Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00361-y

Original Article

Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives

Aniket Anand [1] , Sanjay K. Tiwari [2] , Mritunjay K. Ranjan [3] , Rajashri J. Rikame [4,]

1. *Department of Computer Science & IT, Magadh University, Bodhgaya, IND*
2. *Department of Mathematics, Magadh University, Bodhgaya, IND*
3. *School of Computer Sciences and Engineering, Sandip University, Nashik, IND*
4. *Department of Computer Science, K. K. Wagh Arts, Commerce, Science and Computer Science College, Nashik, IND*

Received: August 29, 2026 | Review began: September 09, 2026 | Review ended: November 06, 2026 | Published: November 09, 2026

© **Copyright** 2026
This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Forensic interrogation narratives are dynamic and evolve over time, across claims, events, locations, and evidential sources, so statement-level analysis alone is inadequate to support the reliability assessment of claimed consistency. The proposed dynamic epistemic-temporal graph reasoning framework with uncertainty calibration for evidence-consistent forensic interrogation narratives aims to provide a robust framework for generating forensic interrogation narratives that are evidence-consistent and incorporate uncertainty. It combines ModernBERT for contextual language representation, DeBERTa-v3 as a natural language inference model to model support and contradiction, a Dynamic Epistemic-Temporal Heterogeneous Graph Transformer for multi-hop reasoning, and conformal prediction to provide uncertainty-aware decision support. Interrogation discourse becomes structured representations of claims, entities, events, temporal expressions, locations, and relations of evidence. These elements are embedded in a dynamic heterogeneous graph that represents semantic, temporal, and evidential relationships between elements of the evolving story. The proposed Dynamic Epistemic-Temporal Heterogeneous Graph Transformer aims to detect inconsistencies that might not be captured by single- or two-part analysis, especially when evidence is spread across various claims, events, time references, places, and sources of evidence. Natural language inference facilitates semantic relation assessment, and temporal reasoning deals with event order, duration, overlap, and timeline compatibility. Conformal prediction is used to separate reliable consistency judgments from cases that are more ambiguous and require human review. It goes beyond the classification of deception and provides a framework of dynamic, evidence-based narrative reasoning designed to complement, not supplant, expert forensic judgment. The evaluation emphasizes evidence-consistency classification, semantic, temporal, evidential, and multi-hop reasoning, as well as uncertainty calibration. The future goals of cross-dataset generalization and real-world robustness are left for future work.

Categories: Cognitive Computing, Machine Learning (ML), Natural Language Processing (NLP)

Keywords: forensic interrogation, epistemic–temporal graph, evidence consistency, dynamic graph reasoning, heterogeneous graph transformer, natural language inference, uncertainty calibration, conformal prediction, multi-hop reasoning, forensic artificial intelligence

Introduction

Forensic interrogation stories vary in statements, events, places, time markers, and sources of evidence. When contradictions rely on temporal, evidential, or multi-hop relationships spread across different parts of the narrative, statement-level analysis might not be adequate [1]. Thus, statements need to be considered in the context of the entire

interrogation process, not individually. Artificial intelligence (AI) and natural language processing (NLP) have recently made strides in the field of understanding context and finding contradictions, which are crucial in investigative discourse [2]. However, most approaches in use today are static or pairwise and do not provide for capturing multiple and evolving evidence structures, temporal dependencies, or inconsistencies that only emerge over multiple interconnected claims. The solution is graph-based reasoning, which can be used to model claims, persons, events, locations, temporal expressions, and evidence as connected objects [3]. Dynamic graph updates support the update of entities with new information, which can affect how entities are related to one another, and multi-hop reasoning is used to detect semantic, temporal, and evidential inconsistencies that span the narrative. In this study, a dynamic epistemic-temporal graph reasoning framework with uncertainty calibration for evidence-consistent forensic interrogation narratives is proposed [4]. The framework integrates contextual language representation, natural language inference (NLI), dynamic heterogeneous graph modeling, multi-hop reasoning, and uncertainty-aware prediction, and evaluates the consistency of evidence in evolving interrogation narratives [5]. It is designed to be a decision-making tool that identifies consistent and inconsistent data but does not indicate guilt, innocence, or "truth" [6], as shown in Figure *1*.

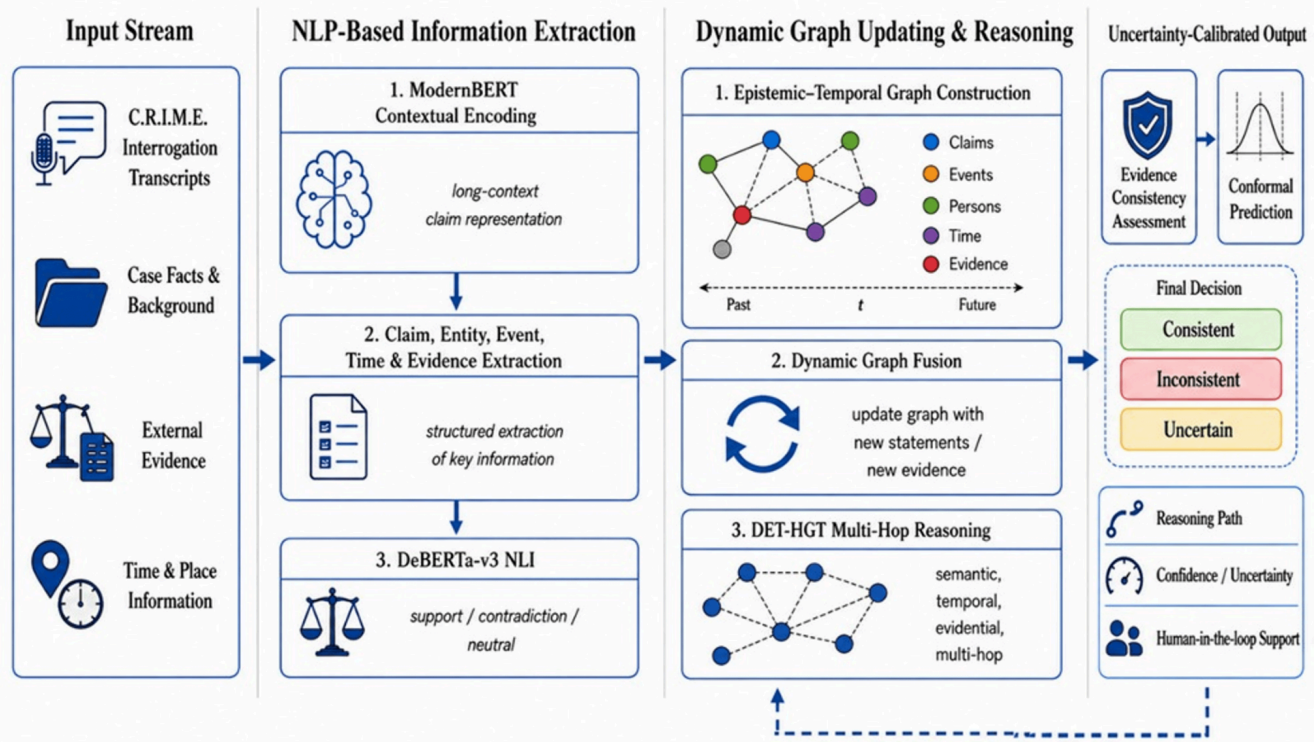


FIGURE 1: Dynamic forensic interrogation graph reasoning architecture overview

BERT, Bidirectional Encoder Representations from Transformers; DET-HGT, Dynamic Epistemic-Temporal Heterogeneous Graph Transformer; NLP, Natural Language Processing; NLI, Natural Language Inference

Significance of the study

This study makes a contribution to forensic interrogation analysis by moving away from statement analysis to a dynamic, evidence-based narrative reasoning approach [7]. The proposed framework illustrates the evolution of claims and evidence during an interrogation and analyzes the consistency of claims and evidence in time, space, semantics, and in the entire narrative. It can detect inconsistencies that arise from multiple interconnected claims, events, and evidence sources, known as multi-hop reasoning. Conformal prediction is used to control predictive uncertainty. Uncertainty is a state of ambiguous prediction sets, which are not to be used as definitive forensic conclusions and are to be referred for human assessment. All of these capabilities offer a systematic and traceable decision-support process in forensic analysis while maintaining human expertise in situations where a high degree of judgment is required.

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

Research objectives

This study aims to design and verify a Dynamic Epistemic-Temporal Heterogeneous Graph Transformer (DET-HGT) for semantic, temporal, evidential, and multi-hop evidence-consistency reasoning of a forensic interrogation narrative; to combine context modeling with NLI to discover the relationships between supporting, contradicting, and neutral statements; and to design a multi-hop graph reasoning mechanism to identify semantic, temporal, and evidential inconsistency relations across the evolving narrative. The study also seeks to embed uncertainty calibration to identify reliable consistency judgments and ambiguous ones that demand human intervention, and to compare the proposed framework with traditional text-based and graph-based methods by applying consistency detection, temporal reasoning, robustness, generalization, and calibration measures.

Literature review

This has led to recent interest in forensic NLP, temporal reasoning, graph-based representation, and uncertainty estimation to enhance the analysis of complex narrative data, as shown in Table *1*. These methods, however, are mostly fragmented, and there is a need for an integrated method that models relationships among semantics, time, evidence, and multi-hop relationships together.

Citation	Domain	Insights	Results	Advantages	Limitations	Significance of the Proposed Work
[8]	Forensic NLP	Applied NLI and contextual features for suspect-statement analysis.	Pairwise NLI showed limited capability for contradiction discovery.	Demonstrates the applicability of NLI to forensic narratives.	Limited by pairwise reasoning and synthetic data.	Supports the need for dynamic, multi-hop narrative reasoning.
[9]	Temporal Graph Generation	Modeled event–time relations through global graph generation.	Improved temporal graph performance and zero-shot generalization.	Captures global temporal structure across events.	Not specifically designed for forensic evidence analysis.	Motivates temporal graph modeling of interrogation events.
[10]	Temporal Relation Extraction	Used hierarchical reasoning to infer event–time relationships.	Improved performance on standard temporal-relation datasets.	Enables structured temporal inference.	Lacks evidential and forensic reasoning mechanisms.	Supports explicit temporal reasoning within the proposed narrative graph.
[11]	Global Temporal Reasoning	Constructed document-level temporal graphs with consistency constraints.	Improved global temporal consistency.	Reduces inconsistencies arising from isolated pairwise predictions.	Primarily focuses on temporal relationships.	Supports global narrative-level graph reasoning and consistency checking.

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

[12]	Temporal Consistency	Investigated LLM-assisted consistency mechanisms for temporal classifiers.	Improved consistency while revealing accuracy–consistency trade-offs.	Highlights the importance of confidence-aware reasoning.	Does not incorporate evidence-grounded forensic reasoning.	Supports uncertainty-aware graph reasoning and confidence calibration.

TABLE 1: Comparative analysis of existing approaches and identified research gaps

LLM, Large Language Model; NLI, Natural Language Inference; NLP, Natural Language Processing

A generative forecasting approach for temporal knowledge graphs leveraging large language models, with a focus on modeling evolving temporal relationships and graph-based reasoning, is proposed by GenTKG [13]. It is helpful for temporal knowledge representation, but its main purpose is prediction, not evidence-consistency analysis in forensic interrogation narratives. In NLP, uncertainty quantification is a model-independent problem that can be addressed using conformal prediction [14]. This is important for the proposed study, as it will help identify cases that are evidence-consistent (confident) and those that are not evidence-consistent (uncertain) and require human review. Structural vulnerability has also been addressed in forensic anthropology to explore how social and contextual issues may impact forensic interpretation [15]. This work is not about computational graph reasoning, but it stresses the need for contextual and responsible interpretation during forensic analysis in the context of the proposed framework that relies on human-centered decision support. In general, the previous research works focus on NLI, temporal reasoning, graph modeling, and uncertainty separately. The proposed framework provides an integration of these capabilities for dynamic, multi-hop, evidence-consistent reasoning over forensic interrogation narratives.

