Orion: A Semantic Uplift Pipeline for Building Persistent, Queryable Knowledge from Egocentric Video

Riddhiman Rana[1*], **Aryav Semwal**[1], **Yogesh Atluru**[1], **Sunith Vallabhaneni**[1†], **Cristian Meo**[1†], **Kevin Zhu**[1†]

[1]Algoverse AI Research

riddhiman.rana@gmail.com, aryavsemwal17@gmail.com, yatluru@gmail.com, sunithv@berkeley.edu, c.meo@tudelft.nl, kevin@algoverseacademy.com

Abstract

Understanding causal dynamics in egocentric video remains a fundamental challenge for grounded language and vision systems. Existing approaches detect objects and actions, but fail to transform perceptual data into coherent, causally structured representations. We introduce **Orion**, a modular system that performs *semantic uplift* - the process of transforming continuous perceptual streams into discrete, symbolic event, and relation graphs. Orion integrates perception (YOLO11x, CLIP), tracking (Hungarian + HDBSCAN), and reasoning (LLM-based event composition) to construct dynamic knowledge graphs from raw video. Using the Video Scene Graph Reasoning (VSGR) dataset, Orion achieves strong triplet accuracy and causal coherence. Our key contribution is a scientifically grounded semantic uplift mechanism supported by justified causal scoring, configuration-aware thresholds, and constrained prompt-based reasoning. This work positions semantic uplift as a bridge between low-level vision and high-level causal language representations, aligning with the goal of GCLR to link perception, reasoning, and symbolic knowledge.

Code — https://github.com/riddhimanrana/orion-research

Introduction

Egocentric video, captured from a first-person perspective, offers rich data for understanding human-object interactions in dynamic environments, with applications in robotics, personal assistants, and augmented reality (Grauman et al. 2022), (Grauman et al. 2024). However, current perception-driven systems, such as object detectors and action recognizers, excel at identifying entities and actions but fail to synthesize these into causally coherent knowledge graphs (Robinson et al. 2023), (Xie et al. 2025). For example, while a system can detect a cup moving, it struggles to infer whether it was picked up, dropped, or knocked over, and how such events are causally linked. This gap hinders intelligent systems' ability to model the *why* and *how* of events, critical for tasks like robotic planning, where understanding causal sequences (e.g., "opening a door causes entry into a room") enables context-aware decision-making (Cibula et al. 2025).

Prior work falls short in addressing this challenge. Temporal reasoning models like LLM-DA (Wang et al. 2023) and DAEMON (Dong et al. 2023) operate on pre-existing symbolic graphs, lacking mechanisms to construct them from raw video. Heuristic-based systems, such as Action Scene Graphs (EASG) (Rodin et al. 2023) and HyperGLM (Nguyen et al. 2025), rely on annotated inputs or brittle spatial rules, missing semantic nuances and causal relationships. Approaches like Video-of-Thought (VoT) (Fei et al. 2024) use unstructured captions, leading to inconsistent reasoning in egocentric settings. CausalVQA (Foss et al. 2025) focuses on question-answering rather than automated graph construction. Consequently, no system fully bridges the gap between raw egocentric video and causally accurate knowledge graphs.

We propose the Orion pipeline, a novel system that transforms raw egocentric video into dynamic knowledge graphs capturing entities, events, and their spatial, temporal, and causal relationships. Orion integrates perception (YOLO11x for detection, CLIP for embeddings), tracking (Hungarian algorithm, HDBSCAN clustering), semantic uplift (Ollama for event composition), and knowledge graph construction (Neo4j storage). Unlike heuristic methods, our pipeline uses structured perceptual logs and constrained LLM reasoning to infer nuanced events (e.g., "Person_1 picks_up Cup_1") and causal edges (e.g., "Event_A CAUSES Event_B"). The system supports flexible configurations (fast, balanced, accurate presets) and secure credential management, making it deployable across local, Docker, and Kubernetes environments.

We evaluate Orion on the Video Scene Graph Reasoning (VSGR) dataset (Nguyen et al. 2025), which provides 1.9M frames with graph-structured annotations, including egocentric clips with causal relationships. Compared to heuristic baselines and HyperGLM, Orion aims to achieve higher Triplet F1 and higher Causal Reasoning Score, validated through ablations and case studies. Our contributions are:

- **Orion Pipeline:** An end-to-end system combining perception, tracking, semantic uplift, and graph construction for causally accurate knowledge graphs from raw video.
- **Empirical Validation on VSGR:** Targets superior triplet accuracy and causal coherence over heuristic and state-

[*]Lead Author

[†]Senior Author

of-the-art baselines.

- **Ablation Insights:** Isolates contributions of clustering, state detection, and LLM reasoning to graph quality.

This work answers: Can an integrated pipeline with LLM-based semantic uplift construct more causally accurate and coherent knowledge graphs from egocentric video than heuristic systems? Our results aim to advance video understanding by bridging perception and symbolic reasoning.

2. Related Work

Our work builds on three axes: temporal knowledge graph reasoning, egocentric video understanding, and causal inference in video.

Temporal Knowledge Graph Reasoning

LLM-DA (Wang et al. 2023) adapts rules for reasoning over pre-existing temporal graphs, predicting future relationships, but assumes clean symbolic inputs, lacking video-to-graph construction. DAEMON (Dong et al. 2023) tracks relationship sequences for prediction but cannot resolve entities from raw video. Both highlight the need for automated graph construction, which Orion addresses through integrated perception and reasoning.

Egocentric Video Understanding

Action Scene Graphs (EASG) (Rodin et al. 2023) extend Ego4D (Grauman et al. 2022) annotations into temporal graphs, but rely on heuristic, annotation-heavy methods, missing causal uplift. HyperGLM (Nguyen et al. 2025) uses a multimodal LLM to build hypergraphs from annotated VSGR frames, capturing multi-way interactions, but its dependence on pre-labeled inputs limits real-time applicability. Video-of-Thought (VoT) (Fei et al. 2024) employs chain-of-thought reasoning on captions, but lacks structured logs for precise causal graphs. Orion's annotation-free pipeline, using the Hungarian algorithm for tracking and Neo4j for structured storage, overcomes these limitations.

Causal Inference in Video

CausalVQA (Foss et al. 2025) provides a benchmark for causal reasoning via QA pairs, but does not automate graph construction. Recent work in causal video understanding (Li et al. 2020) emphasizes physical reasoning, but lacks symbolic representations. Orion uniquely integrates perception (YOLO11x, CLIP), tracking (Hungarian algorithm, HDBSCAN), and LLM reasoning (Ollama) to infer causal edges from raw video, surpassing heuristic and state-of-the-art methods in constructing causally coherent graphs without annotations.

Summary

Unlike prior work, our pipeline constructs causally coherent graphs without annotations, leveraging structured logs and constrained LLM reasoning to surpass heuristic and state-of-the-art methods.

