption is through ESP-focused **regulation** (professional bodies), **legislation** (government) and **treaties** (international).

The concept of 'planetary boundaries' is useful: these are a framework to describe the limits of human activities on the whole Earth system. Crossing these boundaries would lead to our environment tipping into a different - likely harmful - state.



Chapron et al (2017) argue that legal boundaries must be linked directly to planetary boundaries if we are to stay within limits. ESP technology provides a method to build quantitative links.

Capital Markets

The market for sustainable investments that consider environmental, social and corporate Governance (ESG) aspects has grown significantly in recent years. ESP could play a similar, potentially more impactful role, in capital markets - especially if supported by legal frameworks.

For example, ESP twinning technology will allow **modelling of carbon markets** from an integrated perspective, providing quantitative estimates of carbon sequestration opportunities, or carbon emission risk, coupled with economic models, to build far-seeing projections.

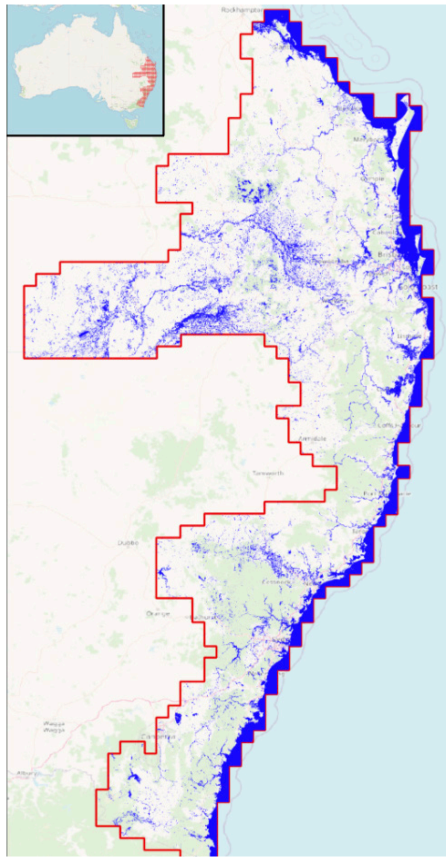
Investors, and investable businesses, will be supported by ESP technology:

Reporting & Compliance:
Companies can use ESP technologies to report on regulation compliance, backed by solid scientific metrics that have been validated by on-ground stakeholders.

Transparency:
ESP should place high value on transparent and validated processes that are adopted into the core of the technology. This will increase the financial integrity of ESP-certified organisations and allow investors to make informed and interpretable judgements.

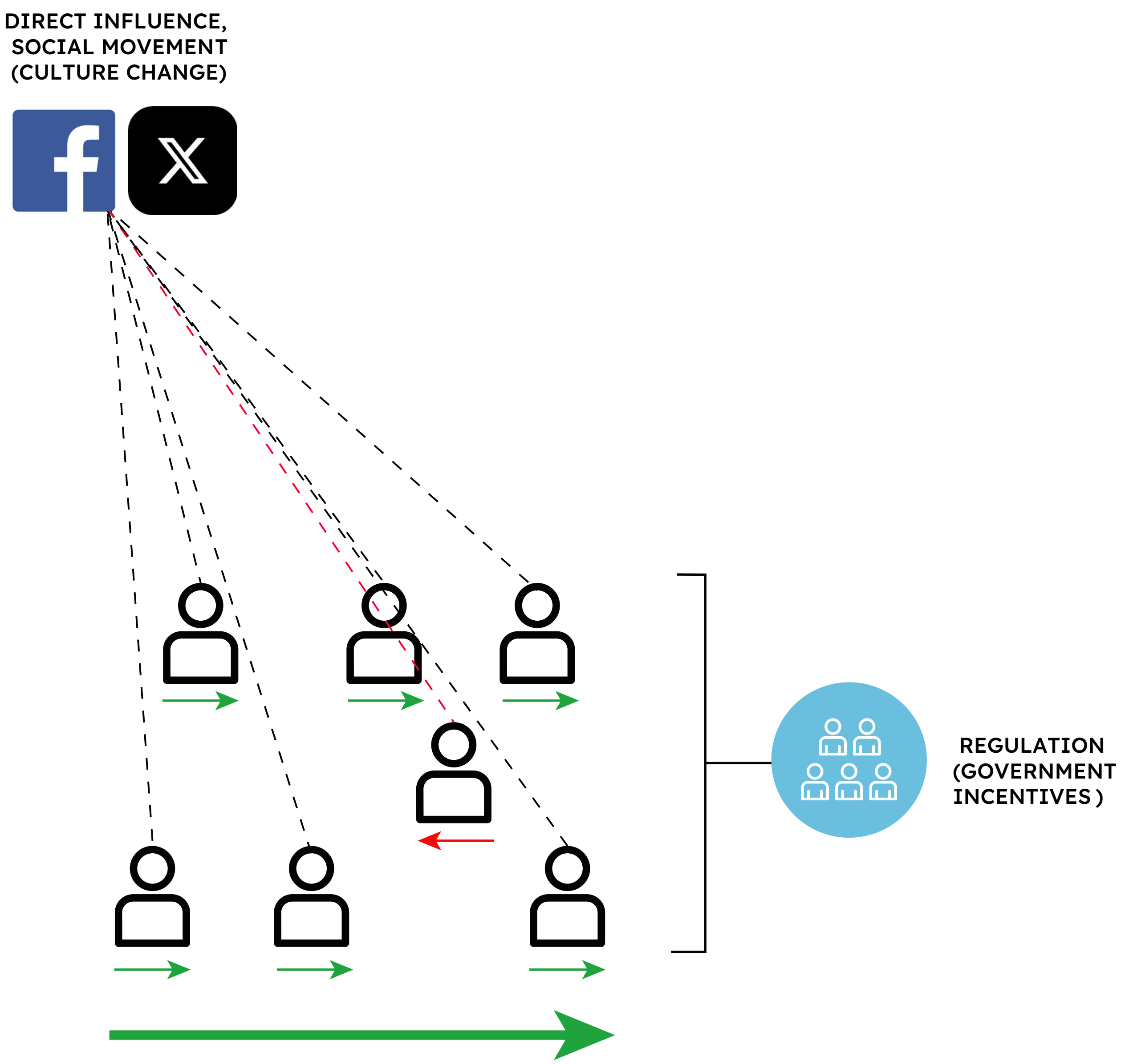
Capital Markets

The insurance industry is increasingly affected by climate change, which is amplifying the frequency and violence of extreme weather events such as floods and wildfires. Droughts and agricultural crises are also becoming more common and severe, leading to significant claims for compensation.



Flooding in eastern Australia during February 2022.

ESP technology will allow better risk assessments by insurers and support better decisions by householders seeking to avoid exposure to climate driven catastrophes.



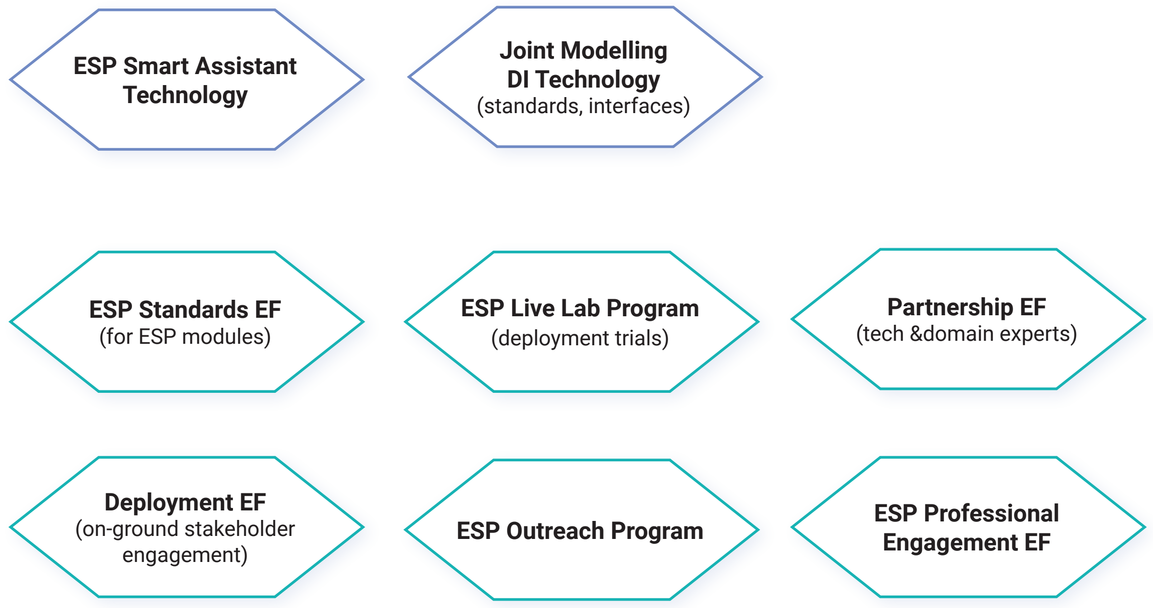
Applied ESP technology aims to foster **systemic change** in the form of small individual changes that accumulate to a significant change in total behaviour.