Problem statement

Current methods for forensic interrogation analysis are primarily based on statement classification or pairwise contradiction detection and cannot effectively capture evolving claims, temporal dependencies, multi-hop inconsistencies, and relationships with supporting evidence [16]. They tend to be a problem with conventional statement-level approaches, which often fail to detect contradictions that only occur when a set of statements, events, locations, and sources of evidence are considered as a whole [17]. Furthermore, conventional models can yield overly confident conclusions even with incomplete and ambiguous data. Hence, there is a need for a dynamic epistemic-temporal graph reasoning framework that allows for the modeling of an evolving forensic narrative, multi-hop reasoning to detect inconsistencies across semantic, temporal, and evidential dimensions, and uncertainty-aware evidence-consistency judgments for dependable forensic decision support [18].

Research Gaps

The study of forensic contradiction detection, temporal reasoning, graph modeling, and uncertainty estimation is largely considered an independent problem in existing studies. They offer partial assistance for the concurrent examination of emerging claims, temporal dependencies, evidence relations, and indirect multi-hop inconsistencies in full interrogation narratives. Forensic scientific research has recently focused on ensuring evidential integrity and traceability in computational workflows, which, in turn, has led to the need for evidence-aware reasoning approaches for complex forensic narratives [19].

Materials And Methods

The proposed approach is a structured pipeline that is used to analyze forensic interrogation narratives using dynamic epistemic-temporal graph reasoning [20]. It starts with data collection and preprocessing of the interrogation, then moves on to encoding the language in the text into its context, identifying the semantic relationships between concepts,

and extracting claims, entities, events, temporal expressions, and evidence. All these elements are represented in a dynamic heterogeneous graph in which multi-hop reasoning is applied to detect semantic, temporal, and evidential inconsistencies. Finally, conformal prediction is used for uncertainty calibration and to provide reliable evidence-consistency judgments for forensic decision support, as shown in Figure *2*. All figures presented are original and were created by the authors specifically for this study; no figure has been reproduced or adapted from an external source.

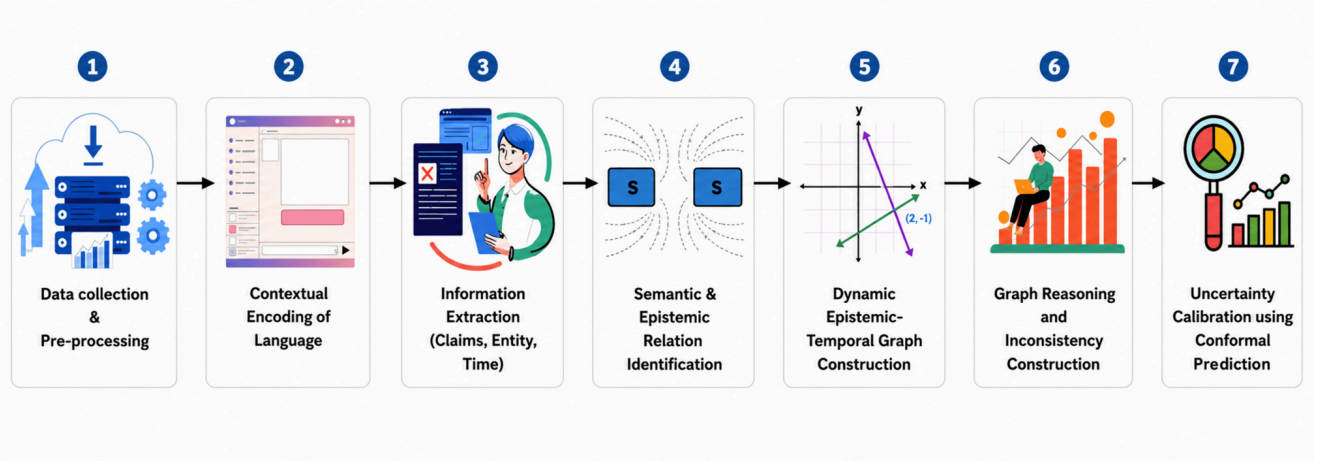


FIGURE 2: Seven-stage methodology for uncertainty-calibrated forensic graph reasoning

The C.R.I.M.E. (Corpus of Recorded Investigative, Media, and Evidence-based Proceedings) corpus is used as the main data source, as it includes real-world criminal justice discourse, such as investigative interviews and police interrogation material [21]. This C.R.I.M.E. corpus was downloaded from Harvard Dataverse (https://doi.org/10.7910/DVN/MLMB6E). Three subsets are chosen for this investigation: Full-Length Criminal Interrogations, Interrogation Raw, and Law & Crime Interrogations, containing approximately 173 interrogation transcripts and 137.41 hours of recorded material, as shown in Table *2*.

In this study, a secondary computational analysis of publicly available, anonymized forensic interrogation transcripts was performed, but no contact with the participants was involved. Forensic scientific research has recently focused on ensuring evidential integrity and traceability in computational workflows, which, in turn, has led to the need for evidence-aware reasoning approaches for complex forensic narratives.

The interrogation-focused subsets of the corpus were used for model development and evaluation, and the original dataset was accessed directly from its official public repository. All experimental codes, preprocessing scripts, split files of datasets, graph construction procedures, model configurations, and scripts for model evaluation can be downloaded from [https://drive.google.com/file/d/1lF4YaNvnBFUO0Vh1tdSahxBZTrmn6S4B/view?usp=sharing].

Dataset	Selected Subsets	Total Transcripts	Duration	Data Type	Role in Study
C.R.I.M.E.	Full-Length Criminal Interrogations; Interrogation Raw; Law & Crime Interrogations	173	137.41 h	Forensic interrogation transcripts	Extraction of claims, entities, events, temporal expressions, and evidence relations

TABLE 2: C.R.I.M.E. dataset configuration for forensic interrogation analysis

C.R.I.M.E., Corpus of Recorded Investigative, Media, and Evidence-based Proceedings

The 173 selected interrogation transcripts were partitioned at the interrogation level to prevent information leakage between related statements. A total of 121 transcripts (69.94%) were assigned to the training set, 26 transcripts (15.03%) to the validation/calibration set, and 26 transcripts (15.03%) to the independent test set. The same partition was used across all evaluated models to ensure a fair and consistent comparison. To ensure a fair comparison between models and prevent information leakage, identical case-level data partitions were utilized in all evaluated models, with the same preprocessing procedures applied. An independent test set was employed to ensure that the test set remained the same for all models. Transcripts focus on a story to be developed, as opposed to a set of individual statements. The selected data are then broken down into segments corresponding to the interviewers' and interviewees' responses and then prepared for claims, persons, events, locations, temporal expressions, and evidential references to be extracted. The claims, persons, events, locations, temporal expressions, and evidence are represented as heterogeneous graph nodes, and the relationships between them are represented as typed edges: semantic, temporal, and evidential. The extracted units and relations are passed through a filter to remove extracted units and relations with low confidence, so as to avoid propagation of errors downstream. The dataset is split at the interrogation/case level, rather than at the individual-statement level, into training, validation, and testing partitions to prevent information leakage.

The C.R.I.M.E. interrogation transcripts are processed using a discourse-preserving pipeline that filters out automatic speech recognition noise, redundant parts of speech, corrupt parts of speech, and other non-speech elements while preserving speaker roles, negation, temporal expressions, discourse order, and timestamps [22]. Each interrogation is anonymized and presented as a series of utterances with timestamps for context-aware forensic analysis. The preprocessing transformation is represented in Eq. (1).

$$\pi^* = F(T, A, S, P, t_q) \tag{1}$$

Here, T is the raw transcript, A is the automatic speech recognition filtering, S is the speaker role segmentation, P is the privacy-preserving anonymization, and t_q is the minimum quality threshold. The retained utterances are represented as in Eq. (2).

$$z_i = [u_i \,\|\, e_{s_i} \,\|\, e_{t_i} \,\|\, e_{c_i}] \tag{2}$$

In this case, u_i is the utterance embedding, e_{s_i} is the speaker embedding, e_{t_i} is the temporal embedding, and e_{c_i} is the local discourse-context embedding. In the symbol, $\|$ represents feature concatenation. To encourage high-quality segments and to punish poor-quality segments, an adaptive quality score is calculated, following Eq. (3).

$$Q_i = \sigma\left(\alpha_1 C_i + \alpha_2 L_i + \alpha_3 R_i + \alpha_4 T_i\right) \tag{3}$$

C_i is the completeness of the transcripts, L_i is the linguistic validity, R_i is the reliability of the speaker's role, T_i is the reliability of the timestamp, and σ(·) maps the score to the range (0, 1). Segments for which $Q_i < \tau_q$ are either removed or marked for manual inspection. Contextual windows are built as in Eq. (4).

$$W_i = \{u_{i-k}, \ldots, u_i, \ldots, u_{i+k}\} \tag{4}$$

The contextual encoder tokenizer is then used to transform them into model-ready sequences of subwords, as shown in Figure *3*. These preprocessing steps preserve information for later claim, event, temporal, evidential, and contradiction reasoning [23].

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

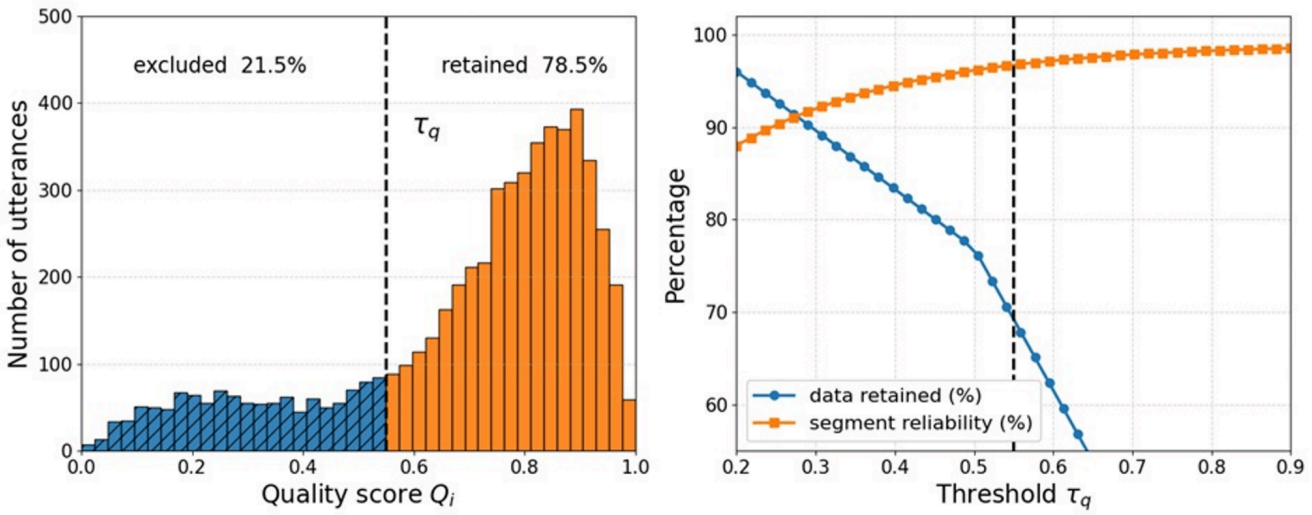


FIGURE 3: Preprocessing quality distribution and threshold sensitivity analysis

ModernBERT's contextual representations, with task-specific extraction heads, are used on the preprocessed interrogation turns to provide structured forensic information [24]. The framework detects claims, persons, locations, events, temporal expressions, and evidence references for each of the utterances and retains the speaker and the context of the discourse. Each extracted claim is represented in Eq. (5).

$$C_i = \{s_i, p_i, o_i, e_i, l_i, t_i, v_i\} \tag{5}$$

Here, s_i denotes the subject, p_i denotes the predicate or action, o_i denotes the object, e_i denotes the associated event, l_i denotes the location, t_i denotes the normalised temporal reference, and v_i denotes the supporting evidential reference. The joint contextual representation is defined in Eq. (6).

$$H_i = \mathrm{ModernBERT}(W_i) \tag{6}$$

where W_i denotes the contextual interrogation window, and H_i denotes the token-level contextual embeddings. Extraction probabilities are obtained, as shown in Eq. (7).

$$P(y_j \mid H_i) = \mathrm{Softmax}\left(W_j H_i + b_j\right) \tag{7}$$

where y_j represents the entity, event, temporal, claim, or evidence category, as shown in Figure 4.

