

A. Knowledge Graph Construction

Here we describe the rules for graph construction for our system, including node and edge type taxonomies, their extraction and relationship formation.

A.1. Entities as nodes

First let's define what object types we consider a node: ENTITY, EVENT, STATE, LOCATION, TIME and ATTRIBUTE.

Entity Type Assignment Entities are classified based on their NER labels (when present) into one of the aforementioned subtypes. When the label is not present or doesn't match any predefined subtype, we default to the ENTITY type.

Node Identification and Deduplication We map entity label to node IDs using the original- and lower-case variants. For entities not beginning with "the", we additionally map "the [entity]" to the same node ID to handle different references to the same entity to ensure consistency.

A.2. Relations as edges

As mentioned earlier in the document, the edge types we focus on are: PERFORMS, EXPERIENCES, TARGETS, LOCATED_IN, HAS_STATE and SAME_AS.

Verbs as events For tokens with VERB part-of-speech, we create EVENT nodes. We establish different relationship types based on following rules:

- `nsubj` (subject) creates a PERFORMS edge from subject to verb,
- `nsubjpass` (passive subject) creates an EXPERIENCES edge from verb to subject,
- `doobj` or `pobj` (direct/prepositional objects) creates TARGETS edges from verb to object.

Prepositions for locations When a verb has a child with `prep` dependency and text "in", "at", or "on", we create a LOCATED_IN edge from the verb to location. Similarly, for tokens with `prep` dependency and text "in" where both head and child exist in the node map, we create a LOCATED_IN edge from head to child.

Attribute and SAME_AS edges If a compound and its head are in the node map and we find the compound phrase exists, we create HAS_STATE edge from the compound to the head entity. And, the SAME_AS edge is used for co-references. This is most likely handled by the node deduplication step.

B. Data Licenses

- **Factify:** CC BY 4.0
- **MMFakeBench:** CC BY-NC-SA 4.0

All data provided under this work is licensed under CC BY-NC-SA 4.0 because the data we originally use are released under this license.

