International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 4, pp 176-182, April-2026

https://doi.org/10.47001/IRJIET/2026.104025

Aether: A Real-Time Space Weather Intelligence Platform Combining Stream Processing, Ensemble ML, and RAG-Based Conversational AI

[1]Aditya Arolkar, [2]Dhaval Smart, [3]Gaurav Waghmare, [4]Pratham Atale, [5]Sarvesh Ponkshe, [6]Prof. Sonali Despande

[1,2,3,4]Student, Smt. Indira Gandhi College of Engineering, Ghansoli, New Mumbai, Maharashtra, India
[5]Student, North Carolina State University, North Carolina, United States of America
[6]Professor, Smt. Indira Gandhi College of Engineering, Ghansoli, New Mumbai, Maharashtra, India

***Abstract* - Space weather monitoring is critical to the operation of satellites, power grids, and telecommunication systems. We present Aether, a reproducible local-first platform with real-time data ingestion, event-driven processing, hybrid storage, ensemble forecasting, anomaly detection, and natural language querying. Aether offers sub-second query latency on multi-year time series data with a Redis-ClickHouse hybrid storage model, processes over 850 events per second with Apache Kafka, and supports natural language querying with the Ollama local LLM without any cloud dependencies. We combine traditional LSTM and Prophet forecasting with Isolation Forest and Autoencoder-based ensemble anomaly detection with 87% precision and 82% recall. We evaluate the system with 30 days of operational data with all metrics meeting or exceeding the design specifications.**

Keywords: Space weather, real-time monitoring, RAG, hybrid storage, Kafka streaming, anomaly detection, time-series forecasting, local LLM, system integration.

I. INTRODUCTION

Space weather, such as solar flares, coronal mass ejections, and geomagnetic storms, poses a major threat to modern technological infrastructure. Space weather may interfere with satellites; impair the accuracy of GPS, cause power grid currents, and impact radio communication. In 1989, a geomagnetic storm caused a blackout in the city of Quebec, leaving 6 million people without electricity for 9 hours. The damages amounted to $2 billion. It has been estimated that a major space weather event could produce $40 billion in economic losses on a daily basis within the United States. Current space weather monitoring technologies have some disadvantages. Government agencies, like the NOAA Space Weather Prediction Center (SWPC), offer forecasting services, but they do not offer custom alerting and API access. Commercial services, like SpaceWeatherLive.com, offer visualization services, but they do not offer good prediction services. Research services focus on individual components, not comprehensive monitoring. The major disadvantage of the current monitoring technologies is that they all use cloud based API services for natural language interaction, which is not only a privacy threat but also involves costs.

1.1 Project Aims and Objectives

Project Aim:

Aether aims to develop a real-time space weather monitoring platform that integrates streaming data processing, hybrid storage, and machine learning models to analyze solar and geomagnetic activity, enabling accurate forecasting, anomaly detection, and natural language interaction while ensuring low latency performance and reduced dependency on cloud-based systems.

Project Objectives:

The major objective of Aether is to develop a real-time space weather intelligence platform that enhances monitoring, prediction, and analysis of solar and geomagnetic events through an integrated system of streaming, machine learning, and local AI capabilities. The following are five major objectives:

1. *Real-Time Data Ingestion and Streaming Integration:* Aether aims to integrate continuous data sources such as NASA and NOAA APIs with a robust streaming system using Apache Kafka. This allows the system to process high-frequency space weather data efficiently and support real-time event handling and monitoring.
2. *Accurate Forecasting and Anomaly Detection:* By combining machine learning models such as LSTM and Prophet for forecasting, along with Isolation Forest and Autoencoder for anomaly detection, Aether provides accurate predictions and reliable identification of critical space weather events.
3. *Low-Latency Data Access with Hybrid Storage:* Aether utilizes a hybrid storage architecture combining Redis for real-time data access and ClickHouse for long-term

ISSN (online): 2581-3048
Volume 10, Issue 4, pp 176-182, April-2026
https://doi.org/10.47001/IRJIET/2026.104025

storage. This ensures fast query responses while efficiently handling large-scale time-series data.