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

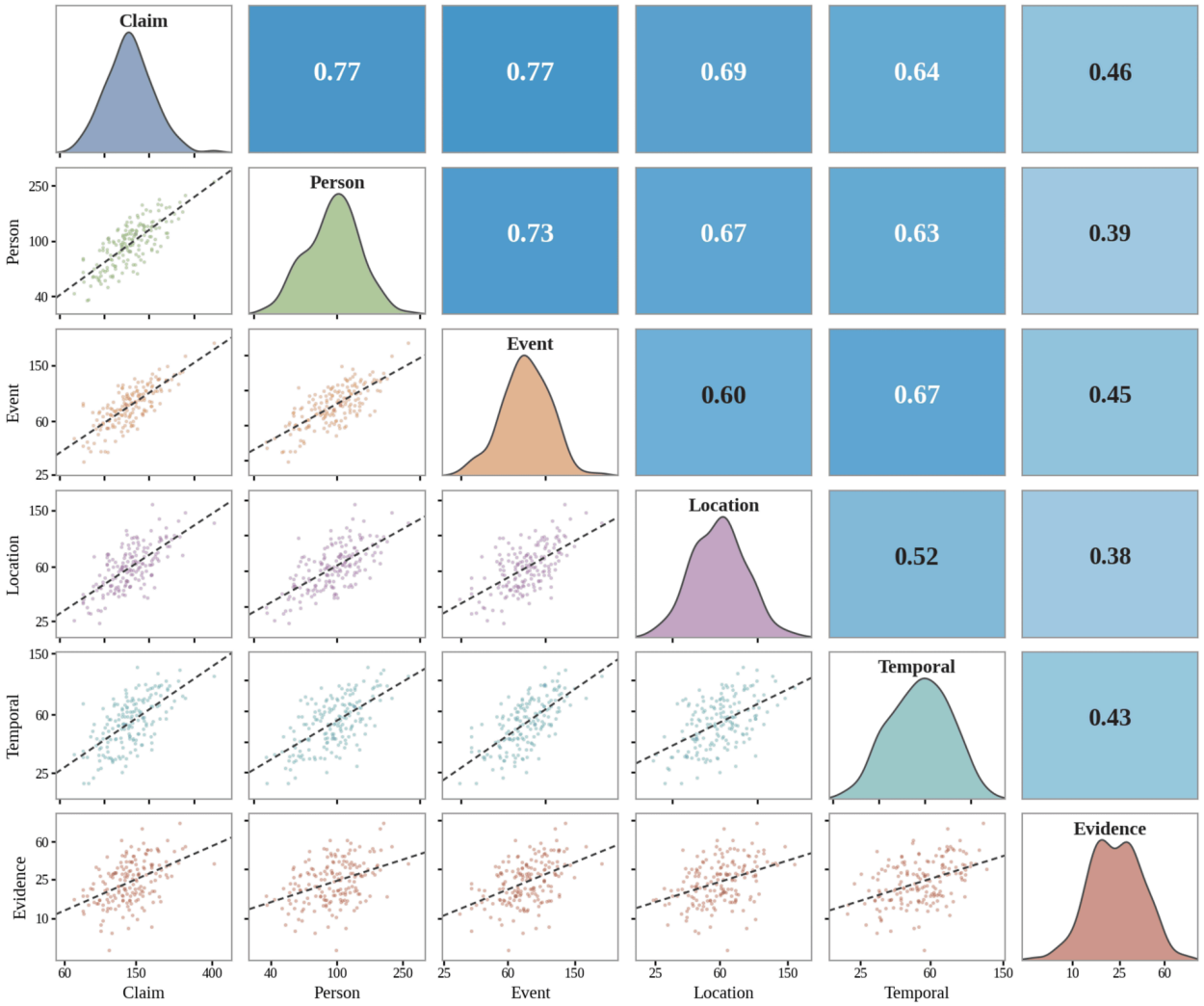


FIGURE 4: Distribution of extracted forensic information units across transcript categories

The temporal expressions are normalized to a common time frame, and the evidence references are linked to the evidence claims. The resulting structured units are used as node and relation inputs for the subsequent semantic relation analysis and dynamic epistemic-temporal graph construction [25].

Using ModernBERT, the units extracted from the interrogation are encoded to account for long-range semantic and discourse relations in changing forensic narratives. Speaker, temporal, and claim-level information required for subsequent relation and graph reasoning is preserved, while each contextual window W_i is converted into token-level representations, as shown in Figure 5.

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

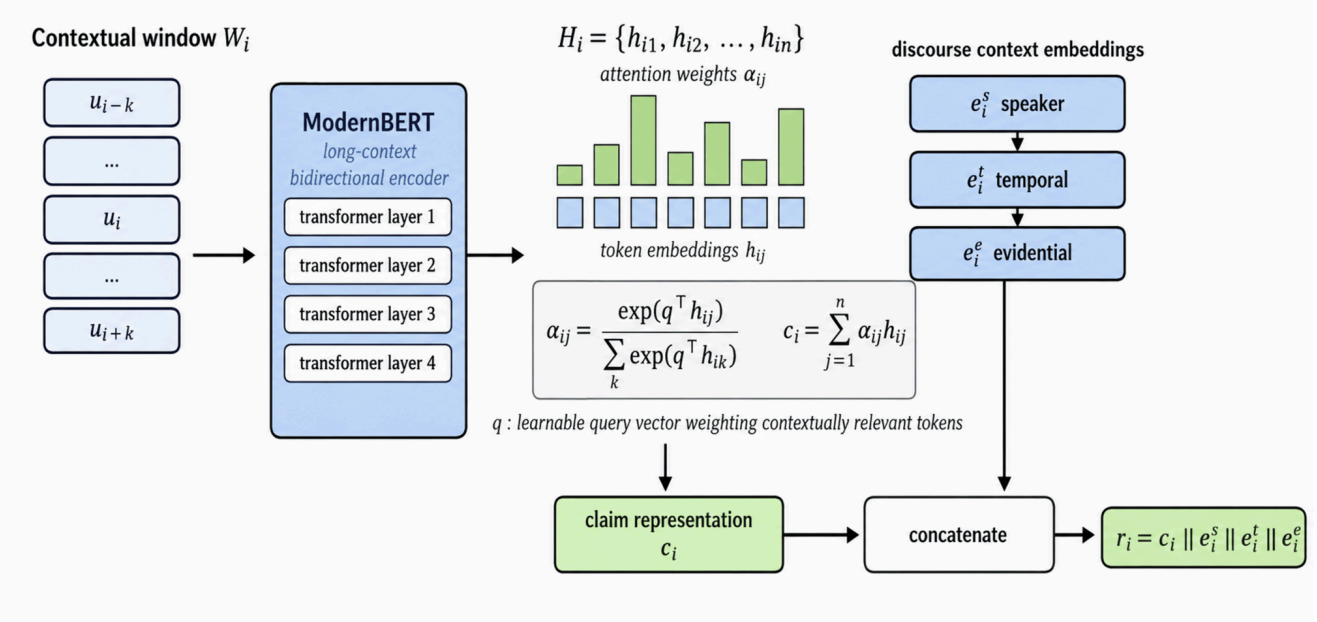


FIGURE 5: Contextual encoding and discourse-aware claim representation

BERT, Bidirectional Encoder Representations from Transformers

The contextual representation is defined as shown in Eq. (8).

$$H_i = \text{ModernBERT}(W_i) = \{h_{i1}, h_{i2}, \ldots, h_{in}\} \tag{8}$$

where H_i is the contextual embedding of the j-th token in interrogation window W_i. The claim-level representation is obtained through attention-weighted aggregation, as shown in Eq. (9).

$$c_i = \sum_{j=1}^{n} \alpha_{ij} h_{ij} \tag{9}$$

where $\alpha_{ij} = \frac{\exp(q^T h_{ij})}{\sum_{k=1}^{n} \exp(q^T h_{ik})}$ and q denotes a learnable query vector that assigns greater importance to contextually relevant tokens. To incorporate discourse and temporal information, the final representation is formed as shown in Eq. (10).

$$r_i = c_i e_{s_i} \| e_{t_i} \| e_{v_i} \tag{10}$$

Here, s_i, t_i, and v_i are the speaker, temporal, and evidential-context embeddings, respectively, as shown in Figure 6. The obtained representation is then passed to the DeBERTa-v3 NLI module for semantic support, contradiction, and neutral relation detection [26].

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

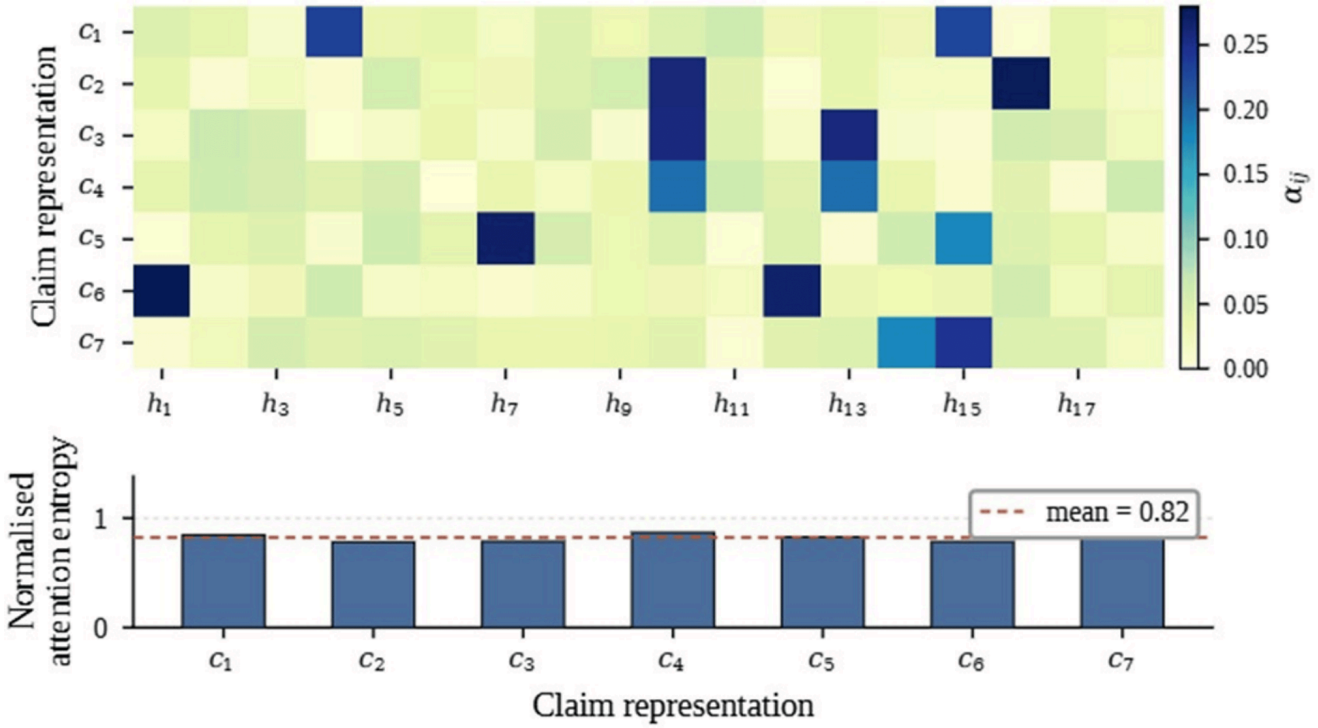


FIGURE 6: Claim-level attention distribution and normalized entropy

We use DeBERTa-v3 NLI to identify semantic relationships between extracted claims [27]. All candidate claims are encoded as pairs of claims (C_i, C_j) to indicate whether or not the second claim supports or contradicts the first claim or is neutral to it, while maintaining surrounding discourse context. The pairwise contextual representation is given in Eq. (11).

$$g_{ij} = \mathbf{DeBERTa\text{-}v3}\left([C_i \parallel C_j \parallel w_{ij}]\right) \tag{11}$$

where w_{ij} is the common discourse context of the two claims. The probability of the semantic relation is calculated as shown in Eq. (12).

$$P(r_{ij} \mid C_i, C_j) = \mathbf{Softmax}\left(W_r g_{ij} + b_r\right) \tag{12}$$

where $r_{ij} \in \{\mathbf{Support}, \mathbf{Contradiction}, \mathbf{Neutral}\}$, and the predicted relation is $r_{ij}^{*} = \mathbf{argmax}_{r_{ij}} P(r_{ij} \mid C_i, C_j)$. To prevent weak relations from entering the graph, an adaptive confidence criterion is applied in Eq. (13).

$$e_{ij} = r_{ij} \text{ if } \max P(r_{ij} \mid C_i, C_j) \geq \tau_r \tag{13}$$

The semantic relations and confidence scores are then used to construct typed, weighted edges for the use of multi-hop consistency reasoning in the construction of the dynamic epistemic-temporal graph [28], as shown in Figure *7* [28].

