`

An Ontology-Based Semantic Model to Identify Risk in Large-Scale Religious Gatherings

Govind Kushwah[1]

Department of Physics and Computer Science, Dayalbagh Educational Institute, Agra, India[1]

govindkushwah1111@gmail.com[1]

Abstract: Mass religious events are very problematic with the extreme density of the crowd, environmental destruction, and social stress. Although data-driven models have been extensively applied to predict such risks, they tend to be less semantically integrated and explainable. This paper proposes an ontology-based framework to model and infer multidimensional stress conditions during mega religious events, using the MahaKumbh as a case study.The framework integrates forecasted crowd, environmental, and social indicators into a formal ontology developed using Web Ontology Language (OWL). Semantic Web Rule Language (SWRL) rules are used to classify the stress levels and make inferences of high-risk events, using logical reasoning. It is based on the MahaKumbh 2025 case study, which shows that the proposed approach can be used to infer the categories of stresses and high-risk events without manual labeling.The findings emphasize the fact that ontology-based reasoning offers understandable, explainable, and reusable knowledge representations that enhance conventional predictive models. The suggested framework provides a decision-support tool that is powerful to the planners and policymakers engaged in the organization of religious meetings of a high scale.

Keywords: Ontology, Semantic Web, SWRL, Risk Assessment, Mass Gatherings, Knowledge Representation

I. INTRODUCTION

Mass religious gatherings are one of the most multifaceted versions of human congregation, with the large-scale inflow of population, dense mobility forms, and long-term socio-environmental pressure. Crowds like MahaKumbh have tens of millions of pilgrims, and their risks increase due to the congestion of crowds, pollution of the environment, social pressure, and risks associated with the safety of people. Recent studies have pointed out that the uncontrolled growth of crowds and environmental degradation during such events can greatly increase the level of safety risks and governance problems [1], [2].

To address these challenges, researchers have turned to more data-oriented approaches, such as machine learning and time-series prediction, which can be used to forecast the behavior of crowds and indicators of stress in advance [3], [4]. Although these models have been used successfully in numerical prediction, they are normally used in isolation and do not explicitly represent semantic relationships between event entities (crowd profiles, environmental factors, and social conditions). Consequently, the decision-makers do not have a single and explainable body of knowledge that can integrate various indicators and automate the reasoning process.

The promising solution is provided by semantic technologies, especially ontologies, which allow knowledge representation of heterogeneous domains in a formal way and logical inference [5]. Ontology-based modeling can be used to fill the quantitative analytics and qualitative decision-making gap in the context of mass gatherings. This paper is motivated by constraints in the previous predictive research, among which we have conducted a previous data-driven study on MahaKumbh stress forecasting. This paper suggests a detailed ontology-based mechanism to predict and model multidimensional stress and risks related to mega religious events.

II. RELATED WORK

The management of mass gatherings has been developed in three significant directions, namely predictive analytics, environmental and social impact assessment, and semantic knowledge modeling.

http://doi.org/10.63766/spujstmr.26.000080

`

The initial research aimed at analyzing the patterns of crowds and safety risks during large events through statistical and simulation-based methods [1]. As the data volumes were expanding, the methods of machine learning have been used to predict crowds of visitors, the intensity of pollution, and the pressure on infrastructure in advance [3]. ARIMA and Prophet are time-series models widely applied to predicting the trends of pollution and waste production during religious and cultural events [4].

Recent works have discussed environmental degradation and social pressure caused by mass gatherings, suggesting the interdependence of the density of crowds, the decline in air and water quality, and the issue of safety concerns of people [2], [6]. Such studies, however, tend to take each area of impact individually and do not have a cohesive representation paradigm.

Ontology-based semantic modeling has been successfully used in smart cities, disaster management, and environmental monitoring to bring together heterogeneous data sources and allow logical inference [5], [7], [11]. However, ontological methods of mega religious events are still scarce, especially in the propagation of stress and risk classification.

In our previous paper, we suggested a data-driven model that predicts ecological and societal stress indicators in MahaKumbh with the help of machine learning and time-series analysis. Although it was useful in forecasting trends of stress, semantic representation and reasoning were not represented, since it was not as interpretable and availed of reuse of knowledge. To fill this gap, the current research presents a model based on ontology that formalizes and gives the event knowledge and allows the automated inference of high-risk conditions using semantic rules.

III. METHODOLOGY

The proposed methodology follows a chronological and reproducible working procedure that includes data-driven stress evaluation and ontology-based semantic reasoning. The entire process comprises six distinct phases, beginning with the data collection and preprocessing through rule-based inference of the high-risk event conditions.

3.1 Data Collection and Preprocessing

The data collected in this research is mainly based on historical and estimated indicators linked to the MahaKumbh event. The data was gathered using publicly available structured datasets on Kaggle and consists of 3 major groups of stress-related indicators: crowd indicators, such as the total number of visitors and domestic–international crowd composition; environmental indicators, including air pollution levels, water quality indices, and waste generation metrics; and social indicators, comprising reported crime incidents and sentiment-related measures.

