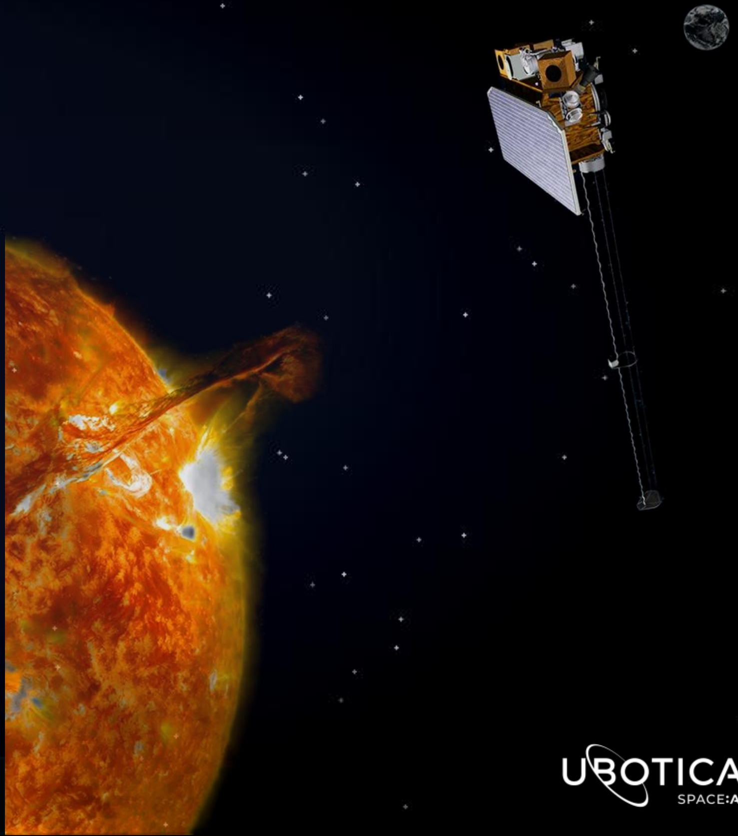
Data Fusion for Space Weather at Scale

Opportunities, Challenges, and AI

Besma Guesmi · ESA Space Weather Days 8-9 June, 2026 ESAC, Spain



Why Space Weather Matters - The Stakes

$2.6T

Estimated economic impact

Carrington-scale event (NAS 2008 upper bound)

40+

Satellites lost or degraded

Cumulative across radiation & drag events

G5

Storm class - May 2024

First extreme event since Halloween 2003

Operational Sectors at Risk

- **Satellite Operations** Orbital decay, surface charging, SEUs, drag modelling
- **GNSS / Navigation** Ionospheric scintillation, TEC fluctuations, positioning errors
- **Pipelines & Railways** Induced currents in long metallic conductor infrastructure
- **Power Grid Infrastructure** GIC-induced transformer failures & cascading outages
- **HF Radio / Aviation** Polar route disruptions, communication blackouts
- **Human Spaceflight** Elevated radiation exposure, ISS crew & deep-space missions

The Space Weather Data Challenge

>100	>1 TB	~13 min	5+
Active monitoring assets	**Data volume per day**	**L1 → Earth transit time**	**Incompatible data formats**
Satellites, ground stations, GNSS networks	Solar imagery, in-situ plasma, ionospheric TEC	Available CME warning window	CDF, HDF5, NetCDF, FITS, ASCII streams

Space-Based Sources

- ACE, DSCOVR, IMAP, L1 solar wind (Bz, Vp, Np)
- SDO/AIA, EUV solar disk imagery (12 channels)
- GOES X-ray & proton flux monitoring
- STEREO-A coronagraph (CME tracking)
- SOHO/LASCO CME catalogues
- Commercial LEO: Spire GPS-RO, Tomorrow.io

Ground-Based Sources

- SuperMAG / INTERMAGNET, 500+ magnetometers
- IGS / EUREF GNSS, global TEC maps
- Digisonde / ionosonde networks
- Incoherent scatter radars (EISCAT, Millstone)
- All-sky cameras & riometers
- Solar radio burst monitors

UBOTICA SPACE:AI

Data Fusion: Core Opportunities

(1) Observational Coverage

Fill spatial and temporal gaps in sparse networks. Satellite + ground fusion delivers near-continuous global ionospheric state estimation, including data-sparse ocean regions.