This could be led in a top-down fashion through government, legislation, international treaties or regulation imposed by professional bodies.

Alternatively, a grass-roots movement could foster the adoption of ESP technology through social networks. This strategy must communicate immediate personal benefits and will likely need to combat significant misinformation.

In reality, both approaches are complementary and necessary.

Linked Roadmap Actions



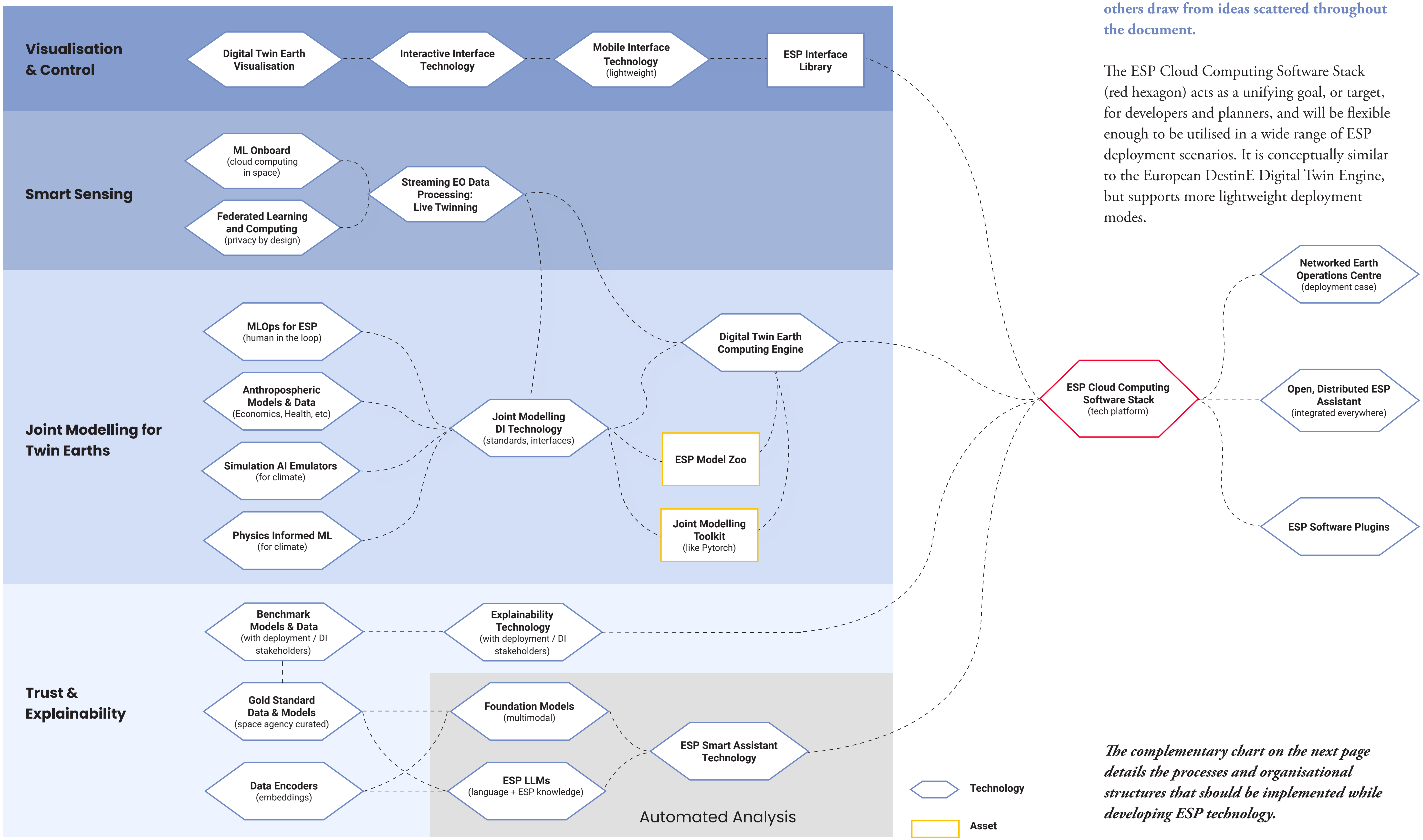
These shapes represent the entries in the ESP Roadmap that are relevant to the current topic. The Roadmap presents technical and executive function maps of ESP, and actions recommended by the ESP community, to move ESP from vision to reality.

AN ESP TECHNICAL ROADMAP

ROADMAP 1:
ESP TECHNICAL ROADMAP

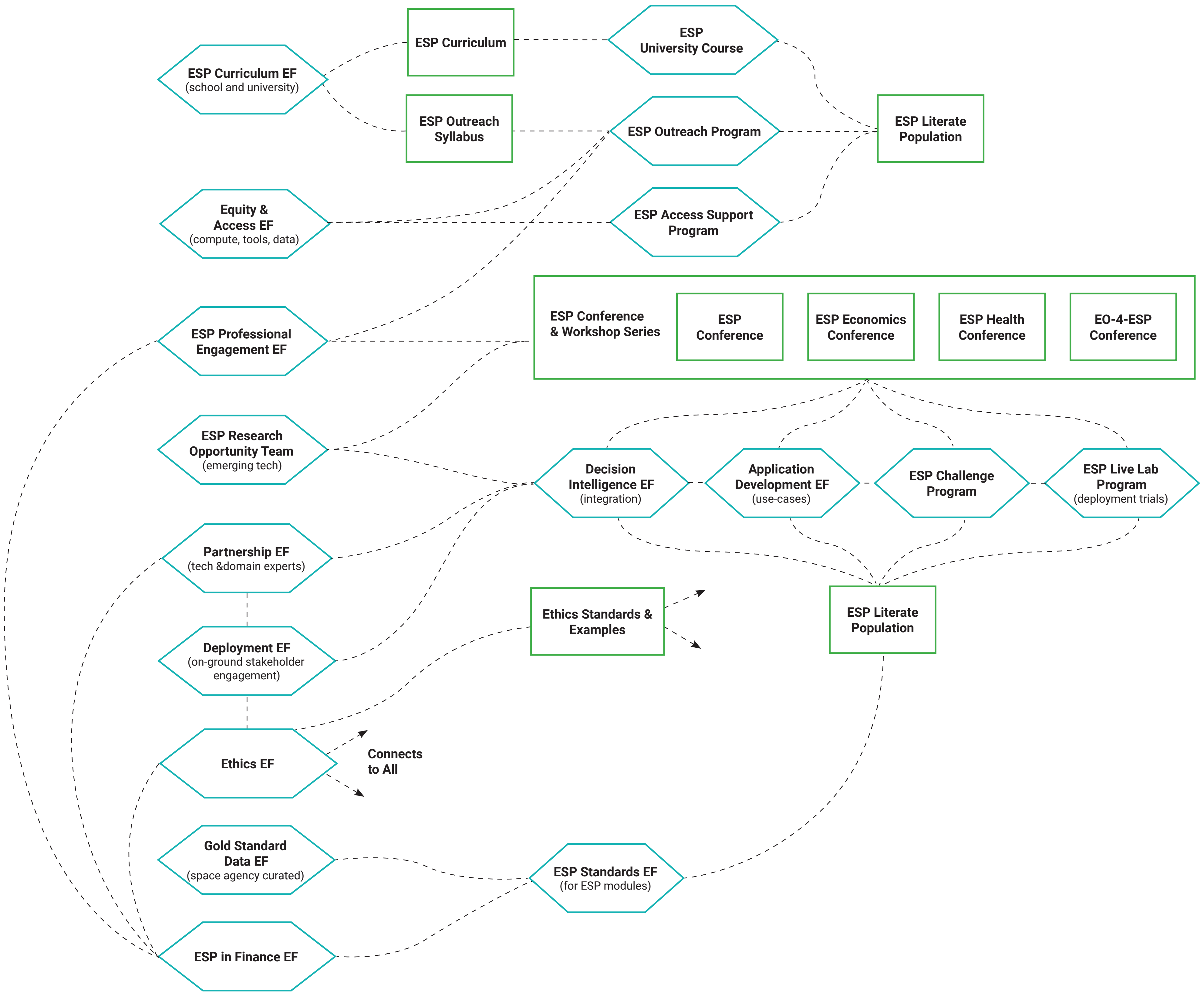
The goal of the ESP Roadmap is to provide context for the reader to take action. This flowchart illustrates the dependencies between the technical components in an integrated ESP system. Some of these are detailed in a single area of the report, while others draw from ideas scattered throughout the document.

