ChatGPT

Overall, ChatGPT exhibited strong capabilities in Named Entity Recognition, relationship extraction, and plot discovery, but with notable limitations. The model’s performance could be explained in this way:

A generally high recognition rate for key entities but occasional misclassification. A fragmented approach to relationship extraction, often failing to track evolving connections accurately. A reasonable understanding of narrative structure but difficulties in correctly placing story elements within the broader plot. A limited ability to recognize multiple story shifts and complex geopolitical relationships.

In contrast, human annotators provided a more precise and nuanced analysis, demonstrating a deeper understanding of context, relationships, and narrative evolution. Their ability to interpret complex themes, recognize subtle transitions, and accurately categorize entities resulted in a more reliable representation of the articles.

**Name entity recognition.**

ChatGPT was good at capturing a broad range of entities. In all the articles it has recognized almost all of the entities with some exceptions. The model successfully identified the central entities as for example Balticconnector, Sweden and the country ministers.

While the model was good overall it did make quite many small mistakes that could make a difference for some other article but at the same they were not that major for these. It did recognize the most important entities but at the same time it had a problem assigning them to the write categories, for example recognizing EU and NATO as a Location but in the articles they are more mentioned as actors.

In contrast, human annotator displayed greater precision and relevance and more understanding of the text. By deeper understanding of the text the human annotator could precisely categorize the entities by not mixing them up. Plus annotators could also recognize more complex entities, especially in event sections, where understanding of the text and the story plays a crucial role.

Additionally, the model had some problems with recognizing timelines. While most of the time it had no problem with recognizing exact dates it has problem with recognizing, for example day, Tuesday and more complex timeline as for example in the next three months.

**Relationship Extraction**

For this relationship extraction the ChatGPT in Latvian language has had quite a few uncertainties. While it does recognize some of the relationships it has had problem of going beyond one point while sometime focusing on just one paragraph. Plus it sometimes has put together the same entities while the relationship has changed but at the same time it still has the same context and relationships.

It also struggled with recognizing the strongest relations between two entities. For example, signed by [siltumenerģijas komersanti, kurināmā iegādes līgumu], this is not relationships that is strong and is important for the article.

In contrast the human annotator had a better understanding of the article which led to better understanding of the relationships in the article and pinpointing them down. While there are quite a few of them, all of them are crucial parts of the story and it would be possible to understand all of the article's main points by getting to know this part.

**Plot discovery**

ChatGPT has recognized the main story line in the article which already proves that it has a bit of understanding of how the story develops. For example, the LA article follows the same aspect as the human annotator has put down. Without some aspect the things that it has recognized are almost the same. And for all the articles it provided a logical sequence of how the story developed.

At the same time had problems with recognizing where the story lines fit. For example in the article of Delfi, it recognized the exposition and rising action the same, although it described it shorter but it did not quite get what the climax of the article was and from then on it made more mistakes. It put more focus on the Balticconnector incident, although it is the initiator of the story it was not the main thing that was mostly talked about which might have been a problem at recognizing the correct climax.

Human annotations provided more clarity and overall understanding of the article. In the Delfi article the whole storyline can be understood from the plot discovery, because the article was not just about the Balticconnector pipeline, but also about energy stability and prices which played a major role in the article. And a human annotator was able to connect how these two thing are more closely related than the model.

**Story Evolution**

Model capabilities in this section is quite shallow as it does recognize one shift but the same time it is also the main plotline why the story was developed. For example for the LSM article it has recognized the story that the story has shifted only in the last part of the article.

At the same time it has good understanding of broader geopolitical concepts and what are the most important shift. But in all of the articles it has problems with recognizing more than one shift. The same thing goes for capsule relations it has focused on the same part of the article as recognizing the shift and it has problems with recognizing more.

The human annotator has no problem with recognizing shift and relation and geopolitical narratives in the articles and has a border understanding of it. For that reason the human annotations provide broader understanding of the article and the layers that could be seen in the article.

Delfi

#### **Named Entity Recognition (NER)**

ChatGPT successfully identified most of the locations mentioned in the *Delfi* article. However, it inexplicably translated "Title Transfer Facility" into English, despite the absence of any English phrases in the original text.

Regarding actor recognition, the model performed comparably to a human annotator but failed to classify *Saeima* as an actor, suggesting a limited understanding of the Latvian language.

In event recognition, ChatGPT identified more than half of the expected events but struggled with those expressed in more complex linguistic structures. For instance, it failed to recognize *"[neierasts spiediena kritums]"*, which requires an advanced understanding of Latvian sentence construction.

For timeline recognition, the model accurately detected explicitly stated dates but failed to interpret temporally vague expressions such as "in the next week."

#### **Relationship Extraction**

ChatGPT identified some basic relationships but significantly fewer than a human annotator. It struggled to detect relationships involving multiple entities and overlooked those that were indirectly relevant to the main topic.

#### **Plot Discovery**

While ChatGPT correctly identified the exposition and climax similarly to the human annotator, it differed in its interpretation of the climax, falling action, and resolution. This suggests that although its understanding was not entirely incorrect, it did not fully grasp the article’s structure.

