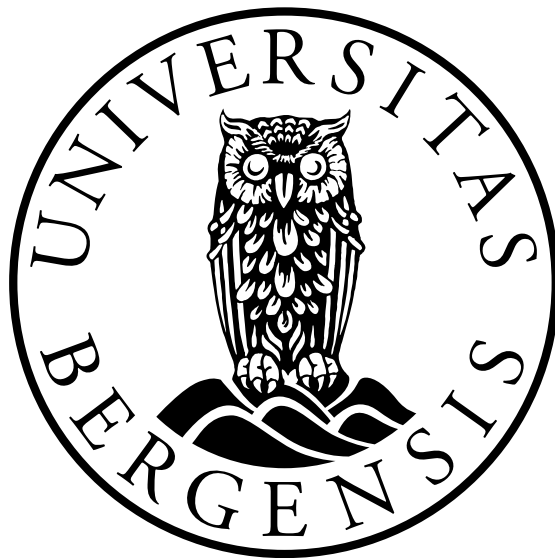


Evaluating Pre-trained Language Model Strategies for Knowledge Graph Extraction: A Comparative Analysis

Martin Skivenesvåg Johannessen

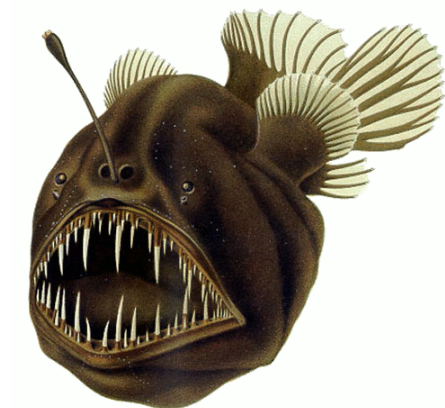


A Master's Thesis
Department of Information Science and Media Studies
University of Bergen

May 3, 2023

Scientific environment

This study is carried out at the Department of Information Science and Media Studies, University of Bergen, and in association with the News Angler project.



Acknowledgements

Thanks to my thesis supervisor Andreas Lothe Opdahl, my project design instructor Csaba Veres, the project template designer Raymond Nepstad, and the ones who made it available on Overleaf, Kjersti Birkeland Daae and Stephan Kral. Thank you Norway for the stipend and education. Thank you OpenAI for Chat GPT. Thank you my parents, and Karoline Øijorden for your love and support.

Martin S. Johannessen
Bergen, May 3, 2023

Abstract

Extracting structured knowledge graphs from natural language text is a critical task in natural language processing, as it facilitates the automatic extraction and organization of information from unstructured sources. This is highly significant for news articles, which provide a wealth of information about the world but can be challenging to analyze and comprehend at a large scale.

In this research, the proficiency of different methods for extracting structured knowledge graphs from news articles is assessed, including large pre-trained language models, datasets for supervised fine-tuning, and artificial intelligence techniques. Large pre-trained transformer models, specifically T5 and BART, are evaluated when fine-tuned on the datasets WebNLG, DART, TEKGEN, and KELM, which are by extension also assessed. AMRBART is assessed for its ability to extract knowledge graphs after converting its output to RDF, and the zero-shot performance of GPT-4 is also evaluated.

To systematically assess knowledge graphs, a set of graph quality metrics is employed as an alternative to gold label comparison.

The results indicate that GPT-4 surpasses other approaches in knowledge graph extraction.

Contents

Scientific environment	i
Acknowledgements	iii
Abstract	v
1 Introduction	1
1.1 Motivation	1
1.2 Research questions	2
1.3 Contribution	2
2 Background	5
2.1 What are knowledge graphs?	5
2.2 Evaluation metrics	6
2.3 Knowledge graph extraction	11
2.3.1 Transformer models	14
2.3.2 T5	15
2.3.3 BART	15
2.3.4 AMR-BART	15
2.3.5 AMR-LD	16
2.3.6 GPT-4	17
3 Methodology and methods	21
3.1 Large pre-trained language models	21
3.1.1 Supervised fine-Tuning	21
3.1.2 Zero-Shot	22
3.2 Datasets	22
3.2.1 Data preprocessing	23
3.3 Knowledge graph evaluation	23
3.3.1 Gold label alternative	23
3.3.2 Metrics	25
3.3.3 Process of evaluation	28
3.3.4 Likert scale	28
3.3.5 Sample input data	31

