

Data Requirements and Supporting Infrastructure



Data Requirements and Supporting Infrastructure

The Session aims to

- **Explore data types and observation strategies:** How do they need to change to optimize AI development?
- **Discuss data governance and policy frameworks:** How do these influence accessibility, quality, and interoperability across public and private sectors?
- **Examine sectoral approaches to data sharing:** From open access to commercial licensing, how can these foster innovation, equity, and operational readiness?
- **Consider data needs of governments/public service:** sustained, reliable, high-quality information for critical forecasting.
- **Consider data needs for local high-resolution prediction:** Especially in developing and least developed countries with significant data gaps.

Data Requirements and Supporting Infrastructure



Ms. Monica Youngman
US National Weather Service



Dr. Naseema Shyju
Space42



Mr. Julian Green
Brightband



Dr. Paolo M. Ruti
EUMETSAT



Mr. Ashish Raval
Synoptic Data



Dr. Pascal Waniha
INFCOM - WMO



Mr. Rei Goffer
Tomorrow.io



Mr. Nir Stav
WMO



U.S. National Oceanic and Atmospheric Administration US National Weather Service (NWS)

NOAA Mission: To understand and predict changes in climate, weather, ocean, and coasts, to share that knowledge and information with others, and to conserve and manage coastal and marine ecosystems and resources.

NWS Vision: A Weather-Ready Nation

Society is prepared for and responds to weather, water, and climate-dependent events.

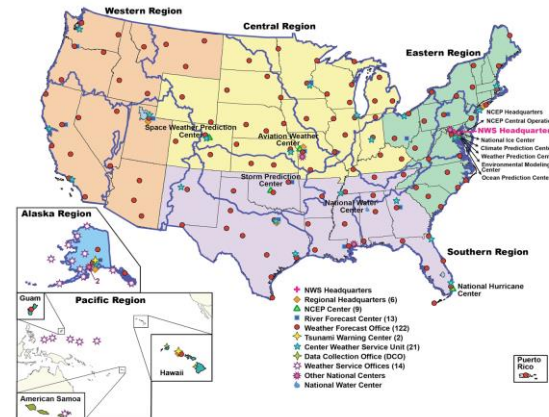


NWS Mission: Providing weather, water, and climate data, forecasts and warnings for the protection of life and property and enhancement of the national economy

Ensuring Ready, Responsive, and Resilient Communities

NOAA is researching and testing AI models with partnerships, including leveraging partnerships and Anemoi

NWS is focused on the operational use of science and technology, especially AI, across the entire weather and water value chain for the mission.



Monica Youngman

**Chief Scientist
AI Lead**

**National Weather
Service
NOAA**





Monica Youngman
National Weather Service

Given AI's remarkable ability to leverage diverse and even lower-quality data, what are the most promising **unconventional data sources** you believe the meteorological community should prioritize for integration into AI models, and what immediate, practical steps can we take to make these sources **usable and impactful**?