3. Methods

The **Orion pipeline** transforms unannotated egocentric video into a dynamically evolving knowledge graph through a structured, five-phase architecture. Each phase—from perception to query—builds on modular, configuration-driven layers that ensure scalability, interpretability, and deployment flexibility. Figure **??** illustrates the complete data flow.

```
VIDEO INPUT
    ↓
[1] PERCEPTION PHASE
    - Object Detection (YOLO11x)
    - Spatial Analysis (bounding boxes)
    - Embedding Generation (CLIP)
    ↓
[2] TRACKING & SEMANTIC UPLIFT PHASE
    - Entity Clustering (HDBSCAN)
    - State Change Detection
    - Temporal Windowing
    - Event Composition (LLM)
    ↓
[3] KNOWLEDGE GRAPH CONSTRUCTION
    - Scene/Entity/Event node creation
    - Spatial relationship analysis
    - Causal reasoning
    - Temporal sequencing
    ↓
[4] STORAGE & INDEXING
    - Neo4j graph persistence
    - Vector indexing
    - Relationship constraints
    ↓
[5] QUERY & Q\&A
    - Knowledge retrieval
    - Contextual reasoning
    - LLM-based answer generation
```

System Architecture Overview

Orion is designed as a *multi-layered video-understanding system* that couples low-level perception with high-level reasoning. The architecture emphasizes:

- **Modularity:** Each stage can run independently or as part of the end-to-end pipeline.
- **Centralized Configuration:** Parameters and credentials are managed through `config.py` and a `ConfigManager` singleton.
- **Hardware Abstraction:** Components operate seamlessly across CUDA, MPS, or CPU backends.
- **Graph-based Representation:** All outputs are stored in Neo4j for structured reasoning and visualization.

Architecture Layers

Layer 1: Configuration & Credential Management **Files:** `config.py`, `config_manager.py`

A three-tier configuration hierarchy ensures reproducibility and security:

```
Environment Variables (.env)
(ORION_NEO4J_PASSWORD, etc.)

                    ↓

ConfigManager Singleton
- Loads env vars & manages credentials
- Provides lazy initialization

                    ↓

OrionConfig Instance
- System & model parameters
- Neo4j / Ollama / CLIP configs
```

Preset modes enable resource-aware tuning:

- **Fast:** YOLO11n + 512-dim CLIP embeddings
- **Balanced:** YOLO11m + 1024-dim embeddings
- **Accurate:** YOLO11x + 2048-dim embeddings

Layer 2: Data Models & Persistence Files: `neo4j_manager.py`, `model_manager.py`

The `Neo4jManager` handles database connections and schema enforcement. Each node type—*Scene*, *Entity*, *Event*, *StateChange*, and *SpatialRelationship*—has a defined property schema and uniqueness constraint. Batch transactions (>1000 entities per commit) ensure efficiency, while vector embeddings are stored for similarity search.

Layer 3: Pipeline Engines

(a) Perception Engine. Decodes video at ≈4 FPS. YOLO11x detects objects, and CLIP embeds each detection into a multimodal feature space. Output: JSON-like perceptual logs containing bounding boxes, class labels, confidences, and embeddings.

(b) Tracking Engine. Maintains entity identity using the Hungarian assignment algorithm combined with HDBSCAN clustering. Composite cost function:

$$C_{ij} = 0.7(1 - \text{cosine\_sim}) + 0.3(1 - \text{IoU})$$

A match is accepted when $C_{ij} < 0.5$. This reduces hundreds of frame-level detections to tens of persistent entities while recording velocity and embedding trajectories.

(c) Semantic Uplift Engine. Semantic uplift is the core contribution of Orion, serving as the bridge between low-level perceptual data and high-level symbolic reasoning. Formally, it transforms a sequence of detections and embeddings into a structured, language-grounded knowledge representation.

Inputs. The semantic uplift module receives:

- $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$: a set of tracked entities, each with bounding boxes, embeddings, and temporal states.
- $\mathcal{S}_t$: scene context for frame window t, including spatial relationships and motion vectors.
- Φ: embedding-based similarity functions for appearance and motion.

Outputs. The module produces:

- $\mathcal{G} = (V, R)$: a scene-specific subgraph with nodes V (entities, events, states) and relations R (spatial, temporal, causal).
- Natural language event descriptions generated via LLM prompting, grounded in observed state changes.

Constraints.

- The mapping $\mathcal{E} \rightarrow \mathcal{G}$ must preserve temporal consistency: entities retain identity across frames.
- Only state changes exceeding the adaptive threshold τ_{state} (derived via sensitivity analysis) trigger event generation.
- Prompts must be grounded in measurable attributes (position, velocity, interaction count) rather than free-text hallucinations.

Algorithm. The overall procedure is summarized below:

```
Algorithm 1: Semantic Uplift
Input: Entities E, Scene Context S_t, Config
Output: Graph G(V, R)

for each temporal window w in video:
    E ← detect_state_changes(E, w, _state)
    events ← compose_events(E, S_t, LLM_prompt
    G.add_nodes(events + entities)
    G.add_edges(temporal_relations(events))
return G
```

Discussion. Unlike prior pipelines that rely solely on object and action recognition, Orion's semantic uplift explicitly operationalizes the transformation from continuous visual states to symbolic representations. The process is deterministic up to event composition, where the LLM operates under explicit input constraints. This allows grounded reasoning while maintaining interpretability.

(d) Knowledge Graph Builder. Instantiates Neo4j nodes and relationships:

- **Spatial:** `SPATIAL_NEAR`, `LEFT_OF`
- **Temporal:** `PRECEDES`, `FOLLOWS`
- **Causal:** `CAUSES`
- **Participation:** `INVOLVED_IN`

(e) Storage & Query Subsystem. Ensures persistence, indexing, and query performance through schema-level constraints (e.g., `entity_id_unique`, `event_timestamp_idx`). Supports flexible deployment (local, Docker, Kubernetes) and secure credentials (`ORION_NEO4J_PASSWORD`).

Pipeline Execution Flow

1. **Initialization:** Load configuration via `ConfigManager`.
2. **Perception:** Extract detections and embeddings.
3. **Tracking:** Assign and cluster entities.
4. **':** Detect states, compose events, and infer causality.
5. **Graph Build & Persist:** Write nodes and edges to Neo4j.

6. **Query:** Retrieve or reason via Cypher or LLM-driven Q&A.

This modular flow allows both isolated benchmarking and full-pipeline execution.

Mathematical Foundations

Core computations rely on cosine similarity, Intersection-over-Union (IoU), the Hungarian algorithm for optimal assignment, HDBSCAN density clustering, and causal-influence scoring. Together, they ensure both spatial–temporal precision and semantic coherence.

Integration Points

Orion exposes standard APIs for embedding generation, event composition, and Neo4j ingestion. All modules share common data-type contracts, enabling substitution of alternative detectors, LLMs, or databases without altering pipeline logic.

Causal Influence Scoring (CIS): Derivation, Validation, and Learning

The Causal Influence Scoring (CIS) module quantifies the strength of causal relationships between detected events by combining multiple normalized factors into a single interpretable metric:

$$\text{CIS} = w_T T_p + w_S S_p + w_E E_o + w_{\text{sem}} S_s$$

where T_p represents temporal proximity, S_p represents spatial proximity, E_o denotes entity overlap, and S_s captures semantic similarity between events. Each component is normalized to the range $[0, 1]$ to ensure comparability across scales. In the standard configuration, initial weights of $w_T = 0.3$, $w_S = 0.3$, $w_E = 0.2$, and $w_{\text{sem}} = 0.2$ were used with a threshold $\tau = 0.6$ to classify event pairs as causally linked.

To establish a principled and reproducible basis for these parameters, we conducted a systematic derivation and validation process consisting of three stages: (1) formalizing the mathematical intent of the metric, (2) performing sensitivity analysis and grid search to optimize weight and threshold selection, and (3) introducing a learnable alternative that removes heuristic dependency. The sensitivity analysis procedure evaluates combinations of $(w_T, w_S, w_E, w_{\text{sem}})$ and thresholds τ using a labeled validation set of event pairs with known causal or non-causal relationships. For each configuration, the CIS is computed and evaluated using precision, recall, and F1-score metrics. The optimal parameters (w^*, τ^*) are selected to maximize F1 while maintaining stability under cross-validation. This process ensures that both the weights and the threshold are grounded in empirical evidence rather than manual selection.