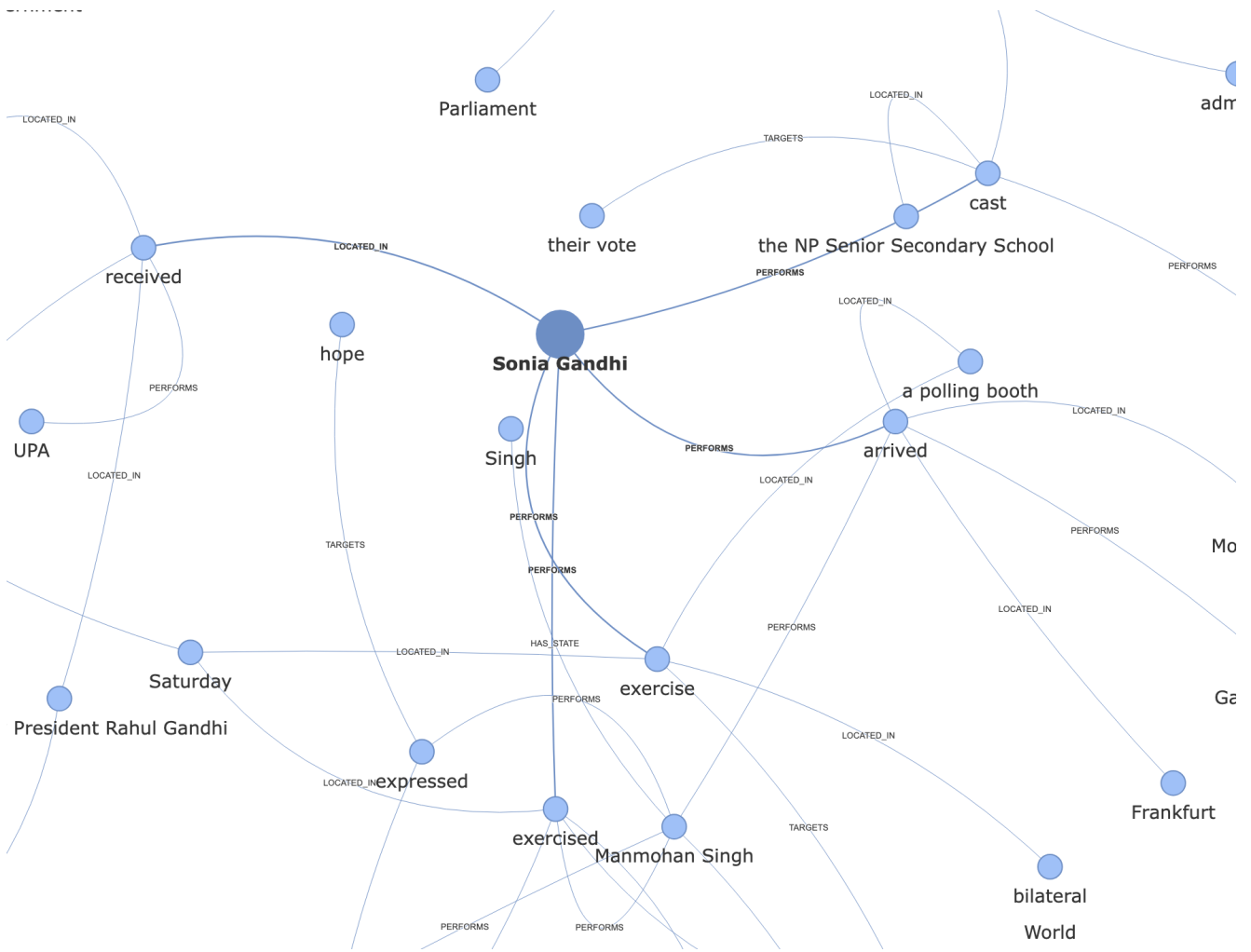


Figure 4. Evidence graph generated by EGMMG for the example in Figure 2

Provide an exhaustive description of the attached image covering all observable elements:

Visual Content:

- Describe all people, objects, structures, and their spatial relationships in detail
- If popular figures or landmarks are identified, make sure to also include them; be specific rather than vague
- Note clothing styles, technology visible, architectural features, vehicles
- Identify any text, signs, logos, watermarks, or written content (transcribe exactly)
- Describe lighting conditions, weather, time of day indicators
- Note crowd sizes, scale of events, and scope of any activities shown

Contextual Analysis:

- What specific type of event, location, or situation does this appear to document?
- Does this look like documentation of a real event or more like stock photography/illustration?
- What time period does this image suggest based on visual cues?
- Are there geographic indicators (license plates, signage, architecture, vegetation)?
- What season/weather conditions are evident?

Technical Assessment:

- Does the image quality, resolution, and compression appear consistent throughout?
- Are there any visual inconsistencies in lighting, shadows, or perspective?
- Are there compression artifacts or quality variations that seem unusual?
- Does this appear to be a screenshot, social media post, or real-world photography?

Plausibility and Consistency:

- Do all elements in the scene make logical sense together?
- Are there any anachronisms or elements that don't fit the apparent context?
- Are there any details that seem inconsistent with the overall narrative the image suggests?

Specificity Level:

- How specific vs. generic is this image? Could it represent multiple different events/locations?
- What unique identifying features, if any, make this image distinctive?

Provide as much detail as possible. Do not make any judgment whatsoever regarding the veracity of the image; just describe the image.

Figure 5. Prompt used to generate VLM description of images.

You are an expert at identifying fact-checking content. Your job is to determine if an article contains fact-checking information or verdicts about claims.

Look for these indicators:

- Explicit verdicts like "FALSE", "TRUE", "MISLEADING", "PARTIALLY TRUE"
- Fact-checking language like "CLAIM:", "VERDICT:", "RATING:", "RULED:"
- Debunking language like "debunked", "misinformation", "false claim"
- References to fact-checking organizations making determinations
- Structured fact-check formats with claims and conclusions

Respond with exactly "YES" if the article contains fact-checking content/verdicts, or "NO" if it's regular news/content.

Analyze this article text:

{article_text}

Does this article contain fact-checking content or verdicts?

Figure 6. Prompt used to filter out evidence documents from fact-checking web portals.

You are a fact-checking assistant tasked with evaluating the accuracy of a claim based on evidence provided. Your goal is to determine whether the claim is true or false based solely on the evidence. Do not consider external knowledge or information not included in the evidence.