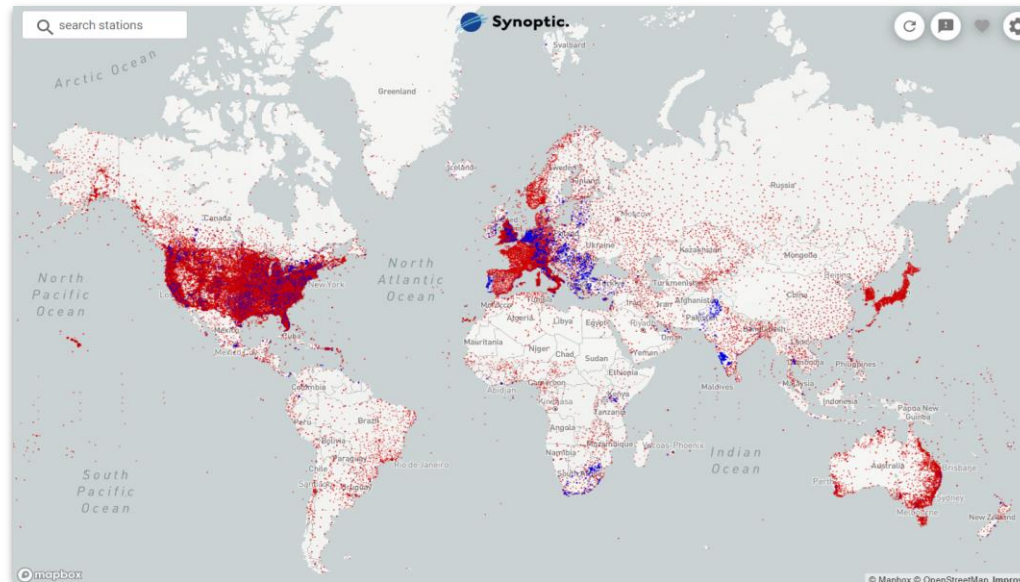
Synoptic Data Public Benefit Corporation

Synoptic is Building the Largest Real-time Environmental Data Platform in the World!

- Data is aggregated from public, private, and academic sources
- Unified data access through API and Data Viewer visualization tools
- Low latency, quality controlled data, and 99% API uptime
- Synoptic awarded multi-year project to operate WIS 2.0 Test Infrastructure
- Data is used by government agencies around the world to enable decisions that save lives, protect property, and improve operations



Mr. Ashish Raval
Synoptic Data



~3 Billion
Annual API Calls

170,000+
Weather Stations

320+
Networks

160+
Environmental Variables

150+
Countries with Data



Mr. Ashish Raval

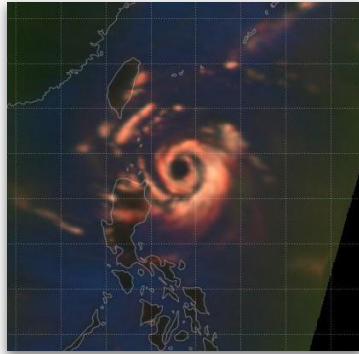
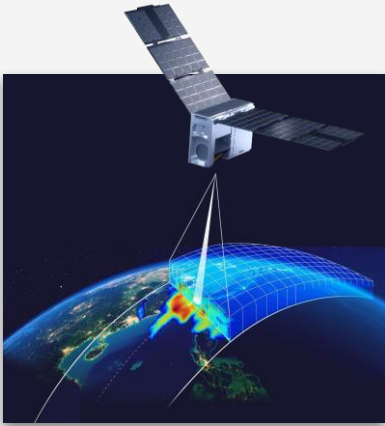
Synoptic Data

What does “**AI-ready**” data actually mean, and how do we get there? What **infrastructure and governance** are required so AI can consume this data at scale—reliably and reproducibly?

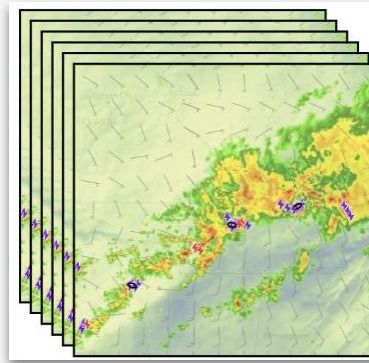
Data Infrastructure Supporting WMO & Members' Full-Stack AI Adoption

Upstream Data Constellation

- Fleet of microwave sounder satellites
- Global coverage, high revisit rate
- Strong impact on NWP/ML forecasts
- Aligned to WMO quality standards
- Public-private partnerships with NMHS