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

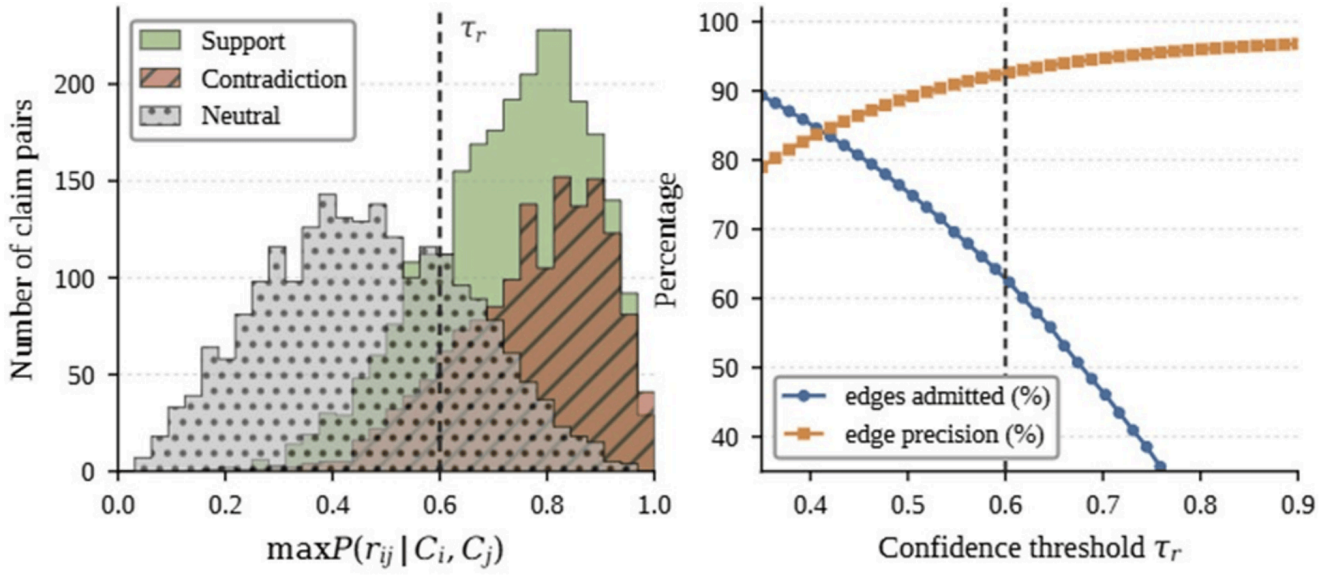


FIGURE 7: Confidence-threshold analysis of forensic semantic relation classification

The forensic information is extracted and represented in a dynamic heterogeneous epistemic-temporal graph, which provides a graphical representation of claims, persons, events, locations, temporal expressions, and evidence, along with their semantic, temporal, and evidential relationships. The graph evolves over time as new interrogation statements and evidence are added. The graph at interrogation step *t* is represented in Eq. (14).

$$G_t = (V_t, E_t, R_t, X_t) \tag{14}$$

Here, V_t represents heterogeneous nodes, E_t represents relational edges, R_t represents the kind of relation, and X_t represents the contextual node features extracted from ModernBERT. The set of nodes is given in Eq. (15).

$$V_t = \{V_{\text{claim}}, V_{\text{person}}, V_{\text{event}}, V_{\text{location}}, V_{\text{time}}, V_{\text{evidence}}\} \tag{15}$$

with relational edges of $R = \{\text{supports}, \text{contradicts}, \text{before}, \text{after}, \text{overlaps}, \text{occursAt}, \text{refersTo}, \text{evidenceFor}\}$

Each edge is assigned a confidence-aware weight of $w_{ij} = \lambda_1 S_{ij} + \lambda_2 T_{ij} + \lambda_3 E_{ij}$

The semantic relation confidence from DeBERTa-v3 NLI is denoted by S_{ij}, the compatibility of the temporal information is captured as T_{ij}, and the evidential association strength is denoted by E_{ij}. The graph is dynamically updated as shown in Eq. (16).

$$G_{t+1} = \text{Update}\,(G_t, \Delta V_{t+1}, \Delta E_{t+1}) \tag{16}$$

This allows new claims/evidence to alter existing relationships. The resulting graph maintains semantic, temporal, evidential, and epistemic dependencies as well as provides structured input to the proposed DET-HGT multi-hop reasoning module, as shown in Table 3 and Figure 8.

Graph Component	Representation	Examples	Role
Nodes	V_t	Claim, Person, Event, Location, Time, Evidence	Represent forensic information
Semantic Edges	S_{ij}	Supports, Contradicts, RefersTo	Capture claim relationships

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

Temporal Edges	T_{ij}	Before, After, Overlaps, OccursAt	Model timeline dependencies
Evidence Edges	E_{ij}	EvidenceFor, EvidenceAgainst	Connect claims with evidence
Node Features	X_t	ModernBERT embeddings	Encode contextual information
Edge Weight	W_{ij}	Semantic + Temporal + Evidence confidence	Quantify relation strength
Dynamic Update	$G_t \rightarrow G_{t+1}$	New claims/evidence	Evolve graph during interrogation
Graph Output	G_t	Structured forensic narrative	Input to DET-HGT reasoning

TABLE 3: Graph components and roles in forensic reasoning

DET-HGT, Dynamic Epistemic-Temporal Heterogeneous Graph Transformer

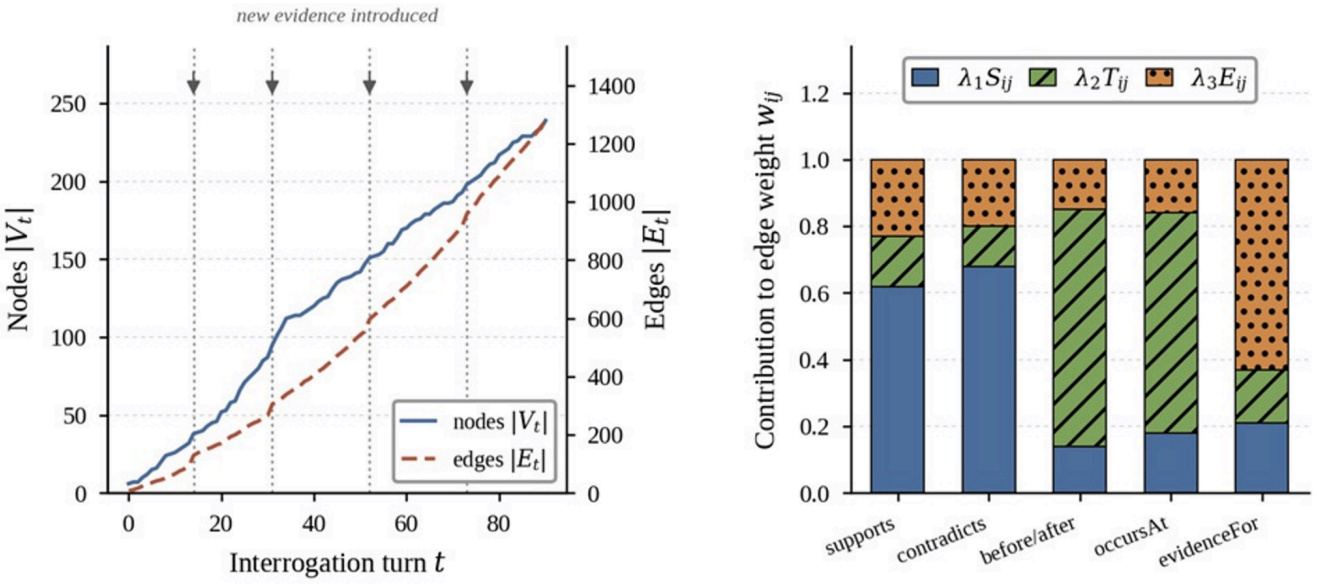


FIGURE 8: Dynamic graph evolution and relation-weight contribution analysis

The resulting epistemic-temporal graph is then processed by the proposed DET-HGT to learn direct and multi-hop interdependencies between claims, events, time, location, and evidence [29]. Inconsistencies that are only uncovered by interwoven elements of the story can be detected by DET-HGT, while they are not visible in pairwise reasoning. The relation-aware aggregation function for node v_i at layer l is defined in Eq. (17).

$$h_i^{(l+1)} = \sigma\left(\sum_r \sum_{j \in N_r(i)} \alpha_{ij}^r W_r h_j^{(l)}\right) \tag{17}$$

where W_r is the relation-specific transformation matrix, and α_{ij}^r represents the attention paid to adjacent node *j* under relation *r*. The attention mechanism is trained with the temporal information, as shown in Eq. (18).

$$\alpha_{ij}^{r} = \mathbf{Softmax}\left(\frac{(Q_r h_i)^T (K_r h_j + \Phi(\Delta t_{ij}))}{\sqrt{d}}\right) \tag{18}$$

where $\Phi(\Delta t_{ij})$ represents the temporal encoding between connected events. For a multi-hop path $P = \{v_1, \ldots, v_k\}$, an inconsistency score is computed as shown in Eq. (19).

$$I(P) = \sigma\left(W_p \cdot \mathbf{Pool}(h_1, \ldots, h_k) + b_p\right) \tag{19}$$

Higher $I(P)$ represents the greater semantic, temporal, or evidential incongruity among the reasoning path. The node and path representations are then sent to the evidence-consistency assessment module to generate the narrative-level decisions.

The DET-HGT representations are employed to evaluate the consistency of every forensic claim based on both semantic and temporal as well as evidential and multi-hop relational evidence. An evidence-consistency score for each claim C_i is calculated using the aggregated graph representation z_i, as shown in Eq. (20).

$$ECS_i = \sigma\left(W_E z_i + b_E\right) \tag{20}$$

The higher the score (from 0 to 1), the more consistent it is with the surrounding text and evidence. To explicitly incorporate different inconsistency sources, the claim-level risk score is defined, as shown in Eq. (21).

$$IR_i = \lambda_s I_{si} + \lambda_t I_{ti} + \lambda_e I_{ei} + \lambda_p I_{pi} \tag{21}$$

where I_{si}, I_{ti}, I_{ei}, and I_{pi} represent semantic, temporal, evidential, and multi-hop inconsistency scores, respectively, with λs + λt + λe + λp = 1, as shown in Figure 9.

