

# Guidelines

## Q1: What rating does the fact-checking site assign to the news?

**Purpose:** Captures the core veracity assessment from professional fact-checkers.

**Options:**

- **0 = False** - The claim is entirely or substantially false
- **1 = Half true & half false** - The claim contains both true and false elements; partially accurate
- **2 = True** - The claim is accurate and verified

**How to Annotate:**

1. Check the verdict section of the fact-checking article
2. Look for labels like "False," "Misleading," "Partly True," "True," "Verified," etc.
3. Map the fact-checker's verdict to our three categories:
  - "False," "Fake," "Misleading," "Fabricated" → **0 (False)**
  - "Partly True," "Mixture," "Half True," "Misleading but contains truth" → **1 (Half true & half false)**
  - "True," "Correct," "Verified," "Accurate" → **2 (True)**

**Examples:**

- A completely fabricated story about a politician → **0 (False)**
- A news item where the event happened, but details are exaggerated → **1 (Half true & half false)**
- A verified news item shared with correct information → **2 (True)**

**Notes:**

- If the fact-checker uses a more granular scale (e.g., "Mostly False," "Mostly True"), use your judgment to map it to the closest category
- When multiple claims exist in one article, base your rating on the primary claim

---

## Q2: Visual data required to judge the verdict?

**Purpose:** Identifies whether the misinformation involves visual manipulation or requires visual evidence for verification.

**Options:**

- **0 = Textual data is enough** - The claim can be verified using only text
- **1 = Visual data required** - Images, videos, or other visual media are necessary for verification

#### How to Annotate:

1. Read the fact-checking explanation
2. Determine if visual analysis was crucial to debunking the claim

#### Select "1 (Visual data required)" when:

- The article involves doctored/edited photos or videos
- Old images are presented as recent events
- Images are taken out of context
- Deepfakes or AI-generated visuals are involved
- The fact-checker used reverse image search or video forensics
- The claim is "this photo shows X" but actually shows Y

#### Select "0 (Textual data is enough)" when:

- The claim is purely textual (false quotes, fabricated statistics, etc.)
- The misinformation can be debunked through documents, official statements, or logical reasoning
- Any accompanying images are authentic but the textual claim is false

#### Examples:

- False claim: "This video shows riots in Hyderabad" (but it's actually from another country) → **1 (Visual data required)**
- False claim: "Minister X said Y" (fabricated quote) → **0 (Textual data is enough)**
- Morphed photo showing a politician in a compromising situation → **1 (Visual data required)**

### Q3: At whom or what is the false news targeted?

**Purpose:** Identifies the subject or entity being affected by the misinformation.

#### Response Format:

- **Text entry** - Name the person, group, organization, or entity targeted
- **"None"** - If there is no specific target

#### How to Annotate:

1. Ask: "Who or what is this news primarily about?"
2. Look for the subject that would be most affected if people believed this false information
3. Be specific but concise

### Types of Targets:

- **Individual persons:** Politicians, celebrities, public figures, activists
- **Groups:** Religious communities, ethnic groups, political parties, social movements
- **Organizations:** Government bodies, companies, NGOs, media outlets
- **Concepts/Systems:** Democracy, healthcare system, education system
- **Events:** Elections, festivals, historical events
- **Places:** Cities, states, countries

### Examples:

- False news claiming a specific politician is corrupt → "[Politician's name]"
- Misinformation about a religious community's practices → "[Religious community name]"
- Fake news about a government scheme → "[Government body/scheme name]"
- General clickbait with no specific target → "None"
- Satire piece with no malicious intent → "None"

### Special Cases:

- If multiple targets exist, list the primary target (the one most prominently featured)
- For complex cases, separate multiple targets with commas
- Use commonly recognized names in Telugu or English

---

## Q4: How does the news affect the target?

**Purpose:** Determines the directional impact of the misinformation on the target.

### Options:

- **0 = No effect** - The news doesn't significantly impact the target's reputation or perception
- **1 = Flattery** - The news portrays the target positively (often false praise or exaggeration)
- **2 = Denigration** - The news portrays the target negatively (attacks, insults, false accusations)

### How to Annotate:

1. Consider: "If people believe this false information, how would they view the target?"
2. Assess whether the portrayal is overwhelmingly positive, negative, or neutral

**Select "0 (No effect)" when:**

- There is no specific target (answered "None" in Q3)
- The misinformation is about general events or phenomena
- The false information doesn't significantly change perceptions of the target

**Select "1 (Flattery)" when:**

- False accomplishments are attributed to the target
- Exaggerated positive claims are made
- Fabricated praise or awards
- The target is portrayed as heroic, exceptional, or virtuous in false ways

**Select "2 (Denigration)" when:**

- False accusations or crimes are attributed
- The target is insulted, mocked, or humiliated
- Fabricated scandals or controversies
- The target's character, competence, or integrity is attacked

**Examples:**

- False claim: "Actor X donated ₹100 crores to flood victims" (when they didn't) → **1 (Flattery)**
- False claim: "Politician Y embezzled funds" (fabricated accusation) → **2 (Denigration)**
- False weather prediction with no specific target → **0 (No effect)**
- Photoshopped image showing a leader in an embarrassing situation → **2 (Denigration)**

**Note:** Even if the target deserves criticism for other reasons, if *this specific* false information attacks them, mark it as "2 (Denigration)."