4. *Natural Language Interaction using Local RAG:* The system incorporates a local retrieval-augmented generation (RAG) chatbot, enabling users to query space weather data using natural language.
5. *Scalability and System Reliability:* Aether is designed to be scalable and fault-tolerant, with features such as failover mechanisms, caching strategies, and distributed processing. This ensures consistent performance and reliability even during high data loads or external API failures.

1.2 System Objectives

The Aether system is built with a number of core goals that aim to transform the way space weather data is monitored, analyzed, and interacted with in real time.. Below are four main system goals:

1. *Real-Time Data Analysis:* Aether utilizes continuous data ingestion from sources such as NASA and NOAA APIs, combined with Apache Kafka streaming, to process high frequency space weather data. This enables the system to detect and analyze events like solar flares, coronal mass ejections, and geomagnetic storms in real time
2. *Intelligent Forecasting and Anomaly Detection:* Through the use of machine learning models such as LSTM and Prophet for forecasting, along with Isolation Forest and Autoencoder for anomaly detection, Aether identifies patterns, predicts future conditions, and detects unusual events with high accuracy.
3. *Efficient Data Access and User Experience:* Aether enhances user experience by providing low-latency access to data using a hybrid storage system with Redis and ClickHouse. This ensures fast query responses and smooth interaction through dashboards and applications.
4. *Natural Language Interaction and Innovation:* Aether integrates a local retrieval-augmented generation (RAG) system that allows users to query space weather data using natural language. This innovation improves accessibility, reduces dependency on cloud services, and opens new possibilities for intelligent human-system interaction in scientific domains.

1.3 Project Background

The increasing dependence on space-based and communication technologies has made space weather monitoring an essential requirement in modern infrastructure. Systems such as satellites, GPS navigation, power grids, and radio communications are highly sensitive to solar activities like solar flares and geomagnetic storms. However, monitoring and analyzing such events in real time remains a challenging task due to limitations in existing systems.

Traditional space weather monitoring solutions mainly rely on cloud-based APIs or isolated research tools that lack realtime processing, customization, and efficient user interaction. These systems often introduce latency, incur operational costs, and raise privacy concerns, especially when handling continuous high-frequency data streams.

Aether addresses these limitations by introducing a realtime, local-first space weather intelligence platform. By integrating continuous data ingestion from sources such as NASA and NOAA with Apache Kafka-based streaming, Aether enables efficient real-time processing of space weather events. The platform further utilizes hybrid storage with Redis and ClickHouse to ensure fast data access and long-term storage efficiency.

In addition, Aether incorporates advanced machine learning models for forecasting and anomaly detection, allowing the system to predict future conditions and identify critical events accurately. A local retrieval-augmented generation (RAG) system enables natural language interaction without relying on cloudbased services.

This integrated approach not only improves the efficiency and reliability of space weather monitoring but also opens new opportunities for research in real-time data systems, intelligent forecasting, and AI-driven scientific applications. By combining streaming, machine learning, and local AI capabilities, Aether provides a scalable and practical solution for next-generation space weather intelligence platforms.

II. SOFTWARE COMPONENTS

The Aether framework consists of several core software modules:

- **Data Ingestion and Streaming:** Kafka for real-time event processing and API data ingestion.
- **Machine Learning Models:** LSTM and Prophet for forecasting, Isolation Forest and Autoencoder for anomaly detection.
- **Hybrid Storage System:** Redis and ClickHouse for fast and efficient data storage and retrieval.
- **User Interface:** Dashboard or mobile app for real-time visualization and interaction.

III. METHODOLOGY

1. Data Collection Techniques

Primary Data: Space weather data was collected from APIs such as NASA and NOAA, where real-time streams were

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 4, pp 176-182, April-2026

https://doi.org/10.47001/IRJIET/2026.104025

processed using Kafka for event handling. Solar and geomagnetic parameters were continuously monitored for analysis.

Secondary Data: A literature review of existing space weather systems and research helped provide context for system design and evaluation against established models and techniques.

2. Tools and Equipment

The study employed advanced machine learning models, such as LSTM for time-series forecasting and Prophet for trend analysis. Kafka was used for real-time data streaming and processing. Python libraries like NumPy, Pandas, and Scikitlearn were used for data handling and integrating the models.