In addition to the grid-search procedure, a data-driven alternative is implemented through logistic regression, where the CIS components $\mathbf{x} = [T_p, S_p, E_o, S_s]$ serve as features. The regression model learns interpretable coefficients corresponding to each factor, effectively producing an adaptive CIS function that maps input features to causal probabilities. These learned coefficients can replace or refine the static weights, providing a statistically justified and reproducible causal scoring mechanism.

Intuition and normalization. Each component of CIS is a normalized score in $[0, 1]$:

- T_p (temporal proximity): normalized inverse time difference between events,
- S_p (spatial proximity): normalized inverse centroid distance (or IoU-based),
- E_o (entity overlap): fraction of shared entities or entity-ID overlap,
- S_s (semantic similarity): cosine similarity between event descriptions or embeddings.

Normalization places all components on comparable scales so linear weighting is meaningful.

Sensitivity analysis and grid search (reproducible procedure). We validate weight/threshold choices via the following reproducible protocol:

1. Prepare a labeled validation set of event-pairs with ground-truth causal/non-causal labels from VSGR (or human annotations).
2. For each candidate weight vector $w = (w_T, w_S, w_E, w_{\text{sem}})$ in a grid (or randomized search), compute CIS for all validation pairs and obtain predicted causal labels $\hat{y} = I(\text{CIS} > \tau)$ for a set of thresholds $\tau \in [0, 1]$.
3. For each (w, τ) compute precision, recall, and F1 against ground truth.
4. Select the (w^*, τ^*) that maximizes F1 (or provides a preferred precision/recall tradeoff).
5. Perform k-fold cross-validation (or bootstrapping) to estimate variance and statistical significance.

We also analyze robustness by perturbing each weight by $\pm 10\%$ and reporting F1 change; this quantifies how sensitive the metric is to each weight.

Threshold selection. Rather than choosing τ heuristically, we (optionally) select it by:

- maximizing F1 on validation data: $\tau^* = \arg\max_\tau \text{F1}(\tau)$, or
- choosing the elbow point on the precision–recall curve (balance point), or
- selecting a threshold to achieve a target precision (e.g., ≥ 0.8) if conservative causal claims are required.

Learning weights (recommended alternative). A principled approach is to learn a classifier that maps the raw features $\mathbf{x} = [T_p, S_p, E_o, S_s]$ to causal probability $P(y = 1 \mid \mathbf{x})$. A logistic regression model provides an interpretable learned linear weighting:

$$\log \frac{P(y = 1 \mid \mathbf{x})}{1 - P(y = 1 \mid \mathbf{x})} = \beta_0 + \beta_T T_p + \beta_S S_p + \beta_E E_o + \beta_{\text{sem}} S_s.$$

The learned β coefficients (after rescaling) can be used as weights w, and the model directly outputs a probability for thresholding.

Reproducible implementation (example). Below is a minimal Python script (scikit-learn) to (a) learn weights, (b) compute PR curve and F1, and (c) run a grid search if desired.

```
# cis_fit.py -- example
import numpy as np
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.metrics import precision_recall_curve, f1_scores

# X: Nx4 array of [T_p, S_p, E_o, S_s], y: Nx1 causal labels (0/1)
X = np.load('cis_features.npy')  # prepare from validation clips
y = np.load('cis_labels.npy')

# Train/test split
Xtr, Xte, ytr, yte = train_test_split(X, y, test_size=0.2, random_state=42)

# Train logistic regression (L2 regularized)
clf = LogisticRegression(penalty='l2', solver='lbfgs', max_iter=1000)
clf.fit(Xtr, ytr)

# Learned weights (coefficients)
print("Intercept:", clf.intercept_)
print("Coefficients:", clf.coef_)

# Predictions and threshold selection via PR
probs = clf.predict_proba(Xte)[:,1]
prec, rec, thr = precision_recall_curve(yte, probs)
f1_scores = 2 * prec * rec / (prec + rec + 1e-12)
best_idx = np.nanargmax(f1_scores)
best_threshold = thr[best_idx] if best_idx < len(thr) else 0.5
print("Best F1 (logistic):", f1_scores[best_idx], "Threshold:", best_threshold)
```

4. Experiments

4.1 Dataset

We evaluate on the Video Scene Graph (VSGR) dataset (Nguyen et al. 2025), comprising 1.9M frames across approximately 100K clips, with 20–30% egocentric footage from ASPIRe/AeroEye. VSGR provides annotations for entities (bounding boxes, classes), events (verb–noun actions), and relationships (spatial, temporal, causal) via its Relation Reasoning (RR) task. We subsample 50 egocentric clips (30–60 s, ∼5K frames) with causal dynamics (e.g., *pick-up* → *place*), using a validation split (10 clips) to tune HDBSCAN parameters. VSGR's graph-structured labels enable direct evaluation of subject–predicate–object (S–P–O) triples and causal edges, surpassing Ego4D's narration-based annotations (Grauman et al. 2022). No training is required; models (YOLO11x, CLIP-ViT-B/32, FastVLM-0.5B, HDBSCAN, Gemma3) are pretrained or unsupervised.

4.2 Evaluation Metrics

- **Triplet Precision/Recall/F1:** Measures S–P–O triple accuracy. A triple is correct if the subject, predicate, and object match VSGR ground truth with timestamps within ± 2 s. To avoid text overlap, the equations are typeset in display math:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

- **Causal Reasoning Score:** Precision/Recall/F1 for causal edges (e.g., "Event_A CAUSES Event_B"), using VSGR's RR annotations.
- **Entity Continuity:** Percentage of events with correct `entity_ids`, mapped via IoU (> 0.5) to ground-truth bounding boxes.

4.3 Baselines

- **HyperGLM (?):** Adapts hyperedge generation to produce S–P–O triples and causal edges on VSGR, leveraging its multimodal LLM and procedural rules.
- **Heuristic Uplift Engine:** Uses the same perception log but applies rules (e.g., proximity → "IS_NEAR," state change → "CAUSED").
- **LLM-Only Captions:** Feeds FastVLM captions (1 FPS) to Gemma3 without structured logs, mimicking VoT (**?**).

4.4 Experimental Setup

Each VSGR clip is processed to generate a perception log (YOLO11x detections, CLIP embeddings, FastVLM descriptions, HDBSCAN clusters). Logs are fed to the Semantic Uplift Engine, Heuristic Uplift Engine, HyperGLM, and LLM-Only Captions. Outputs are compared to VSGR ground truth. Ablations remove CLIP clustering, state detection, or structured prompting. Visualizations include UMAPs of embeddings and predicate confusion matrices. Statistical significance is tested via the Wilcoxon signed-rank test ($p < 0.05$).

5. Results and Analysis

Placeholder: Results pending. We hypothesize that our Semantic Uplift Engine achieves higher Triplet F1 and higher Causal Reasoning Score than baselines, driven by LLM reasoning and structured logs. Expected outcomes:

- **Table 1:** Triplet F1 and Causal Reasoning Scores on VSGR's egocentric subset, showing our pipeline outperforms HyperGLM and heuristic baselines due to nuanced causal inference.
- **Ablations:** Removing CLIP clustering, state detection, or prompting reduces F1 by ∼0.05–0.10, confirming component contributions.
- **Case Studies:** 3–5 clips illustrate LLM successes (e.g., "picked_up" vs. "knocked_over").