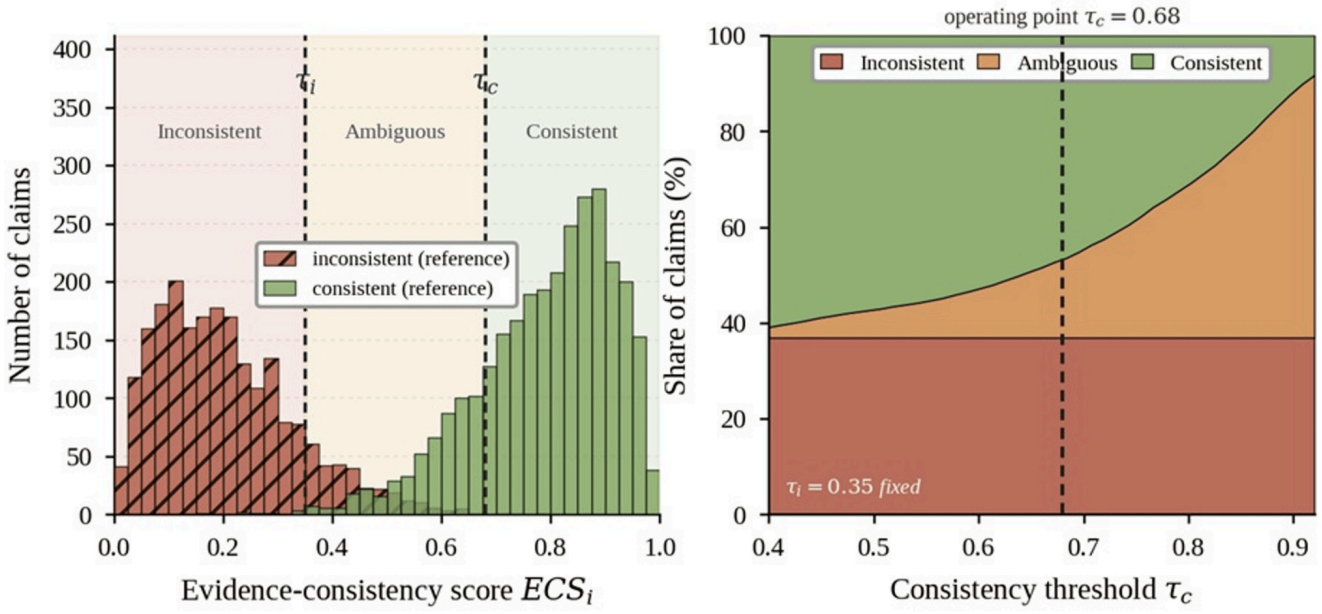


FIGURE 9: Evidence-consistency score distribution and threshold sensitivity analysis

The preliminary consistency decision is obtained, as shown in Eq. (22).

$$Y_i = \mathbf{Consistent\ if}\ ECS_i \geq \tau_c;\ \mathbf{Inconsistent\ if}\ ECS_i \leq \tau_i;\ \mathbf{Ambiguous\ otherwise} \tag{22}$$

The consistency score and decision are then passed to the conformal prediction module to learn the uncertainties and generate reliable forensic output.

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

The output probabilities of DET-HGT are used as input to a split conformal prediction algorithm to deal with uncertainty in evidence-consistency decisions [30]. Each class's reliability is estimated using a held-out calibration set before making the final forensic decision. The nonconformity score for sample *i* is defined in Eq. (23).

$$s_i = 1 - p_i(y_i) \tag{23}$$

Here, $p_i(y_i)$ is the predicted probability assigned to the true class y_i. The conformal threshold is found as in Eq. (24).

$$\hat{q}_\alpha = \text{Quantile}_{\frac{\lceil (n_{\text{cal}}+1)(1-\alpha) \rceil}{n_{\text{cal}}}} \left(\{ s_1, s_2, \ldots, s_n \} \right) \tag{24}$$

The prediction set for a new claim x is defined in Eq. (25).

$$\Gamma_\alpha(x) = \{ y : 1 - p(y \mid x) \leq \hat{q}_\alpha \} \tag{25}$$

If a single label is in the prediction set, it means that the decision is reliable, while a prediction set with multiple labels is considered Uncertain and is referred for human examination. The exchangeability assumption leads to the coverage property being the target of the conformal procedure, as shown in Eq. (26).

$$P\left(Y \in \Gamma_\alpha(X)\right) \geq 1 - \alpha \tag{26}$$

The final output of the system is Consistent, Inconsistent, or Uncertain. This reduces the overconfidence of forensic conclusions and provides the advantage of uncertainty control.

The last step is to take the calibrated reasoning and convert the results into an interpretable forensic decision-support output [31]. The system returns the evidence-consistency status for every claim and every part of a narrative, the type of inconsistency found, the reasoning path followed to reach the conclusion, and the level of confidence in the conclusion. The final result is shown in Eq. (27).

$$O_i = \{ Y_i, ECS_i, U_i, P_i \} \tag{27}$$

where Y_i denotes the final decision, ECS_i denotes the evidence-consistency score, U_i denotes the calibrated uncertainty, and P_i denotes the supporting graph-based reasoning path.

The decision categories are defined as follows: consistent = there is sufficient evidence and narrative relations; inconsistent = a significant semantic, temporal, or evidential conflict is found; uncertain = not enough evidence to reach a reliable conclusion. The framework includes traceable reasoning paths, links to evidence, consistency scores, and uncertainty estimates to allow human experts to examine the logic behind each automated assessment. The framework is therefore a human-in-the-loop decision-support system and not an automatic truth/rejection mechanism or a determination of guilt.

System architecture

The proposed system architecture uses an NLP, graph-reasoning, and uncertainty-aware pipeline to process forensic interrogation narratives [32]. The first step in preprocessing C.R.I.M.E. transcripts is to convert them into claims, entities, events, temporal expressions, and evidence references. ModernBERT produces contextual representations, and DeBERTa-v3 NLI recognizes support, contradiction, and neutral relationships between claims. The structured elements are represented in a dynamic epistemic-temporal heterogeneous graph, which is analyzed by the proposed DET-HGT model to detect semantic, temporal, evidential, and multi-hop inconsistencies, as shown in Table *4* and Figure *10*. Then, evidence-consistency scores are generated and calibrated using conformal prediction to produce Consistent, Inconsistent, or Uncertain outputs, along with reasoning paths and confidence estimates for human-centered forensic decision support [33].

Layer	Main Components	Function

Data Layer	C.R.I.M.E. interrogation transcripts	Provides forensic narrative data
NLP Processing Layer	Preprocessing + claim/entity/event/time/evidence extraction	Converts raw transcripts into structured forensic information
Context & Semantic Layer	ModernBERT + DeBERTa-v3 NLI	Generates contextual representations and detects Support/Contradiction/Neutral relations
Graph Intelligence Layer	Dynamic Epistemic–Temporal Graph	Represents claims, persons, events, locations, time, evidence, and their relations
Reasoning Layer	DET-HGT	Performs temporal, semantic, evidential, and multi-hop inconsistency reasoning
Decision Layer	Evidence Consistency + Conformal Prediction	Generates Consistent/Inconsistent/Uncertain decisions with calibrated confidence

TABLE 4: Layered architecture and functional roles of the framework

BERT, Bidirectional Encoder Representations from Transformers; DET-HGT, Dynamic Epistemic-Temporal Heterogeneous Graph Transformer; NLP, Natural Language Processing; NLI, Natural Language Inference

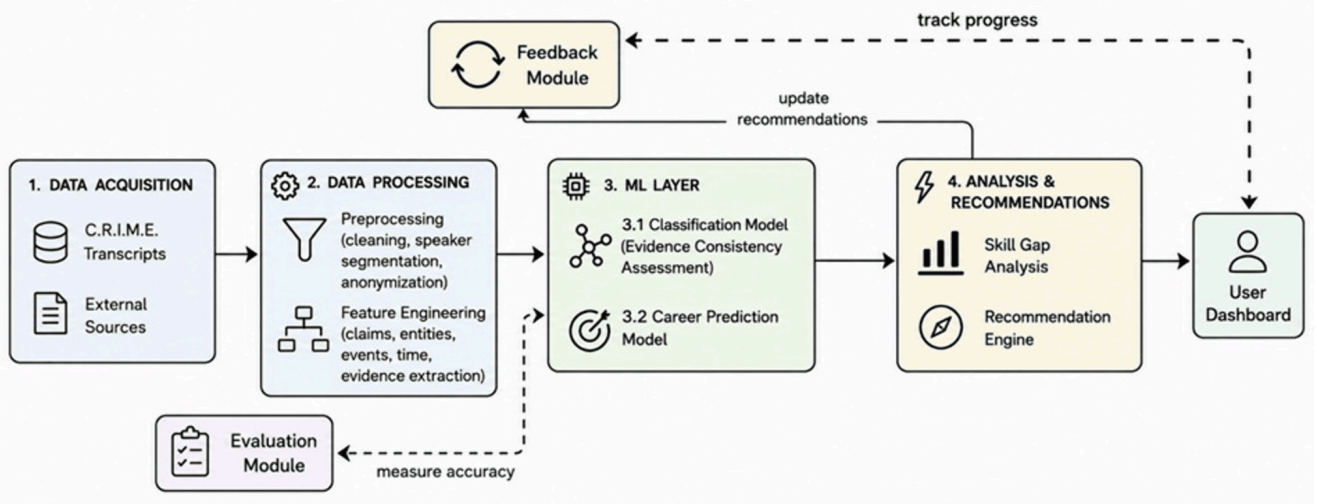


FIGURE 10: Integrated forensic interrogation analysis and feedback architecture

ML, Machine Learning

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

Results And Discussion

The proposed framework examines and structurally captures forensic interrogation narratives and evaluates their evidence consistency in the case of predictive uncertainty [34]. The confusion matrix of the proposed DET-HGT model shows its capability to identify Consistent and Inconsistent forensic claims. The high diagonal values represent good classification performance, and the low numbers of false-positive and false-negative predictions provide good evidence-consistency discrimination. The matrix complements the overall precision, recall, F1-score, and Matthews correlation coefficient (MCC) results, and it gives error analysis for each class, as shown in Figure *11*.

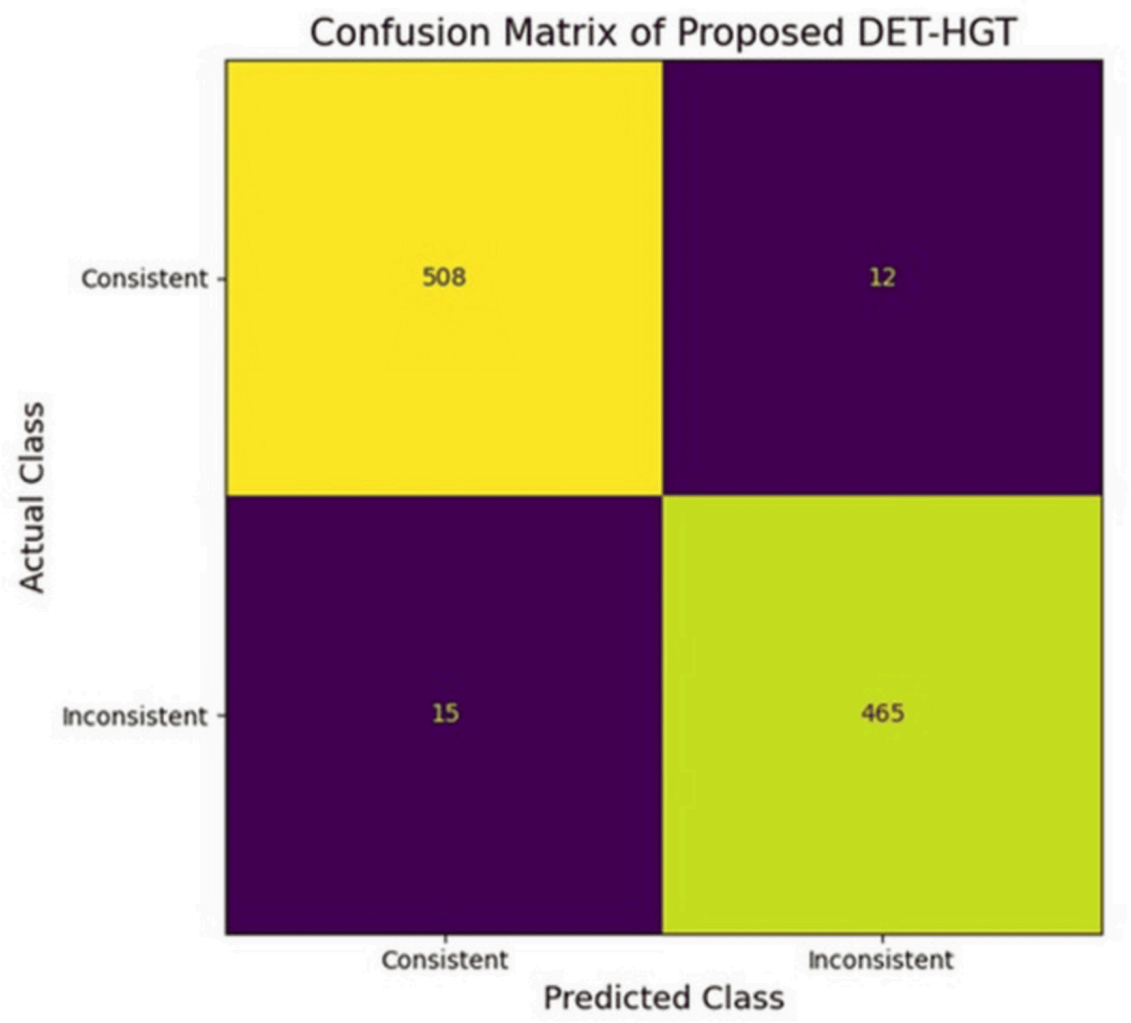


FIGURE 11: Confusion matrix of the proposed DET-HGT model

DET-HGT, Dynamic Epistemic-Temporal Heterogeneous Graph Transformer

The results of the evaluated models are summarized in the table and graph. ModernBERT achieved a precision of 93.4%, a recall of 92.8%, an F1-score of 93.1%, and an MCC of 0.92, while DeBERTa-v3 NLI obtained a precision of 94.6%, a recall of 93.9%, an F1-score of 94.2%, and an MCC of 0.93. The conventional HGT model further enhanced the results, achieving a precision of 95.5%, a recall of 94.8%, an F1-score of 95.1%, and an MCC of 0.94. The proposed DET-HGT model was able to obtain the maximum observed values among all the evaluated models, with a precision of 97.4%, a recall of 96.8%, an F1-score of 97.1%, and an MCC of 0.95. The results show that the introduction of dynamic epistemic-temporal graph reasoning and multi-hop relational modeling could be beneficial for evidence-consistency classification compared with