4 Results	33
4.1 Results	33
4.2 Discussion	33
4.2.1 Limitations	33
4.2.2 Syntax	34
4.2.3 Conciseness	34
4.2.4 Consistency	35
4.2.5 Precision and Recall	35
4.2.6 Density and Size	36
4.2.7 Functionality	36
4.3 Overall Performance	36
4.3.1 GPT-4 is the winner	36
4.3.2 The method of evaluation is useful	36
5 Conclusion	39
5.1 Summary of the results	39
5.2 Implication and Recommendation	40
5.3 Future Research Directions	40
A Code for repeatability	43
A.1 Preprocessing of datasets for fine-tuning	43
A.1.1 WebNLG	43
A.1.2 DART	44
A.1.3 TEKGEN	44
A.1.4 KELM	46
A.2 Models	47
A.2.1 GPT-4	47
A.2.2 AMRBART	49
A.2.3 T5 and BART	51
B Evaluation data	55
B.1 Russian official makes new claim	55
B.2 AI pioneer quits Google to warn about the technologys dangers	56
B.3 Palestinian Khader Adnan dies in Israel jail after 86 days on hunger strike	58
B.4 Erling Haaland scored his 50th goal of the season in all compe- titions	60
B.5 Smith’s Donation Motivation	61
C Output of GPT-4	63
C.1 First output	63
C.2 First regenerate output	64
C.3 Second generate output	66
C.4 GPT-4 evaluating itself	68
D Output of AMR-BART with AMR-LD	73

E	Outputs of T5 fine tuned on different data sets	109
E.1	T5-DART	109
E.2	T5-WebNLG	110
E.3	T5-KELM	112
E.4	T5-TekGen	113
F	Preliminary outputs of T5 and BART	115
F.1	T5	115
F.2	BART	131
F.3	Intermediate evaluation	152

Chapter 1

Introduction

Knowledge graphs have been successfully applied in diverse domains such as healthcare, finance, and e-commerce. In these industries, they facilitate data integration, analytics, and decision-making processes.

One such area is search engines, where knowledge graphs enable more accurate and relevant search results. They achieve this by understanding the relationships between different entities and their attributes, helping users find the information they need quickly and efficiently.

Recommendation systems are another domain where knowledge graphs have a significant impact. By leveraging the connections between users, items, and their respective features, knowledge graphs can provide personalized and context-aware recommendations. This makes it easier for users to discover new content, products, or services that align with their interests and preferences.

In question-answering systems, knowledge graphs help improve response accuracy by capturing the underlying semantic structure of the information. This allows for a better understanding of user queries and provides more precise answers to their questions.

Journalism is yet another field that benefits from knowledge graphs. By uncovering hidden connections between news events, people, and organizations, knowledge graphs enable journalists to find new angles and find more stories.

Given their widespread use and significant impact, research that advances knowledge graph extraction techniques is of considerable importance. By improving these techniques, we have the potential to benefit various industries and applications, making our world more connected and data-driven.

1.1 Motivation

Knowledge graphs are a powerful tool for organizing and structuring information, however, creating knowledge graphs manually is a time-consuming and expensive process, and may not be feasible for large-scale applica-

tions. Current methods of automatic knowledge graph extraction have made progress, but there is still room for improvement. Recent advancements in pre-trained language models have shown promising results in natural language processing tasks, and have the potential to enhance the automated extraction of knowledge graphs from unstructured text data.

Representing information in a knowledge graph can be done in multiple ways, making it difficult to objectively measure the effectiveness of a generative model by using a hard gold label of expected triples, as opposed to classification models where possible outputs are predetermined. For example, if the gold label is [James, hasFather, John] and the model gives the output [John, hasSon, James], then neither subject, predicate, nor object matches. The inverse relation of [James, hasSon, John] is semantically equivalent and should be considered correct. A naive similarity metric can capture that they share two out of three terms for a recall of 0.66, which does not adequately convey that the output could have been an alternative gold label.

There is a need to evaluate and compare the effectiveness of different pre-trained language models at the task of knowledge graph extraction and to develop methods for fine-tuning these models to improve their performance on this task. Additionally, reliable evaluation metrics for extracted knowledge graphs are required, which can help researchers compare the performance of various models and techniques. By addressing these research questions, we can advance the field of natural language processing and create more efficient and accurate methods for knowledge graph extraction.

1.2 Research questions

- How do different architectures of pre-trained language models, specifically the encoder-decoder models BART and T5, and the decoder-based model GPT-4, compare in terms of their ability to extract knowledge graphs from natural language text?
- How can we evaluate the quality of extracted knowledge graphs and compare the performance of different knowledge graph extraction methods?

1.3 Contribution

The contribution of this paper is twofold: it introduces a set of evaluation metrics that comprise a novel framework for assessing automatically generated knowledge graphs, and it provides an evaluation of state-of-the-art language models at the task of lifting knowledge graphs from text. This work aims to advance the understanding of the capabilities and limitations

of these models and to offer guidance for future research in the area of knowledge graph extraction.