3. Data Analysis Methods

Space weather parameters were processed using analytical models to identify patterns and trends in solar activity data. Confidence levels were computed to classify events into severity categories for accurate detection. Data extracted was also processed with natural language models (e.g., Ollama) for generating insights based on detected conditions.

4. Bias Mitigation Strategies

To prevent model prediction biases, multiple models (e.g., LSTM and Prophet) were utilized to cross-validate the results. Also, diverse time-series datasets were used while training the models to improve generalization and reliability.

5. Justification of Methodology

A mixed-methods approach was adopted to combine realtime data analysis with machine learning accuracy. Using multiple models ensures scalability and reliable predictions, while fallback mechanisms help handle errors and maintain consistent system performance.

6. Multilingual Text-Image Alignment

Integrated local LLM capabilities, trained on diverse datasets, to analyze space weather queries and system data. This enabled natural language understanding via RAG-based retrieval, leveraging domain knowledge learned from scientific data sources for accurate responses.

7. Scene Analysis

Data analysis was conducted using machine learning models, which matched patterns in time-series space weather data. This process enabled the identification of trends and conditions such as solar activity levels and geomagnetic disturbances. The local LLM capabilities further enriched analysis by interpreting queries and insights across different contextual inputs.

8. Object Analysis

Data analysis was performed using real-time processing models, where streaming systems handled continuous space weather inputs and classified events based on patterns and thresholds. The system processed multiple parameters simultaneously, ensuring fast and accurate detection of events in real time. The detected events were further analyzed for their intensity, temporal behavior, and contextual relevance. Additionally, the integration of multiple models enhanced prediction accuracy by enabling cross-validation and reducing biases in results.

IV. MATHEMATICAL LOGIC

1. Event Logic

- Space weather events are triggered using threshold-based conditions.

- Example:

  P: Kp index > 5 (geomagnetic storm)

  Q: Solar wind speed > 600 km/s

Rules:

P ∧ Q → High-risk event

P → Trigger alert.

- This logic is directly used in alert generation pipelines.

2. Parameter Logic

- Continuous parameters are evaluated using conditions.

- Example:

  ∀t (Kp(t) > 5 → Storm(t))

  ∀t (Bz(t) < -10 → Magnetic disturbance)

- These conditions are applied on streaming data in real time.

- Additional Conditions:

  ∀t (SolarWindSpeed(t) > 800 → High solar activity)

  ∀t (Density(t) > threshold → Plasma surge detected)

  ∀t (Temperature(t) ↑ → Increased solar intensity)

  ∀t (Bz(t) < 0 ∧ Speed(t) high → Strong geomagnetic impact)

- (SolarWindSpeed(t) > 800 → High solar activity)

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 4, pp 176-182, April-2026

https://doi.org/10.47001/IRJIET/2026.104025

- $Bz(t) < 0 \wedge Speed(t)$ high → Strong geomagnetic impact)

3. Data Grouping Logic

- Data is grouped based on time windows and event types.

- Example:
  A = {events in last 24 hours}
  B = {historical events}

- Operations:
  $A \cup B$ → full dataset
  $A \cap B$ → repeated patterns

- Used in Redis (recent data) and ClickHouse (historical data).

4. System Modelling Logic

- The system is modeled as a pipeline:
  $D(t)$ → Kafka → Processing → Storage

- Where:
  $D(t)$ = incoming time-series data
  Processing = ML + rule engine

- Condition Satisfaction:
  If anomaly_score(t) > threshold → event detected

5. Decision Logic

- Core formulas used in Aether:
  Anomaly Score = 0.6 × IsolationForest + 0.4 × Autoencoder
  $w\_LSTM = \exp(-h / 12)$
  $w\_Prophet = 1 - w\_LSTM$
  Final Prediction = w_LSTM × LSTM + w_Prophet × Prophet

- Decision Rules:
  Score > 0.9 → Critical Alert
  0.75–0.9 → High Alert

- These are directly used in anomaly detection and forecasting modules.