The stress forecasting, preprocessing, and data collection approach is already formulated and tested in our previous study of MahaKumbh stress analysis [8]. In that study, the machine learning and time-series forecasting methods are used to predict future patterns of crowd pressure, environmental degradation, and social stress on the basis of Kaggle-based datasets [9].

In the current research, the predicted stress variables and significant contributing variables calculated in the previous predictive model are reutilized as validated input indicators. Instead of recalculating numerical predictions, it is about semantic formalization and ontology-based reasoning, which makes it possible to infer the high-risk event predicates in an explainable way.

3.2 Quantification of Event Stress (Prior Work Integration)

The prior predictive model reported by Kushwah and Singh [8] used multiple stress variables that were normalized and aggregated to calculate a composite Event Stress Index (ESI), representing the cumulative level of stress for a given event year. The ESI is defined as:

$$ESI_t = \sum_{i=1}^{n} w_i \times S_i \tag{1}$$

WhereS_i represents the normalized stress indicator (e.g., pollution, crowd size, crime), and w_i denotes the relative importance weight learned using machine learning models.This quantitative measure provided an opportunity to rank the years of events by the severity of stress prognoses. Nevertheless, this method is good at numerical comparison, but it is impossible to interpret the results semantically and explicitly model the interaction of components of stress, which inspired the ontology-based extension of this work.

3.3 Ontology Development and Concept Formalization

Based on the identified dimensions of stress, a domain ontology is constructed in OWL to give a formal representation of the MahaKumbh event ecosystem. The ontology introduces MahaKumbhEvent as the primary concept that has semantic connections with three key dimensions of stress, namely CrowdProfile, EnvironmentalImpact, and SocialStressImpact. Stress dimensions are further broken down into meaningful sub classes (eg. DomesticCrowd and InternationalCrowd under CrowdProfile, and AirQuality and WaterQuality under EnvironmentalImpact).

Object properties such as hasCrowdProfile, hasEnvironmentalImpact and hasSocialStressImpact are established to capture the semantic relationships between the event and stress factors that are associated to it. The data properties are used to match the quantitative values obtained through the previous predictive analysis with the individuals. Figure 1 shows the supposed conceptual scheme of the ontology, whereby the major stress domains are interconnected without inferential or personal instances.

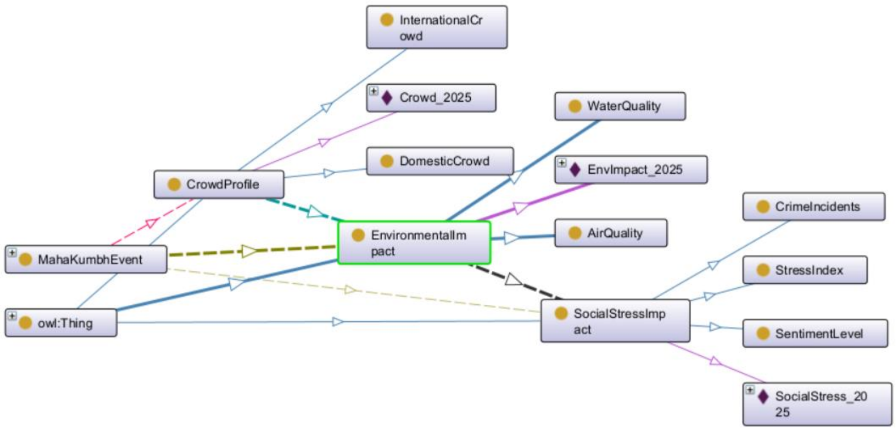

Fig.1. Conceptual ontology structure for modeling multi-dimensional stress factors.

3.4 Individual Instantiation from Forecasted Data

Individuals are instantiated in the ontology using forecasted and observed indicators in order to make semantic reasoning about real-world event scenarios. As an example, Mahakumbh2025PrimaryEvent is the instance of the MahaKumbh event in 2025, and Crowd2025, EnvImpact2025, and SocialStress2025 are the instances of the crowd, environmental, and social stress, respectively.

Data properties connect quantitative values of the Event Stress Index items and personal stress indicators with these cases. This step of instantiation fills the gap between numerical prediction output and semantic knowledge representation to allow further logical inference. The instantiated ontology before inference is shown in figure 2 which shows how predicted data are embedded semantically in the ontology structure.

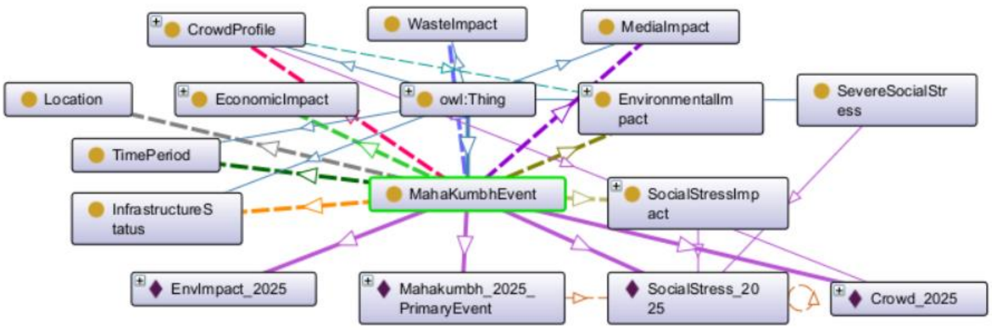

Fig.2. Ontology instantiation using forecasted crowd, environmental, and social stress data.