Chapter 2

Background

2.1 What are knowledge graphs?

Knowledge graphs are constituted by a set of facts, where each fact, also known as an RDF statement or relation, is represented by a semantic triple, composed of three atoms: a subject, a predicate, and an object. In the context of knowledge graphs, these three components are sometimes referred to as head, relationship, and tail, respectively. Subjects and objects act as nodes in the graph, while predicates serve as relational edges between them, forming a network of entities and concepts that contextualize each other.

Knowledge graphs offer several advantages, including the ability to represent complex relationships, facilitate knowledge discovery, and enable reasoning over the stored information. For example, a simple semantic triple in a knowledge graph could be [Albert_Einstein, nationality, German], where "Albert_Einstein" is the subject, "nationality" is the predicate, and "German" is the object.

To be in accordance with the Resource Description Framework (RDF), a standard for data interchange on the Internet, all components of a semantic triple should be represented as IRI (Internationalized Resource Identifier) resources. These IRIs serve as unique identifiers for every concept, and ideally, they should resolve to a website that further specifies and defines the resource. Although an IRI might not resolve to an actual web page, it will still function as a unique identifier for the concept.

```
@prefix dbo: http://dbpedia.org/ontology/ .
```

```
@prefix dbr: http://dbpedia.org/resource/ .
```

```
dbr:Albert_Einstein dbo:nationality dbr:Germany .
```

In this example, "dbr:Albert_Einstein" is the subject, "dbo:nationality" is the predicate, and "dbr:Germany" is the object. The prefix format is for convenience, so we can avoid typing out namespace addresses redundantly:

"dbr:Albert_Einstein" will be interpreted as "http://dbpedia.org/resource/Albert_Einstein".

However, not all methods of KG extraction assessed in this paper result in actual KGs. Some methods produce proto semantic triples that are not comprised of IRIs but can still be evaluated, and can be resolved to IRIs post-process.

2.2 Evaluation metrics

There is sparse literature on the subject of evaluating automatically generated KGs. *Chen et al. (2019a)* conducted a study of existing literature relevant to knowledge graph quality evaluation and proposed a framework for doing so consisting of the criteria below.

The references informing each criterion are included, as they were needed to understand the various criteria that were not sufficiently clear. Even when following the references, it was difficult to procure the necessary explanations or justifications for some of the criteria.

1. "Knowledge graph does not contain redundant triples (*Deng et al., 2019*)."

Is a good criterion to evaluate a knowledge graph, but for the purpose of evaluating a lifter by its outputted knowledge graph it is not necessary because the information could be stated redundantly in the source text, and it can be contended that a lifter's task should be complicated by duplicate removal when it is a trivial task to do in post process.

2. "Knowledge graph can be updated dynamically (*Muppalla et al., 2017*)."

It is tricky to address this criterion and following the reference does not resolve what is meant exactly. It stands to reason that beyond what is measured by other metrics the enhancements of and limits to dynamism would be governed by ontology loosely and whether it fits the domain. Using an ontology that fits the domain could help provide instant context, knowledge enrichment, semantics, and reasoning capability, with little effort to newly introduced instances. On the other side, an ontology that does not sufficiently encompass the domain it is applied to with the update could make it more work to fit it together sensibly. Knowledge graphs do not have to use a domain ontology in the strict sense. By default, knowledge graphs are dynamically adaptable.

3. "Data for constructing a knowledge graph should include different types and from different resources (*Ping et al., 2017; Wang et al., 2018*)."

4. "Data source should be multi-field (*Huang et al.*, 2019; *Lukovnikov et al.*, 2017; *Wang et al.*, 2018)."

Regardless of its merits as a metric, is irrelevant for evaluating lifted graphs that are supposed to represent the information from only a specific source text.

5. "The scalability with respect to the KG size (*Wang et al.*, 2019b)."

The reference is about the scalability of the knowledge-based recommendations system approaches with respect to KG size, and not the scalability of the KG. It seems the scalability of KGs themselves is an inherent property of the technology at large, and less so tied to the specific knowledge graph.

6. "Knowledge graph should be publicly available and proprietary (*Dietz et al.*, 2018)."

This is the ideal underlying the RDF paradigm, but not relevant for evaluating a lifted graph; it's not the automated lifter's responsibility.

7. "Knowledge graph should be authority (*Zhao et al.*, 2018)."

It is not clear from the reference what this means. Evaluating trustworthiness based on reputation is not a luxury we can afford in our context. The provenance is not relevant unless it is extractable from the data, which then falls under the recall umbrella.

8. "Knowledge graph should be concentrated (*Zhao et al.*, 2018)."

From the reference *Zhao et al.* (2018): "The knowledge graph can be divided into general field and vertical field according to its knowledge range. General field knowledge graph is characterized by large amount of data, strong versatility and wide range, but vertical field knowledge graph has the characteristics of relatively small amount of data, high quality, high density, high concentration, high authority and high knowledge accuracy".