6. Limitations

Handling Novel and Out-of-Distribution Objects. While Orion's current implementation is designed around a closed-set detection model, the system incorporates

mechanisms that allow for limited generalization to novel or out-of-distribution (OOD) objects. During inference, objects that are not confidently classified by the primary detector are still processed through the CLIP encoder, which provides open-vocabulary embeddings in a shared image–text space. These embeddings are clustered independently of class labels, allowing the system to preserve unique entity identities and reason about their spatial or causal relationships even when categorical alignment is unknown. This effectively enables zero-shot representation of unseen objects, a property demonstrated empirically but not explicitly formalized in the original description. However, this mechanism remains constrained by the representational coverage and semantic grounding of the embedding model; completely novel visual concepts may fail to form coherent clusters or obtain accurate textual grounding. Future work will extend this mechanism through dynamic vocabulary expansion and continual embedding adaptation to improve robustness across unseen object domains.

7. Conclusion

Placeholder: To be expanded post-results. Our Semantic Uplift Engine advances egocentric video understanding by constructing causally accurate knowledge graphs from raw data, leveraging LLM reasoning and structured perception. Evaluations on VSGR aim to demonstrate superior triplet and causal accuracy, with implications for robotics and AR.

Introduction

Congratulations on having a paper selected for inclusion in an AAAI Press proceedings or technical report! This document details the requirements necessary to get your accepted paper published using PDFLaTeX. If you are using Microsoft Word, instructions are provided in a different document. AAAI Press does not support any other formatting software.

The instructions herein are provided as a general guide for experienced LaTeX users. If you do not know how to use LaTeX, please obtain assistance locally. AAAI cannot provide you with support and the accompanying style files are **not** guaranteed to work. If the results you obtain are not in accordance with the specifications you received, you must correct your source file to achieve the correct result.

These instructions are generic. Consequently, they do not include specific dates, page charges, and so forth. Please consult your specific written conference instructions for details regarding your submission. Please review the entire document for specific instructions that might apply to your particular situation. All authors must comply with the following:

- You must use the 2026 AAAI Press LaTeX style file and the aaai2026.bst bibliography style files, which are located in the 2026 AAAI Author Kit (aaai2026.sty, aaai2026.bst).
- You must complete, sign, and return by the deadline the AAAI copyright form (unless directed by AAAI Press to use the AAAI Distribution License instead).
- You must read and format your paper source and PDF according to the formatting instructions for authors.
- You must submit your electronic files and abstract using our electronic submission form **on time.**
- You must pay any required page or formatting charges to AAAI Press so that they are received by the deadline.
- You must check your paper before submitting it, ensuring that it compiles without error, and complies with the guidelines found in the AAAI Author Kit.

Copyright

All papers submitted for publication by AAAI Press must be accompanied by a valid signed copyright form. They must also contain the AAAI copyright notice at the bottom of the first page of the paper. There are no exceptions to these requirements. If you fail to provide us with a signed copyright form or disable the copyright notice, we will be unable to publish your paper. There are **no exceptions** to this policy. You will find a PDF version of the AAAI copyright form in the AAAI AuthorKit. Please see the specific instructions for your conference for submission details.

Formatting Requirements in Brief

We need source and PDF files that can be used in a variety of ways and can be output on a variety of devices. The design and appearance of the paper is strictly governed by the aaai style file (aaai2026.sty). **You must not make any changes to the aaai style file, nor use any commands, packages, style files, or macros within your own paper that alter that design, including, but not limited to spacing, floats, margins, fonts, font size, and appearance.** AAAI imposes requirements on your source and PDF files that must be followed. Most of these requirements are based on our efforts to standardize conference manuscript properties and layout. All papers submitted to AAAI for publication will be recompiled for standardization purposes. Consequently, every paper submission must comply with the following requirements:

- Your .tex file must compile in PDFLaTeX — (you may not include .ps or .eps figure files.)
- All fonts must be embedded in the PDF file — including your figures.
- Modifications to the style file, whether directly or via commands in your document may not ever be made, most especially when made in an effort to avoid extra page charges or make your paper fit in a specific number of pages.
- No type 3 fonts may be used (even in illustrations).
- You may not alter the spacing above and below captions, figures, headings, and subheadings.
- You may not alter the font sizes of text elements, footnotes, heading elements, captions, or title information (for references and mathematics, please see the limited exceptions provided herein).
- You may not alter the line spacing of text.

- Your title must follow Title Case capitalization rules (not sentence case).
- LaTeX documents must use the Times or Nimbus font package (you may not use Computer Modern for the text of your paper).
- No LaTeX 209 documents may be used or submitted.
- Your source must not require use of fonts for non-Roman alphabets within the text itself. If your paper includes symbols in other languages (such as, but not limited to, Arabic, Chinese, Hebrew, Japanese, Thai, Russian and other Cyrillic languages), you must restrict their use to bit-mapped figures. Fonts that require non-English language support (CID and Identity-H) must be converted to outlines or 300 dpi bitmap or removed from the document (even if they are in a graphics file embedded in the document).
- Two-column format in AAAI style is required for all papers.
- The paper size for final submission must be US letter without exception.
- The source file must exactly match the PDF.
- The document margins may not be exceeded (no overfull boxes).
- The number of pages and the file size must be as specified for your event.
- No document may be password protected.
- Neither the PDFs nor the source may contain any embedded links or bookmarks (no hyperref or navigator packages).
- Your source and PDF must not have any page numbers, footers, or headers (no pagestyle commands).
- Your PDF must be compatible with Acrobat 5 or higher.
- Your LaTeX source file (excluding references) must consist of a **single** file (use of the "input" command is not allowed.
- Your graphics must be sized appropriately outside of LaTeX (do not use the "clip" or "trim" command) .

If you do not follow these requirements, your paper will be returned to you to correct the deficiencies.

What Files to Submit

You must submit the following items to ensure that your paper is published:

- A fully-compliant PDF file.
- Your LaTeX source file submitted as a **single** .tex file (do not use the "input" command to include sections of your paper — every section must be in the single source file). (The only allowable exception is .bib file, which should be included separately).
- The bibliography (.bib) file(s).
- Your source must compile on our system, which includes only standard LaTeX 2020 TeXLive support files.
- Only the graphics files used in compiling paper.
- The LaTeX-generated files (e.g. .aux, .bbl file, PDF, etc.).

Your LaTeX source will be reviewed and recompiled on our system (if it does not compile, your paper will be returned to you. **Do not submit your source in multiple text files.** Your single LaTeX source file must include all your text, your bibliography (formatted using aaai2026.bst), and any custom macros.

Your files should work without any supporting files (other than the program itself) on any computer with a standard LaTeX distribution.

Do not send files that are not actually used in the paper. Avoid including any files not needed for compiling your paper, including, for example, this instructions file, unused graphics files, style files, additional material sent for the purpose of the paper review, intermediate build files and so forth. **Obsolete style files.** The commands for some common packages (such as some used for algorithms), may have changed. Please be certain that you are not compiling your paper using old or obsolete style files.

Final Archive. Place your source files in a single archive which should be compressed using .zip. The final file size may not exceed 10 MB. Name your source file with the last (family) name of the first author, even if that is not you.

Using LaTeX to Format Your Paper

The latest version of the AAAI style file is available on AAAI's website. Download this file and place it in the TeX search path. Placing it in the same directory as the paper should also work. You must download the latest version of the complete AAAI Author Kit so that you will have the latest instruction set and style file.

Document Preamble

In the LaTeX source for your paper, you **must** place the following lines as shown in the example in this subsection. This command set-up is for three authors. Add or subtract author and address lines as necessary, and uncomment the portions that apply to you. In most instances, this is all you need to do to format your paper in the Times font. The helvet package will cause Helvetica to be used for sans serif. These files are part of the PSNFSS2e package, which is freely available from many Internet sites (and is often part of a standard installation).

Leave the setcounter for section number depth commented out and set at 0 unless you want to add section numbers to your paper. If you do add section numbers, you must uncomment this line and change the number to 1 (for section numbers), or 2 (for section and subsection numbers). The style file will not work properly with numbering of subsubsections, so do not use a number higher than 2.