Key technique: Optimal interpolation · kriging · neural assimilation

(2) Holistic Situational Awareness

From isolated readings to a coherent system state. Event chain tracking , solar eruption → CME propagation → magnetospheric impact → ionospheric response, becomes computationally tractable.

Key output: Integrated space weather state vector

(3) Improved Forecast Skill

Multi-source models outperform single-source baselines. Fused solar wind + ionospheric inputs could reduce Kp/Dst forecast RMSE in ensemble frameworks.

Key metric: Skill Score vs. persistence and climatology baselines

(4) Operational Decision Support

Satellite drag corrections, HF outage windows, GNSS integrity flags, grid GIC risk indices. Confidence-weighted fused inputs drive operator-ready, sector-specific notifications.

Key users: Satellite ops · grid managers · aviation · GNSS providers

AI & ML: Core Approaches

Transformer / Attention Models

Capture long-range dependencies in multivariate solar wind series. Encoder-decoder architectures for Dst, Kp, AL prediction.

e.g. SWFormer, STORM-AI, ForeFlux

Graph Neural Networks

Encode irregular magnetometer network topology. Message-passing propagates geomagnetic disturbances spatially for GIC risk mapping.

e.g. HelioGNN, MagNet-Graph

Convolutional Networks

Process SDO EUV/X-ray imagery for flare detection and classification. CME front extraction from coronagraph sequences.

e.g. FLARECAST, SDO-ML, HELCATS-DL

Physics-Informed Surrogates

Neural emulators of TIEGCM, OpenGGCM, WACCMX. Orders-of-magnitude speedup enabling real-time ensemble generation.

e.g. DL-TIEGCM, ML-WACCM

Bayesian & Ensemble UQ

MC Dropout, Deep Ensembles, Conformal Prediction for calibrated forecast uncertainty. Confidence intervals, not just point estimates.

Critical operational requirement

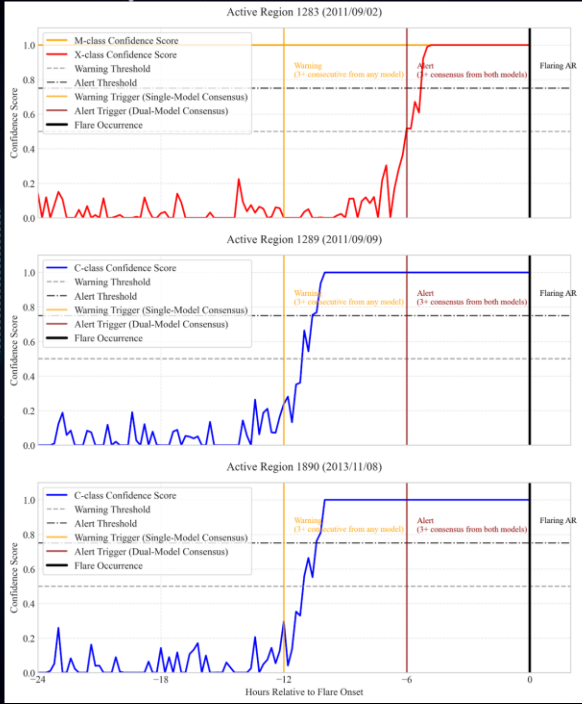
Anomaly Detection

Unsupervised detection of sensor faults, data outages, and novel events. VAEs and isolation forests on streaming telemetry.

e.g. OMNI-QA, SWx-Monitor

UBOTICA SPACE:AI

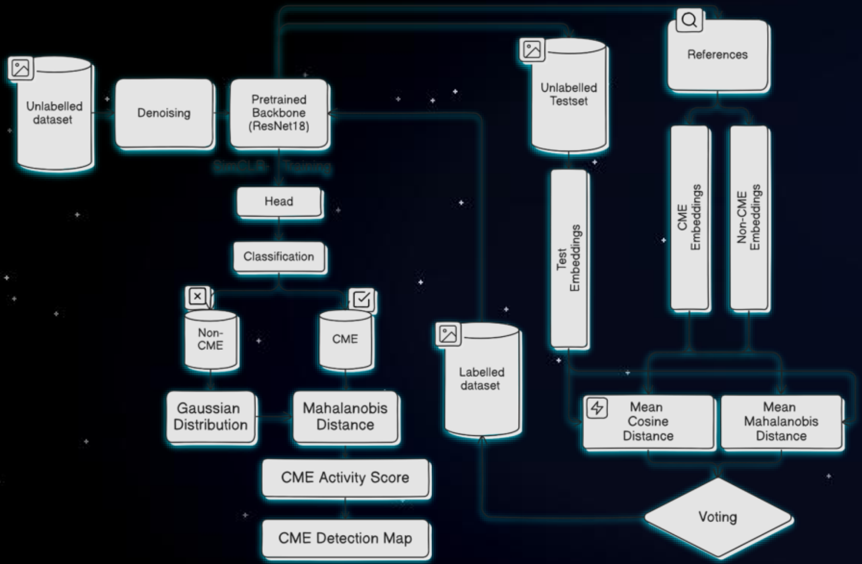
AI & ML: Core Approaches: EoFTCNets Framework