The criterion does not obviously follow from the citation, which characterizes two types of knowledge graphs by juxtaposition and "concentrated" a characteristic of "vertical" knowledge graphs, which are not meant to be taken as superior.

9. "The triples should not contradict with each other (*Huang et al.*, 2019; *Wang et al.*, 2019b)."

This is good advice in general but does not necessarily work for the purpose of evaluating a lifter based on its knowledge graph output. The triples should contradict each other if they are corresponding to contradictory information from the source data.

10. "For domain specific tasks, the knowledge graph should be related to that field (*Chen et al.*, 2019b; *Deng et al.*, 2019; *Zhao et al.*, 2018)."
11. "Knowledge graph should contain the latest resources to guarantee freshness (*Zheng et al.*, 2018; ?)."

The limitations outlined above, highlight a need for more work on the topic of evaluating automatically generated knowledge graphs. However, some of the criteria are useful to work off of:

1. "Triples should be concise (*Deng et al.*, 2019; *Zhao et al.*, 2018)."

Following the references of "triples should be concise", the word concise isn't mentioned by *Zhao et al.* (2018) and it is difficult to see where the requirement is inferred from, and only mentioned in passing by *Deng et al.* (2019), where it is linked to redundancy problems, and dealt with by pre-processing the news articles to remove adjectives and adverbs. In the way that they dealt with inconciseness, it is inferred that it is about the packing of unnecessary information into semantically compound atom names or information that should instead be separated into other triples. For example the hypothetical atom "singsLoudly" should instead be "sings", and "loudly", if necessary at all, should be its own atom.

It is noteworthy that the conciseness of triples in this sense is at the cost of the conciseness of the graph at large since mashing meanings together in atoms could greatly reduce the complexity of a graph by reducing the need for blank nodes or other more elaborate structures. However, the conciseness of the graph is not an evaluation metric, and conflicts somewhat with other metrics. Semantically compound atoms would also come at the cost of reusability since they could hardly be linked to an actual open resource which is the point of RDF in the first place.

2. "Contextual information of entities should be captured (*Deng et al.*, 2019)."

By following the reference this concern is contextualized by seeing that the authors enriched their graphs by importing the entities' first-degree neighbors from their linked open data knowledge base, in order to improve the graph embedding of the entities after vectorization. For the purpose of evaluating the captured information from a news text, the top score will be limited to capturing all the relevant information in the text only, and hallucinating even correct information is probably not desirable.

3. "Entities should be densely connected (*Keller*, 2019; *Radhakrishnan et al.*, 2018; *Zhao et al.*, 2018)."

"Entities should be densely connected" means that there are more relations and more types of relations between entities, and more classes defining entities. This was how it was used in the references, and some other criteria below seem to be specifications of this one.

4. "Relations among different types of entities should be included (*Wang et al.*, 2018)."

Can be understood as a specification of "Entities should be densely connected".

5. "The attributes of the entities should not be missed (*Ostapuk et al.*, 2019)."

Can be understood as a specification of "Entities should be densely connected".

6. "Synonyms should be mapped and ambiguities should be eliminated to ensure reconcilable expressions (*Huang et al.*, 2019; *Lukovnikov et al.*, 2017; *Wang et al.*, 2018)."

This is highly relevant for evaluating lifters based on their outputted knowledge graph. It is important for graph consistency.

7. "Knowledge graph should be organized in structured triples for easily processed by machine (*Wang et al.*, 2018)."

This is the most basic and objective criterion. If a fact does not follow this syntax then it can not be part of a knowledge graph.

"Entities should be densely connected", "contextual information captured", and "relations among different types of entities should be included" can be summarised as "recall", since the scope is only the source text and enrichment beyond that is out of scope.

"Triples should be concise", "knowledge graph should be organized in structured triples", "synonyms should be mapped and ambiguities should be eliminated", and "for domain-specific tasks, the knowledge graph [structure/vocabulary/ontology] should be related to that field" are all good concerns of evaluation with regard to graph syntax, consistency, and quality in general.

Huaman (2022) performed a literature review on the quality assessment of data, information, linked data, and KGs. They proposed a "general-purpose, customizable to a domain or task, and practical quality assessment framework for assessing the quality of KGs".