The Following Must Appear in Your Preamble

```
\documentclass[letterpaper]{article}
% DO NOT CHANGE THIS
\usepackage{aaai2026} % DO NOT CHANGE THIS
\usepackage{times} % DO NOT CHANGE THIS
\usepackage{helvet} % DO NOT CHANGE THIS
\usepackage{courier} % DO NOT CHANGE THIS
\usepackage[hyphens]{url} % DO NOT CHANGE THIS
\usepackage{graphicx} % DO NOT CHANGE THIS
```

```
\urlstyle{rm} % DO NOT CHANGE THIS
\def\UrlFont{\rm} % DO NOT CHANGE THIS
\usepackage{graphicx}  % DO NOT CHANGE THIS
\usepackage{natbib}  % DO NOT CHANGE THIS
\usepackage{caption}  % DO NOT CHANGE THIS
\frenchspacing % DO NOT CHANGE THIS
\setlength{\pdfpagewidth}{8.5in} % DO NOT CHANGE THIS
\setlength{\pdfpageheight}{11in} % DO NOT CHANGE THIS
%
% Keep the \pdfinfo as shown here. There's no need
% for you to add the /Title and /Author tags.
\pdfinfo{
/TemplateVersion (2026.1)
}
```

Preparing Your Paper

After the preamble above, you should prepare your paper as follows:

```
\begin{document}
\maketitle
\begin{abstract}
%...
\end{abstract}
```

If you want to add links to the paper's code, dataset(s), and extended version or similar this is the place to add them, within a *links* environment:

```
\begin{links}
  \link{Code}{https://aaai.org/example/guidelines}
  \link{Datasets}{https://aaai.org/example/datasets}
  \link{Extended version}{https://aaai.org/example}
\end{links}
```

You should then continue with the body of your paper. Your paper must conclude with the references, which should be inserted as follows:

```
% References and End of Paper
% These lines must be placed at the end of your paper
\bibliography{Bibliography-File}
\end{document}
```

```
\begin{document}\\
\maketitle\\
...\\
\bibliography{Bibliography-File}\\
\end{document}\\
```

Commands and Packages That May Not Be Used

There are a number of packages, commands, scripts, and macros that are incompatable with aaai2026.sty. The common ones are listed in tables 1 and 2. Generally, if a command, package, script, or macro alters floats, margins, fonts, sizing, linespacing, or the presentation of the references and citations, it is unacceptable. Note that negative vskip and vspace may not be used except in certain rare occurances, and may never be used around tables, figures, captions, sections, subsections, subsubsections, or references.

Page Breaks

For your final camera ready copy, you must not use any page break commands. References must flow directly after the text without breaks. Note that some conferences require references to be on a separate page during the review process. AAAI Press, however, does not require this condition for the final paper.

Paper Size, Margins, and Column Width

Papers must be formatted to print in two-column format on 8.5 x 11 inch US letter-sized paper. The margins must be exactly as follows:

- Top margin: .75 inches
- Left margin: .75 inches
- Right margin: .75 inches
- Bottom margin: 1.25 inches

The default paper size in most installations of LaTeX is A4. However, because we require that your electronic paper be formatted in US letter size, the preamble we have provided includes commands that alter the default to US letter size. Please note that using any other package to alter page size (such as, but not limited to the Geometry package) will result in your final paper being returned to you for correction.

Column Width and Margins. To ensure maximum readability, your paper must include two columns. Each column should be 3.3 inches wide (slightly more than 3.25 inches), with a .375 inch (.952 cm) gutter of white space between the two columns. The aaai2026.sty file will automatically create these columns for you.

Overlength Papers

If your paper is too long and you resort to formatting tricks to make it fit, it is quite likely that it will be returned to you. The best way to retain readability if the paper is overlength is to cut text, figures, or tables. There are a few acceptable ways to reduce paper size that don't affect readability. First, turn on \frenchspacing, which will reduce the space after periods. Next, move all your figures and tables to the top of the page. Consider removing less important portions of a figure. If you use \centering instead of \begin{center} in your figure environment, you can also buy some space. For mathematical environments, you may reduce fontsize **but not below 6.5 point**.

Commands that alter page layout are forbidden. These include \columnsep, \float, \topmargin, \topskip, \textheight, \textwidth, \oddsidemargin, and \evensizemargin (this list is not exhaustive). If you alter page layout, you will be required to pay the page fee. Other commands that are questionable and may cause your paper to be rejected include \parindent, and \parskip. Commands that alter the space between sections are forbidden. The title sec package is not allowed. Regardless of the above, if your paper is obviously "squeezed" it is not going to to be accepted. Options for reducing the length of a paper include reducing the size of your graphics, cutting text, or paying the extra page charge (if it is offered).

Type Font and Size

Your paper must be formatted in Times Roman or Nimbus. We will not accept papers formatted using Computer Modern or Palatino or some other font as the text or heading type-

\abovecaption	\abovedisplay	\addevensidemargin	\addsidemargin
\addtolength	\baselinestretch	\belowcaption	\belowdisplay
\break	\clearpage	\clip	\columnsep
\float	\input	\input	\linespread
\newpage	\pagebreak	\renewcommand	\setlength
\text height	\tiny	\top margin	\trim
\vskip{-	\vspace{-		

Table 1: Commands that must not be used

authblk	babel	cjk	dvips
epsf	epsfig	euler	float
fullpage	geometry	graphics	hyperref
layout	linespread	lmodern	maltepaper
navigator	pdfcomment	pgfplots	psfig
pstricks	t1enc	titlesec	tocbind
ulem			

Table 2: LaTeX style packages that must not be used.

face. Sans serif, when used, should be Courier. Use Symbol or Lucida or Computer Modern for *mathematics only.*

Do not use type 3 fonts for any portion of your paper, including graphics. Type 3 bitmapped fonts are designed for fixed resolution printers. Most print at 300 dpi even if the printer resolution is 1200 dpi or higher. They also often cause high resolution imagesetter devices to crash. Consequently, AAAI will not accept electronic files containing obsolete type 3 fonts. Files containing those fonts (even in graphics) will be rejected. (Authors using blackboard symbols must avoid packages that use type 3 fonts.)

Fortunately, there are effective workarounds that will prevent your file from embedding type 3 bitmapped fonts. The easiest workaround is to use the required times, helvet, and courier packages with LaTeX2e. (Note that papers formatted in this way will still use Computer Modern for the mathematics. To make the math look good, you'll either have to use Symbol or Lucida, or you will need to install type 1 Computer Modern fonts — for more on these fonts, see the section "Obtaining Type 1 Computer Modern.")

If you are unsure if your paper contains type 3 fonts, view the PDF in Acrobat Reader. The Properties/Fonts window will display the font name, font type, and encoding properties of all the fonts in the document. If you are unsure if your graphics contain type 3 fonts (and they are PostScript or encapsulated PostScript documents), create PDF versions of them, and consult the properties window in Acrobat Reader.

The default size for your type must be ten-point with twelve-point leading (line spacing). Start all pages (except the first) directly under the top margin. (See the next section for instructions on formatting the title page.) Indent ten points when beginning a new paragraph, unless the paragraph begins directly below a heading or subheading.

Obtaining Type 1 Computer Modern for LaTeX. If you use Computer Modern for the mathematics in your paper (you cannot use it for the text) you may need to download type 1 Computer fonts. They are available without charge from the American Mathematical Society: http://www.ams.org/tex/type1-fonts.html.

Nonroman Fonts. If your paper includes symbols in other languages (such as, but not limited to, Arabic, Chinese, Hebrew, Japanese, Thai, Russian and other Cyrillic languages), you must restrict their use to bit-mapped figures.