3.5 Rule Formulation for Stress Classification

Quantitative stress indicators are converted into qualitative stress categories in the ontology by using Semantic Web Rule Language (SWRL) rules. The reasoning strategy is threshold based whereby stress conditions are deduced with the help of domain informed cutoff values that are based on the analytical conclusions and empirical findings that we have provided in our previous study [8]. The SWRL rules directly represent expert knowledge and allow showing a transparent, explainable classification of stress conditions [10]. The summaries of representative rules are as follows.

High Crowd Strness Rule:

According to this rule, a crowd profile is considered a highly stressed profile when the number of visitors surpasses some predefined limit:

$$CrowdProfile(x) \wedge visitorCount(x) > \theta_c \rightarrow HighCrowdStress(x) \qquad (2)$$

where θ_c denotes the crowd density threshold.

High Environmental Stress Rule:

Environmental stress is inferred based on pollution indicators exceeding acceptable limits:

$$EnvironmentalImpact(y) \wedge PollutionIndex(y) > \theta_e \rightarrow HighEnvironmentalStress(y) \quad (3)$$

where θ_e represents the environmental pollution threshold.

Severe Social Stress Rule:

SevereSocialStress is inferred when social disruption indicators—such as crime incidence and negative public sentiment—exceed predefined critical thresholds, indicating heightened social instability during the event:

$$SocialStress(s) \wedge CrimeIncidents(s) > \theta_{crime} \rightarrow SevereSocialStress(s) \quad (4)$$

Where θ_{crime} represents the crime incidence threshold.

High-Risk Event Rule:

An event at MahaKumbh is a HighRiskEvent when several dimensions of stress are at the same time at a critical level:

$$MahaKumbhEvent(z) \wedge hasCrowdProfile(z, x) \wedge HighCrowdStress(x) \wedge hasEnvironmentalImpact(z, y) \wedge HighEnvironmentalStress(y) \wedge hasSocialStressImpact(z, s) \wedge SevereSocialStress(s) \rightarrow HighRiskEvent(z)$$

This rule integrates crowd, environmental, and social stress dimensions to infer overall event risk.

The ontology provides automatic and explainable inference of more complicated risk situations that cannot be described by numerical prediction alone through the formalization of stress classification logic with SWRL. The application of SWRL rules in the Protégé system, as shown in Figure 3, indicates that stress classification and inference of high-risk events are done using executable semantic rules as opposed to conceptual definitions only.

Fig.3. Implementation of SWRL rules in the Protégé environment for stress classification and high-risk event inference.

3.6 Semantic Reasoning and Inference Execution

The ontology and SWRL rules are used to perform semantic reasoning using a description logic reasoner. When reasoning, implicit knowledge is derived when asserted facts are automatic. Instances of events that meet predetermined stress conditions are sorted under the higher-level risk types, like HighRiskEvent. This inference mechanism eliminates the need for manual risk labeling and enhances explainability compared to purely data-driven models.

3.7 Ontology Quality Evaluation Metrices

The ontology was evaluated using structural, logical, and inference-oriented quality checks. The purpose of this evaluation was to verify whether the ontology can consistently represent the MahaKumbh stress domain, execute SWRL rules without contradiction, and generate traceable risk classifications from asserted indicator values.

TABLE I

ONTOLOGY QUALITY EVALUATION SUMMARY FOR THE PROPOSED SEMANTIC RISK MODEL

Evaluation metric	**Validation method**	**Genuine reported outcome**
Consistency check	Reasoner execution in Protégé after class, property, individual, and SWRL rule definition	The ontology was treated as valid only after successful reasoning without contradictory class assertions or conflicting inferred risk classes.
Completeness assessment	Mapping of the selected indicators to crowd, environmental, and social stress concepts	The selected case-study indicators were represented through ontology classes, object properties, data properties, and individuals.
Rule coverage	Verification of SWRL rules against the three stress dimensions and the composite event-risk condition	Rules covered crowd stress, environmental stress, social stress, and final HighRiskEvent classification.
Inference traceability	Inspection of inferred class hierarchy, DL Query result, and SWRL execution view	The HighRiskEvent inference can be traced to HighCrowdStress, HighEnvironmentalStress, and SevereSocialStress conditions.
Reasoning execution status	Manual verification of rule execution in the Protégé environment	Reasoning was completed for the prototype case-study ontology; no hardware-dependent numerical runtime benchmark is claimed.
Reproducibility of decision logic	Use of explicit OWL classes/properties and SWRL rules instead of manually assigned risk labels	Risk classification is repeatable because the rule antecedents and inferred classes are explicitly represented.