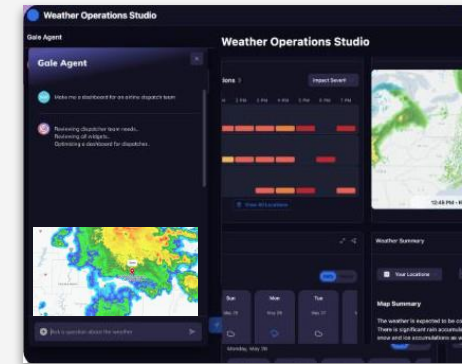


Mr. Rei Goffer
Tomorrow.io



Downstream AI Models & Platforms

- Fusion of satellite and ground datasets
- AI/ML enhancing Member-led prediction
- Reliable, explainable models at scale
- Decision-support aligned with NMHS needs
- Expert support for WMO-led initiatives





Mr. Rei Goffer
Tomorrow.io

How do you see the future of weather observations evolving in the era of **proliferated commercial observation** systems? How would these datasets feed into the public and private prediction systems? **Will all benefit from them?**



Julian Green

CEO, Co-Founder

6th startup.
AI moonshots @
Google X.



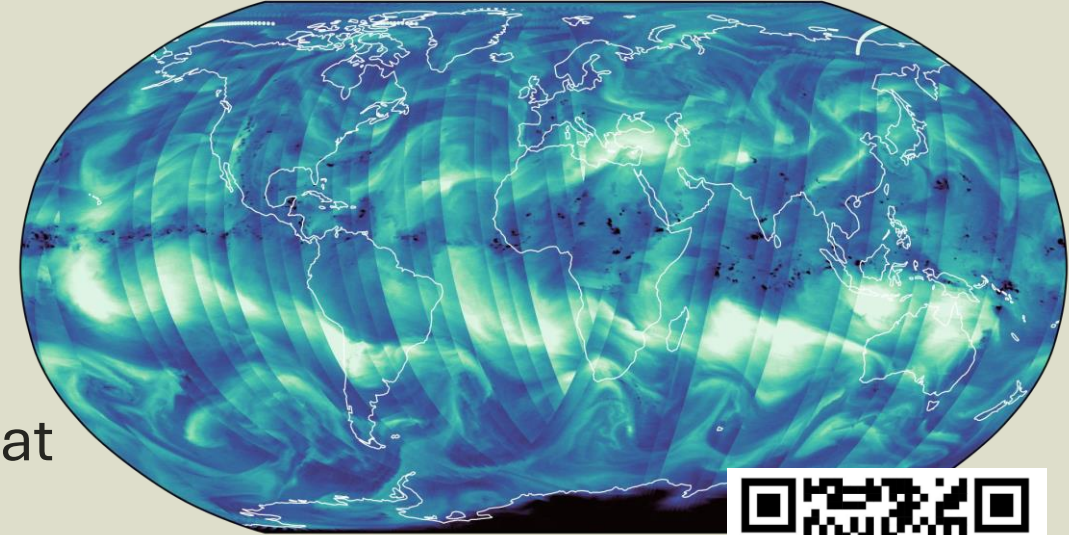
Brightband - Public Benefit Corporation AI weather forecasting tools for custom forecasts, open-source benchmark datasets, metrics

- Forecasts from real-time feed of public & private weather observations

- NNJA-AI v1 >50 TiB of curated observations from NOAA & NASA in Cloud-native AI-ready format (OpenSource)

- Extreme Weather Bench community weather evaluation framework (OpenSource)