Title and Authors

Your title must appear centered over both text columns in sixteen-point bold type (twenty-four point leading). The title must be written in Title Case according to the Chicago Manual of Style rules. The rules are a bit involved, but in general verbs (including short verbs like be, is, using, and go), nouns, adverbs, adjectives, and pronouns should be capitalized, (including both words in hyphenated terms), while articles, conjunctions, and prepositions are lower case unless they directly follow a colon or long dash. You can use the online tool https://titlecaseconverter.com/ to double-check the proper capitalization (select the "Chicago" style and mark the "Show explanations" checkbox).

Author's names should appear below the title of the paper, centered in twelve-point type (with fifteen point leading), along with affiliation(s) and complete address(es) (including electronic mail address if available) in nine-point roman type (the twelve point leading). You should begin the two-column format when you come to the abstract.

Formatting Author Information. Author information has to be set according to the following specification depending if you have one or more than one affiliation. You may not use a table nor may you employ the \authorblk.sty package. For one or several authors from the same institution, please separate them with commas and write all affiliation directly below (one affiliation per line) using the macros \author and \affiliations:

```
\author{
    Author 1, ..., Author n\\
}
\affiliations {
    Address line\\
    ... \\
    Address line\\
}
```

For authors from different institutions, use \textsuperscript {\rm x } to match authors and affiliations. Notice that there should not be any spaces between the author name and the superscript (and the comma should come after the superscripts).

```
\author{
    AuthorOne\equalcontrib\textsuperscript{\rm 1,\rm2},
    AuthorTwo\equalcontrib\textsuperscript{\rm 2},
    AuthorThree\textsuperscript{\rm 3},\\
    AuthorFour\textsuperscript{\rm 4},
    AuthorFive\textsuperscript{\rm 5}}
}
\affiliations {
    \textsuperscript{\rm 1}AffiliationOne,\\
    \textsuperscript{\rm 2}AffiliationTwo,\\
    \textsuperscript{\rm 3}AffiliationThree,\\
    \textsuperscript{\rm 4}AffiliationFour,\\
    \textsuperscript{\rm 5}AffiliationFive\\
    \{email, email\}@affiliation.com,
    email@affiliation.com,
    email@affiliation.com,
    email@affiliation.com
}
```

You can indicate that some authors contributed equally using the \equalcontrib command. This will add a marker after the author names and a footnote on the first page.

Note that you may want to break the author list for better visualization. You can achieve this using a simple line break (\\).

LaTeX Copyright Notice

The copyright notice automatically appears if you use aaai2026.sty. It has been hardcoded and may not be disabled.

Credits

Any credits to a sponsoring agency should appear in the acknowledgments section, unless the agency requires different placement. If it is necessary to include this information on the front page, use \thanks in either the \author or \title commands. For example:

> \title{Very Important Results in AI\thanks{This work is supported by everybody.}}

Multiple \thanks commands can be given. Each will result in a separate footnote indication in the author or title with the corresponding text at the botton of the first column of the document. Note that the \thanks command is fragile. You will need to use \protect.

Please do not include \pubnote commands in your document.

Abstract

Follow the example commands in this document for creation of your abstract. The command \begin{abstract} will automatically indent the text block. Please do not indent it further. Do not include references in your abstract!

Page Numbers

Do not print any page numbers on your paper. The use of \pagestyle is forbidden.

Text

The main body of the paper must be formatted in black, ten-point Times Roman with twelve-point leading (line spacing). You may not reduce font size or the linespacing. Commands that alter font size or line spacing (including, but not limited to baselinestretch, baselineshift, linespread, and others) are expressly forbidden. In addition, you may not use color in the text.

Citations

Citations within the text should include the author's last name and year, for example (Newell 1980). Append lowercase letters to the year in cases of ambiguity. Multiple authors should be treated as follows: (Feigenbaum and Engelmore 1988) or (Ford, Hayes, and Glymour 1992). In the case of four or more authors, list only the first author, followed by et al. (Ford et al. 1997).

Extracts

Long quotations and extracts should be indented ten points from the left and right margins.

> This is an example of an extract or quotation. Note the indent on both sides. Quotation marks are not necessary if you offset the text in a block like this, and properly identify and cite the quotation in the text.

Footnotes

Use footnotes judiciously, taking into account that they interrupt the reading of the text. When required, they should be consecutively numbered throughout with superscript Arabic numbers. Footnotes should appear at the bottom of the page, separated from the text by a blank line space and a thin, half-point rule.

Headings and Sections

When necessary, headings should be used to separate major sections of your paper. Remember, you are writing a short paper, not a lengthy book! An overabundance of headings will tend to make your paper look more like an outline than a paper. The aaai2026.sty package will create headings for you. Do not alter their size nor their spacing above or below.

Section Numbers. The use of section numbers in AAAI Press papers is optional. To use section numbers in LaTeX, uncomment the setcounter line in your document preamble and change the 0 to a 1. Section numbers should not be used in short poster papers and/or extended abstracts.

Section Headings. Sections should be arranged and headed as follows:

1. Main content sections
2. Appendices (optional)
3. Ethical Statement (optional, unnumbered)
4. Acknowledgements (optional, unnumbered)
5. References (unnumbered)

Appendices. Any appendices must appear after the main content. If your main sections are numbered, appendix sections must use letters instead of arabic numerals. In LaTeX you can use the `\appendix` command to achieve this effect and then use `\section{Heading}` normally for your appendix sections.

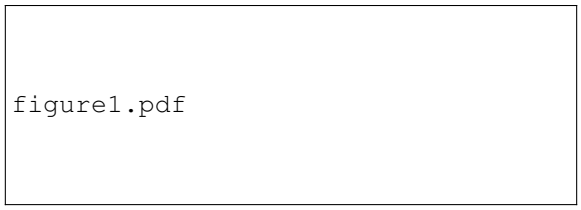

Figure 1: Using the trim and clip commands produces fragile layers that can result in disasters (like this one from an actual paper) when the color space is corrected or the PDF combined with others for the final proceedings. Crop your figures properly in a graphics program – not in LaTeX.

Ethical Statement. You can write a statement about the potential ethical impact of your work, including its broad societal implications, both positive and negative. If included, such statement must be written in an unnumbered section titled *Ethical Statement*.

Acknowledgments. The acknowledgments section, if included, appears right before the references and is headed "Acknowledgments". It must not be numbered even if other sections are (use `\section*{Acknowledgements}` in LaTeX). This section includes acknowledgments of help from associates and colleagues, credits to sponsoring agencies, financial support, and permission to publish. Please acknowledge other contributors, grant support, and so forth, in this section. Do not put acknowledgments in a footnote on the first page. If your grant agency requires acknowledgment of the grant on page 1, limit the footnote to the required statement, and put the remaining acknowledgments at the back. Please try to limit acknowledgments to no more than three sentences.

References. The references section should be labeled "References" and must appear at the very end of the paper (don't end the paper with references, and then put a figure by itself on the last page). A sample list of references is given later on in these instructions. Please use a consistent format for references. Poorly prepared or sloppy references reflect badly on the quality of your paper and your research. Please prepare complete and accurate citations.

Illustrations and Figures

Your paper must compile in PDFLaTeX. Consequently, all your figures must be .jpg, .png, or .pdf. You may not use the .gif (the resolution is too low), .ps, or .eps file format for your figures.

Figures, drawings, tables, and photographs should be placed throughout the paper on the page (or the subsequent page) where they are first discussed. Do not group them together at the end of the paper. If placed at the top of the paper, illustrations may run across both columns. Figures must not invade the top, bottom, or side margin areas. Figures must be inserted using the \usepackage{graphicx}. Number figures sequentially, for example, figure 1, and so on. Do not use minipage to group figures.