These validation checks strengthen the reliability of the framework because the inferred risk output is not only a numerical prediction, but a logically explainable conclusion derived from formally represented entities, properties, and SWRL rules.

IV. RESULTS AND CASE STUDY ANALYSES

This section presents the results of the implementation of the proposed ontology-based framework on the MahaKumbh case study. The assessment aims to show how semantic reasoning could convert forecasted indicators of stress into high-level risk classifications by rule-based inference.

4.1 Case Study Description

The MahaKumbh 2025 event is used as a representative of a large-scale religious gathering to test the proposed methodology. Predicted crowd, environmental, and social stress metrics based on the previous predictive analysis are implemented in the ontology as individual entities. These instances include the primary event **(**Mahakumbh_2025_PrimaryEvent**)** and its associated crowd, environmental, and social stress profiles.

After the individual instantiation, the ontology reasoning process is implemented to infer the severity of stress categories and the general risk of events using pre-specified SWRL rules. The ontology instantiation of the MahaKumbh 2025 case study before semantic inference is shown in figure 4, showing how the event is related to its respective stress profiles.

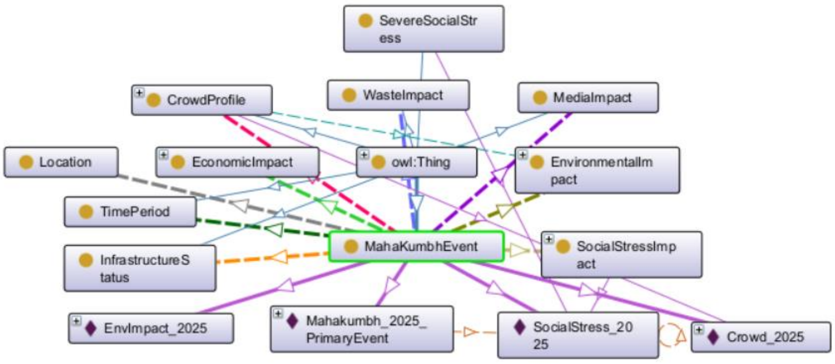

Fig.4. Ontology instantiation for the MahaKumbh 2025 case study using forecasted crowd, environmental, and social stress indicators prior to semantic inference.

4.2 Inference of Stress Conditions

Once the forecasted event data has been instantiated, a description logic reasoneris used with SWRL rules to perform semantic reasoning. This step is aimed at automatically deriving qualitative stress conditions based on quantitative crowd, environmental, and social indicators without any manual labeling.

According to the set threshold-based rules, the reasoner categorized stress-associated individuals into higher-level stress factors. In particular, the individual Crowd2025 is inferred as an instance of HighCrowdStress when the indicators of crowd intensity surpassed the predefined limits. On the same note, EnvImpact2025 and EnvtStress2025 are automatically inferred as HighEnvironmentalStress due to the high pollution and environmental degradation levels. SocialStress2025, in the social dimension, is inferred under SevereSocialStress, as the stress related to crime and sentiment significantly increased.

These stress classifications are generated automatically through semantic inference and are not explicitly asserted in the ontology. Inferred class memberships can only be seen in the Inferred class hierarchy view, and thus, it is clear that the stress conditions are the product of rule execution and logical reasoning instead of manual assignment. The reasoning derived the inferred stress conditions presented in Figure 5(a-c), which shows the categorization of the crowd, environmental, and social stress dimensions respectively. The findings reveal the usefulness of the ontology in converting numerical stress indicators into semantic, high-level knowledge that can be explained.