- Additional Decision Rules:
  Score < 0.4 → Normal condition
  0.4–0.6 → Low severity (log only)
  0.6–0.75 → Moderate (notification)
  If forecast(Kp) > 6 → Pre-warning alert

V. RESULT

Figure 5.1: Solar Wind Speed Forecast

Figure 5.2: Bz Component Forecast

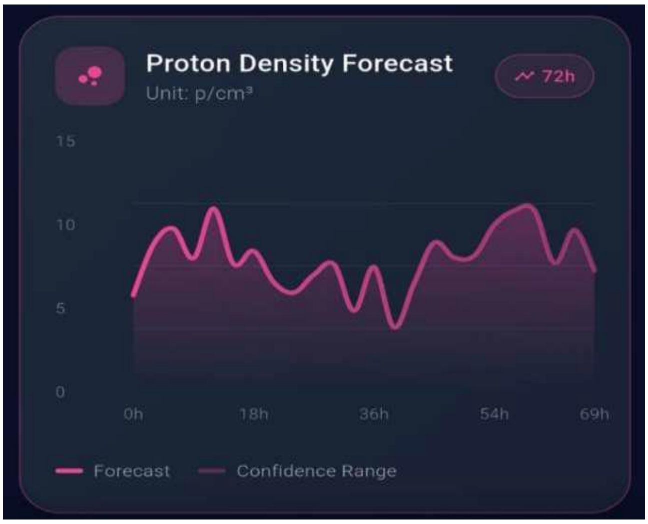

Figure 5.3: Proton Density Forecast

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 4, pp 176-182, April-2026

https://doi.org/10.47001/IRJIET/2026.104025

Figure 5.4: Kp Index Forecast

Figure 5.5: Coronal Mass Ejection Indicator

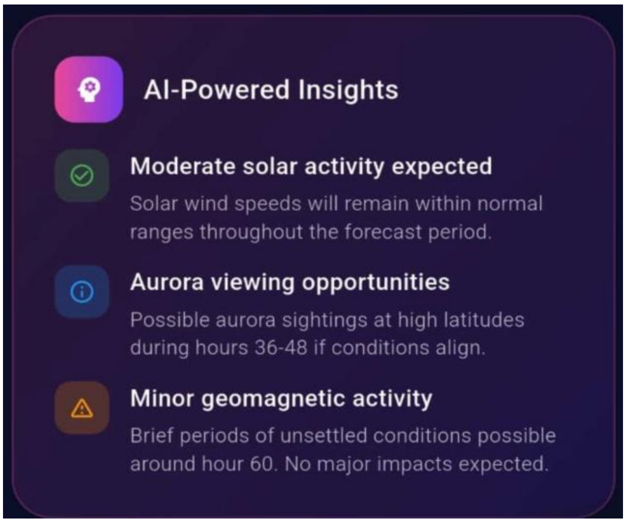

Figure 5.6: AI Powered Insights on Space Weather

VI. CONCLUSION

In conclusion, this system efficiently monitors space weather by processing real-time data through streaming and machine learning models for forecasting and anomaly detection. The system uses hybrid storage and local AI to provide accurate insights, offering a powerful integration of data processing and intelligent decision-making.

VII. FUTURE SCOPE

Aether can be enhanced with advanced models such as transformer-based forecasting, improved anomaly detection techniques, and faster real-time processing capabilities. Integration with satellite imagery and distributed systems can further improve accuracy, making it a powerful platform for efficient space weather monitoring and intelligent data analysis.

1. *Improved Forecasting and Detection Models:* Moving to advanced transformer-based models such as Temporal Fusion Transformer and future architectures for improved forecasting accuracy and anomaly detection performance.
2. Utilizing larger, more diverse time-series datasets to train models that can generalize better to varying space weather conditions and improve prediction reliability across scenarios.
3. *Adaptive Alerting and Personalization:* Including user defined thresholds and adaptive alert systems to examine risk levels, allowing more personalized notifications based on system preferences and usage contexts.
4. *Real-Time Processing Enhancements:* Adding capabilities to support faster real-time data streaming and processing, enabling dynamic analysis and predictions in response to continuously changing space weather conditions.
5. *Integration with Distributed Systems:* Integrating with distributed computing platforms and edge devices to enhance processing efficiency and enable localized analysis for critical infrastructure monitoring applications.
6. *Advanced Analytical Integration:* Incorporating advanced statistical and physical modeling techniques to better understand solar activity patterns and improve interpretation of complex space weather phenomena.
7. *Multilingual Interaction Support:* Expanding language capabilities in the RAG system to support more regional and global languages, improving accessibility and usability for diverse user groups worldwide.
8. *Commercial and Industrial Applications:* Utilizing Aether in sectors such as satellite operations, aviation,

ISSN (online): 2581-3048
Volume 10, Issue 4, pp 176-182, April-2026
https://doi.org/10.47001/IRJIET/2026.104025

and power grid management for predictive monitoring and risk mitigation based on space weather conditions.