ATMS Brightness Temperature, Channel 22, 2024-08-13



nnja-ai.readthedocs.io



Mr. Julian Green
Brightband

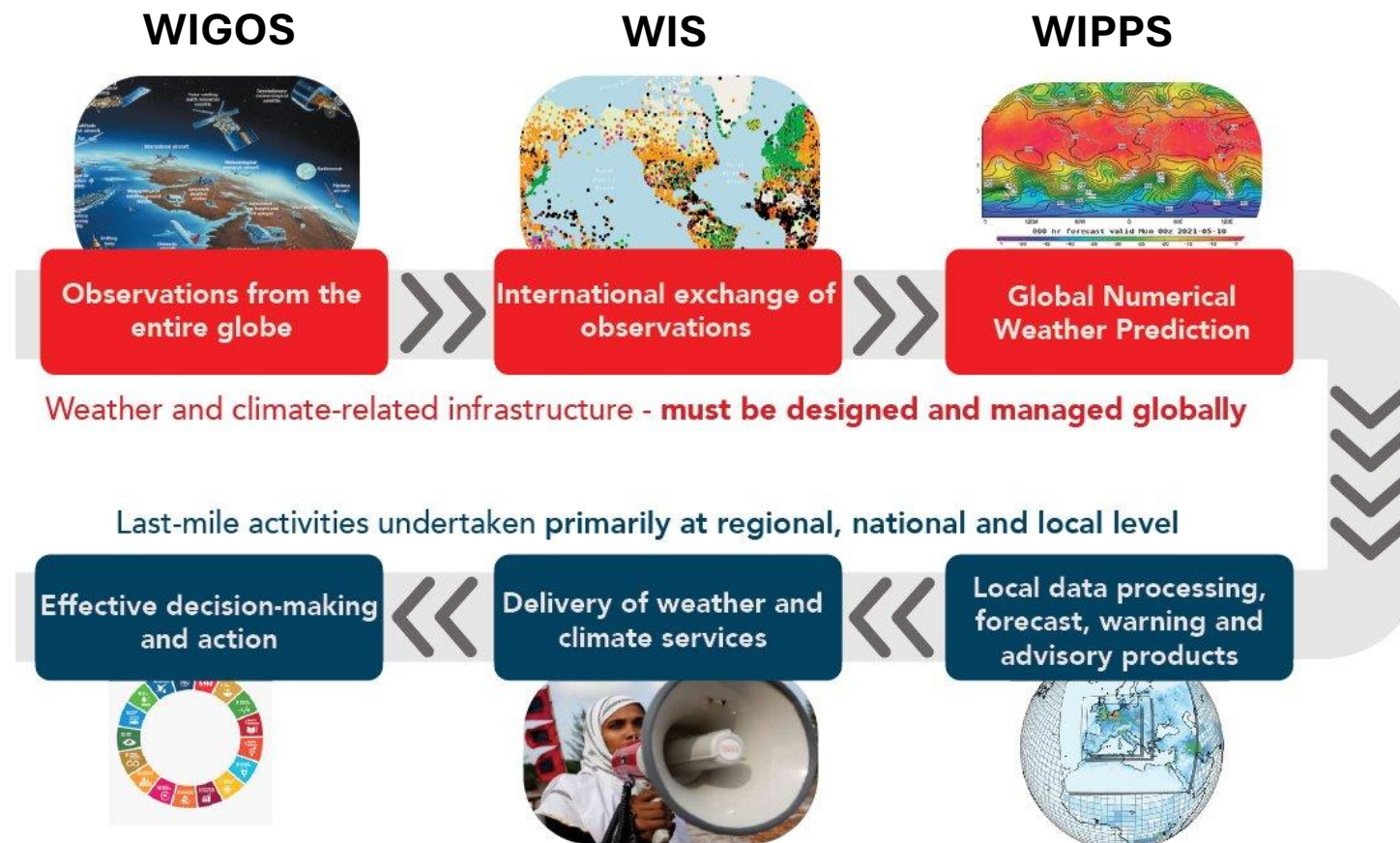
What are the **emerging needs of private sector** meteorology firms that the WMO and its Members should be aware of?
How can we **make weather data available** to the private sector to encourage the most AI innovation?