If you normally create your figures using pgfplots, please create the figures first, and then import them as pdfs with proper bounding boxes, as the bounding and trim boxes created by pfgplots are fragile and not valid.

When you include your figures, you must crop them **outside** of LaTeX. The command \includegraphics*[clip=true, viewport 0 0 10 10]... might result in a PDF that looks great, but the image is **not really cropped.** The full image can reappear (and obscure whatever it is overlapping) when page numbers are applied or color space is standardized. Figures 1, and 2 display some unwanted results that often occur.

If your paper includes illustrations that are not compatible with PDFTeX (such as .eps or .ps documents), you will need to convert them. The epstopdf package will usually work for eps files. You will need to convert your ps files to PDF in either case.

Figure Captions. The illustration number and caption must appear *under* the illustration. Labels and other text with the actual illustration must be at least nine-point type. However, the font and size of figure captions must be 10 point roman. Do not make them smaller, bold, or italic. (Individual words may be italicized if the context requires differentiation.)

Tables

Tables should be presented in 10 point roman type. If necessary, they may be altered to 9 point type. You must not use `\resizebox` or other commands that resize the entire table to make it smaller, because you can't control the final font size this way. If your table is too large you can use `\setlength{\tabcolsep}{1mm}` to compress the columns a bit or you can adapt the content (e.g.: reduce the decimal precision when presenting numbers, use shortened column titles, make some column duble-line to get it narrower).

Tables that do not fit in a single column must be placed across double columns. If your table won't fit within the margins even when spanning both columns and using the above techniques, you must split it in two separate tables.

Table Captions. The number and caption for your table must appear *under* (not above) the table. Additionally, the font and size of table captions must be 10 point roman and must be placed beneath the figure. Do not make them smaller, bold, or italic. (Individual words may be italicized if the context requires differentiation.)

Low-Resolution Bitmaps. You may not use low-resolution (such as 72 dpi) screen-dumps and GIF files—these files contain so few pixels that they are always blurry, and illegible when printed. If they are color, they will become an indecipherable mess when converted to black and white. This is always the case with gif files, which should never be used. The resolution of screen dumps can be increased by reducing the print size of the original file while retaining the same number of pixels. You can also enlarge files by manipulating them in software such as PhotoShop. Your figures should be 300 dpi when incorporated into your document.

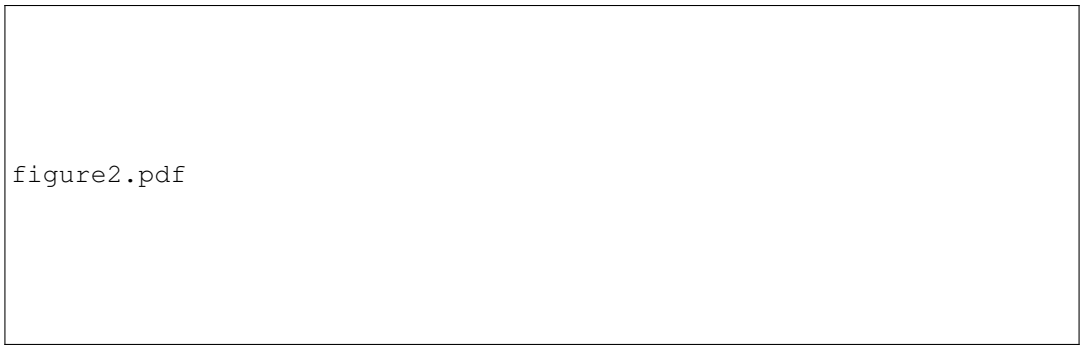

Figure 2: Adjusting the bounding box instead of actually removing the unwanted data resulted multiple layers in this paper. It also needlessly increased the PDF size. In this case, the size of the unwanted layer doubled the paper's size, and produced the following surprising results in final production. Crop your figures properly in a graphics program. Don't just alter the bounding box.

LaTeX Overflow. LaTeX users please beware: LaTeX will sometimes put portions of the figure or table or an equation in the margin. If this happens, you need to make the figure or table span both columns. If absolutely necessary, you may reduce the figure, or reformat the equation, or reconfigure the table. **Check your log file!** You must fix any overflow into the margin (that means no overfull boxes in LaTeX). **Nothing is permitted to intrude into the margin or gutter.**

Using Color. Use of color is restricted to figures only. It must be WACG 2.0 compliant. (That is, the contrast ratio must be greater than 4.5:1 no matter the font size.) It must be CMYK, NOT RGB. It may never be used for any portion of the text of your paper. The archival version of your paper will be printed in black and white and grayscale. The web version must be readable by persons with disabilities. Consequently, because conversion to grayscale can cause undesirable effects (red changes to black, yellow can disappear, and so forth), we strongly suggest you avoid placing color figures in your document. If you do include color figures, you must (1) use the CMYK (not RGB) colorspace and (2) be mindful of readers who may happen to have trouble distinguishing colors. Your paper must be decipherable without using color for distinction.

Drawings. We suggest you use computer drawing software (such as Adobe Illustrator or, (if unavoidable), the drawing tools in Microsoft Word) to create your illustrations. Do not use Microsoft Publisher. These illustrations will look best if all line widths are uniform (half- to two-point in size), and you do not create labels over shaded areas. Shading should be 133 lines per inch if possible. Use Times Roman or Helvetica for all figure call-outs. **Do not use hairline width lines** — be sure that the stroke width of all lines is at least .5 pt. Zero point lines will print on a laser printer, but will completely disappear on the high-resolution devices used by our printers.

Photographs and Images. Photographs and other images should be in grayscale (color photographs will not reproduce well; for example, red tones will reproduce as black, yellow may turn to white, and so forth) and set to a minimum of 300 dpi. Do not prescreen images.

Resizing Graphics. Resize your graphics **before** you include them with LaTeX. You may **not** use trim or clip options as part of your \includegraphics command. Resize the media box of your PDF using a graphics program instead.

Fonts in Your Illustrations. You must embed all fonts in your graphics before including them in your LaTeX document.

Algorithms. Algorithms and/or programs are a special kind of figures. Like all illustrations, they should appear floated to the top (preferably) or bottom of the page. However, their caption should appear in the header, left-justified and enclosed between horizontal lines, as shown in Algorithm 1. The algorithm body should be terminated with another horizontal line. It is up to the authors to decide whether to show line numbers or not, how to format comments, etc.

In LaTeX algorithms may be typeset using the `algorithm` and `algorithmic` packages, but you can also use one of the many other packages for the task.

Listings. Listings are much like algorithms and programs. They should also appear floated to the top (preferably) or

Algorithm 1: Example algorithm

Input: Your algorithm's input
Parameter: Optional list of parameters
Output: Your algorithm's output

```
Let t = 0.
while condition do
    Do some action.
    if conditional then
        Perform task A.
    else
        Perform task B.
    end if
end while
return solution
```

Listing 1: Example listing `quicksort.hs`

```
quicksort :: Ord a => [a] -> [a]
quicksort []     = []
quicksort (p:xs) = (quicksort lesser) ++
     [p] ++ (quicksort greater)
  where
    lesser  = filter (< p) xs
    greater = filter (>= p) xs
```

bottom of the page. Listing captions should appear in the header, left-justified and enclosed between horizontal lines as shown in Listing 1. Terminate the body with another horizontal line and avoid any background color. Line numbers, if included, must appear within the text column.