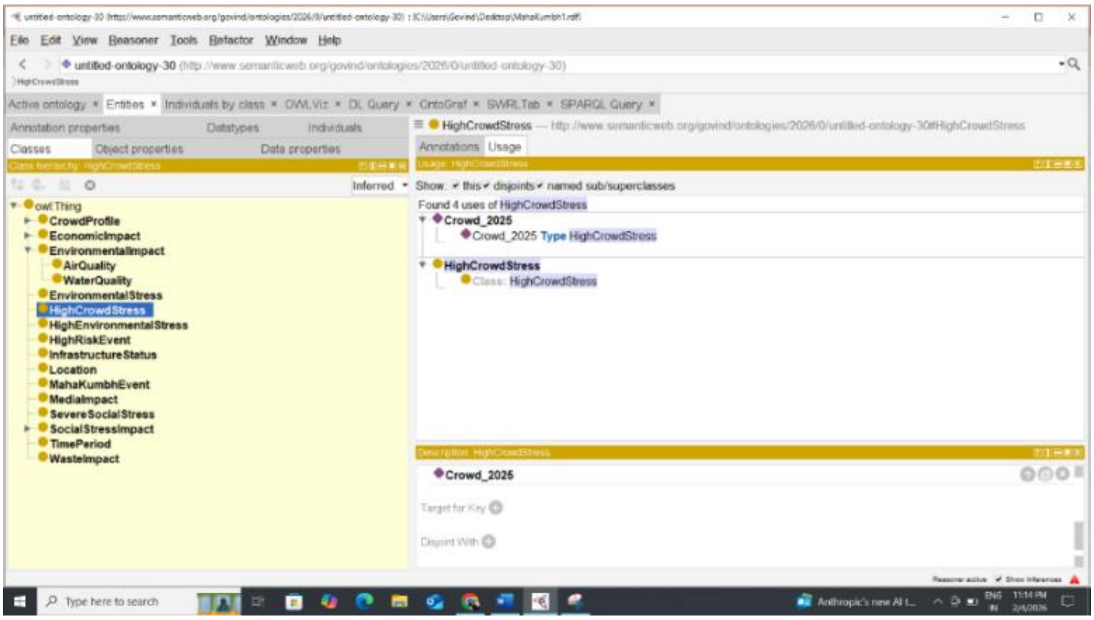

Fig.5(a). Inferred classification of crowd-related stress, where the individual Crowd_2025 is automatically inferred as an instance of HighCrowdStress through semantic reasoning based on crowd intensity indicators.

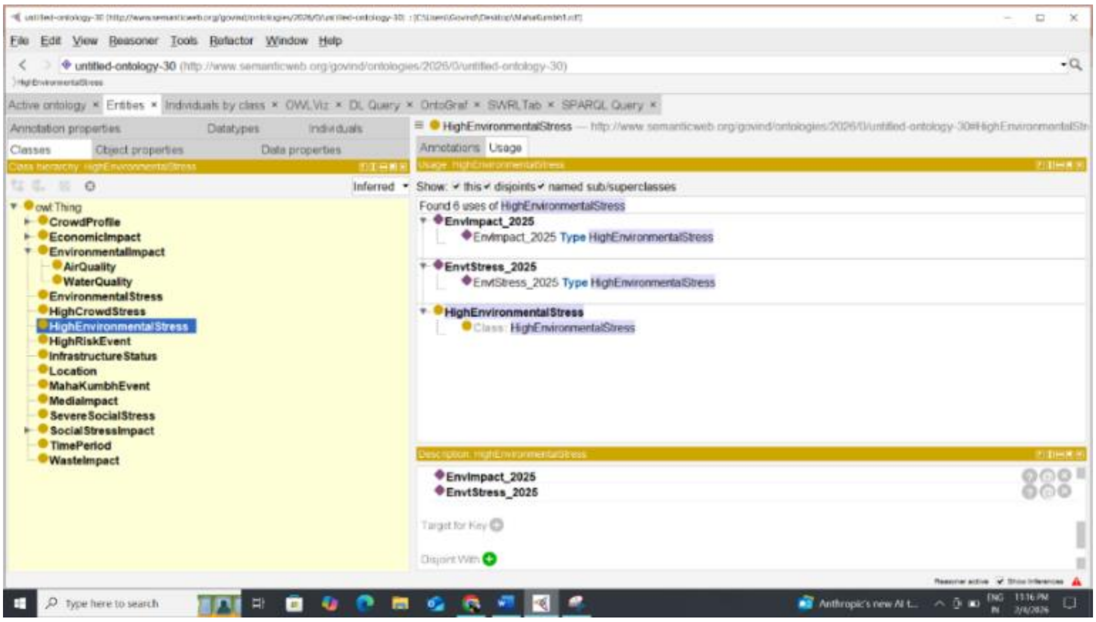

Fig.5(b). Inferred classification of environmental stress, illustrating the automatic reasoning-based assignment of EnvImpact_2025 and EnvtStress_2025 to the HighEnvironmentalStress class due to elevated environmental degradation indicators.

`

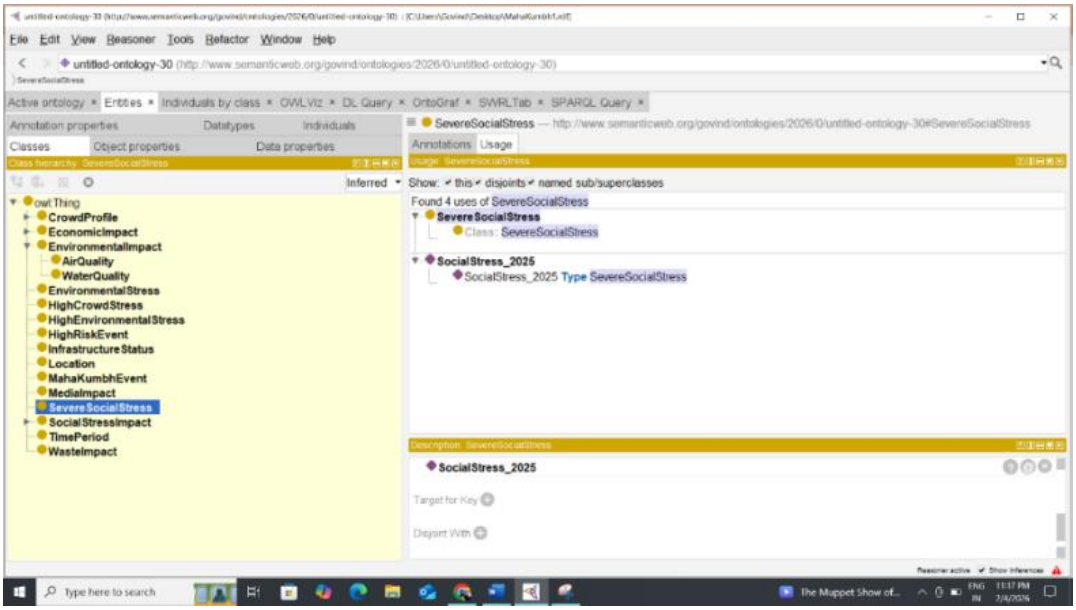

Fig.5(c). Inferred classification of social stress, showing the semantic inference of *SocialStress_2025* as an instance of *SevereSocialStress* based on crime incidence and sentiment-related stress indicators.

