RoMemes Annotation Guide

1. Memes retrieval

Memes are to be collected from different websites and social media platforms as long as the content is freely available online. This project aims at gathering only content that is produced with the specific purpose to be shared online. No content will be gathered as long as it is created to be hosted onto a single website.

Memes of interest are those containing both image and text. However, it is possible to have a small sample of memes without text.

For each meme image, a source indication and an URL will be kept.

2. Classification

The collected memes will be classified according to 5 criteria, as detailed below.

2.1. Complexity

Memes can be:
- Simple: Single image
- Combination of images

2.2. Real/Fake

The following classes are envisaged:
- Real: the image seems to be a real image repurposed for the meme
- Fake: the image is a clear fake
- DeepFake: the image is fake, but this is hard to tell without additional world knowledge. It looks like a real image, but the annotator knows it cannot be true.

2.3. Sentiment1

We distinguish amongst 3 classes:
- Positive: the meme conveys a positive sentiment (generally speaking, or towards a person or group)
- Negative: the meme conveys a negative sentiment
- Neutral: the meme conveys no clear positive or negative sentiment

2.4. Sentiment2

We consider the 6 emotion categories described in (Parrot, 2001):
- Joy, love, fear, anger, surprise, sadness

Of course all the memes are supposed to be amusing. However, the annotation will take into account the emotion that is conveyed by the actual content of the meme.

2.5. Political

This will identify political memes (related to political parties or individuals involved in politics).

3. Annotation process

Each annotator will receive access to a Google Drive folder (see Figure 1). This folder will be used to upload the meme images. In addition it contains a Google Sheets file (see Figure 2) used to store the memes classification, as described above. The annotation process will proceed in batches of 100 files.
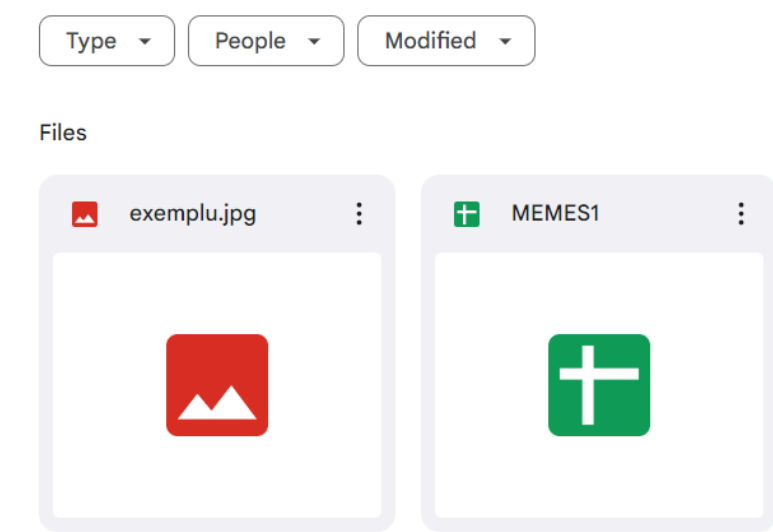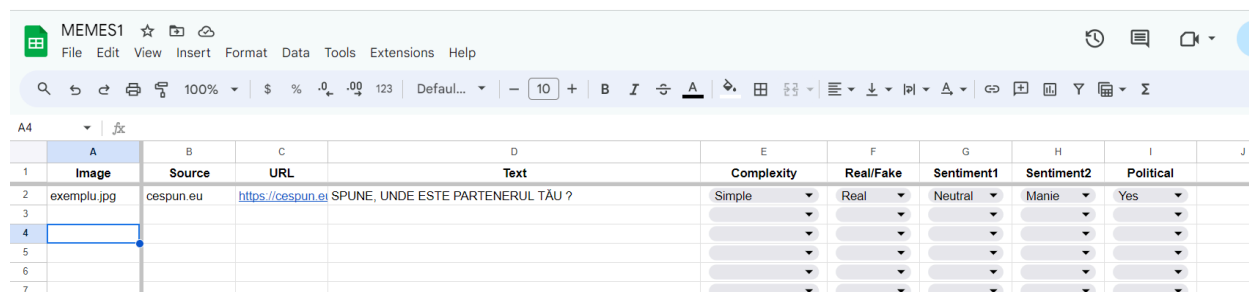

Figure 1. Example Google Drive folder

Figure 2. Example Google Sheets file

For students: Annotation Process Guide

This section provides step-by-step instructions for students on how to collect and annotate Romanian memes for the RoMEMEs corpus, ensuring consistency with the methodology described in the paper and the annotation guidelines.

1. Meme Collection (Retrieval)

- ## Source Requirements:
  - Collect memes only from public social media platforms (e.g., Facebook, Instagram, Twitter) or meme-sharing websites.
  - Ensure memes are freely shareable and contain Romanian-specific content (e.g., language, local figures, political references).
  - Exclude memes created for single-platform use (e.g., private groups).
- ## Format:
  - Memes must include both image and text, but a small subset without text is acceptable.
  - Save each meme as an image file (JPEG/PNG/WEBP) and record the source URL.