the evaluated text-based and conventional graph-based models, as shown in Figure *12* and Table *5*. Bootstrap resampling at the interrogation level was used to estimate the statistical uncertainty of the reported model performance for the key metrics used in the principal evaluation. All interrogations were resampled with replacement to account for the dependency among claims from the same interrogation rather than treating individual claims as independent observations. The 2.5th and 97.5th percentiles of the bootstrap distributions were used to calculate the lower and upper confidence limits, respectively. The corresponding point estimates of precision, recall, F1-score, MCC, receiver operating characteristic-area under the curve, and average precision are also reported, together with the resulting confidence intervals for these metrics. The best observed performance was achieved by DET-HGT, with a precision of 97.4%, a recall of 96.8%, an F1-score of 97.1%, and an MCC of 0.95 among the evaluated models. These results show a general positive trend compared with the baseline models, and statistical significance is determined independently using paired comparative testing.

Model	Precision	Recall	F1-score	MCC
ModernBERT	93.4%	92.8%	93.1%	0.92
DeBERTa-v3 NLI	94.6%	93.9%	94.2%	0.93
HGT	95.5%	94.8%	95.1%	0.94
DET-HGT	97.4%	96.8%	97.1%	0.95

TABLE 5: Comparative performance of baseline and proposed models

BERT, Bidirectional Encoder Representations from Transformers; DET-HGT, Dynamic Epistemic-Temporal Heterogeneous Graph Transformer; MCC, Matthews Correlation Coefficient; NLI, Natural Language Inference

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

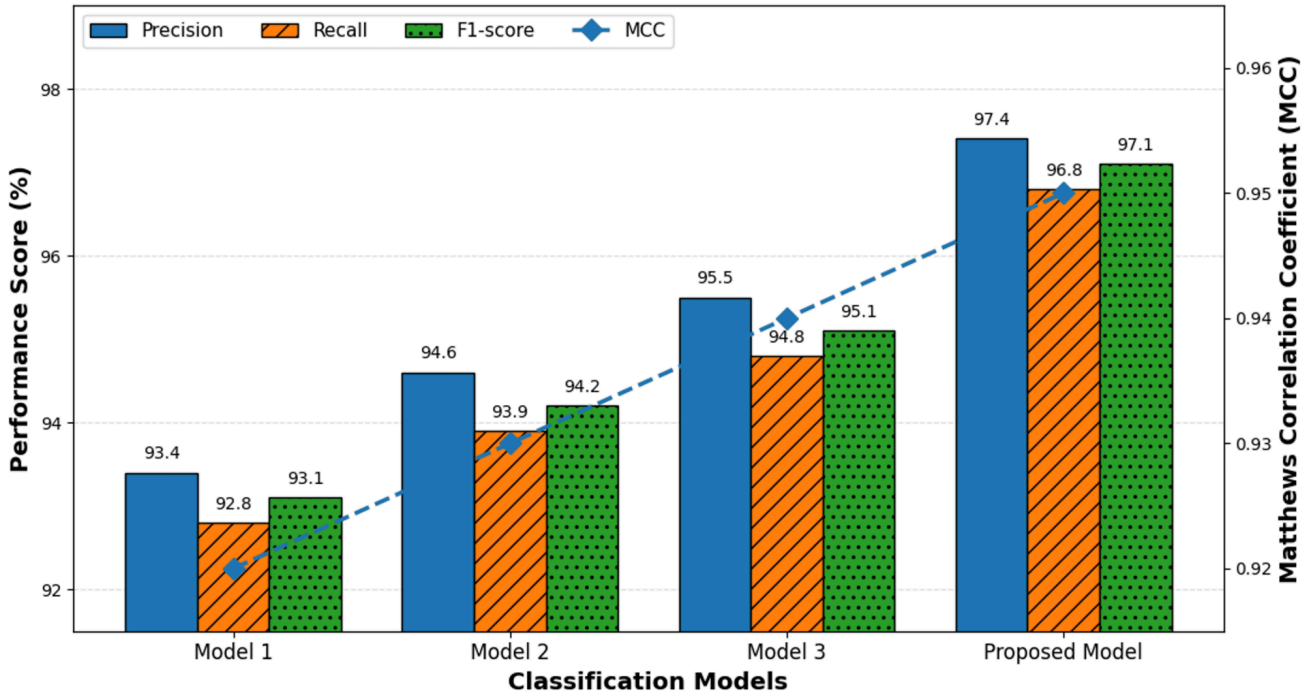


FIGURE 12: Comparative performance of baseline and proposed models

The proposed DET-HGT framework showed consistently good F1-scores across all the evaluated inconsistency categories. Semantic inconsistencies achieved the highest F1-score (97.3%), followed by temporal inconsistencies (96.5%) and evidential inconsistencies (95.7%). Multi-hop inconsistencies scored an F1-score of 94.8% and were the most difficult reasoning category, requiring the propagation of information across multiple interlinked pieces of evidence, claims, and temporal relationships. Although performance gradually declined as the complexity of reasoning increased, the findings suggest that DET-HGT maintained a high ability to detect inconsistencies in both direct and multi-hop reasoning, as shown in Table *6* and Figure *13*.

Inconsistency Type	F1-Score
Semantic	97.3%
Temporal	96.5%
Evidential	95.7%
Multi-Hop	94.8%

TABLE 6: Inconsistency-specific F1-score performance of the proposed DET-HGT framework

DET-HGT, Dynamic Epistemic-Temporal Heterogeneous Graph Transformer

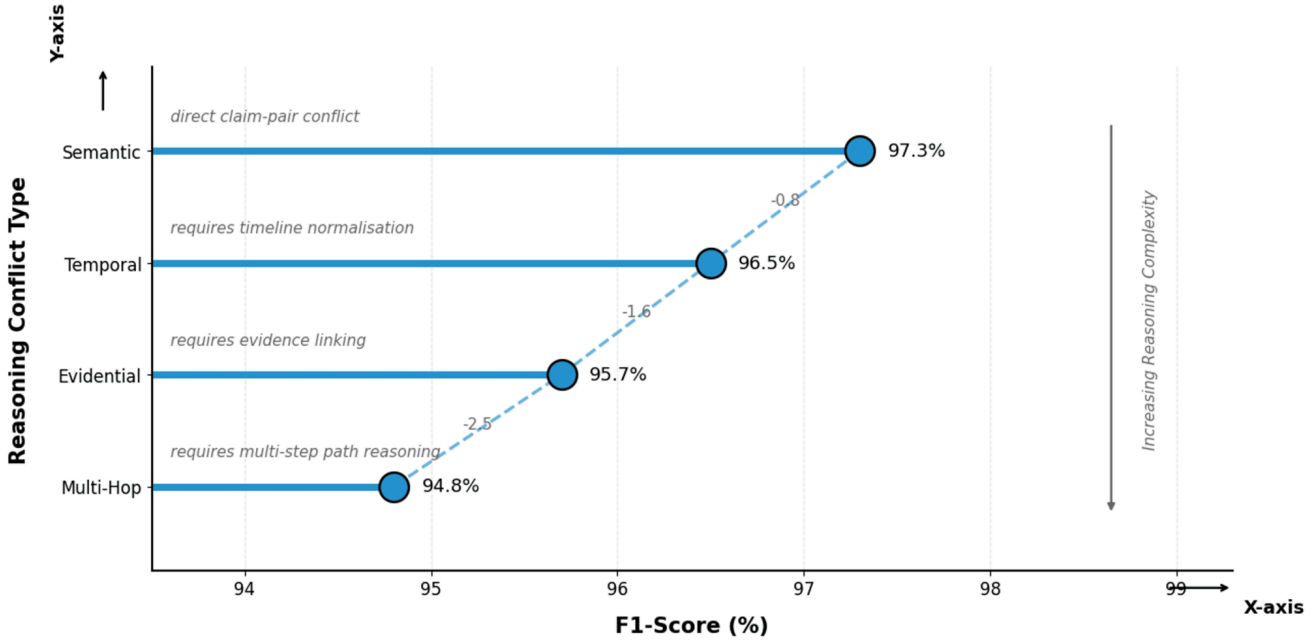


FIGURE 13: Inconsistency-specific F1-score performance across increasing reasoning complexities

To test the accuracy of the multi-hop reasoning function, the performance of DET-HGT was measured as the graph depth increased. For 1-hop, 2-hop, and 3-hop reasoning, the model's accuracy was 97.0%, 95.6%, and 93.4%, respectively. This equates to step-wise decreases of 1.4 and 2.2 percentage points, respectively. The gradual decrease is due to the fact that information needs to be passed among multiple interconnected claim-evidence relationships, and the maintained accuracy of greater than 93% shows that DET-HGT can be used to enable deeper multi-hop forensic reasoning, as shown in Figure *14*.