4.3 HighRiskEvent Classification

After the individual stress conditions have been inferred, the composite HighRiskEvent rule is then applied. When crowd stress, environmental stress, and severe social stress are present at the same instance of an event, the reasoner automatically tags the event as a HighRiskEvent.

In the case of MahaKumbh 2025, the reasoning engine manages to infer Mahakumbh2025_PrimaryEvent as a HighRiskEvent. This outcome shows the usefulness of the ontology to the requirements of capturing interdependency among heterogeneous stress factors and generating high-level risk knowledge without human intervention.

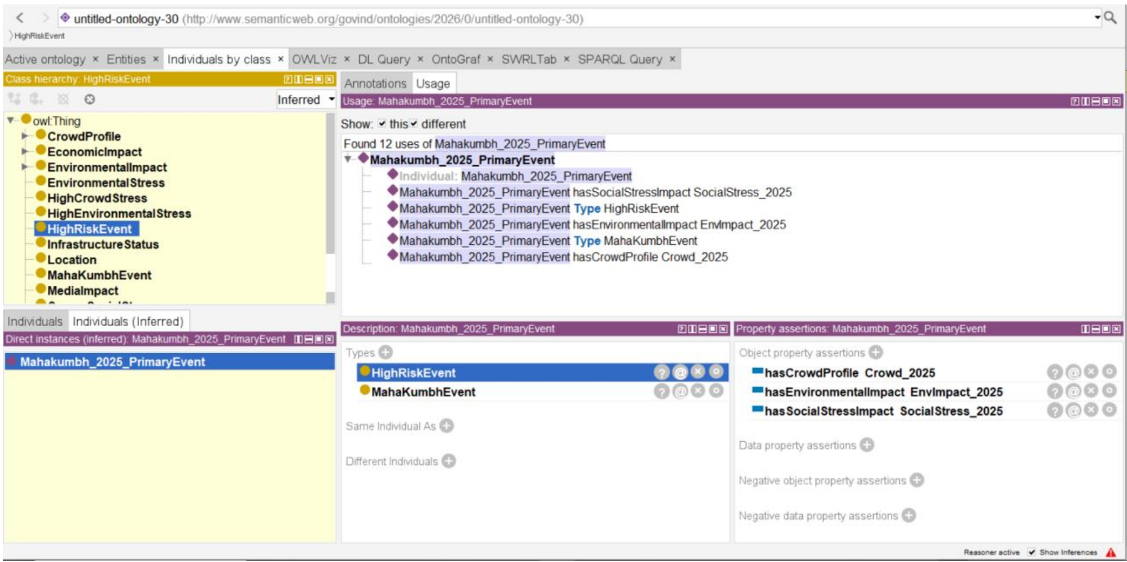

Fig.6. Semantic inference of Mahakumbh_2025_PrimaryEvent as a HighRiskEvent based on inferred high crowd stress, high environmental stress, and severe social stress conditions.

4.3 Explainability of Results

In contrast to black-box predictive models, the suggested ontology-based framework facilitates a transparent and explainable decision-making process with the help of clear semantic rules and logical inferences. Each inferred classification in the ontology can be directly traced back to the underlying conditions of stress and the similar SWRL rules that pushed the inference.

For instance, how Mahakumbh2025PrimaryEvent is defined as a HighRiskEvent is that three independently inferred stress conditions are met at the same time, and these are HighCrowdStress, HighEnvironmentalStress, and SevereSocialStress. The DL Query interface confirms the inferred high-risk status of the event, and the SWRL execution view proves that the assigned status is based on rules that can be executed instead of assigned manually.

This is a rule-based reasoning process that allows the stakeholders to not only know what risk level is inferred, but also why it is inferred. This explainability is especially useful to policymakers and event managers because it aids them in making informed decisions, in addition to making auditing and having confidence in automated risk assessment systems.