Model	Configuration	Acc. (%)	Rec. (%)	Prec. (%)	FAR (%)	TSS (%)	Params	Inf. Time (ms)
EoFPhyNet (3DTCN) - X-class								
	(3, [4,8,16], 120, No)	92.78	92.49	100.0	100.0	85.26	186K	4.2
	(3, [4,8,16], 60, No)	94.26	94.03	100.0	100.0	88.30	186K	6.4
	(3, [32,64,128], 120, Yes)	**98.33**	**98.27**	100.0	100.0	**96.60**	187K	6.3
	(5, [4,8,16,32,64], 60, No)	96.17	100.0	**96.17**	**0.00**	96.17	334K	7.0
EoFPhyNet (3DTCN) - M-class								
	(3, [4,8,16], 120, No)	88.42	87.95	100.0	100.0	76.34	186K	4.2
	(3, [4,8,16], 60, No)	90.16	89.72	100.0	100.0	79.83	186K	6.4
	(3, [32,64,128], 120, Yes)	**96.10**	**98.11**	90.70	15.79	**88.22**	187K	6.3
	(5, [4,8,16,32,64], 60, No)	94.27	99.20	93.85	2.15	93.65	334K	7.0
EoFPhyNet (3DTCN) - C-class								
	(3, [4,8,16], 120, No)	85.94	85.32	100.0	100.0	71.28	186K	4.2
	(3, [4,8,16], 60, No)	88.02	87.45	100.0	100.0	75.16	186K	6.4
	(3, [32,64,128], 120, Yes)	**93.70**	**97.85**	88.42	11.58	**85.74**	187K	6.3
	(5, [4,8,16,32,64], 60, No)	91.15	98.60	90.25	5.40	90.80	334K	7.0
EoFGeoNet (3DTCN Variant) - X-class								
	(3, [4,8,16], 120, No)	91.50	91.20	100.0	100.0	83.50	186K	4.3
	(3, [4,8,16], 60, No)	93.80	93.50	100.0	100.0	87.10	186K	6.5
	(3, [32,64,128], 120, Yes)	97.80	97.60	100.0	100.0	95.50	187K	6.4
	(5, [4,8,16,32,64], 60, No)	95.30	99.90	95.40	0.10	95.30	334K	7.1
EoFGeoNet (3DTCN Variant) - M-class								
	(3, [4,8,16], 120, No)	87.60	87.10	100.0	100.0	75.20	186K	4.3
	(3, [4,8,16], 60, No)	89.85	89.40	100.0	100.0	78.90	186K	6.5
	(3, [32,64,128], 120, Yes)	**95.40**	**97.50**	89.85	16.25	**86.75**	187K	6.4
	(5, [4,8,16,32,64], 60, No)	93.65	98.80	93.10	3.90	92.85	334K	7.1
EoFGeoNet (3DTCN Variant) - C-class								
	(3, [4,8,16], 120, No)	85.21	84.60	100.0	100.0	70.42	186K	4.3
	(3, [4,8,16], 60, No)	87.60	87.05	100.0	100.0	74.30	186K	6.5
	(3, [32,64,128], 120, Yes)	**92.70**	**96.90**	87.65	12.35	**84.15**	187K	6.4
	(5, [4,8,16,32,64], 60, No)	90.40	97.80	89.50	6.50	89.20	334K	7.1
Baselines - X-class								
	ResTCN (Window=120)	66.11	64.74	100.0	100.0	30.85	11.3M	24.9
	ConvLSTM (2 Layers)	58.20	55.50	61.30	38.70	12.50	2.43M	3.8
Baselines - M-class								
	ResTCN (Window=120)	62.50	60.85	100.0	100.0	26.70	11.3M	24.9
	ConvLSTM (2 Layers)	54.69	52.10	58.40	41.60	9.80	2.43M	3.8
Baselines - C-class								
	ResTCN (Window=120)	60.94	59.20	100.0	100.0	24.40	11.3M	24.9
	ConvLSTM (2 Layers)	52.08	49.50	56.25	43.75	7.50	2.43M	3.8
EoFTCNets (Fusion)								
	X-class Flare							
	Fusion (EoFPhyNet+EoFGeoNet)	**99.20**	**99.10**	99.80	0.20	**98.20**	373K	12.7
	M-class Flare							
	Fusion (EoFPhyNet+EoFGeoNet)	**97.50**	**98.80**	95.20	4.80	**95.30**	373K	12.7
	C-class Flare							
	Fusion (EoFPhyNet+EoFGeoNet)	**95.80**	**98.20**	92.10	7.90	**92.50**	373K	12.7