Dr. Pascal Waniha
INFCOM - WMO



Meteorological Value Cycle & WMO Global Infrastructure





Dr. Pascal Waniha
INFCOM - WMO

What are the main physical and institutional **barriers to data sharing for AI systems in the developing world? What should be the main efforts of WMO to support developing countries **feed the AI systems local data** for training, operational use and verification?**



Dr. Naseema Shyju
Space42

Harness the Power of AI for
Geospatial Intelligence

gIQ leverages state-of-the-art artificial intelligence algorithms to process massive datasets, helping you uncover hidden patterns, identify trends, and gain a comprehensive understanding of your spatial data. Our advanced machine learning models deliver accurate predictions, optimize workflows, and enable you to make data-driven decisions with confidence.

<https://giq.ae/>

AI Powered Geospatial Data Analytics & Visualization Platform (gIQ) =

Seamless Data Ingestion

gIQ centralizes data from diverse sources to foster efficient collaboration within and beyond organizations.

Satellite Imagery Acquisition

Access high-resolution satellite imagery from various data providers for diverse applications.

Data Analytics, Visualization and Reporting

Leverage annotations, segmentation and visualization tools to interpret complex data and generate clear, actionable reports for decision-makers.

Marketplace and Collaboration

Host and share AI models and apps on gIQ Marketplace to foster innovation and collaboration across users.



Dr. Naseema Shyju

Space42

**Can AI/ML weather models be effectively applied
in **data-sparse regions** such as the Middle East
and Africa ?**



Current EUMETSAT satellites

www.eumetsat.int

SENTINEL-3A & -3B (98.7° incl.)

Low Earth, sun-synchronous orbit

Copernicus satellites delivering marine data services from 814km altitude

JASON-3 (63° incl.)

Low Earth, non-synchronous orbit

Copernicus ocean surface topography mission (shared with CNES, NOAA, NASA and Copernicus)

Sentinel-6 Michael Freilich (66° incl.)

Low Earth, non-synchronous orbit

Copernicus ocean surface topography mission (shared with NASA, NOAA, ESA and Copernicus with support from CNES)

Dr. Paolo M. Ruti
EUMETSAT

METEOSAT-10, -11

Geostationary orbit

Meteosat Second Generation

Two-satellite system

Full disc imagery mission (15 mins)
(Meteosat-11 (0°))

Rapid scan service over Europe (5 mins)
(Meteosat-10 (9.5° E))

METEOSAT-9 (45.5° E)

Geostationary orbit

Meteosat Second Generation
providing Indian Ocean
data coverage

METOP-B & -C (98.7° incl.)