References

The AAAI style includes a set of definitions for use in formatting references with BibTeX. These definitions make the bibliography style fairly close to the ones specified in the Reference Examples appendix below. To use these definitions, you also need the BibTeX style file "aaai2026.bst," available in the AAAI Author Kit on the AAAI web site. Then, at the end of your paper but before \enddocument, you need to put the following lines:

\bibliography{bibfile1,bibfile2,...}

Please note that the aaai2026.sty class already sets the bibliographystyle for you, so you do not have to place any \bibliographystyle command in the document yourselves. The aaai2026.sty file is incompatible with the hyperref and navigator packages. If you use either, your references will be garbled and your paper will be returned to you.

References may be the same size as surrounding text. However, in this section (only), you may reduce the size to *small* (9pt) if your paper exceeds the allowable number of pages. Making it any smaller than 9 point with 10 point linespacing, however, is not allowed.

The list of files in the \bibliography command should be the names of your BibTeX source files (that is, the .bib files referenced in your paper).

The following commands are available for your use in citing references:

cite: Cites the given reference(s) with a full citation. This appears as "(Author Year)" for one reference, or "(Author Year; Author Year)" for multiple references.

shortcite: Cites the given reference(s) with just the year. This appears as "(Year)" for one reference, or "(Year; Year)" for multiple references.

citeauthor: Cites the given reference(s) with just the author name(s) and no parentheses.

citeyear: Cites the given reference(s) with just the date(s) and no parentheses.

You may also use any of the *natbib* citation commands.

Proofreading Your PDF

Please check all the pages of your PDF file. The most commonly forgotten element is the acknowledgements — especially the correct grant number. Authors also commonly forget to add the metadata to the source, use the wrong reference style file, or don't follow the capitalization rules or comma placement for their author-title information properly. A final common problem is text (expecially equations) that runs into the margin. You will need to fix these common errors before submitting your file.

Improperly Formatted Files

In the past, AAAI has corrected improperly formatted files submitted by the authors. Unfortunately, this has become an increasingly burdensome expense that we can no longer absorb). Consequently, if your file is improperly formatted, it will be returned to you for correction.

Naming Your Electronic File

We require that you name your LaTeX source file with the last name (family name) of the first author so that it can easily be differentiated from other submissions. Complete file-naming instructions will be provided to you in the submission instructions.

Submitting Your Electronic Files to AAAI

Instructions on paper submittal will be provided to you in your acceptance letter.

Inquiries

If you have any questions about the preparation or submission of your paper as instructed in this document, please contact AAAI Press at the address given below. If you have technical questions about implementation of the aaai style file, please contact an expert at your site. We do not provide technical support for LaTeX or any other software package. To avoid problems, please keep your paper simple, and do not incorporate complicated macros and style files.

AAAI Press
1101 Pennsylvania Ave, NW Suite 300
Washington, DC 20004 USA
Telephone: 1-202-360-4062
E-mail: See the submission instructions for your particular conference or event.

Additional Resources

LaTeX is a difficult program to master. If you've used that software, and this document didn't help or some items were not explained clearly, we recommend you read Michael Shell's excellent document (testflow doc.txt V1.0a 2002/08/13) about obtaining correct PS/PDF output on LaTeX systems. (It was written for another purpose, but it has general application as well). It is available at www.ctan.org in the tex-archive.

Reference Examples

* Formatted bibliographies should look like the following examples. You should use BibTeX to generate the references. Missing fields are unacceptable when compiling ref-

erences, and usually indicate that you are using the wrong type of entry (BibTeX class).

Book with multiple authors Use the `@book` class.
em:86.

Journal and magazine articles Use the `@article` class.
r:80.
hcr:83.

Proceedings paper published by a society, press or publisher Use the `@inproceedings` class. You may abbreviate the *booktitle* field, but make sure that the conference edition is clear.
c:84.
c:83.

University technical report Use the `@techreport` class.
r:86.

Dissertation or thesis Use the `@phdthesis` class.
c:79.

Forthcoming publication Use the `@misc` class with a `note="Forthcoming"` annotation.

```
@misc(key,
  [...]
  note="Forthcoming",
)
```

c:21.

ArXiv paper Fetch the BibTeX entry from the "Export Bibtex Citation" link in the arXiv website. Notice it uses the `@misc` class instead of the `@article` one, and that it includes the `eprint` and `archivePrefix` keys.

```
@misc(key,
  [...]
  eprint="xxxx.yyyy",
  archivePrefix="arXiv",
)
```

c:22.

Website or online resource Use the `@misc` class. Add the url in the `howpublished` field and the date of access in the `note` field:

```
@misc(key,
  [...]
  howpublished="\url{http://...}",
  note="Accessed: YYYY-mm-dd",
)
```

c:23.

For the most up to date version of the AAAI reference style, please consult the *AI Magazine* Author Guidelines at https://aaai.org/ojs/index.php/aimagazine/about/submissions#authorGuidelines

Acknowledgments

AAAI is especially grateful to Peter Patel Schneider for his work in implementing the original aaai.sty file, liberally using the ideas of other style hackers, including Barbara Beeton. We also acknowledge with thanks the work of George Ferguson for his guide to using the style and BibTeX files — which has been incorporated into this document — and Hans Guesgen, who provided several timely modifications, as well as the many others who have, from time to time, sent in suggestions on improvements to the AAAI style. We are especially grateful to Francisco Cruz, Marc Pujol-Gonzalez, and Mico Loretan for the improvements to the BibTEX and LATEX files made in 2020.

The preparation of the LATEX and BibTEX files that implement these instructions was supported by Schlumberger Palo Alto Research, AT&T Bell Laboratories, Morgan Kaufmann Publishers, The Live Oak Press, LLC, and AAAI Press. Bibliography style changes were added by Sunil Issar. \pubnote was added by J. Scott Penberthy. George Ferguson added support for printing the AAAI copyright slug. Additional changes to aaai2026.sty and aaai2026.bst have been made by Francisco Cruz, Marc Pujol-Gonzalez, and Mico Loretan.

Thank you for reading these instructions carefully. We look forward to receiving your electronic files!

References

Cibula, K.; et al. 2025. Learning Low-Level Causal Relations Using a Simulated Robotic Arm. In *International Conference on Artificial Neural Networks (ICANN)*.

Dong, E.; et al. 2023. Adaptive Path-Memory Network for Temporal Knowledge Graph Reasoning. In *International Joint Conference on Artificial Intelligence (IJCAI)*.

Fei, L.; et al. 2024. Video-of-Thought: Step-by-Step Video Reasoning from Perception to Cognition. In *International Conference on Machine Learning (ICML)*.

Foss, A.; et al. 2025. CausalVQA: A Physically Grounded Causal Reasoning Benchmark for Video Models. *arXiv preprint arXiv:2506.09943*.

Grauman, K.; et al. 2022. Ego4D: Around the World in 3,000 Hours of Egocentric Video. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Grauman, K.; et al. 2024. EgoExo4D: Understanding Skilled Human Activity from First- and Third-Person. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Li, W.; et al. 2020. Causal Discovery in Physical Systems from Videos. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Nguyen, T.; et al. 2025. HyperGLM: HyperGraph for Video Scene Graph Generation and Anticipation. *arXiv preprint arXiv:2411.18042v2*.

Robinson, A.; et al. 2023. Robotic Vision for Human-Robot Interaction and Collaboration: A Survey and Systematic Review. *ACM Transactions on Human-Robot Interaction (THRI)*.

Rodin, J.; et al. 2023. Action Scene Graphs for Long-Form Understanding of Egocentric Videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Wang, H.; et al. 2023. LifelongMemory: Leveraging LLMs for Answering Queries in Long-form Egocentric Videos. *arXiv preprint arXiv:2312.05269*.

Xie, Y.; et al. 2025. A Comprehensive Survey on Video Scene Parsing: Advances, Challenges, and Prospects. *arXiv preprint arXiv:2506.13552*.