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

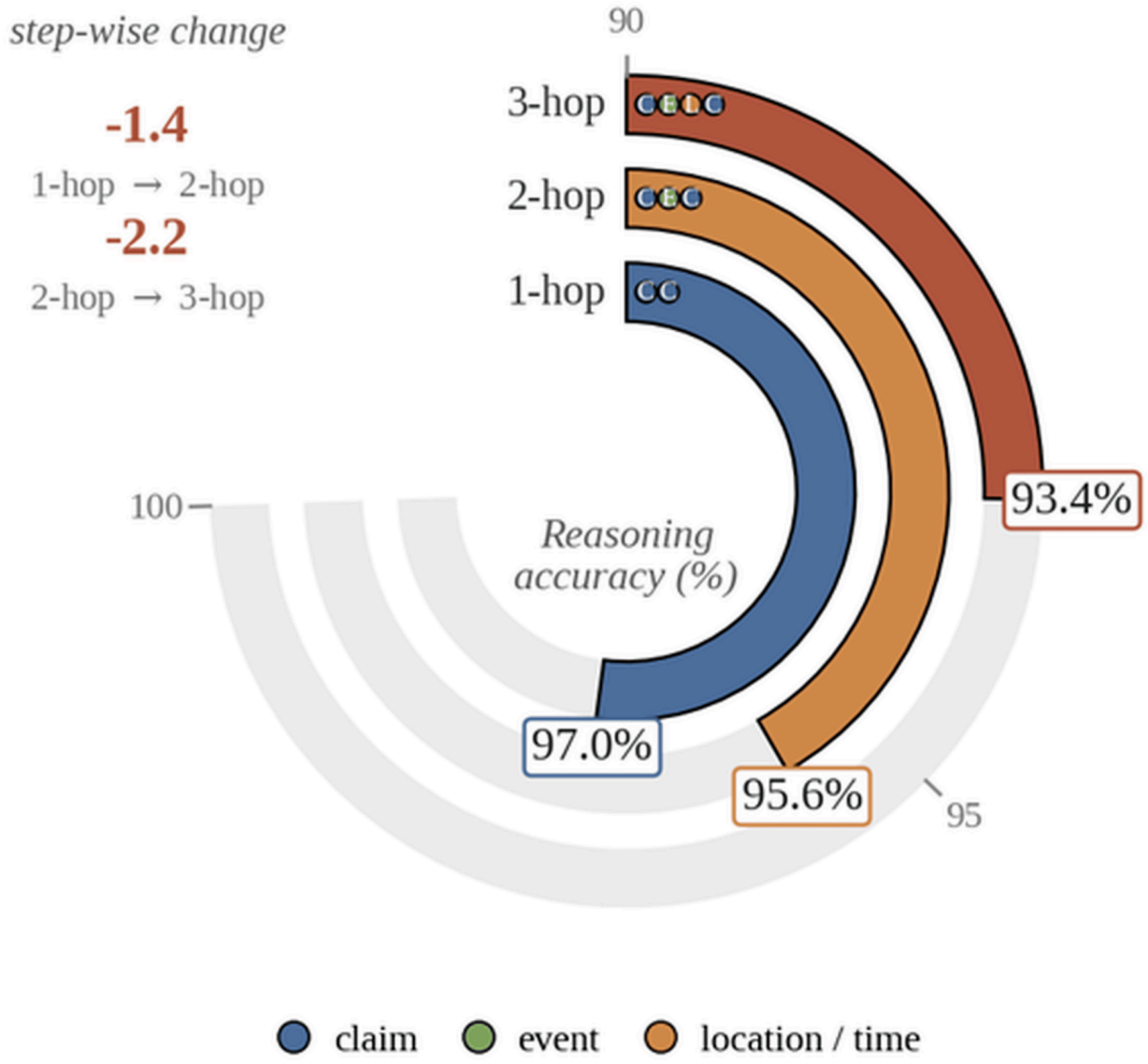


FIGURE 14: Multi-hop reasoning accuracy across increasing graph depths

Calibration resulted in a significant decrease in expected calibration error from 0.074 to 0.026 and in the Brier score from 0.081 to 0.039, with an empirical coverage of 95.2%. These results reveal better agreement between predictive confidence and actual results, as well as the required uncertainty-coverage requirement for evidence consistency assessment, as shown in Table *7* and Figure *15*.

Measure	Before Calibration	After Calibration
ECE	0.074	0.026
Brier Score	0.081	0.039

Coverage	—	95.2%

TABLE 7: Calibration performance before and after conformal prediction

ECE, Expected Calibration Error

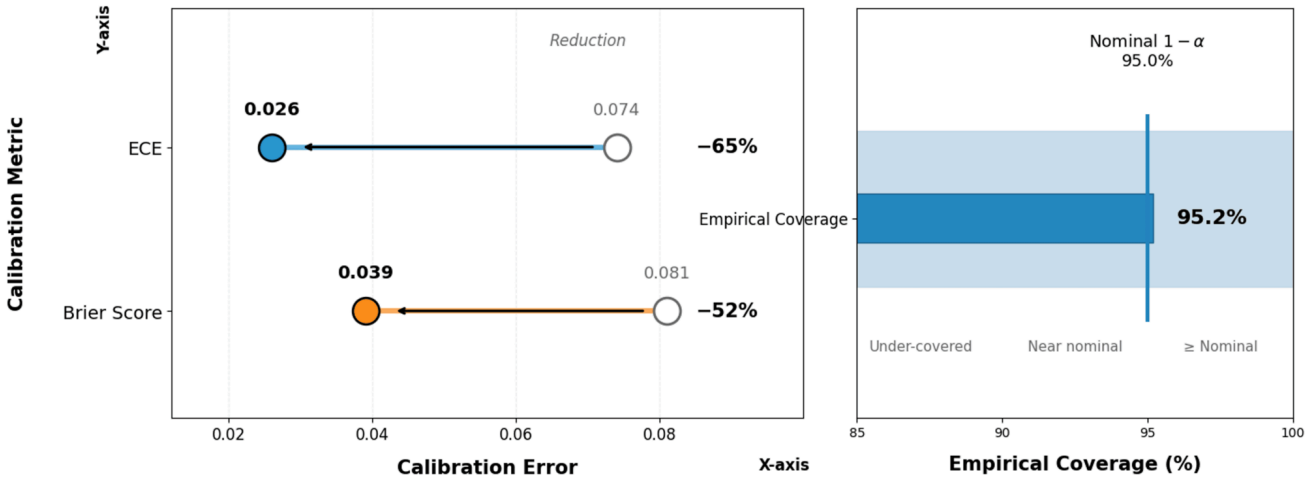


FIGURE 15: Calibration error reduction and empirical coverage analysis

ECE, Expected Calibration Error

The comparative receiver operating characteristic and precision-recall curves indicate that the proposed DET-HGT consistently outperforms ModernBERT, DeBERTa-v3 NLI, and conventional HGT, with the highest area under the curve of 0.998 and average precision of 0.996. The results in the low false-positive region are considered an experimental operating characteristic and should not be applied as a validated forensic threshold without independent domain-specific evaluation. The risk-coverage analysis reveals that the uncertainty-aware selective prediction yields an error rate of 0.91% for 95.8% of claims, as shown in Figure *16*, thus providing significant benefit in terms of the reliability of evidence-consistency judgments at much lower levels of error rate than at full coverage.

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

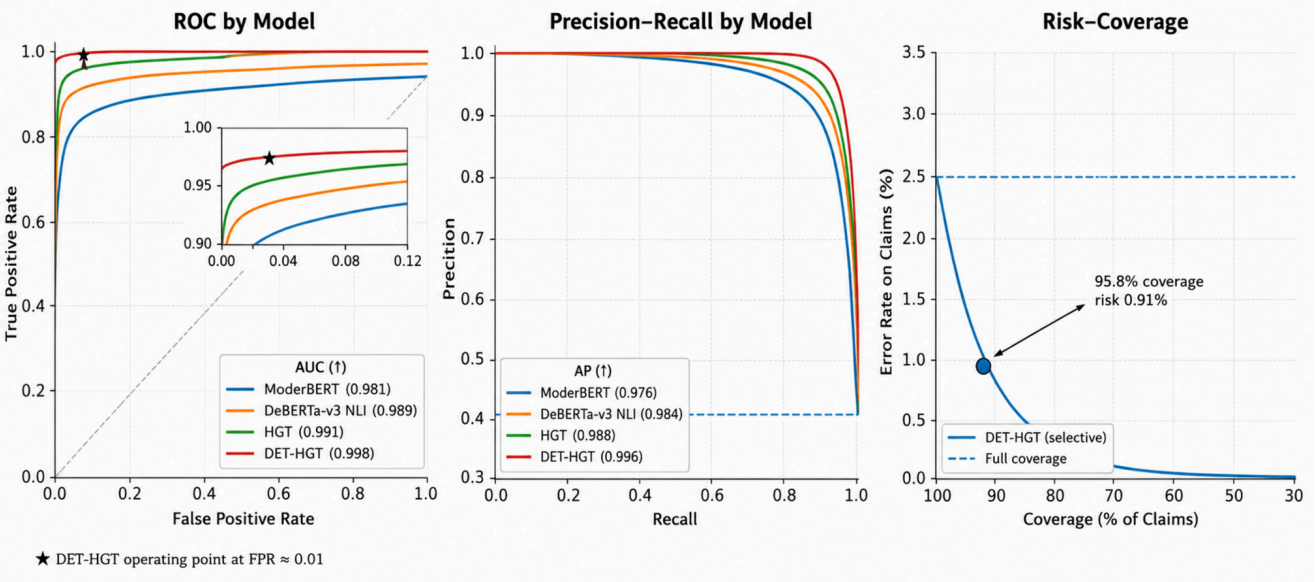


FIGURE 16: Comparative model ROC, precision–recall, and risk–coverage analysis

AUC, Area Under the Curve; BERT, Bidirectional Encoder Representations from Transformers; DET-HGT, Dynamic Epistemic-Temporal Heterogeneous Graph Transformer; FPR, False Positive Rate; ROC, Receiver Operating Characteristic

Ablation analysis

The ablation analysis also showed that each component had its own contribution in the overall performance. The overall F1-score of the full DET-HGT framework was 97.1%, and the F1-score decreased to 95.6%, 94.9%, and 93.5% when temporal relations, evidential edges, and multi-hop graph reasoning were removed, respectively. The most significant performance loss occurred when heterogeneous graph reasoning was removed, thereby establishing its importance in the fusion of distributed evidence in the interrogation narrative.

Discussion

Finally, the results show that the forensic interrogation consistency analysis can be improved by combining contextual NLP and dynamic epistemic-temporal graph reasoning. ModernBERT is able to grasp long-range narrative context, DeBERTa-v3 NLI models semantic support and contradiction, and DET-HGT captures semantic, temporal, and evidential dependency to detect both direct and indirect inconsistencies. The high performance across all four tasks (semantic, temporal, evidential, and multi-hop) indicates that structured graph reasoning is particularly beneficial in cases where there are multiple turns in which the evidence for a contradiction lies. Uncertainty calibration also bolsters the framework by avoiding ambiguous cases being considered as fully reliable decisions. The system is thus seen as a forensic decision-support system, rather than a machine that automatically identifies facts. The framework results in consistency scores and contradiction types, graph-based reasoning paths, and uncertainty estimates calibrated by humans, and final interpretation is left to human experts.

Experimental setup

Three subsets of the C.R.I.M.E. corpus were used for the experimental evaluation, all of which were focused on interrogation and included 173 interrogation transcripts. The dataset was not split randomly per utterance but instead split into the training, validation, and test sets at the level of interrogation/case to avoid information leakage between related statements in the same case. The contextual representation of interrogation narratives was performed with ModernBERT, and the detection of Support, Contradiction, and Neutral relationships between the claims was carried out using DeBERTa-v3 NLI. A baseline graph-based approach, a conventional HGT, was employed, and the proposed approach DET-HGT was tested for semantic, temporal, evidential, and multi-hop reasoning on the constructed

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

interrogation graphs. Precision, recall, F1-score, and MCC were used to evaluate the models for the classification of evidence consistency. The performance of semantic, temporal, evidential, and indirect reasoning was assessed using inconsistency-specific F1-score and multi-hop reasoning accuracy. Conformal prediction was used to quantify the predictive uncertainty by splitting the held-out set of new points into a calibration set, and calibration quality was assessed using expected calibration error, Brier score, and empirical coverage. The same case-level partitions were used across all models for a consistent and fair comparison.

To determine whether the observed differences in performance between the proposed DET-HGT framework and the baseline models were statistically significant, statistical significance testing was conducted. All models were evaluated using the same independent test cases; therefore, paired comparisons were utilized. Differences in model performance were evaluated using paired interrogation-level bootstrap/permutation testing, and the differences in binary classification errors, where applicable, were evaluated using McNemar's test. The significance level adopted was $\alpha = 0.05$, and the Holm correction was used when several pairwise comparisons were made to control the family-wise error rate.

Implementation Details

The experiments were implemented using Python 3.11, PyTorch 2.4.1, Hugging Face Transformers 4.44.2, PyTorch Geometric 2.5.3, and Scikit-learn 1.5.1. ModernBERT and DeBERTa-v3 NLI were trained for 10 epochs using a linear learning-rate schedule (10% warmup), AdamW (learning rate = 2×10^{-5}, weight decay = 0.01), and gradient clipping with a maximum norm of 1.0. For ModernBERT, the maximum sequence length was 2,048 tokens, while for DeBERTa-v3 NLI, it was 512 tokens. Early stopping with a patience of 3 epochs was used based on the F1-score on the validation set.

The DET-HGT model consisted of 4 heterogeneous graph-transformer layers, 8 attention heads, a hidden dimension of 256, and a dropout rate of 0.20. AdamW with a learning rate of 1×10^{-3}, weight decay of 1×10^{-4}, a graph batch size of 8, and a maximum of 50 epochs was used to train DET-HGT, with early stopping patience set to 5 epochs. The final evidence-consistency classification was done using binary cross-entropy loss, and Support, Contradiction, and Neutral relation classification was done using cross-entropy loss. The training and evaluation were done with a fixed random seed of 42, and split conformal prediction was calibrated with $\alpha = 0.05$.