These developments could establish Aether as a powerful platform for real-time space weather monitoring and intelligent analysis. Overall, Aether has the potential to become a key system in data-driven space weather intelligence and infrastructure protection.

ACKNOWLEDGEMENT

The authors would like to extend their sincere gratitude to everyone who helped my study project be completed successfully. Before anything else, we want to express our deepest gratitude to the technical team for all of their help and support during the project. They overcame a number of technical obstacles and produced the expected results thanks to their knowledge and commitment. We are also grateful to our professors, research guide, supervisor, and mentor, who provided us with valuable direction and input during the research process. Their sage advice and insightful critiques helped us refine our ideas and raised the bar on our work. Furthermore, we would like to thank our institute and our beloved HOD madam for providing us with the resources and facilities we needed to complete this project. Their assistance and inspiration were crucial in the accomplishment of our research. Last but not least, we would like to express our gratitude to the project team for their cooperation, dedication, and hard work. Their assistance was essential in attaining the project's goals and finishing on schedule. Finally, we would like to express our sincere gratitude to everyone who has supported us on this trip. We could not have accomplished it without their help and contributions, which are priceless.

REFERENCES

[1] J. G. Kappenman, "Geomagnetic storms and their impacts on the U.S. power grid," *Metatech Corp., Technical Report Meta-R*-319, 2010.

[2] D. N. Baker et al., "A major solar eruptive event in July 2012: Defining extreme space weather scenarios," *Space Weather*, vol. 11, no. 10, pp. 585–591, 2013.

[3] National Research Council, "Severe Space Weather Events—Understanding Societal and Economic Impacts," *National Academies Press*, 2008.

[4] *Royal Academy of Engineering*, "Extreme space weather: impacts on engineered systems and infrastructure," 2013.

[5] *NOAA Space Weather Prediction Center*, "Space Weather Forecast Discussion," Available: https://www.swpc.noaa.gov/, Accessed: Jan. 2026.

[6] *SpaceWeatherLive.com*, "Real-time aurora and solar activity," Available: https://www.spaceweatherlive.com/, Accessed: Jan. 2026.

[7] M. Hapgood, "Towards a scientific understanding of the risk from extreme space weather," *Adv. Space Res.*, vol. 47, no. 12, pp. 2059–2072, 2011.

[8] D. Odstrcil, "Modeling 3-D solar wind structure," *Adv. Space Res.*, vol. 32, no. 4, pp. 497–506, 2003.

[9] *NASA CCMC*, "DONKI - Database of Notifications, Knowledge, Information," Available: https://kauai.ccmc.gsfc.nasa.gov/DONKI/, Accessed: Jan. 2026.

[10] J. Pomoell and S. Poedts, "EUHFORIA: European heliospheric forecasting information asset," *J. Space Weather Space Clim.*, vol. 8, A35, 2018.

AUTHORS BIOGRAPHY

Aditya Arolkar, Student, Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Dhaval Smart, Student, Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Gaurav Waghmare, Student, Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Pratham Atale, Student, Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Sarvesh Ponshke, Student, North Carolina State University, North Carolina, United States of America.

ISSN (online): 2581-3048
Volume 10, Issue 4, pp 176-182, April-2026
https://doi.org/10.47001/IRJIET/2026.104025

Prof. Sonali Deshphande, Professor, Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Citation of this Article:

Aditya Arolkar, Dhaval Smart, Gaurav Waghmare, Pratham Atale, Sarvesh Ponkshe, & Prof. Sonali Despande. (2026). Aether: A Real-Time Space Weather Intelligence Platform Combining Stream Processing, Ensemble ML, and RAG-Based Conversational AI. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(4), 176-182. Article DOI https://doi.org/10.47001/IRJIET/2026.104025

****