#### **Story Evolution**

For shift detection, ChatGPT correctly identified the central event around which the article is structured but did not effectively detect shifts within the narrative.

Regarding causal relationship extraction, the model correctly identified one causal link; however, it failed to recognize additional relevant causal connections present in the text.

—

LA

### **Named Entity Recognition**

For Named Entity Recognition it has challenges in distinguishing between entity types, particularly in classifying organizations as locations. In this case, NATO and the EU were incorrectly recognized as locations rather than actors, which affects the accuracy of relationship extraction and narrative structuring.

Additionally, while the system correctly identifies most actors and events, it fails to recognize some that require contextual understanding.

System successfully identifies explicit time such as *“naktī uz svētdienu”* (the night leading into Sunday) but fails to recognize relatively simple timelines like *“pieci mēneši”* (five months). This gap in recognizing non-explicit time markers affects the timeline reconstruction and understanding of event sequencing.

### **Relationship Extraction**

The system performs relatively well in identifying direct relationships between entities but fails to capture all potential connections within the text. Some relationships that were not noted by human annotators are correctly recognized, such as the link between Latvia’s supply and Balticconnector’s closure. However, the system also misses several important relationships that contribute to a deeper understanding of the narrative, indicating a need for improved contextual analysis.

Additionally, while the system identifies relationships, it does not always clarify the nature of these connections, leading a room to interpretation.

### **Plot Discovery**

Model effectively identifies key events and maintains consistency with human annotators’ interpretations. It successfully constructs the primary storyline, demonstrating a structured understanding of the text’s core narrative. However, minor details are occasionally missed, which, while not altering the overall plot, could affect the depth of analysis and the perception of event importance.

Another limitation is that the system does not always prioritize events according to their significance within the narrative. Some events are given equal weight despite their differing levels of importance, which could lead to an inaccurate representation of the story’s focus.

### **Story Evolution**

The system encounters difficulties in correctly identifying narrative shifts, often mistaking the central theme of the article for a major shift in the story. This misinterpretation prevents an accurate depiction of how the narrative evolves, as it fails to distinguish between continuity and significant developments.

Although the system successfully extracts some causal relationships, it overlooks others that are crucial for understanding how events influence one another.

—

LSM

### **Named Entity Recognition**

Model effectively identifies geographical locations but fails to recognize non-geographical places like "laboratory." This limitation suggests a need for improved classification of functional locations. In actor recognition, it correctly identifies NATO and the EU as actors, improving upon previous misclassifications.

While the system recognizes most events, it struggles with those requiring contextual understanding rather than keyword detection. Temporal recognition is inconsistent, identifying broader timelines like "in a week" but missing specific days.

### **Relationship Extraction**

The system accurately extracts key relationships but places excessive focus on NATO-related connections. While this emphasis is not incorrect. A more balanced approach would ensure all relevant relationships receive appropriate weight.

Additionally, the system could improve how it contextualizes relationships, clarifying their nature rather than simply listing them. Refining its methodology would enhance the clarity and relevance of the extracted connections.

### **Plot Discovery**

The system correctly identifies the main plot points but note notice additional aspects noted by human annotators. While this can provide deeper insights, it also narrows the focus, potentially overlooking broader narrative elements.

To improve, the system should balance detailed analysis with a comprehensive representation of the full story. A more flexible approach would ensure that both key themes and contextual nuances are properly captured.

### **Story Evolution**

Model disproportionately focuses on NATO in the final section, limiting its ability to track the full evolution of the narrative. While the information it extracts is correct, it overlooks other thematic shifts, leading to an incomplete representation of the article.

Model should to detect multiple story developments rather than fixating on a single prominent topic. Strengthening its ability to track diverse narrative shifts would enhance its overall accuracy.

**—**

TVNET

### **Named Entity Recognition**

The system fails to recognize Finland as a location while incorrectly classifying Latvia in a context where it functions as an actor. Additionally, it continues to misidentify the EU as a location rather than an actor. These inconsistencies suggest a need for improved contextual understanding in entity classification.

Event recognition remains partial, capturing only a few key events but missing those that require a deeper contextual grasp.

However, timeline recognition is accurate, aligning with human annotations.

### **Relationship Extraction**

The system identifies some of the same relationships as human annotators but also introduces alternative connections. While the entities remain the same, the nature of their relationships is sometimes presented differently. This suggests that while the extraction process is functional, it lacks consistency in how relationships are interpreted.

A more structured methodology would enhance the clarity and reliability of extracted relationships.

### **Plot Discovery**

Model correctly identifies significant elements of the article but starts its analysis from the climax, leading to an incomplete representation of the overall narrative. While it captures certain details effectively, it fails to convey the full scope of the article's progression.

more structured approach to sequencing and emphasizing key events would provide a clearer and more comprehensive narrative structure.

### **Story Evolution**

The system does not effectively recognize narrative shifts, instead identifying only the article's ending as a potential shift. While this conclusion is not entirely incorrect, it overlooks other critical transitions throughout the text.

Additionally, while model extracts causal relationships, there is room for improvement in identifying more connections that contribute to the story’s progression.