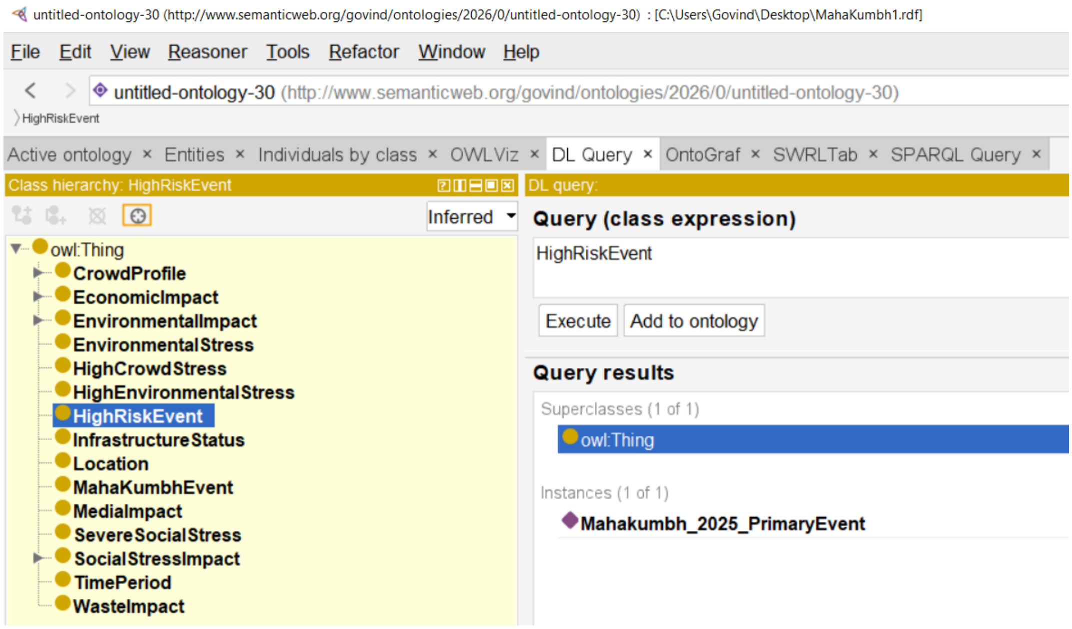

Fig.7a. DL Query result confirming inferred HighRiskEvent instance.

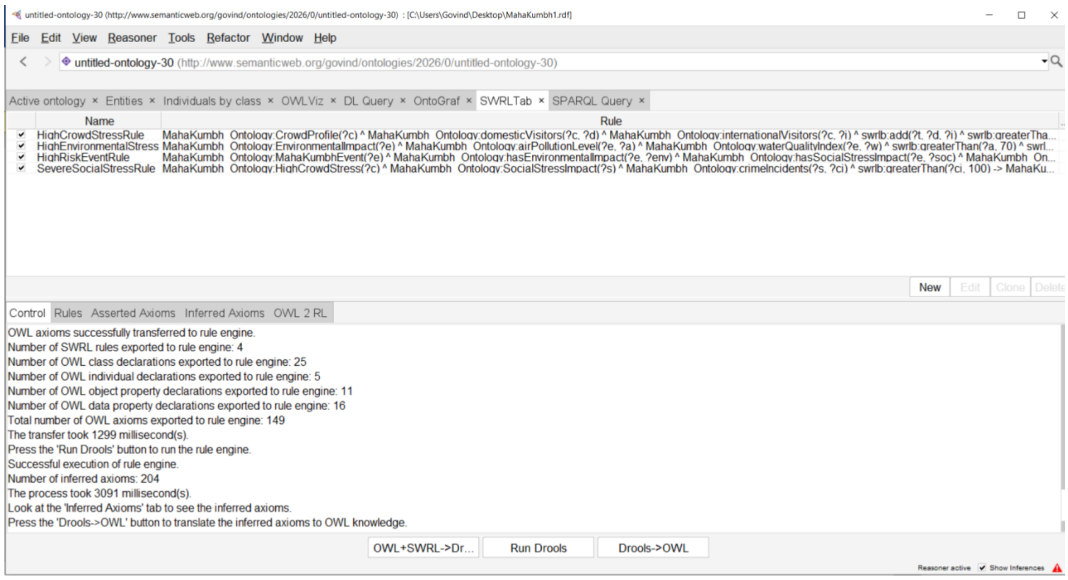

Fig.7b. SWRL rule execution demonstrating rule-based stress and risk classification in Protégé.

4.4 Comparison with Non-Ontology Approaches

The proposed approach was compared with common non-ontology alternatives used for risk assessment. The aim of this comparison is not to claim higher prediction accuracy, because the numerical forecasting is inherited from the prior data-driven study [8], but to clarify how ontology-based reasoning adds semantic representation, explainability, rule traceability, and reusable domain knowledge.