Table 2.1: Goals, Questions, Metrics, and Types (QN/QL)

Goal	Question, Metrics, and Type (QN/QL)
Accessibility	Is the KG (or at least part of it) available (QN), provides an SPARQL endpoint (QN), retrievable (QN), supports content negotiation (QN), and contain a license (QN)?
Accuracy	Is the KG reliable and correct, e.g., syntactically (QN) and semantically (QN)?
Appropriate amount	Does the KG contain an appropriate amount (QN) of instances for a specific task?
Believability	Does the KG provide provenance information (QN), is trustworthy (QL), and has not unknown nor empty values (QN)?
Completeness	At which degree is the KG complete regarding data (QN), population (QN), and interlinking (QN)?
Concise representation	Is the KG concisely represented by avoiding blank nodes (QN) and reification (QN)?
Consistent representation	Is the KG consistently represented, e.g., disjoint inconsistencies of classes (QN), inconsistent inverse functional property values (QN), schema restrictions (QN)?
Cost-effectiveness	Does the KG require extra data at any cost (QL)?
Ease of manipulation	At which level does the KG provide documentation (QL)?
Ease of operation	Is it possible to update (QN), download (QN), and integrate (QN) the KG?
Ease of understanding	Is the KG represented using self-descriptive URIs (QN) and in various languages (QN)?
Free of error	Does the KG provide correct property values (QN)?
Interoperability	Is the KG interoperable, e.g., openly available (QN) and uses standard vocabularies (QL)?
Objectivity	Is the KG objective, e.g., unbiased (QL) and declares provenance information (QN)?

Continued on next page

Continued from previous page

Goal	Question, Metrics, and Type (QN/QL)
Relevancy	Is the KG relevant for the task at hand, e.g., at which level the KG provides knowledge for a specific domain or use case (QL)?
Reputation	Is the KG well rated? E.g is the KG well positioned in explicit ratings (QL).
Security	Does the KG provide security mechanisms like digital signature (QN) and KG authentication (QL)?
Timeliness	At which degree is the KG up to date (QN) and fresh (QN)?
Traceability	Does the KG provide mean to verify its provenance (QL) and authenticity (QL)?
Variety	At which degree the KG integrates various domain sources (QL)?

^{QN} Quantitatively measured metrics. ^{QL} Qualitatively measured metrics.

Because it is not specifically a framework for evaluating automatically generated knowledge graphs, some goals and questions are more relevant for evaluating the quality of KGs in general, and less so for evaluating a lifter based on its outputted knowledge graph for the same reasons as above. Accessibility, cost-effectiveness, ease of operation, reputation, believability, security, ease of manipulation, timeliness, traceability, and variety are important factors when considering the quality of a knowledge graph, but they do not directly pertain to the performance of a lifter.

In 3.3.2 and 3.3.2, the relevant metrics of the frameworks above are mapped to the corresponding metrics used in this thesis in which the pertaining questions are incorporated.

2.3 Knowledge graph extraction

The task of mining a knowledge graph (KG) in natural language processing (NLP) has been primarily divided into two main sub-tasks: named entity extraction (NEE) and relation extraction (RE). NEE can be further decomposed into named entity recognition (NER), named entity disambiguation (NED), and named entity linking (NEL). As

highlighted by *Al-Moslmi et al. (2020)*, the development of these sub-tasks has progressed along two primary axes:

Firstly, there has been a shift from a step-wise pipeline process towards an integrated end-to-end approach. Secondly, entities that were previously analyzed in isolation are now being processed in parallel within the context of one another. This allows for better utilization of interdependencies and contextual information between tasks that might otherwise be lost when transitioning from one step to the next (*Miwa and Sasaki, 2014*).

Al-Moslmi et al. (2020) also emphasized the necessity for a standardized method of evaluating end-to-end NEE systems. However, the culmination of this trend seems to involve the extraction of KG from natural language (NL) in one joint task, which encompasses both NEE and RE, rather than just a subset of the task. Recent approaches have started integrating NER (but not full NEE) with RE for similar reasons, such as the work of *Sui et al. (2020)*, which currently leads in the New York Times and WebNLG datasets with their set prediction approach for joint entity and relation extraction, and the study by *Giorgi et al. (2022)* that showed that an end-to-end approach for document level entity and relation extraction outperformed a pipeline approach.

In recent years, natural language processing (NLP) has experienced a significant increase in utility, primarily due to the emergence of large pre-trained transformer-based language models that have disrupted various fields. One such notable example is ChatGPT which has garnered public awareness. These large language models (LLMs) are widely employed for NER and consistently top the leaderboards of benchmarking datasets, such as XGLUE, which features NER as its primary task (*Liang et al., 2020*), and SuperGLUE (*Wang et al., 2019a*), a benchmarking dataset aimed at evaluating artificial language understanding. This improved understanding of natural language holds immense potential for automating the extraction of knowledge from text, making it an invaluable capability in the field of NLP.

Attention heads in transformer models have been found to capture both semantic and syntactic information (*Jo and Myaeng, 2020*). *Wang et al. (2020)* discovered latent features within the attention

matrices that are useful for entity linking. They utilized the matching degree between attention scores of noun-based candidate entities from the text to construct semantic triples directly from the attention matrices of models like BERT and GPT-2. These semantic triples served as candidate facts, which were then mapped to Linked Open Data (LOD) facts using secondary entity and relation linking methods.