The ESP Cloud Computing Software Stack (red hexagon) acts as a unifying goal, or target, for developers and planners, and will be flexible enough to be utilised in a wide range of ESP deployment scenarios. It is conceptually similar to the European DestinE Digital Twin Engine, but supports more lightweight deployment modes.



The complementary chart on the next page details the processes and organisational structures that should be implemented while developing ESP technology.

ROADMAP 2:
ESP EXECUTIVE FUNCTIONS

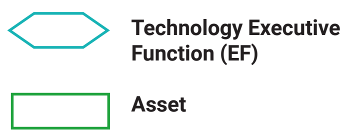


Any future ESP system should be developed with thoughtful human oversight for both ethical and safety reasons, but also to ensure ESP investments are fit-for-purpose and remain useful.

Taking inspiration from the human brain, we envision a future distributed ESP network requiring a similarly distributed 'executive function' needed to manage different aspects during development, deployment and maintenance.

In the following section we briefly describe the technical and executive components of the ESP Roadmap and present lists of actions recommended by the ESP Forum participants.

This list is only the beginning: we invite you to recommend actions, offer corrections and suggestions, and engage with the burgeoning ESP community. Help us build a brighter future!



ROADMAP 3: ACTIONS FOR EACH ESP COMPONENT

Technical Actions

Digital Twin Earth Visualisation

Visualisation technology for Digital Twin Earths must support a wide range of display sizes and types: megatron, laptop and phone screens; virtual and augmented reality displays; audio and language outputs. Useful for deployment stakeholders making decisions and developers monitoring ESP system performance.

- Create interactive 'virtual planet' visualisations linked directly to 'in-game' choices so that people can see the effect of their decisions and change their behaviour accordingly. These experiences should be compelling and immersive.
- Build models, visualisations and narratives that support conversations to address the non-linearity of climate solutions and build non-linear thinking skills.
- Determine how to choose the right metrics or variables to present to stakeholders and decision makers for maximum understanding and positive impact.
- Work with stakeholders to understand domain-specific visualisations; we may be missing easy-to-use visualisations that will convey meaning and emotion to change behaviours.

Interactive Interface Technology

Interactive controls technology linked to visualisations and live ESP predictive models that support an immersive narrative experience. Useful in educational and decision making settings.

- Build sophisticated interactive dashboard systems that leverage LLMs and present both visual and textual output.
- Take advantage of emerging display technologies like augmented reality (AR) and virtual reality (VR) to highlight the invisible consequences of actions (e.g., aerosol pollution, cloud cover changes) when visualising EO data and simulation results. Use the ability to 'fast forward' in time, and 'zoom' in space to make their impact on Earth systems visible.

Mobile Interface Technology

Technology that allows users of mobile devices to interact with local or cloud-based ESP systems in a responsive way. Useful for making ESP technology accessible to a broad range of users, especially in the developing world.

- Create a lightweight visualisation software stack that is optimised for low bandwidths and small displays so that ESP and EO visualisations work well on mobile devices, such as tablets, phones and smart watches.

ML Onboard

Intelligent systems on-board spacecraft that analyse and interpret remote-sensing and systems data in-situ. Useful for transmitting reduced information packets and supporting autonomous decision-making.

- Develop intelligence that applies data validation onboard satellites.
- Develop analysis routines that can process sensor data into insight (e.g., maps of land use) on the spacecraft, transmitting sparse data to the ground, to the correct end-users.

Federated Learning and Computing

Decentralised machine learning technology that allows the exchange of knowledge without moving large data or big models. Useful when privacy is of paramount importance, or communication bandwidth is low, or when collating different types of data would be difficult.

- Federated learning technology could help satellites from different operators communicate meaningful information. Develop federated learning systems for use during emergencies, for example, in support of the "Space and Major Disasters International Charter".

Streaming EO Data Processing: Live Twinning

Onboard and ground-based rapid processing for remote sensing observations. Useful for feeding the latest information into digital Earth twins for near-real-time awareness.

- Define standardised interfaces from EO systems into decision models so that the results of digital twin simulation can be integrated into decision-to-action processes.

MLOps for ESP

Software, standards and procedures for managing the development and deployment of ESP systems, including training models, monitoring performance, benchmarking and incorporating human feedback.

- Incorporate human feedback into MLOps loops.
- Use reinforcement learning with human feedback (RLHF) to integrate knowledge from domain experts effectively.
- Make validation a first-class citizen of ESP Earth observation toolkits to allow easy comparison of calibrated satellite data and derived parameters to on-ground measurements.
- Within the DI framework, build tools to integrate human feedback that are informed by social context and aligned with wider goals.

Anthropospheric Models & Data

Simulations and data from non-Earth system domains that should be incorporated to ESP joint decision-making (for example, anthroposphere models such as economics, health, supply chains etc.).

- Analyse what external data and models should be incorporated into ESP systems, and how to prioritise the integration.

Simulation AI Emulators

Machine learning models that learn directly from physics-based models of weather and climate. These accelerate prediction times by thousands of times, making ensemble analysis and scenario analysis possible. Also includes downstream modelling of impacts.

- Design and build more climate emulator models with AI instead of PDEs, allowing for faster, larger-scale models that can generate ensembles to quantify uncertainty and risk of climate extremes.
- Prioritise research into creating accurate medium-high resolution policy emulators (like MIT EN-ROADS) that not only show global changes in response to policy decisions but are capable of showing regional impacts. Existing emulators are low-resolution and cannot show what local impacts will be (e.g.,"I can see the global temperature is getting higher, but what does that mean for my fishing boat?") and are therefore unrelatable to local decision and policy makers.
- Accelerate the dynamics of emulators. Prioritise research into faster and more accurate down-scaling (converting low-res predictions to high-res using prior knowledge). These algorithms are used to sharpen fuzzy impact maps to show local results but outcomes can be highly inaccurate, resulting in a lack of trust.
- Accelerate the operation of impact models. Creating impact models (e.g., "What does this mean for crop yields?") can be very slow, and adjacent models need to be accelerated in tandem.

Ensure that modules that implement emulators are maximally reusable between different applications and rely on an open core (e.g., an open repository of conceptual models).

Physics Informed ML

Fast hybrid ML-and-physics based climate and weather models. These can be as reliable as physical models, but also have the potential to learn new physics from observational data.

- Support the development of physics-aware AI-driven global climate models at moderate resolutions of 10-50 km.
- Design AI-driven climate models that accept EO data for calibrating predictions. This could be accomplished with ensemble Kalman methods that are used in weather prediction.
- Address the challenges associated with hybrid models trained on EO data at high-resolution. These challenges include coupling, tuning and out-of-distribution behaviour.

Joint Modelling DI Technology

Technology rooted in decision intelligence that allows easy integration between predictive models and data processing systems from disparate knowledge domains. The key to stitching multiple concerns into a decision framework.

- Develop standardised APIs for data fusion, data integration and data provenance.
- Create a shared framework for integrating EO elements (data, models and derived insights) into decision workflows, organised around the ways decision makers use EO elements.

- Adhere to FAIR principles and Open Architectures for long-term sustainability.
- Publish a glossary of common terms and definitions with translations to similar concepts in each subject domain.
- Build a system-of-systems where climate actions are local but their aggregate is global, integrating connections between local policies and global impact.