Low Earth, sun-synchronous orbit

EUMETSAT Polar System (EPS)/
Initial Joint Polar System

METOP-12

Geostationary orbit

Meteosat Third Generation imaging mission

MTG-S2

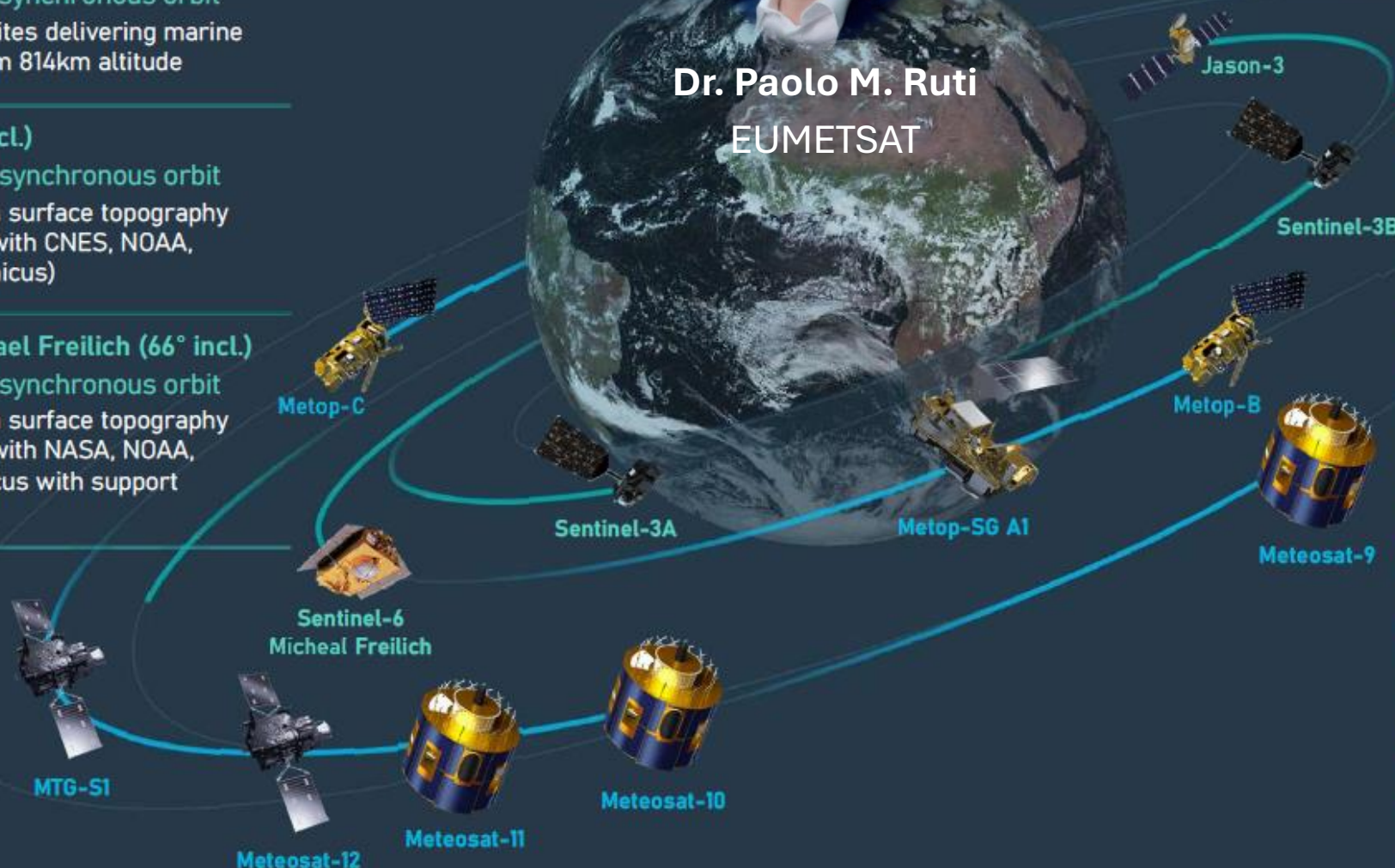
Geostationary orbit

Meteosat Third Generation sounder
mission, currently in commissioning phase

METOP-SG A 1

Low Earth, sun-synchronous orbit

EUMETSAT Polar System (EPS)/
Second Generation – commissioning phase





Dr. Paolo M. Ruti
EUMETSAT

In the AI prediction era, do we need to change the paradigm of observational network design used up to now? Will we still need **backbone sustained global high quality observation, or do we need to consider also other options such as **concentrated effort to provide “data cubes” for model training?****

Questions to the panel:

- What are the structural or conceptual **barriers** that prevent truly **universal, AI-ready data sharing**?
- What is needed to **unlock the full potential** of global and regional data for AI training?
- What mechanisms could foster greater trust and **collaborative data-sharing ecosystems** in the AI era?

Questions to the panel:

- With **private companies rapidly developing observational networks** in addition to the governmental/public efforts, how can we ensure that the **global public good** is prioritized while fostering private sector innovation?

Questions to the panel:

- Given AI's ability to derive significant value from **"good enough" data**, challenging the WMO's long-standing emphasis on "golden standard" observations and rigid quality requirements, **how can WMO adjust** to embrace this paradigm shift without compromising public trust or the **reliability** needed for mission-critical forecasting?

Data Requirements and Supporting Infrastructure