Instructions:

1. Carefully read the evidence document, which consists of excerpts from multiple news articles.
2. Analyze the claim provided and compare it to the evidence.

CLAIM: "{caption}"

EVIDENCE:
{evidence}

Task: Determine if the claim is TRUE or MISINFORMATION based solely on the evidence provided.

Provide your response in the following JSON format: { "verdict": "true" or "misinformation", "confidence": 0.0 to 1.0, "reasoning": "brief explanation of your decision" }

Response:

Figure 7. Prompt used to evaluate misinformation detection performance of LLMs.



Claim: Finance Minister Nirmala Sitharaman announced that the government is committed to the goal of doubling farmer income.

Text: Fm Nirmala Sitharaman Budget 2020 What Is There For Education And Skills 2025 LIVE — 2025 Union Budget, India Central Budget FM Nirmala Sitharaman Speech on Kendra budget in Kannada. Read latest and breaking news of pm modi 4.0 budget on income tax slab, income tax returns, ITR filing, analysis and highlights at Vijaya Karnataka.

****Visual Content:****

The image depicts a formal setting, likely a legislative or parliamentary chamber. The central figure is a woman seated at a desk, dressed in a bright yellow saree with a subtle pattern. She wears glasses with a reddish-brown frame, a bindi on her forehead, and a simple necklace. Her hair is neatly tied back. She appears to be focused on reading or writing, with a document in front of her. The desk is equipped with multiple microphones, suggesting she is addressing or preparing to address the assembly.

To her left (viewer's right), a man with a bald head and a white mustache is visible. He is wearing a light purple shirt and appears to be observing the proceedings. Behind him, another man with gray hair and a white beard is partially visible, dressed in a dark suit. The background shows rows of wooden desks and chairs, typical of a parliamentary setup, with other individuals seated, though their faces are not clearly visible.

The lighting in the room is bright and even, indicative of indoor artificial lighting. The atmosphere appears formal and orderly, with no signs of disorder or casual behavior. There are no visible windows, weather indicators, or outdoor elements.

****Contextual Analysis:****

The setting strongly suggests a legislative or parliamentary session, likely in a country where sarees are part of traditional attire, such as India. The presence of microphones and documents indicates an official proceeding, possibly a debate, discussion, or presentation. The attire of the central figure and the formal setup align with the norms of such institutions.

(1817 more characters without spaces)

Original label: Support_Multimodal
EGMMG prediction: Not misinformation

Figure 8. A consistently correctly predicted image-claim-text sample, across several runs.



Claim: Punjab passengers are stranded as flight operations at Amritsar Airport have been suspended.

Text: ****Visual Content:****

The image depicts a bustling outdoor scene in front of a building labeled "AAI & BCAS INTEGRATED OFFICE COMPLEX" in both English and Hindi. The sign is prominently displayed in the upper right corner of the image. The crowd consists of approximately 30-40 individuals, primarily men, with a few women and children. The people are dressed in a mix of traditional and casual attire, with some wearing head coverings such as turbans, caps, and scarves.

In the foreground, a woman in a yellow top and dark pants is pushing a luggage cart loaded with a red suitcase, a pink and white checkered bag, and other items. Beside her, a man in a blue and white checkered shirt is pulling a black suitcase on wheels. To the left, a security officer in a khaki uniform and black beret is walking away from the camera, his back turned to the viewer.

The crowd appears to be waiting in line or gathering near the entrance of the building. Some individuals are standing, while others are seated on a low wall or curb. The atmosphere seems orderly but busy, with people engaged in conversation or looking around.

The background features a mix of greenery, including trees and shrubs, and a clear sky, suggesting it is daytime. The lighting is bright, indicating it is likely midday or early afternoon. The weather appears to be clear and sunny.

****Contextual Analysis:****

The image appears to document a real event, likely at an airport or a related facility, given the reference to AAI (Airports Authority of India) and BCAS (Bureau of Civil Aviation Security). The presence of luggage and the diverse crowd suggest this could be a check-in area, arrival point, or a security checkpoint.

The clothing styles and the mix of traditional and modern attire suggest the image is contemporary, likely taken in the 21st century. The presence of Hindi and English on the sign indicates the location is in India. The season appears to be warm, possibly spring or summer, given the clothing and the clear weather.