B. Guesmi et al., "EoFNets: EyeonFlare Networks to predict solar flare using Temporal Convolutional Network (TCN)," 2024 10th International Conference on Control, Decision and

AI & ML: Core Approaches: CMEGNets Framework



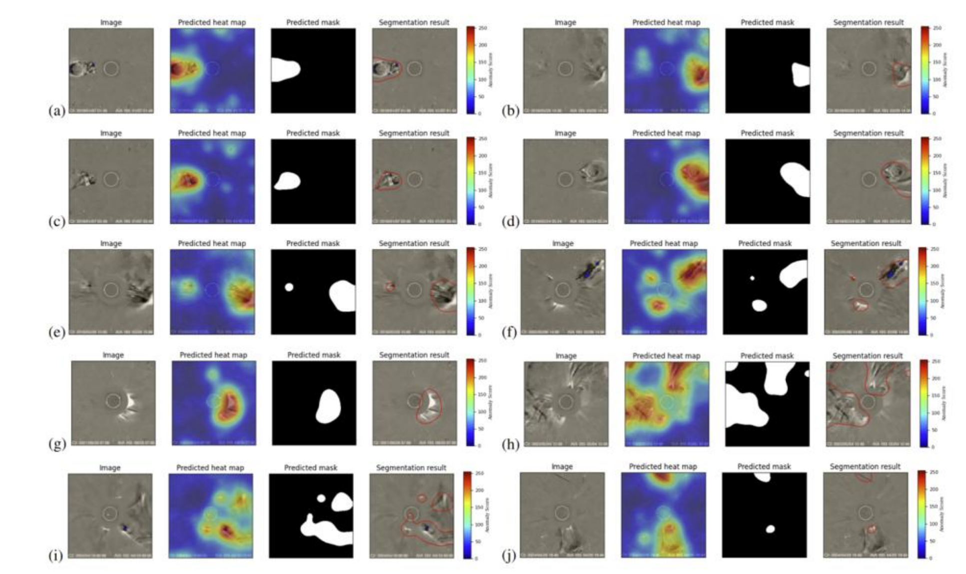
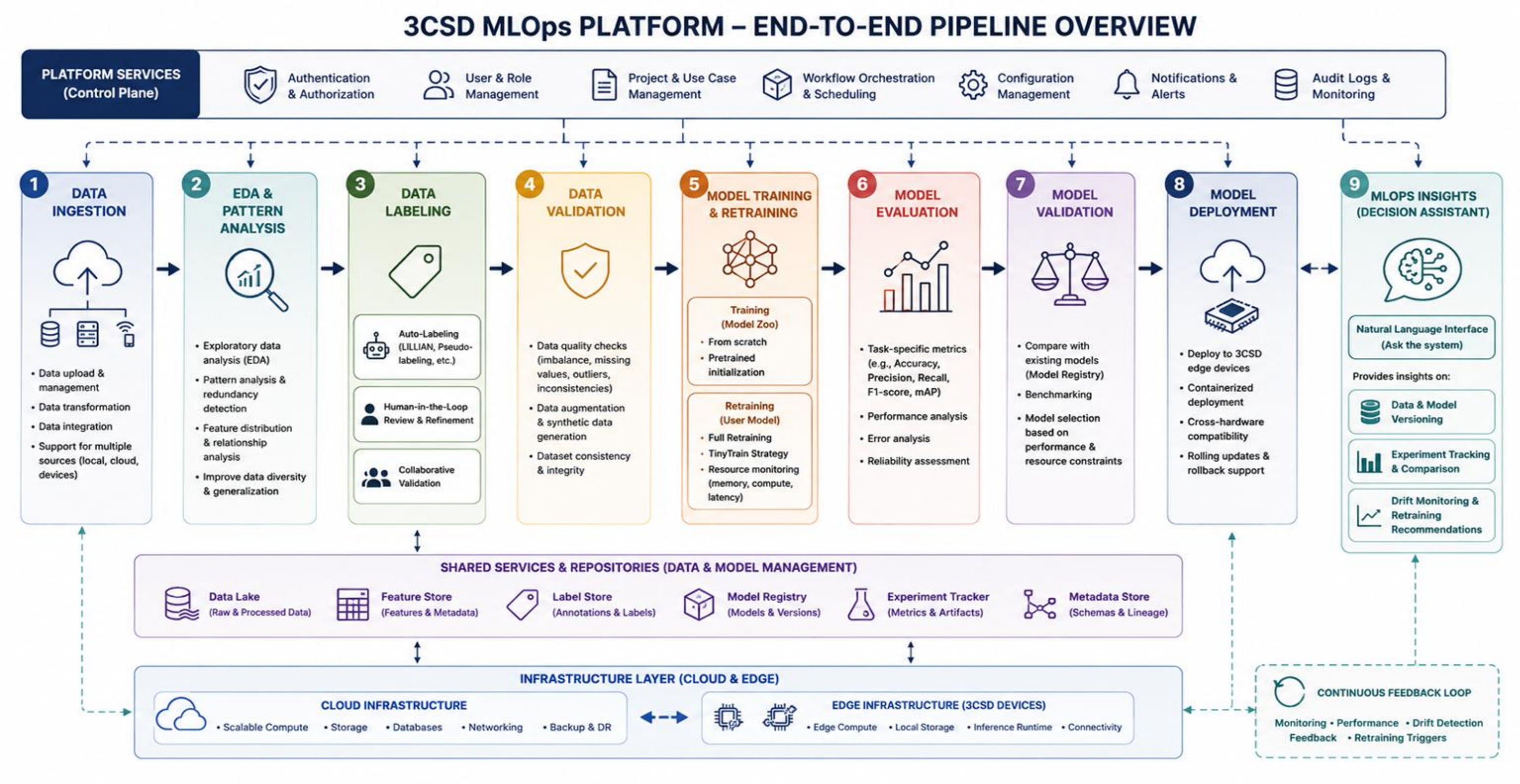