---

**Q5: To what extent is the news harmful to society?**

**Purpose:** Assesses the severity of potential harm if the misinformation is believed and spreads.

**Options:**

- **1 = Harmless** - Minimal to no negative societal impact
- **2 = Sadness/distress/anxiety** - Causes emotional distress, worry, or fear
- **3 = Threat/terror/violence** - Could incite violence, hatred, or pose serious safety threats

**How to Annotate:**

1. Ask: "What would happen if many people believed this?"

2. Consider both immediate and potential downstream effects
3. Think about vulnerable populations who might be most affected

**Select "1 (Harmless)" when:**

- The misinformation has trivial consequences
- Purely entertaining false information with no societal impact
- Minor errors that don't affect important decisions
- Harmless celebrity gossip or entertainment news

**Select "2 (Sadness/distress/anxiety)" when:**

- False health information that causes unnecessary worry
- Misinformation that creates panic about safety issues
- False information about disasters, accidents, or tragedies
- Content that generates social anxiety or fear
- Misinformation affecting community trust or cohesion
- False economic information causing financial anxiety

**Select "3 (Threat/terror/violence)" when:**

- Could directly incite violence against individuals or groups
- Hate speech or content promoting discrimination
- False information that could lead to mob violence or lynching
- Dangerous health misinformation (fake cures that could cause death)
- Content that could trigger communal riots
- Inflammatory misinformation during sensitive periods (elections, religious festivals)
- Terrorism-related false information
- False accusations that could lead to vigilante justice

**Examples:**

- False claim about a celebrity's personal life → **1 (Harmless)**
- False warning about contaminated water supply causing panic → **2 (Sadness/distress/anxiety)**
- Fake news accusing a community of child kidnapping (could trigger mob violence) → **3 (Threat/terror/violence)**
- Misinformation about COVID-19 treatments that could be dangerous → **3 (Threat/terror/violence)**

**Important Considerations:**

- Consider the context: The Same content might be more harmful during tense social situations
- Harm level increases if the misinformation targets vulnerable or marginalized groups
- When in doubt between 2 and 3, ask: "Could this directly lead to physical violence?"

---

## Q6: Category of harm?

**Purpose:** Classifies the specific type of societal harm caused by the misinformation.

**Options:**

1. **Harmless** - No significant harm
2. **Confusion and anxiety about society** - Causes confusion about social norms, institutions, or general reality
3. **Threat to honor and trust in people and companies** - Damages the reputation of individuals or organizations
4. **Threat to correct understanding of politics and social events** - Distorts political knowledge or awareness of current events
5. **Health** - Affects health decisions or medical understanding
6. **Prejudice against country and race** - Promotes discrimination, xenophobia, or ethnic hatred
7. **Conspiracy Theory** - Promotes unfounded conspiracy narratives
8. **Not sure** - Difficult to categorize or multiple categories apply equally

**How to Annotate:**

1. Identify the primary domain of harm
2. If multiple categories apply, choose the most severe or prominent
3. Use "Not sure" only when genuinely uncertain

**Detailed Category Guidance:**

**1. Harmless**

- Trivial false information with no discernible societal impact
- Entertainment-focused misinformation with no real-world consequences
- Examples: False release date for a movie, incorrect celebrity age

**2. Confusion and anxiety about society**

- Creates general confusion about how society functions
- Undermines trust in social institutions (courts, police, media)
- False information about laws, rights, or social systems
- Misinformation about technological changes affecting society
- Examples: False claims about new laws, fake social trends, misinformation about digital platforms

**3. Threat to honor and trust in people and companies**

- False accusations against individuals
- Fabricated scandals about public figures
- Fake product failures or company malfeasance
- Defamatory content targeting specific people or organizations
- Examples: False corruption charges, fake product reviews, fabricated celebrity scandals

#### **4. Threat to the correct understanding of politics and social events**

- Electoral misinformation
- False claims about government policies
- Fabricated political statements or positions
- Misinformation about historical events
- Fake diplomatic or international relations news
- Examples: Fake quotes from politicians, false claims about election fraud, misinformation about government schemes

#### **5. Health**

- False medical advice or treatments
- Vaccine misinformation
- Fake disease outbreaks
- Dangerous health practices presented as beneficial
- Examples: Fake COVID-19 cures, false vaccine side effects, fabricated disease statistics

#### **6. Prejudice against country and race**

- Content promoting ethnic, racial, or national hatred
- Stereotyping communities in harmful ways
- False claims designed to create inter-community tension
- Misinformation targeting religious groups
- Examples: False claims about religious practices, fabricated stories about minority communities, hate-inciting content

#### **7. Conspiracy Theory**

- Promotes elaborate, unfounded conspiracy narratives
- Claims of secret plots or hidden agendas without evidence
- Often involves shadowy groups or orchestrated schemes
- Examples: False flag operation claims, global elite manipulation theories, fabricated secret government projects

#### **8. Not sure**

- The harm category is genuinely ambiguous
- Multiple categories apply equally
- You cannot determine the primary harm with confidence

## Decision Tree for Difficult Cases:

If the misinformation involves:

- **Political figures + false accusations** → Likely **3** (Honor and trust) or **4** (Politics), depending on whether it's personal character attack (3) or policy-related (4)
- **Health + conspiracy** → Choose **5** (Health) if immediate health impact is primary concern
- **Multiple communities + tension** → Likely **6** (Prejudice) if it promotes inter-group hatred
- **Government + secret plots** → Likely **7** (Conspiracy Theory) if it's elaborate and unfounded

## Examples:

- False claim that a politician took bribes → **3** (Threat to honor and trust)
  - Misinformation about election rigging without evidence → **7** (Conspiracy Theory) or **4** (Politics)
  - False home remedy for diabetes that could be dangerous → **5** (Health)
  - Fabricated story about Muslims abducting Hindu children → **6** (Prejudice against country and race)
  - Fake news about a new education policy → **4** (Threat to correct understanding of politics)
  - False claim about 5G towers causing diseases → **7** (Conspiracy Theory) or **5** (Health)
-