2. Annotation Workflow

Students annotated each meme using five classification criteria, as outlined in the paper and the *Ghid_adnotare_ROMEMES_v2.docx*.

2.1. Complexity

- Simple: Single image (e.g., a photo with overlaid text).
- Combination: Multiple images merged (e.g., collages or comparative memes).

2.2. Real/Fake

- Real: Authentic image repurposed for the meme.
- Fake: Clearly manipulated (e.g., photoshopped elements, drawings).
- DeepFake: Highly realistic fake (requires external knowledge to detect).

2.3. Sentiment Polarity

- Positive: Praises a person/group or conveys optimism.
- Negative: Criticizes or mocks a subject.
- Neutral: No clear emotional tone.

2.4. Emotion

Based on Parrot's (2001) categories:

- *Joy, Love, Fear, Anger, Surprise, Sadness*.

● Note: Even humorous memes may convey underlying emotions (e.g., anger in political satire).

2.5. Political Relevance

● Yes: References politicians, parties, or electoral events.
● No: Focuses on culture, humor, or non-political topics.

3. Practical Steps for Students

1. Template Use:
   ○ Annotate memes in the provided Google Sheets template (see *Figure 2* in the guide).
   ○ Each row corresponds to one meme, with columns for text extraction, classification, and source URL.
2. Text Extraction:
   ○ Extract only the meme's core message (ignore logos/clothing text unless relevant).
   ○ Preserve special characters and diacritics (e.g., *„Țăranul modern"*).
3. Quality Control:
   ○ Avoid duplicates (use reverse image search if unsure).
   ○ Flag ambiguous cases (e.g., unsure if *DeepFake*) for lecturer review.

4. Submission & Corpus Integration

● Batch Upload: Submit memes in batches of 10 via the shared Google Drive folder (*Figure 1*).
● File Naming: Use the format Meme[Number].jpg (e.g., Meme1.jpg).
● Post-Processing:
   ○ Duplicates were removed using perceptual hashing (see paper, *Section III*).
   ○ Annotations were converted to a machine-readable format (.tsv + .txt) for the public corpus.

Below are the first five annotated memes from the dataset collected/annotated by first-year History students during the 2024–2025 academic year, demonstrating the annotation process and content diversity. All memes adhere to the refined annotation schema.

ID	text	Complexity		Real/Fake			Polarity			Sentiment						Political	
		simple image	multiple images	Real	Fake	DeepFake	Positive	Negative	Neutral	joy	love	fear	anger	surprise	sadness	Yes	No
Meme1.jpeg	Academia Română dând cu flit lui Cărtărescu/Cărtărescu prin vacante cu soția după ce a ținut un curs la Colombia Uni		x		x				x						x	x	
Meme2.jpg	3 miliarde de ani mai devreme, în Ciorba Rădăuțeană primordială		x		x				x					x			x
Meme3.jpg	Măi Iohannis măi, măi băiatule	x		x					x		x						x
Meme4.jpg	Brandul strămoșești la prețuri mici	x			x				x						x		x
Meme5.jpg	See, we're quite safe if we keep quiet	x		x				x							x		x
Meme6.jpg	După Crin Antonescu și Victor Ponta, Traian Băsescu își anunță și el candidatura		x	x					x	x							x
Meme7.jpg	Dacă s-ar desființa pensiile normale, s-ar putea dubla cele speciale, dar nu se vrea	x			x			x						x	x	x	
Meme8.jpg	Reduceri mari la Lidl, grăbiți-vă!	x			x			x							x		x
Meme9.jpg	Dați repede pe History Chanel, e grav	x		x		x			x	x				x			x
Meme10.jpg	Riga Crypto și Lepona Enigel	x			x		x			x							x

Figure 3. An example of Romanian meme annotation – History students

1. **Meme1.jpg** – *[Brief description: e.g., Political satire featuring a well-known Romanian politician with sarcastic text.]*
2. **Meme2.jpg** – *[Brief description: e.g., Humorous cultural reference combining a popular Romanian TV show character with a viral catchphrase.]*
3. **Meme3.jpg** – *[Brief description: e.g., Fake image mocking a recent social trend, labeled as "Fake" in the corpus.]*
4. **Meme4.jpg** – *[Brief description: e.g., Emotionally charged meme (anger) targeting a public event, with negative polarity.]*
5. **Meme5.jpg** – *[Brief description: e.g., Surreal meme blending local folklore with internet humor, tagged as "Combination" for image complexity.]*

Annotations included: Text extraction, sentiment polarity (positive/neutral/negative), emotion (joy, anger, etc.), image authenticity (real/fake/deepfake), and political relevance.

Meme1.jpg   Meme2.jpg   Meme3.jpg   Meme4.jpg   Meme5.jpg

References

Parrott, W. G. (Ed.). (2001). Emotions in social psychology: Essential readings. Psychology Press.