Figure: Exemplar CME events detected by CMEGNets but omitted from SEEDS (b,d,e), CDAW (a,f,g,h,j), or both (c, i). SEEDS omissions: (b) 2019-03-20 14:36:05 (sub-threshold intensity ratio 1.2), (d) 2019-02-24 05:24:05 (pseudo-streamer blowout), (e) 2019-03-20 15:05:46 (high-background event). CDAW omissions: (a) 2019-01-07 01:48:05 (narrow, angular width: 30° as measured by SEEDS) (f) 2022-02-06 14:00:07 (single-frame), (g) 2021-09-28 07:00:00 (stealth CME), (h) 2023-05-04 13:00:00 (jet-accelerated narrow CME), (j) 2024-04-20 19:48:05 (complex background). Both omissions: (c) 2019-01-07 03:48:06 (stealth CME), (i) 2024-04-19 00:00:05 (streamer blowout mini-CME).

CMEGNets: A self-supervised framework for coronal mass ejection detection & region segmentatio, Besma Guesmi, Jinen Daghrir, David Moloney, Jose Luis Espinosa-Aranda, Elena Hervas-Martin, Advances in Space Research (2026, Vol. 77, No. 6, pp. 7455–7483) — DOI: https://doi.org/10.1016/j.asr.2026.01.061 . URL:

End-to-End AI Fusion Pipeline



Challenge 1 & 2: Scalability & Interoperability

Scalability

Volume Growth

Commercial LEO constellations will add thousands of streaming sources this decade. Current batch architectures were not designed for this throughput.

Latency Requirements

End-to-end latency <5 min from observation to product. Streaming frameworks (Kafka, Flink) required, batch pipelines cannot meet this SLA.

Edge vs. Cloud

Pushing inference on-orbit reduces downlink bottlenecks. Cloud-native autoscaling (Kubernetes) handles ground-tier bursts.

Compute Heterogeneity

Space-grade NPUs/FPGAs differ from cloud hardware. Model portability and quantisation are active challenges.