Eberts and Ulges (2019) revealed that embeddings also contain features useful for entity linking. Both approaches involve performing a single forward pass on the corpus and exclusively using their respective types of latent representation, along with an explicit search algorithm. This suggests that the algorithmic interpretation of select latent features may overlook other relevant features present in the model. To potentially utilize all pertinent latent features in an LLM, it may be more effective to employ the models themselves in the process.

Fine-tuning large pre-trained language models typically involves using supervised learning on a labeled dataset. This process takes advantage of the extensive computational resources expended during the self-supervised pre-training of the model to achieve strong performance on downstream tasks. However, recently there has been a notable shift towards reinforcement learning (RL) based fine-tuning (*Ouyang et al., 2022; Ziegler et al., 2019*). This approach has gained significant prominence, particularly through OpenAI's ChatGPT model, which is a variant of their InstructGPT concept that utilizes human feedback-based RL fine-tuning (*Ouyang et al., 2022*).

In the context of knowledge graph extraction, RL-based fine-tuning holds the potential to enhance performance by enabling the model to learn from feedback on its output quality, rather than relying solely on labeled data. This approach is especially beneficial for knowledge graph extraction, where obtaining high-quality labeled data is often challenging. Utilizing RL-based fine-tuning, the model can optimize its performance on knowledge graph extraction based on feedback from domain experts or other sources.

However, there are a few challenges associated with using RL-based fine-tuning for knowledge graph extraction. One challenge is defining an appropriate reward function that accurately reflects the

quality of the extracted knowledge graph. This task can be difficult, as the quality of a knowledge graph may depend on various factors, such as completeness, coherence, and domain relevance. Another challenge is addressing the vast search space of potential knowledge graphs, which can render RL-based fine-tuning computationally expensive.

Despite these challenges, RL-based fine-tuning demonstrates promise in improving the performance of models for knowledge graph extraction. By enabling the model to learn from feedback and optimize its performance, RL-based fine-tuning could potentially lead to more accurate and useful knowledge graphs that capture the nuances and complexities of natural language text.

In order to employ machine learning for constructing knowledge graphs, a training signal that reflects the quality of the constructed knowledge graph is necessary. One potential basis for such a signal is back translation. By implementing a secondary, reverse machine learning algorithm that back translates from the knowledge graph to text, it may be possible to use the similarity between the original input and the back-translated output to create a training signal for both algorithms. This approach allows the models to learn iteratively by comparing the reconstructed text to the original input, thereby encouraging the generation of more accurate and coherent knowledge graphs.

2.3.1 Transformer models

Transformer models have revolutionized the field of natural language processing (NLP) with their ability to process and generate text in parallel, rather than sequentially (Vaswani *et al.*, 2017). This parallel processing capability allows for faster and more efficient training, leading to improved performance on a wide range of NLP tasks (Devlin *et al.*, 2018; Raffel *et al.*, 2019). Some of the most prominent transformer-based models include BERT (Devlin *et al.*, 2018), T5 (Raffel *et al.*, 2019), BART (Lewis *et al.*, 2019), and GPT-4 (OpenAI, 2023).

2.3.2 T5

T5, or Text-to-Text Transfer Transformer, is a transformer model developed by Google Research (*Raffel et al., 2019*). The key innovation of T5 is its text-to-text framework, which unifies various NLP tasks into a single format: converting input text to output text. This approach allows T5 to be trained on a diverse set of tasks, including translation, summarization, and question-answering. By leveraging transfer learning and pre-training on large text corpora, T5 has achieved state-of-the-art performance on numerous NLP benchmarks.

2.3.3 BART

BART, or Bidirectional and Auto-Regressive Transformers, is a model developed by Facebook AI Research (*Lewis et al., 2019*). BART combines the strengths of auto-regressive models like GPT and auto-encoding models like BERT by training on a denoising auto-encoding objective. This training objective involves reconstructing corrupted text, which forces the model to learn both bidirectional and unidirectional representations of the input text. As a result, BART is well-suited for both text generation and understanding tasks, such as summarization, translation, and sentiment analysis.

2.3.4 AMR-BART

AMR-BART is a variant of the BART model specifically designed for Abstract Meaning Representation (AMR) parsing and generation (?). AMR is a graph-based representation of meaning in natural language, which captures the relationships between entities and their properties. AMR-BART is trained on parallel corpora of text and AMR graphs, enabling it to learn the mapping between natural language text and its corresponding AMR representation. This model has demonstrated strong performance in tasks like AMR parsing, where it identifies the AMR structure from text, and AMR generation, where it generates text from a given AMR graph.