Conclusions

This study designed a dynamic epistemic-temporal graph reasoning framework with uncertainty calibration for evidence-consistent analysis of forensic interrogation narratives. The framework combines the state-of-the-art contextual representation model ModernBERT, the semantic relation model DeBERTa-v3 NLI, and the proposed DET-HGT for heterogeneous temporal and multi-hop reasoning over claims, events, locations, time, and evidence. Conformal prediction also enhances this structure by going beyond the rules of consistent judgments and leaving room for uncertain cases that must be examined by human experts. DET-HGT performs better than the baselines evaluated (ModernBERT and DeBERTa-v3 NLI) when tested on the selected C.R.I.M.E. subsets, and further forensic generalizability needs to be tested on an external dataset. The framework thus moves from a purely classificatory perspective in forensic computational analysis to a more narrative approach, based on evidence. Most importantly, the proposed system is not a system that determines the truth or falsity of a case, but is a human-in-the-loop system that offers traceable reasoning paths, evidence-consistency scores, and calibrated uncertainty to aid the expert interpretation of the case.

Future research will involve seeking to validate the proposed framework across the forensic and legal interrogation datasets and extend it to other contexts through evaluations of generalizability, to other types of forensic and legal datasets. The framework can be extended to multimodal forensic reasoning with integration of audio, visual, and behavioral cues to text, and to multilingual analysis for cross-linguistic interrogation settings. Also, if updates to the graph are dynamic, then in real-time, new statements and evidence can arrive during the course of an interrogation, which will also enable continuous reasoning over the graph. Furthermore, human-AI validation should be performed to assess the practical usefulness, interpretability, trust calibration, and effectiveness of the framework as a decision support tool for real investigation settings, while involving forensic experts as the human component.

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Rajashri J. Rikame, Aniket Anand, Sanjay K. Tiwari , Mritunjay K. Ranjan

Acquisition, analysis, or interpretation of data: Rajashri J. Rikame, Aniket Anand, Sanjay K. Tiwari , Mritunjay K. Ranjan

Drafting of the manuscript: Rajashri J. Rikame, Aniket Anand, Sanjay K. Tiwari , Mritunjay K. Ranjan

Critical review of the manuscript for important intellectual content: Rajashri J. Rikame, Sanjay K. Tiwari , Mritunjay K. Ranjan

Supervision: Rajashri J. Rikame

Disclosures

Human subjects: Consent was waived for all participants in this study. This study involved secondary computational analysis of publicly available and anonymized forensic interrogation transcripts. No direct interaction with human participants occurred, and no attempt was made to re-identify individuals. Informed consent was therefore waived/not required in accordance with the applicable institutional determination for secondary analysis of publicly available anonymized data.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following:

- **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work.
- **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work.
- **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.

References

1. Farrugia L, Oxburgh G, Gabbert F: Evaluation of forensic interviews and interrogations: a thematic analysis of law-enforcement views. Journal of Criminal Psychology. 2024, 16:269-281. 10.1108/jcp-02-2024-0015
2. Fede MS, Compagnucci P, Montana A, Dello Russo A, Giorgetti R, Busardò FP: Forensic perspectives on postmortem CIED interrogation: A systematic review and meta-analysis. Forensic Science International. 2024, 359:112001. 10.1016/j.forsciint.2024.112001
3. Tataroğlu Özbulak GA, Shrestha YR, Calbimonte JP: Bridging dynamics and semantics: A unified perspective on explainable Graph Neural Networks based stream reasoning. Knowledge-Based Systems. 2026, 351:116662. 10.1016/j.knosys.2026.116662
4. Liang Q, Li HN, Zhu J, Feng K, Hu JW, Cheng YH, Huang CG: Hierarchical domain-physical topology enhanced Fully-Connected Spatial-Temporal Graph Neural Network for remaining useful life prediction. Reliability Engineering & System Safety. 2026, 275:112848. 10.1016/j.ress.2026.112848
5. Yang C, Zhang XG, Zhang GH, Wang C, Fan JS: Towards reliable deep excavation monitoring through graph recurrent neural network-based spatio-temporal imputation. Reliability Engineering & System Safety. 2026, 270:112149.

10.1016/j.ress.2025.112149

6. Zhang X, Yan L, Ma Z: Multi-hop reasoning of knowledge chain over fuzzy spatiotemporal knowledge graphs with box embeddings. Engineering Applications of Artificial Intelligence. 2026, 173:114443. 10.1016/j.engappai.2026.114443
7. Zhang J, Yuan N, Liu Y, Wu X, Fang S, Peng Y, Gong Y: Multi-temporal UAV rice AGB estimation with dynamic biomass allocation and interpretable deep learning. European Journal of Agronomy. 2026, 177:128096. 10.1016/j.eja.2026.128096
8. Adkins J, Al Bataineh A, Khanal A: A psycholinguistic NLP framework for forensic text analysis of deception and emotion. Frontiers in Artificial Intelligence. 2025, 8:1669542. 10.3389/frai.2025.1669542
9. Cheng K, Ye J, Lu X, Sun L, Du B: Temporal graph network for continuous-time dynamic event sequence. Knowledge-Based Systems. 2024, 304:112452. 10.1016/j.knosys.2024.112452
10. Yuan C, Xie Q, Ananiadou S: Temporal relation extraction with contrastive prototypical sampling. Knowledge-Based Systems. 2024, 286:111410. 10.1016/j.knosys.2024.111410
11. Huang M, Du J, Liu Q, Xian W: Enhanced temporal knowledge graph reasoning through TKG-LPEHD: LLM-driven dynamic fusion of graph and rule-based methods. Data & Knowledge Engineering. 2026, 165:102623. 10.1016/j.datak.2026.102623
12. Warner B, Chaffin A, Clavié B, et al.: Smarter, better, faster, longer: a modern bidirectional encoder for fast, memory efficient, and long context finetuning and inference [PREPRINT]. arXiv. 2024, 10.48550/arXiv.2412.13663
13. Liao R, Jia X, Li Y, Ma Y, Tresp V: GenTKG: generative forecasting on temporal knowledge graph with large language models. Findings of the Association for Computational Linguistics: NAACL 2024. 2024, 4303-4317. 10.18653/v1/2024.findings-naacl.268
14. Campos M, Farinhas A, Zerva C, Figueiredo MAT, Martins AFT: Conformal prediction for natural language processing: a survey. Transactions of the Association for Computational Linguistics. 2024, 12:1497-1516. 10.1162/tacl_a_00715
15. Dwyer I, Justinvil D, Cunningham A: Caribbeanist casualties: Interrogating the application of structural vulnerability to forensic anthropology. Forensic Science International: Synergy. 2023, 6:100327. 10.1016/j.fsisyn.2023.100327
16. Hribnak L, Dubravova H: Use of large language models to automate the extraction of interrogation data within joint investigation teams. Procedia Computer Science. 2025, 263:200-207. 10.1016/j.procs.2025.07.025
17. Krishna S, Ahuja P: Forensic implications of organic gunshot residue (OGSR) degradation in delayed evidence collection scenarios: An assessment of environmental stressors on OGSR stability using Raman spectroscopy. Vibrational Spectroscopy. 2026, 144:103918. 10.1016/j.vibspec.2026.103918
18. Panda A, Dogra DP, Dey PP: Physics-constrained dynamic graph and multi-resolution temporal modeling for traffic forecasting. Information Sciences. 2026, 757:123911. 10.1016/j.ins.2026.123911
19. Tarttelin Hernández P, Haines A, Mousseau V, Elbourne M, Hanna T: Generative artificial intelligence in forensic science education: Student perspectives on AI-supported crime scene decision-making. Forensic Science International: Synergy. 2026, 13:100723. 10.1016/j.fsisyn.2026.100723
20. Cohen PR, Weinberg JM, Hoenig LJ, Parish LC: Forensic dermatology and the crime scene investigation (CSI) effect: Are jurors influenced by forensic dermatologic investigations that they view on television, printed media, and the internet?. Clinics in Dermatology. 2026, 44:650-668. 10.1016/j.clindermatol.2026.03.007
21. Acsiniuc A, Manfredi B, Menéndez-Pérez J, Branfield S, Washington AV: Interrogating a compound library in search of an inhibitor for TREM-like transcript-1 to fibrinogen binding. The Journal of Pharmacology and Experimental Therapeutics. 2025, 392:100009. 10.1124/jpet.124.002086
22. Liu P, Chen S, Liu J, Xu Y, Cao F, Wu G: An efficient contradiction separation based automated deduction algorithm for enhancing reasoning capability. Knowledge-Based Systems. 2023, 261:110217. 10.1016/j.knosys.2022.110217
23. Zhu T, Yang Y, Weng Z, Sun H, Ma Z, Chen G: ProGrasp: Storage-efficient provenance graph compression for APT forensics via structure prediction and attribute aggregation. Knowledge-Based Systems. 2026, 343:115986. 10.1016/j.knosys.2026.115986
24. Kawaguchi K, Maeda M, Oda F, Nakashima Y, Fukuda H: Development and validation of a machine learning model for frailty screening using claims data in Japan: the Longevity Improvement & Fair Evidence Study. Experimental Gerontology. 2026, 215:113050. 10.1016/j.exger.2026.113050
25. Rehman MZU, Raghuvanshi D, Jain U, Bansal S, Kumar N: A multimodal-multitask framework with cross-modal relation and hierarchical interactive attention for semantic comprehension. Information Fusion. 2026, 126:103628. 10.1016/j.inffus.2025.103628

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y

26. Kannojiya V, Dixit V: Zero-shot text classification: harnessing pretrained language models for label-free document categorization. Procedia Computer Science. 2026, 283:2735-2742. 10.1016/j.procs.2026.06.342

27. Feng J, Li W, Zhao X, Yang L, Liu B, Liu M, Zhang M: Adaptive multi-hop reasoning with type-aware calibration for process knowledge graph completion. Information Sciences. 2026, 754:123643. 10.1016/j.ins.2026.123643

28. Jiang H, Ding L, Chen S, Ji J, Zhu J: Improving the modeling reliability of domino effect spatial-temporal evolution process via a novel prediction method considering aleatory and epistemic uncertainties. Process Safety and Environmental Protection. 2025, 204:108038. 10.1016/j.psep.2025.108038

29. Parreño SJ: Uncertainty-aware sepsis survival prediction using conformal XGBoost on minimal clinical features under Sepsis-3 criteria. Computer Methods and Programs in Biomedicine. 2026, 283:109429. 10.1016/j.cmpb.2026.109429

30. Farley M, Yim AD: An exploratory analysis of decision-making in population affinity estimation among forensic anthropology practitioners in the United States. Science & Justice. 2026, 66:101500. 10.1016/j.scijus.2026.101500

31. Mason M: Police Interrogation, Language, and the Law: The Invocation Game. Cambridge University Press, Cambridge, UK; 2023. 10.1017/9781009052047

32. Mazzolini D, Mignone P, Pavan P, Vessio G: An easy-to-explain decision support framework for forensic analysis of dynamic signatures. Forensic Science International: Digital Investigation. 2021, 38:301216. 10.1016/j.fsidi.2021.301216

33. Du Y, Yan R, Hou Y, Pei Y, Han H: Adversarial entity graph convolutional networks for multi-hop inference question answering. Expert Systems with Applications. 2024, 258:125098. 10.1016/j.eswa.2024.125098

34. Noriega M: The application of artificial intelligence in police interrogations: An analysis addressing the proposed effect AI has on racial and gender bias, cooperation, and false confessions. Futures. 2020, 117:102510. 10.1016/j.futures.2019.102510

How to cite this article:
Anand A, Tiwari S K, Ranjan M K, et al. (November 09, 2026) Dynamic Epistemic–Temporal Graph Reasoning With Uncertainty Calibration for Evidence-Consistent Forensic Interrogation Narratives. Cureus J Comput Sci 3 : es44389-026-00361-y. DOI https://doi.org/10.7759/s44389-026-00361-y