Interoperability

Format Fragmentation

CDF (NASA), HDF5 (ESA), NetCDF (atmospheric), FITS (solar), ASCII (legacy). No single standard dominates. Automated translation pipelines are essential.

Metadata Inconsistency

SPASE standard exists but adoption is incomplete. Calibration history, quality flags, and provenance are inconsistently populated.

Data Latency & Gaps

Real-time feeds have variable latency; archives have gaps. Fusion models must be robust to missing and irregularly sampled inputs.

Access & Governance

Commercial data licencing conflicts with open-science mandates. Federated learning offers a pathway to shared models without raw data sharing.

Challenge 3 & 4: Uncertainty and Trust

Uncertainty Quantification (UQ)

Calibration is a Hard Requirement

Stated confidence intervals must be statistically valid. Miscalibrated AI outputs create false security in safety-critical decisions.

Methods

Deep Ensembles, MC Dropout, Conformal Prediction, Bayesian NNs. Each trades computational cost for calibration quality differently.

Full Pipeline UQ

Uncertainty must propagate from sensor measurement through feature extraction and model inference, not just added at the output layer.

Extreme Event Coverage

Training data is sparse for rare extreme storms. Models must flag elevated epistemic uncertainty when conditions fall outside the training distribution.

Trust & Explainability

Black-Box AI Not Yet Operational

Safety-critical operators (satellite companies, grid operators, aviation authorities) require explainable, auditable model behaviour before deployment.

XAI Methods

SHAP feature attribution, attention map visualisation, saliency maps for solar imagery. Translate model reasoning into operator-interpretable language.

Validation Against History

Models must be tested on benchmark events: Oct 2003 Halloween storms, March 1989, Sept 2017 X9 flare, May 2024 G5 storm.

Model Governance

Version control, statistical drift monitoring, retraining triggers. Operational AI requires the same lifecycle rigour as safety-critical software.

UBOTICA™
SPACE:AI

Ubotica Technologies: Edge AI for Space Weather

Ubotica develops AI/ML inference platforms for resource-constrained edge environments, including satellite payloads. The 3CSD platform provides NPU-accelerated, space-grade AI inference, enabling real-time onboard processing independent of ground infrastructure. Demonstrated on commercial LEO missions.

Onboard Event Detection & Triage

- Real-time anomaly detection on satellite processors
- Space weather event flagging without ground contact
- Alert latency from hours → seconds
- Supports onboard autonomous safe-mode triggering

Federated Learning Across Constellations

- Collaborative model training across constellation nodes
- No raw data centralisation, preserves data sovereignty
- Reduces uplink cost, respects commercial licences
- Models improve continuously as the constellation grows

Intelligent Telemetry Prioritisation

- AI-driven selection of highest-value data to downlink
- Critical during geomagnetic storms, windows are limited
- Ensures observational continuity at peak event periods
- Compressed representations reduce bandwidth 60-80%

Aligned with ESA Space Safety Programme · SWE Service Network requirements

UBOTICA SPACE:AI

Next-Generation Space Weather Architecture

EDGE TIER

On-orbit & Sensor-Proximate

- Onboard inference
- Event detection & triage
- Compressed telemetry streams
- Anomaly flagging
- Autonomous safe-mode

FUSION TIER

Cloud / Distributed Ground

- Multi-source data ingestion
- Format harmonisation & QA
- ML fusion & forecast models
- Ensemble UQ frameworks
- Space weather product gen.

USER TIER

Operational Decision Support

- Sector dashboards & APIs
- Alert & notification services
- Confidence-weighted products
- XAI / explainability layer
- Model governance & audit

Data Inputs: Space-based sensors · Ground observation networks · Commercial constellations · Model reanalysis

UBOTICA SPACE:AI

Key Takeaways

01 Heterogeneous data fusion is foundational, no single source provides adequate coverage or forecast skill for operational space weather services.

02 AI/ML is the only practical route to real-time fusion at scale, methods must be matched to the physics

03 Scalability and interoperability are the unsolved engineering challenges, not the AI algorithms. Format harmonisation and streaming architectures must be community priorities.

04 Uncertainty quantification and explainability are prerequisites for operational adoption, calibrated confidence intervals and transparent model reasoning are non-negotiable.

05 Edge AI (on-orbit inference) offers a step-change in alert latency and data efficiency, Ubotica is building the bridge between satellite observation and ground-based fusion.

Thank you · Besma Guesmi, besma.guesmi@ubotica.com · ESA Space Weather Days Spain, Madrid

UBOTICA SPACE:AI