Digital Twin Earth Computing Engine

The computing system and architecture that underlies digital twin Earths and the ESP cloud computing stack.

- Build digital twins (DTE) with spatial and temporal accuracy, full explainability and integration maps to DI systems. Explore hybrid DTE architectures, such as combining CPU/GPU and building bridges between Fortran and Python.
- Determine how to choose the correct position in the trade-off between speed and accuracy (if we know the uncertainty) when responding to extreme events.
- Build links to the Earth Virtualisation Engine (EVE) project that is developing a concept for the ML and HPC computing engine to *"operationalise a global climate prediction and information system across regional partners or nodes"*.

Benchmark Models & Data

Published algorithms, systems-code, models, data and performance metrics that serve as a reference for future development. Useful for measuring the performance of new research, or the practicality of a system for deployment, or for ongoing monitoring.

- Publish new models and code in support of reproducibility.
- Build the credibility of AI-driven climate models by developing benchmark datasets and models like ClimateBench.
- Generate better reference datasets that use physics to validate outputs.
- Coordinate the creation and promotion of open benchmarking datasets and EO-processing models to spur the development of robust tools that can be made deployment-ready on accelerated timescales.
- Benchmarking and validation metrics should be linked directly to use cases and decision requirements should guide model accuracy and benchmarking goals.
- Define trust and reliability metrics for models and data in the context of downstream decisions.
- Use independent third party validation partners that are separate from industry or political conflicts of interest. Explicitly identify and control for possible biases and conflicts of interest when collaborating with industry or political organisations.

Explainability Technology

Technology that supports an understanding of how and why models produce a result, their confidence in predictions, and sensitivity to changes in inputs or assumptions. Useful for comparing adjacent models, making weighted decisions, or diagnosing issues with intelligent systems.

- Build upon Google's Pathways Language Models (PaLM) to do chain-of-thought, deductive reasoning and validation, minimising hallucinations in the process.
- Develop a library of science-based factuality queries for PaLM-style models, which are linked to human validation workflows.
- Research and improve LLM's capacity for causal reasoning in climate science.
- Create an experiment process that uses LLMs to support hypothesis testing workflows. The LLMs should ask useful questions, help design experiments using best practices and uncover underutilised resources in systems.
- Integrate the latest explainability technologies into AI-driven climate and weather modelling efforts.

Gold Standard Data & Models

Analysis ready data and validated models that are actively maintained in support of mission-critical applications. These must be highly accessible, support fast downloads or deployments, and have minimal down-time.

- There is a need for a 'gold standard' in data and predictive models, with a role for space agencies to be the provider of the very accurate predictions. Perhaps there is room for an ecosystem of smaller, or more entrepreneurial, providers that provide a second pair of eyes.
- Improve the quality of remote-sensing data for calibrating climate models. This data should be sampled globally, frequently refreshed, available in ML-ready format, available in data-lakes and validated for quality using recognised tests.
- Make data easier to access over long timescales using open data tools like Pangaea and CKAN.

Data Encoders

Data-loader code and algorithms that allow specific datasets to be encoded in compressed machine-readable formats, while still preserving necessary information. Useful for reading complex data into LLMs and other ML models.

- Build a library of open-access data-loaders for remote-sensing instruments that allow the data to be used with modelling efforts. These encoders should produce ML-ready data.

Foundation Models

Very large models, associated training data and deployment modules that can be adapted to a broad range of use-cases. Ideally supports multimodal data, while being robust enough to accommodate malformed data.

- Support the creation of open-access scientific foundation models with thorough documentation and example use cases. These models should provide well-documented pipelines for training and deployment to facilitate knowledge-transfer into the ESP community and capability enhancement across the board.

ESP LLMs

Large Language Models (LLMs) that integrate deep knowledge of Earth systems and information from domain areas linked through ESP concepts. LLMs trained to understand, and assist with, ESP analysis workflows.

- Support new research into how LLMs (and AI in general) can be applied to ESP. LLMs are at the driver's seat to the powerful under-the-hood technology we are building.
- Build LLMs to be concept engines that are continuously updated with contextual information.
- Build knowledge systems to support expert-focused reinforcement learning with human feedback (RLHF), using LLM technology to engage with the experts.

ESP Smart Assistant Technology

Technology that leverages the capabilities of LLMs and foundation models to perform analysis of EO data, build ESP workflows and more.

- Investigate the creation of semantic layers for knowledge translation. Often, jargon is a barrier for communication between disciplines and LLM-driven semantic layers could offer skilled translations.
- Explore how LLMs should be used in the last mile by teaching concepts and using continuous learning.

ESP Cloud Computing Stack

Unifying software framework and hardware platform for ESP.

- ESP systems and processes should support the processing of EO data into meaningful insights that lead to effective actions; specialist skills and experience are needed to derive insight from EO data and on-ground sensor data.
- Provide validated tools that encode the requisite knowledge and streamline the end-to-end workflow
- Provide expertise (e.g., in the form of interactive support from humans or AI-assistants) for accessing and analysing data.

Networked Earth Operations Centre

Earth operations centres provide hubs for deploying ESP technology (EO, AI, DI) and digital Earth twins in support of planning, decision-making, and crisis management. A network of Earth Operations Centres could better scale to continental or international scales.

- Build a Mission Control Center for Earth systems data, with accessible, real-time, high-resolution EO data.

Open, Distributed ESP Assistant

An ML-powered assistant with a natural language interface that can be instructed to perform advanced analysis related to ESP and Earth observation workflows.

- Build a LLM-powered assistant that has knowledge of Earth system science and intersecting domains in ESP. Should be skilled at building and orchestrating analysis pipelines that make use of GIS data.

ESP Software Plugins

Modular ESP 'plugins' for ubiquitous software packages like browsers, document editors and spreadsheets. These add ESP analysis workflows to ordinary tasks.

- Identify ESP capabilities that would be useful for a range of businesses and government agencies. Match these capabilities with common software packages used within these domains and design ESP plugins that would fit within existing workflows.

Executive Function (EF) Actions

ESP Curriculum EF

Curriculum, lesson plans and teacher training resources for a modular course in Earth Systems Predictability.

- Work with schools and education organisations to work out where ESP links into the current school curriculum.
- Develop comprehensive teacher training resources, including professional development courses, lesson plans, online teaching resources (e.g., like for MOOCs).
- Engage government education departments to include ESP in curriculum updates.

ESP University Course

A university-level practical ESP course. Could be aligned to science and arts departments.

- Work with leading university departments to develop ESP courses that fit into existing undergraduate degree programs.
- Work with industry partners to advise on the design of university courses that meet their ESP and ESG needs.

ESP Outreach Program

Activities to engage the general public and special interest groups to educate them about ESP. Includes citizen science programs that feed back into live ESP applications and engagement with industry groups.

- Galvanise citizen buy-in and generate collective demand through outreach.
- Run a climate decision-mapping workshop with no technology required for participation.
- Identify simple, inexpensive exercises to drive value and interest as the first steps in the larger initiative.
- Build a simple LLM-driven website loaded with ESP knowledge that can engage citizens and be a market intelligence tool to understand what people want to know.

Create an education/marketing campaign to overcome barriers, illustrating simple entry ramp use cases that provide a lot of value with little effort.

Equity & Access EF

Executive function to manage and support equal access to ESP technology, programs and services, with a particular focus on disadvantaged and underrepresented populations.

- Lower the cost of deploying scientific foundation models; this can be done by directly subsidising cloud-computing costs or by training skillful student models that are more easily deployed on commodity hardware.
- Undertake to support a general goal of transparency, accountability and equity in technology, as well as data access. Support pathways for third party validation of decisions and data.

ESP Access Support Program

Real-world programs to support disadvantaged groups in accessing ESP technologies for active decision making.

- Support open access to data, models and development frameworks that can be used to develop sophisticated decision support systems.

ESP Professional Engagement EF

Executive function to engage professional organisations who would benefit from ESP technology, who need to make significant operational changes for climate change reasons.

- Run decision mapping workshops for specific job roles and work functions (e.g., CEOs, CSR officers).
- Provide services that assist organisations with building in-house ESP tools.
- Run a series of ESP workshops and conferences aimed at embedding ESP in new industries.