2.3.5 AMR-LD

AMR-LD (*Burns and Ambite, 2014*) is a Python library designed for mapping Abstract Meaning Representation graphs to linked data formats, such as RDF and JSON-LD. AMR-LD provides a means to transform the graph structure of AMR into RDF using *rdflib*, a popular Python library for working with RDF. By applying simple heuristics to map namespaces and generate valid RDF for the AMR, AMR-LD bridges the gap between natural language meaning representation and the world of linked data. It is heavily derived from Saphra's 'AMERICA' library and the SMATCH tool:

AMERICA

AMERICA (AMR Inspector for Cross-language Alignments) is a versatile tool for aligning and visually representing AMRs, both in bilingual contexts and for monolingual inter-annotator agreement (*Saphra, 2014*). It extends and builds upon the Smatch system, which is used for determining AMR interannotator agreement. With AMERICA, users can visualize manual alignments they have edited or compiled themselves.

SMATCH

SMATCH is an evaluation tool designed specifically for Abstract Meaning Representation (*Cai and Knight, 2013*). It calculates the Smatch score, a metric that measures the similarity between two AMR graphs in terms of their matching triples (edges). SMATCH finds a variable (node) mapping that maximizes the count, M , of matching triples. Using this metric, the tool computes:

M , the number of matching triples T , the total number of triples in the first AMR G , the total number of triples in the second AMR
Precision (P), defined as $P = M/T$ Recall (R), defined as $R = M/G$

The Smatch score is then calculated as the F-score: $F = 2 * (P*R)/(P+R)$. This score serves as an important evaluation metric for comparing AMR graphs and assessing the performance of AMR parsing and generation models.

RDFLib

RDFLib (*Boettiger et al., 2020; Jones et al., 2022*) is a popular Python library designed for working with RDF. It provides users with a comprehensive set of tools for parsing, serializing, storing, querying, and manipulating RDF data in various formats, such as XML, Turtle, N-Triples, and JSON-LD.

One of the key features of RDFLib is its support for SPARQL, a powerful query language and protocol for RDF data. This enables users to perform complex queries and extract specific information from RDF graphs with ease. Additionally, RDFLib offers a plugin system for customizing and extending its functionality, making it highly adaptable to various use cases and applications.

2.3.6 GPT-4

GPT-4, or Generative Pre-trained Transformer 4, is the latest iteration of the GPT series developed by OpenAI (*OpenAI, 2023*). Like its predecessors, GPT-4 is an auto-regressive model that generates text by predicting the next word in a sequence, given the previous words. GPT-4 builds upon the success of GPT series of language models by further scaling up model size and training data, which has led to improvements in performance across a wide range of NLP tasks. Notably a decoder-only model (*Openai et al., 2018*), GPT-4 excels in text generation tasks, such as story generation, summarization, and translation, but can also be fine-tuned for other NLP tasks, including classification, and question-answering and hopefully KG extraction.

RLHF

Reinforcement Learning from Human Feedback (RLHF) involves using human feedback to create a reward model that guides the training process (*Ziegler et al., 2019*), and is a training method used to improve the performance of language models like GPT-4 (*Ouyang et al., 2022*):

The model is first fine-tuned on a dataset of human-generated examples, which include the desired input-output pairs. This step provides the model with an understanding of the task it will be optimized for.

A set of input-output pairs is collected, where the model generates multiple alternative completions for a given input. Human evaluators then rank these completions based on their quality, and this information is used to create a dataset of comparison data.

A reward model is trained to predict the human-provided rankings for the generated completions. This model is used to assign a score to different model-generated responses.

The model is fine-tuned using Proximal Policy Optimization (PPO), an algorithm that optimizes the model's policy based on the reward model. This process continues iteratively, with the model generating new completions, receiving feedback from the evaluators, and being fine-tuned accordingly.

The RLHF approach allows GPT-4 to improve its performance by learning from human preferences and feedback. This method helps mitigate issues such as biased or unsafe content generation and assists in generating more useful and coherent responses.

Can human involvement be automated? Possible alternatives could be:

Back-translation

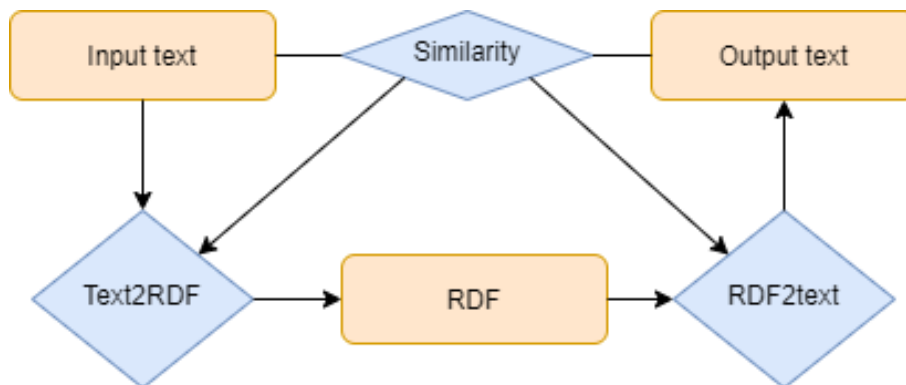


Figure 2.1: an overview of the basic idea.