TABLE II

COMPARISON OF THE PROPOSED ONTOLOGY-BASED FRAMEWORK WITH NON-ONTOLOGY RISK ASSESSMENT APPROACHES

Criterion	ML-only predictive model	Simple threshold-based rule system	Proposed ontology-based SWRL reasoning
Main input	Historical and forecasted indicators	Indicator values and if-then thresholds	Forecasted indicators mapped to OWL individuals and properties
Main output	Forecasted value, stress score, or predicted label	Risk label based on coded conditions	Inferred stress classes and HighRiskEvent classification
Semantic relationship modeling	Weak; relationships are usually implicit in features	Limited; conditions are not connected to a formal domain model	Strong; event, crowd, environment, social stress, and risk classes are explicitly represented
Explainability	Limited in black-box models	Moderate, but rules may remain application-specific	High through SWRL antecedents, inferred class hierarchy, and DL Query results
Knowledge reuse	Limited to trained model or specific feature set	Limited to manually coded rules	Higher reuse through OWL classes, properties, and extensible rule base
Manual labeling need	Often required for supervised classification	Risk labels depend on manually written conditions	Reduced because final risk class is inferred from explicit semantic rules
Decision-support value	Prediction-oriented	Condition-checking oriented	Knowledge-driven, explainable, queryable, and auditable

The comparison shows that the contribution of the proposed model lies in semantic integration and explainable inference rather than in replacing numerical forecasting. This makes the framework suitable for decision-support contexts where stakeholders must understand why an event is classified as high risk.

4.5 Summary of Results

The paper shows that an ontology-based framework is effective in modeling, classifying, and reasoning about multidimensional stress situations during large-scale religious events. The semantically instantiated forecasted crowd, environmental, and social indicators for the MahaKumbh 2025 case study were analyzed using SWRL-based reasoning. The ontology inferred individual stress categories and classified the event as a HighRiskEvent without direct manual assignment of the final risk label. The added ontology quality evaluation and comparison with non-ontology approaches further show that the proposed framework improves semantic integration, inference traceability, explainability, and reusable decision-support knowledge.

V. CONCLUSION AND FUTURE WORK

5.1 Conclusion

In this paper, a holistic ontology-based model of stress measurement and risk classification of mega religious events is introduced. The proposed approach will help fill the increasing gap between quantitative prediction and qualitative decision-making by combining the forecasted stress indicators with semantic modeling and SWRL-based reasoning. The MahaKumbh ontology is an ontological modeling of heterogeneous dimensions of stress, encompassing crowd phenomena, environmental deterioration, and social stress, and it allows to infer high-risk situations automatically. The experimental assessment of the MahaKumbh 2025 case study shows that the framework is effective in the inferential assessment of stress categories and the composite risk states in a transparent and explainable format. The findings show how semantic technologies can be used to improve event management, safety management, and policy development.

5.2 Future work

The proposed future work will be to extend the ontology to accommodate real-time data flows of IoT sensors, social media, and surveillance systems. Another direction that can be promising is the incorporation of spatio-temporal reasoning and uncertainty handling mechanisms. Also, the ontology can be further instructed to support the decision-support dashboards and emergency response systems to promote their potential in the context of man-management of large-scale events

REFERENCES

[1] Verma, R. (2025). *Security and Surveillance at Large Religious Gatherings: Crowd Management Strategies for the 2025 MahaKumbhMela. Synergy:*International Journal of Multidisciplinary Studies, 2(3), 47-57.

[2] Singh, S., Shaloo, K., Sinha, S., Kumar, P., &Nath, B. (2025). *Delivering emergency care at the religious mass gathering of India: an amalgamation of technology and spirituality—an institutional experience of managing pilgrims at MahakumbhPrayagraj.* Frontiers in Public Health, 13, 1695241.

[3] Flambard, V., Benamrouz-Vanneste, S., &Karami, A. B. (2023). *Current Trends and Future Prospects on Smart and Sustainable Cities.* Smart Cities.

[4] Gu, K., Liu, H., Liu, J., Yu, X., Shi, T., &Qiao, J. (2022). *Air pollution prediction in mass rallies with a new temporally-weighted sample-based multitask learner*. IEEE Transactions on Instrumentation and Measurement, 71, 1-15.

[5] Stănescu, G., &Oprea, S. V. (2025). *Recent Trends and Insights in Semantic Web and Ontology-Driven Knowledge Representation Across Disciplines Using Topic Modeling*. Electronics, 14(7), 1313.

[6] Yang, X., Sun, X., & Tang, L. (2024). *Multifactor Mathematical Modeling and Analysis of the Impact of Extreme Climate on Geological Disasters.* Water, 16(15), 2211.

[7] De Nicola, A., &Villani, M. L. (2025). *Actionable Semantic Patterns in the Crisis Management Lifecycle: The TERMINUS Ontology.* Smart Cities, 8(5), 179.

[8] Kushwah, G., & Singh, P. (2026). *Forecasting and quantifying event-induced stress during the PrayagrajMahaKumbh using machine learning and a composite event stress index*. International Research Journal of Innovations in Engineering and Technology (IRJIET), 10(1), 113–122. https://doi.org/10.47001/IRJIET/2026.101013.

[9] AdvaitaSen. (2025). MahaKumbh Complete Dataset. Kaggle. Available at: https://www.kaggle.com/datasets/advaitasen/mahakumbh-complete-dataset?resource=download (accessed 6 June 2026).

[10] Jiang, S., & Cao, X. (2025). *Rule Inferring for Engineering Quality Risk Management Based on Ontology in Housing Construction.* Sustainability, 17(4), 1643.

[11] Liu, R., & Fischer, M. (2026). *Ontology-Based Modeling and Reasoning for Construction Field Context. Journal of Management in Engineering, 42*(1), 04025059.