(1329 more characters without spaces)

Original label: Support_Multimodal
EGMMG prediction: Not_misinformation

Figure 9. A consistently correctly predicted image-claim-text sample, across several runs.



Claim: The Tanzanian government approved a herbal medicine called Covidol to cure patients suffering from COVID.

Text: ****Visual Content:****

The image depicts a collection of brown plastic bottles filled with a brown liquid, each labeled with the brand name "COVIDOL" in bold, white uppercase letters on a yellow and green label. The label also includes the volume "500 ml" and a graphic of a banana, suggesting the product may be banana-flavored or contain banana as an ingredient. The bottles are arranged in two rows: the front row has three bottles, and the back row has four bottles, all placed on a white surface. The bottles in the front row have white plastic caps, while the caps of the bottles in the back row are not clearly visible.

The background features green foliage, including large leaves that resemble those of tropical plants, such as banana or palm trees. The bottles are positioned on a woven, brown surface that appears to be a wicker or rattan tray or mat, adding a natural, rustic element to the presentation.

****Contextual Analysis:****

The image appears to document a product display, likely for marketing or sales purposes. The product, "COVIDOL," is presented in a way that emphasizes its natural or tropical qualities, as indicated by the banana graphic and the green foliage in the background. The name "COVIDOL" could suggest a product related to health or wellness, possibly playing on the word "COVID" (as in COVID-19) combined with "DOL," which might imply a soothing or medicinal quality. However, without additional context, the exact nature of the product remains speculative.

The setting does not resemble a specific geographic location but rather a staged environment designed to evoke a natural, tropical vibe. The use of green foliage and a woven surface suggests an attempt to associate the product with nature and freshness.

(1413 more characters without spaces)

Original label: Refute

EGMMG prediction: Not misinformation

Figure 10. A consistently incorrectly predicted image-claim-text sample, across several runs.



Claim: A video shows an eagle carrying a shark in its claws while flying over the sea.

Text: ****Visual Content:****

The image captures a dynamic moment of a bird of prey, likely an eagle or a large hawk, in mid-flight over a body of water. The bird is the central focus, with its wings fully extended, showcasing its impressive wingspan. Its feathers are detailed, with darker shades on the upper wings and lighter, possibly white or gray, feathers on the underside. The bird's talons are firmly gripping a fish, which appears to be a salmon or a similar species, judging by its streamlined body and silvery color. The fish is suspended below the bird, its body slightly curved as if struggling against the grip.

The background consists of a vast expanse of water, likely an ocean or a large lake, with visible waves and ripples. The water's surface reflects a grayish tone, suggesting overcast weather conditions. The waves are not overly turbulent but indicate a steady breeze or mild wind. The horizon line is faintly visible, blending into a misty or foggy atmosphere, which adds a sense of depth and scale to the scene.

There are no other discernible objects, structures, or human-made elements in the image. The scene is entirely natural, focusing solely on the interaction between the bird and its prey.

****Contextual Analysis:****

The image appears to document a real-life moment in nature, specifically a predatory event where a bird of prey has successfully caught a fish. The absence of human-made structures or other artificial elements suggests a remote or wilderness setting, possibly near a coastal area or a large inland body of water.

The bird's species and the type of fish provide clues about the geographic location. Eagles, such as the bald eagle or the osprey, are commonly found near water bodies in North America, Europe, and Asia. The fish's appearance is consistent with salmon, which are native to the North Atlantic and Pacific Oceans.

The overcast weather and the misty atmosphere suggest a cool, possibly temperate climate. The lack of foliage or other seasonal indicators makes it difficult to pinpoint the exact season, but the overall tone leans towards late autumn or early spring.

****Technical Assessment:****

The image quality is slightly grainy, indicating it might be a photograph taken with a telephoto lens or in low-light conditions. The resolution is moderate, with enough detail to distinguish the bird's feathers and the fish's scales, but not so high as to reveal finer textures. There are no obvious compression artifacts, but the overall sharpness is slightly reduced, possibly due to the distance of the subject or the camera's focus settings.

(1427 more characters without spaces)

Original label: Refute

EGMMG prediction: Not misinformation

Figure 11. A consistently incorrectly predicted image-claim-text sample, across several runs.