ESP Conference & Workshop Series

ESP workshops and conferences dedicated to developing ESP as a multidisciplinary discipline, and promoting ESP decisions making in government and business.

- Run a series of professional conferences and practical workshops to organise ESP research and development.
- Organise workshops with deployment stakeholders to show how ESP software tools can be integrated into daily workflows

ESP Research Opportunity Team

Multidisciplinary team that 'scans the horizon' for emerging technologies and develops ideas on how they might function in the ESP framework.

- Rethink cities to make them more easily monitorable from an EO/climate perspective.

Partnership EF

Executive function to establish and manage partnerships with research and industry stakeholders, including academia, technology companies and philanthropic organisations.

- Make ESP a core principle to incorporate direct stakeholder voices, particularly the voices of those with little bargaining power.
- Ensure knowledge tools are co-created by all significant stakeholders.
- Develop new projects using DI action-to-outcome simulation methodologies. Work with partners to run end-to-end simulations on state-of-the-art computing hardware.

Deployment EF

Executive function to engage with deployment stakeholders, supporting the exchange of information between potential users and ESP developers.

- Provide incentives for using EO data, including economic incentives in different arenas, proven business models for collection, dissemination and management of EO data and validated case studies quantifying the value creation of that economic impact.
- For each EO use case, work with end users to understand their needs and communicate how EO data helps them. We need to consider which of those use cases would require the support, buy-in and trust of community stakeholders. How do we develop and maintain that trust?
- Build an understanding of the decision makers who will use our systems; start with who the decision makers are and what their problems are, not with the data (e.g. via concurrent design methodology).

Decision Intelligence EF

Executive in charge of integration for ESP: including interfaces between models, data, applications and standards.

- Use decision intelligence (DI) methods to situate in-development predictive models within an action-to-decision-outcome simulation. This will allow the development of formal methods to assess decision quality sensitivity to model accuracy and data cleanliness.
- Promote greater understanding and adoption of DI methods in the scientific community.
- When building large decision systems for ESP, create a graduated roadmap of integration priorities, starting with quick wins and culminating in a comprehensive ESP and DI ecosystem.
- Co-create causal decision diagrams (CCDs) in consultation with a diverse range of possible stakeholders to anticipate unintended consequences.
- Integrate SROI more deeply with DI so that social outcomes are included in all decision models, as ranked by SROI methods.

Application Development EF

Executive function to manage the development of real-world ESP applications and systems.

- Support the creation of a planetary health (Φ) metric for products and services linked to open modelling and transparent assumptions.
- Build accessible visualisations of decision levers that can change a Φ score and support organisations to make positive changes.

ESP Challenge Program

Accelerated research sprint program to develop core ESP technology, including machine learning capabilities, integrated workflows and downstream applications.

- Perform applied multi-disciplinary research that demonstrates the utility of advanced technology to solve challenging real-world problems. Involve deployment stakeholders at all stages of the research process.
- Run reproducibility challenges (e.g., like this one for climate AI) to advance high-quality open science.

ESP Live Lab Program

Program to run live trials of ESP technologies in collaboration with deployment stakeholders.

- Build innovation acceleration and experimental development hubs.
- Develop decision maker-centric projects which begin with the authorities (levers) and responsibilities (outcomes) of specific decision maker roles within the planetary stewardship ecosystem. Work through the DI processes with experts and stakeholders to build systems that improve climate outcomes.

Ethics EF

Executive function to embed ethics principles in ESP development, management and application functions.

- Require an advisory board for projects that create, update or manage AI knowledge systems. Mandate regular monitoring and reporting, including validation by independent experts.
- Deliberately incorporate non-traditional and non-human viewpoints into knowledge models, because non-humans make up a large part of the planet.
- Determine if we need to have a constitution or other governing model for AI models that have agency. Create a process to build and evaluate this.

Gold Standard Data EF

Executive function that manages the development of 'gold standard' data sources and reference models.

- Work with national and international deployment stakeholders to certify algorithms, model implementations and datasets for specific on-ground use-cases.

ESP Standards EF

Executive function that develops and manages the standards for joint modelling interfaces, data exchange formats, metadata and more.

- Synchronise with the emerging https://opendi.org/ and https://aousd.org/ organisations and similar initiatives to define shared standards and application programming interfaces for mutual benefit.
- Be adaptable and open to new, and existing, ideas from outside organisations to avoid fragmentation.
- Build a universal, codified 'pre-flight checklist' for deploying AI systems in specific domains. These would set quantitative standards for performance, safety and monitoring over time inspired by certification systems in the airline and medical industries.
- Define the gold standards for data provenance and reproducibility in scientific research and for deployed AI systems.
- Create transparency on the governance, modelling process, assumptions and levels of uncertainty associated with a planetary health metric (Φ).
- Investigate the emerging ecosystem of organisations providing DI components who are integrating them into a common architecture. Work with them to ensure the creation and adoption of standards that support interoperability, and provide platforms that stimulate new DI research.

ESP In Finance EF

Executive function that manages the development of ESP technology, standards and certification for the finance industry.

- When using ESP technology to support corporate environmental, social, and corporate governance (ESG), make sure to identify the incentives that work against positive climate and environmental outcomes. This will help build a fuller picture to guide investment.
- Develop incentives (economic, political, human wellbeing) for businesses to innovate on climate mitigation processes and not just "greenwash".
- Use SROI valuation to enable people who are traditionally excluded from having a voice in business decisions and policy to get their experience of value seen and heard in a language that can travel all the way up the value chain.
- Build bridges with the fields of impact management and social value/impact valuation.
- Ensure that decision models (specifically those based on CDDs leading from actions to outcomes) support SROI calculation, and SROI measurement creation becomes a standard step within the DI processes (such as in the 9-step model of www.dihandbook.com).

ROADMAP 4: ESP MEETS DESTINATION EARTH (DestinE)

ESP is aligned with the European Destination Earth (DestinE) digital twin Earth platform, but aims to be more accessible by virtue of being more distributed and more integrated with everyday life, while still delivering 'good enough' decisions.

The ESP Cloud Computing Software Stack shares much in common with DestinE's Digital Twin Engine and Service Platform, but extends the concept to the following areas:

- Rapid joint modelling through a decision intelligence framework (Earth systems plus anthropospheric systems).
- 'Good enough' accuracy, as determined by the requirements of the decisions being modelled.
- AI technology to facilitate cheap and accessible Earth systems predictions, even on low-cost devices with limited bandwidth.
- Intelligent interfaces that easily allow ordinary citizens to incorporate Earth system modelling into their workflows (e.g., using natural language commands).
- Interactive visualisations that educate and inform.
- The typical user of ESP technology would be a non-technical person in government, business or at home, seeking to make climate-friendly decisions.

Advantages of developing ESP technology in parallel with DestinE

ESP technology can be developed alongside DestinE by co-creating with the user community.

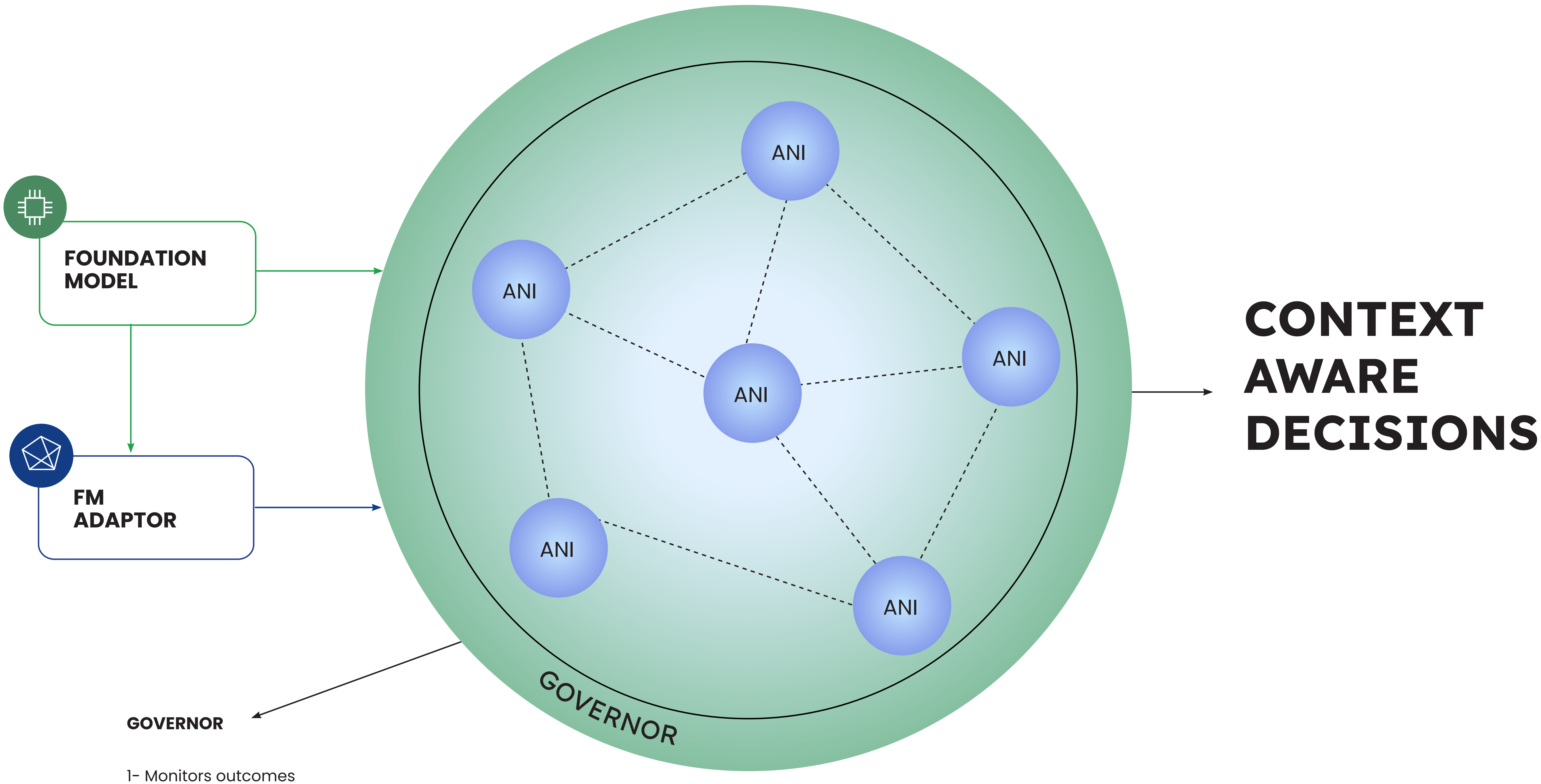
- Rapid co-creation of novel use-cases for digital twin Earths, which can be later migrated to DestinE.
- Technology developed for distributed computing and joint modelling can be incorporated into DestinE Core Service Platform.
- ESP will provide an introduction to digital twin Earths for audiences that would not otherwise consider DestinE.
- ESP will build a strong community of users connected to experts, who will become champions of decision making supported by digital twins, including DestinE
- ESP will have a strong educational and citizen science component.

ESP can be considered a 'skunkworks' project, to test cutting-edge technologies, innovative ideas and trailblaze community engagement with digital twin Earths.

The similarities **(green)** and differences **(blue)** between ESP and DestinE are tabulated here.

DestinE	Features	ESP
Specialist digital twin Earths	Scope	Joint Earth-anthropospheric system models
Physics and AI-based digital replica	Design approach	Massively integrated narrow AI models and foundation models
Historic and live EO data	Data	Historic and live EO data
HPC and Cloud	Compute framework	In orbit, Cloud, desktop and edge, mobile (distributed)
Active machine learning	Learning	Active machine learning
On demand	Speed to insight	On demand
Very high	Accuracy	Variable (demand driven)
User validated	Safety & Ethics	User validated
Hours - decades	Temporal scope	Minutes - decades
Government, businesses	Users	Government, businesses, communities, individuals
Expert defined	Use-case development	User co-created
Portal, APIs	Access	Multiple APIs
EU Project	Governance	Open collaboration
"Evidence based decisions"	Design Philosophy	"Good enough decisions"

From monolithic models to massively integrated AI services



GOVERNOR

1- Monitors outcomes
2- Trottles compute services
3- Manages energy
4- Providers reporting to Humans

ANI Artificial Narrow Intelligence

ROADMAP 5:
CALLS TO ACTION

Creators of Digital Twin Earths

We call on the organisations building operationalised Digital Twin Earths to engage with the ESP roadmap, especially around **making DTEs accessible and useful in a distributed framework**, and on constrained computing platforms like mobile devices. Engaging visualisations and interactivity should be at the core of DTEs.

Deep Simulations Specialists

We call on the developers of simulations, including climate, weather, economic and social models, to **embrace the concept of joint modelling in a decision intelligence framework.** This also entails deeper engagement with deployment stakeholders, using their decision making requirements to build benchmarks and continuous validation workflows.

Scientists and Educators

We challenge all scientists to understand how their discipline fits into an ESP framework, and to **develop programs to educate people about the long-term consequences of their individual decisions on Earth's systems**. Engage in two-way conversations, listening to ordinary needs, and leveraging compelling narratives supported by interactive digital technology.

Legislative and Policy Experts

We call on politicians, legislators, civil servants and policy experts to focus on long-term climate consequences of decisions by **creating incentives and regulations that catalyse everyday use of ESP informed decisions across industry and government.** Continuous validation and impact monitoring should feed into flexible decision-making that can adapt to real-world consequences and unexpected outcomes.

Space Agencies and National Labs

We call on space agencies, national laboratories and research institutes to support cross-disciplinary research through dedicated programs linked to ESP. Space agencies should be the creators and custodians of 'gold standard' Earth science datasets, including benchmarks and validation programs, working with deployment stakeholders. **Government agencies should build partnerships with private industry to deliver and regulate open-access foundation models and language models for Earth observation and Earth science as national assets.**

Commercial Technology Developers

We call on technology companies and industry bodies to engage in the development of ESP technology and to **drive the democratisation of ESP systems by supporting open standards and APIs**. Develop business cases built on ESP foundations and support positive legislation and regulation.

NGOs and Civil Organisations

We call on non-governmental and civil society organisations to work with on-ground stakeholders to lead the **development of ethical frameworks around AI and ESP, and determine the needs of the community**, which are then communicated to government and regulation bodies.

AFTERWORD

A PLANETARY NERVOUS SYSTEM

Our century is marked by intersecting crises: geopolitical and energy security uncertainties, despeciation, pollution, water stress and extreme weather. The impact of these disproportionately affects the world's poor, as well as all the species we share our planet with. If it is not obvious, we have exceeded the carrying capacity of everything that sustains us. Moreover, the longer it takes for us to change, multiplies the time it takes for things to heal.

It is now becoming apparent that this isn't a climate crisis anymore. It's a habitability one.

However, we now have the toolbox to build a positive, thriving future that benefits all. But it will take a change of mindset about how we account for value when we make decisions.

A key problem has been that up until now we haven't been able to see or quantify the value of these life sustaining systems. As management guru Peter Drucker famously commented, "What isn't measured isn't done" and in the case of Earth system services, what hasn't been measured, has been ignored.

In this document, it is our hope that the seeds of a new way of seamlessly assessing value at the point of decsions will remove our collective blinkers, making the invisible, visible. While these kinds of impact assessments have been possible for some time in academia, the key change is speed, cost, veracity and the ability to place this insight into the hands of everyone dynamically at the point of decision making.

The 2021 Space for Net Zero report from the World Economic Forum's global Future Council catalysed our efforts and we are incredibly grateful for the impetus. The Report proposed the **Earth Operations Center concept** aligning Earth system modelling with just-in-time information from space to facilitate informed climate decision-making.

Our contribution extends on this idea by proposing an 'Earth Operations Network' where intelligent decsion making infrastructure becomes a globally distributed, fully integrated nervous system for the planet - decentralised and democratised, for the benefit of all.

Solving the challenges of the coming decades will require every single human being making better decisions than we have in the past. If we can do that, maybe, with a bit of luck we can achieve a pathway to a future where the human race survives and thrives.

I'll leave the final word to Canadian philosopher Marshall McLuhan: *'There are no passengers on Spaceship Earth, we are all crew.'*

We sincerely hope our work on ESP inspires others to carry the torch.

We look forward to hearing what you do.

James,

London
November 2023

James Parr
CEO Trillium Technologies
Founder Frontier Development Lab

JANUARY - MARCH 2023

9 THEMED WORKSHOPS

Short workshops to deepen the understanding and scope of each key question.

9 EXPERIENCED LEADERS

Leaders in each of the themes, invited from renowned institutions, and working with global experts and practitioners.

MAY 8TH - 10TH 2023

	TWINING & SIMULATION	INTEGRATING KNOWLEDGE	DECISION INTELLIGENCE
SPEED & SCALE	Scaling Simulations	Scaling Knowledge	Multi-Disciplinary Decision Modelling
PROJECTIONS	Forecasting With ML & Physics	Integrating Context & Experience	Joint Reasoning To Recommend Action
TRUST & INSIGHT	Trusted & Explainable Simulations	Continuous Validation	Decisions We Can Trust

3-Day Intensive Forum

150+ selected expert participants, united to address questions on each intersecting theme. The Forum facilitates knowledge exchange between ML experts, data custodians, Earth scientists and decision makers.

SEPTEMBER 2023

SYNTHESIZE & VALIDATE

Synthesize ideas and identify links between topics. Iterate and validate with technical experts and decision makers.

ROADMAP MEMOS

NEW PROJECTS

COLLECTIVE KNOWLEDGE

OUTCOMES

Release memos and resources to the community. Establish ongoing affinity groups.

The results in this document are synthesised from **three days of presentations and discussions at the ESP Forum,** in addition to **nine focused workshops** involving small **groups of curated experts.**

ESP CONTRIBUTORS

ALISON LOWNDES
Senior Scientist, Global AI, NVIDIA both intrapreneur building the space business within NVIDIA and social entrepreneur running her charity in Kenya.

ANAMARIA BEREA
Associate Professor at George Mason University

ANNE SPALDING
Program Director, Trillium Tech

ATHANASIOS VLONTZOS
AI specialist in causal inference and learning at Spotify and specialist in counterfactuals.

BJÖRN JÜTJENS
Postdoc at Massachusetts Institute of Technology

CORMAC PURCELL
Chief Scientist, Trillium Tech

CHEDY RAISSI
Data Science Senior Manager, Riot Games

CHRISTIAN SCHROEDER DE WITT
Postdoctoral Researcher, University of Oxford

CRISTIAN ROSSI
Geospatial Science Lead,
Satellite Applications Catapult

DAVID BRIN
Seminal "Hard SF" author and raconteur. His books Existence and Earth predict our modern era in stark detail.

DAVID HALL
Senior Data Scientist at NVIDIA who is consistently ahead of the curve with AI for EO.

EDOARDO NEMNI
Lead Geospatial Data Scientist at BeZero Carbon

ED GILLESPIE
Futurenaut, Director Greenpeace UK, Forward Institute Senior Facilitator.

GIUSEPPE BORGHI
Head of the Phi-Lab Division at ESA, currently leading ESA's push to accelerate the future of Earth Observation by means of transformational technologies.

GREG RENARD
NLP expert, entrepreneur and AI advisor to the French government.

HARUKO WAINWRIGHT
Assistant Professor at Massachusetts Institute of Technology

JACQUELINE LE MOIGNE
Manager of NASA's Earth Science AIST (Advanced Information Systems Technology)

JAMES PARR
CEO, Trillium Technologies

JOEL FREEDMAN
Head of Innovation Services, Satellite Applications Catapult

JOHN ELKINGTON
Arguably the father of modern sustainability, who believes it's time to put AI in sustainability.

JONATHAN BAMBER
Remote sensing professor who specialises in the cryosphere and its interactions with the Earth system.

JONATHAN KNOWLES
Scientist and explorer serving as Exploration Director at FDL.AI and Science and Exploration Fellow at Mission Blue: The Sylvia Earle Alliance

JOSH VEITCH-MICHAELIS
Award winning applied AI scientist with a focus on social good, conservation and disaster response.

KAROL BOT GONÇALVES
Postdoctoral Researcher at INESC Technology and Science

KIRSTEN DUNLOP
CEO of Climate-KIC, Europe's leading organisation for climate mitigation and adaptation and all around force of nature.

LARA SUZUKI
Technical Director at Google Cloud, Engineer of the year 2021 and all round polymath.

LORIEN PRATT
Chief Scientist, Quantellia, inventor of Transfer Learning and Decision Intelligence. Simply one of the greats.

LUCAS KRUITWAGEN
AI4EO expert who developed the world's audit of solar panels and now is CTOtechnology lead and co-founder of Transition Zero.

MARKUS WOLTRAN
Programme Officer at United Nations Office for Outer Space Affairs

MASSY MASCARO
Distinguished Technical Director leading the data science / AI team at Google Cloud, with expertise in deployment at scale, neuroscience, computer vision and LLMs.

MINOO RATHNASABAPATHY
Research engineer at MIT and project fellow on the Future of Space at the World Economic Forum that puts her work at the heart of ESP.

NICKI MCGOH
Senior Director and Co-lead at Caribou Space.

PAULA HARDER
PhD Researcher at Fraunhofer ITWM

PIERRE PHILIPPE MATHIEU
Head of Phi-Lab Explore Office, ESA ESRIN, leading the charge in the application of AI to EO and ultimately decisions on the ground.

RAHUL RAMACHANDRAN
Senior research scientist at NASA IMPACT leading the charge on Earth System Informatics

SANTIAGO GOWLAND
CEO of the Rainforest alliance. Looking after our planet's lungs.

SARA OLSEN
Impact management pioneer of measuring social and environmental value and social capital.
SVT Group

SEAN MCGREGGOR
Founder of the AI incident database and advocate of the responsible use of AI.

THAMMY EVANS
Member of the Global Strategy Programme at RCDS;
Senior Fellow at Atlantic Council GeoTech Center;
Senior Research Associate CCIP

VALENTINA SALVATELLI
Senior Applied Researcher, MicrosoftResearch

YARIN GAL
Professor at Oxford University and progenitor of Bayesian Deep Learning - fundamental for navigating uncertainties in ML predictions.

Our heartfelt thanks to all our contributors, ESP Forum participants and reviewers, who generously donated their **time and effort** to build the vision of **Earth Systems Predictability.**

ESP CONTRIBUTORS

Andres Luyts
Adrian Guzman
Adrienne Hoarfrost
Alethea Gollan
Alistair Francis
Alix Leroy
Amartya Hatua
Anca Anghelea
Anna Jungbluth
Ash Hoover
Aurelie (Lily) Jean
Avinash Mahech
Beichen Zhang
Belina Raffy
Benjamin Smith
Bertrand Le Saux
Boris Petrovic
Carlos López-Martínez
Catherine Faber
Christophe Bisson
Christopher Burr
Constantin Weisser
Dana Ngo
Devin Cornell
Diego Arenas
Douglas Rao
Dylan Tull
Elena Degli Innocenti
Emily Elhacham

Fanny Lehmann
Florian Lalande
Frank Geisel
Gabriele De Canio
Gabriele Murchio
Gennadii Donchyts
Gerbrand Koren
Grace Lindsay
Guillermo Jimenez
Hannes Felix Muller
Heather Papworth
Helen Russell
Helga Jordão
Ioannis Prapas
Jakub Nalepa
Jayabalambika Ramakrishnan
Jennie-Marie Larsen
Jian Gong
Jie Zhao
Johanna Afrodita Zuleta
John Ashley
Jonas Weiss
José David Bermúdez
Juan Sebastian Villegas Santos
Karthik Mukkavilli
Kathiravan Thangavel
Keli Wang
Ken Pucker
Kirk McKinzie

Loveneesh Rana
Luca Brocca
Luís Simões
Luke Nicholson
Lyubomyr Demkiv
Mahesh Jampani
Marco Zaccaria Di Fraia
Maria João Sousa
Mateo Petel
Michael Obersteiner
Michelle St Jane
Miguel-Ángel Fernández-Torres
Mihai Datcu
Moctar Dembele
Murray Kerr
Nadine Malcolm
Najla Said
Naylor Peter
Nikolaos Dionelis
Noah Kasmanoff
Oscar Pellicer-Valero
Patrick Griffiths
Philip Stier
Qingsong Xu
Rafael Kaufmann
Ravi Kapur
Raviv Turner
Richard Galvez
Robert Zimmerman

Rochelle Schneider
Roelien Timmer
Rummi Ganguly
Sally Goerner
Sanja Scepanovic
Sankalp Gilda
Saso Dzeroski
Shan Zhao
Shashi Jain
Simon Ilyushchenko
Simon Jackman
Srija Chakraborty
Stefan Kueppers
Stefanie Lumnitz
Sylvester Kaczmarek
Tara Parhizkar
Thanh-Tam Le
Thomas Brunschwiler
Tim Tensen
Vincenzo Schiano Di Cola
Vinutha Magal Shreenath
Viola Steidl
Vit Ruzicka
Yu Zhou
Yue Zeng
Zoltan Szantoi

TRILLIUM TECH

ABOUT
TRILLIUM TECHNOLOGIES

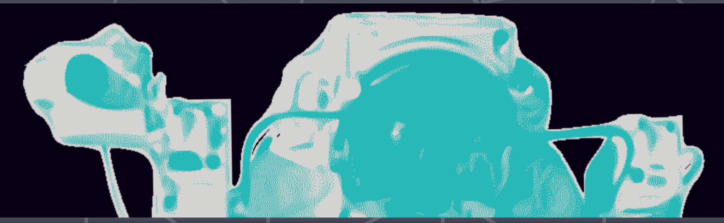
Trillium Technologies is a research and development company with a focus on intelligent systems and collaborative communities for planetary stewardship, disaster response, space exploration and human health.

We are an international team of scientists, developers and AI specialists motivated by impact for human good through the application of technology.
Our work is driven by the scientific method and we believe in open development, transparent processes and the ability to clearly explain outcomes.

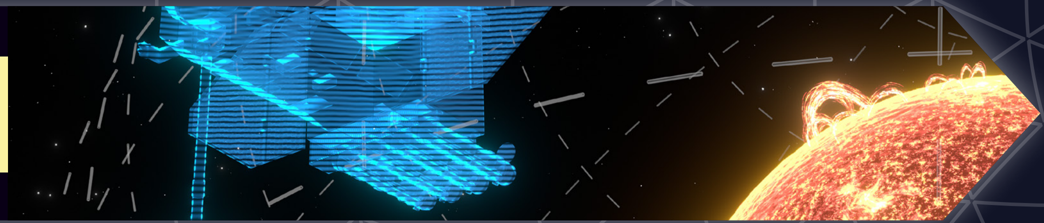
We have deep experience of developing high-performance AI systems in partnership with global space agencies, Silicon Valley tech companies, leading universities and a community of 100+ interdisciplinary PhDs.

We believe that intelligent systems require collaboration and understanding of the human components to ensure solutions are trusted and adopted. Let's talk about the systems in your work and how we can help you achieve your goals.

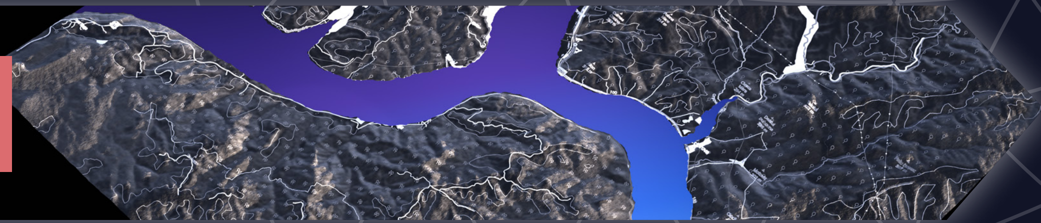
AI Systems for Medicine and Human Health



Heliophysics and Space-weather Digital Twins

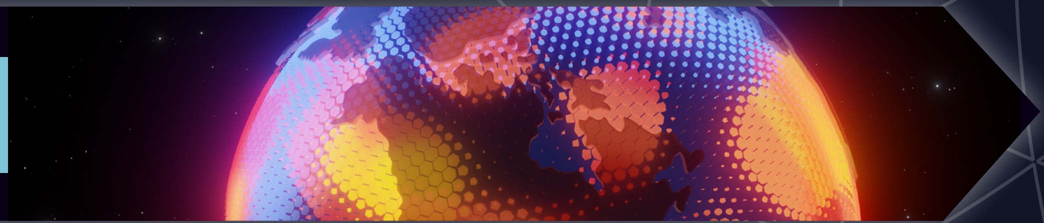


IntelligentSsystems for Disaster Management and Mitigation



Satellites and Exploration

Earth System Models for Planetary Stewardship



team@trillium.tech

RESOURCES AND LINKS

ESP Vision

What is Earth Systems Predictability?
CERN Knowledge Transfer Report
Dissemination of technologies developed at CERN
EO-Use-cases by ESA, Proliferation of commercial use cases for AI
AI for Space Report, Artificial Intelligence for Space: use cases and commercial opportunities
#futuregen: Lessons from a Small Country by Jane Davidson

Building Blocks

The Latest Data from Earth and Space
State of AI for Earth Observation by the Satellite Applications Catapult
Advanced Information Systems Technology (AIST) group NASA's Earth Science Technology Office (ESTO)

Google Earth Engine A planetary-scale platform for Earth science data & analysis

HPC2020 ECMWF's new High-Performance Computing Facility
DUG-Cool data centre cooling solution
UN MARS Methane Alert and Response System
OpenEEW - open earthquake sensor network
ESA'S ANNUAL SPACE ENVIRONMENT REPORT

An AI Capability Paradigm Shift
Google, Transformer Models: "Attention Is All You Need"
FME by Safe Software data integration platform.
Fourier Neural Operators
Adaptive FNOs & SFNOs
Torch Harmonics Spherical Harmonic transform in PyTorch
DeepONet nonlinear operators for identifying differential equations
Differentiable programming for Earth system modelling (paper) & [talk]
Pangaea Data Publisher for Earth & Environmental Science
Apache Beam batch and streaming data processing
CKAN open source data management system

Accelerated Weather and Climate Models
FourCastNet is a data-driven, global, medium-range weather forecast model.
Pangu-Weather, introduced AI model for global weather forecasting
WeatherBench, a benchmark dataset for weather forecasting
ClimaX, a foundation model for weather and climate
FloodDAM, flood prediction tool - prototype digital twin
LLaMA Open and Efficient Foundation Language Models
Visual Instruction Tuning LLM tuning methods
Community Earth System Model global climate simulations of the Earth's past, present, and future climate states.
E3SM: Energy Exascale Earth System Model Harnessing AI and computing to advance climate modelling and prediction, (Schneider et al, 2023)
GraphCast, ML-based weather prediction model

Visualisations and Interfaces for ESP
Bushfire.io, a natural disaster map of weather and fire variables
Earth.nullschool.net, a configurable visualisation of global weather variables
En-ROADS emulator, created by MIT for climate solutions
Earth-2, interactive visualisation provided by NVIDIA
Destination Earth, European Commission digital twin initiative
Earth Intelligence Engine for flood risk visualisation
Omniverse, platform for creating and operating metaverse applications
World3 Simulator, MIT simulator of human population growth
Deep Learning Weather Prediction tools for predicting the gridded atmosphere
WorldGame design science approach to the world problems (historical)

ESP And Generative AI
Neo4j, formal knowledge graph software
RARR, Google LLM for fixing hallucinations and providing citations

Operations

Integration Through Decision Intelligence
Floodbrain, automated flood reporting tool
TransitionZero, data and modelling for climate change

Cooperation and Co-creation
MIT Media Labs, interdisciplinary research group
JustDiggit, re-greening nonprofit
Rainforest Alliance, nonprofit centered on biodiversity, climate, and human rights
Climate-KIC climate innovation community
EU Smart Cities, European Union mission
AIID, artificial intelligence incident database

Making ESP Trusted and Understandable
Ada Lovelace Institution, research institute
Lorien Pratt and Quantellia future powered by decision intelligence

Deployment

Earth Operations Centres and Digital Twins
SAS Analytics for data analysis
European ALMA Regional Centre (ARC) telescope array and support network
Restor network of restoration and conservation sites
OpenAerialMap open collection of drone imagery

ESP in Economics and Finance
Social Value International Social Value Analysis and SROI
RIDDL, social value measurement software

Who Needs ESP?
Climate-Neutral and Smart Cities from Climate-KIC
Country Centred Design Indigenous knowledges in the design of place

TRILLIUM EUROPE