Back-translation and the re-purposing of the resulting cross-entropy loss signal for fine-tuning the original translator would capture information about the quality of the knowledge graph. If the knowledge graph does not represent the information in the source text, then the back-translator would not be able to reconstruct the original text, and the loss would be high. If the knowledge graph accurately represents the source text, then the back translator can

more accurately recreate the original article, and the loss would be low.

Another possible way of using back-translation is to use the cosine similarity between the original source text embedding and a back-translator's reconstructed text embedding to calculate a loss function.

Self-critique

GPT-4 is able to follow instructions well. Given sufficient instructions such as an evaluation survey, GPT-4 may be able to fine-tune itself C.4.

Chapter 3

Methodology and methods

This study employs a mixed-methods approach to evaluate the effectiveness of advanced language models in generating knowledge graphs from unstructured text. The approach combines quantitative techniques, namely custom performance metrics and Likert scales, with qualitative metrics. This methodology aims to provide a nuanced understanding of the models' performance and potential areas for improvement.

3.1 Large pre-trained language models

3.1.1 Supervised fine-Tuning

The models T5 and BART are evaluated by using the evaluation metrics to measure performance on a set of news-article paragraphs, chunked to match the training data. The steps taken to produce the models are for reproducibility recorded as code in an appendix and discussed in the methods section.

The training was done using various cloud computing back-ends from Google's Colab notebook, and a "pro" subscription, with powerful GPU's such as for instance A100-SXM4-40GB. Runs could be anywhere from around 40 minutes to around 17 hours, the latter was a gigantic waste of compute with no performance benefit, and a waste of energy. Google colab also has TPUs available, but it requires some changes to the code compared to using GPU/CPU, so they were not used.

Figure 3.1 shows that experimental models generally improved and converged.

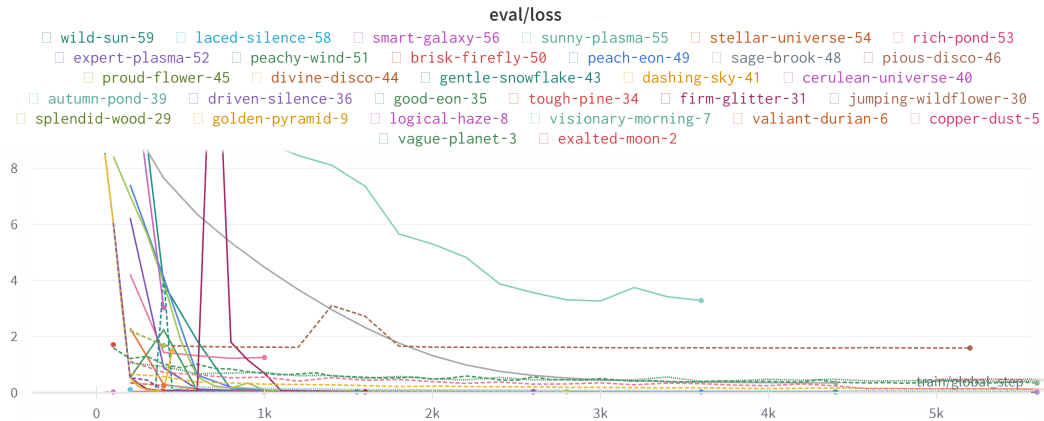


Figure 3.1: A detailed evaluation of the loss function.

Batch size was decided by the GPU memory, which was 40 GB. For DART and WebNLG, the memory allowed a batch size of 8, and only 4 for KELM and TEKGEN because of their larger instances, anything more would result in memory crash.

The learning rate followed a warmup schedule for ten percent of the total steps, and the learning rate ranged during runs from $5E-6$ to $5E-4$.

3.1.2 Zero-Shot

GPT is evaluated with the same evaluation metrics so that the results are comparable to the fine-tuned models, however, the advantage of being able to use longer text as input and thereby utilize long-range dependencies in the text is exploited. Standardizing input across all language models would artificially limit GPT, and would invalidate results rather than add validity. As an artifact, it is tested as-is without tempering, but the prompts used to trigger the display of capability could be considered artifacts as well for the purpose of adherence to formal research methodology language. The prompts are for reproducibility recorded in appendix and discussed in the methods section.

3.2 Datasets

WebNLG, DART, TEKGEN, and KELM datasets are evaluated and compared based on their evaluation metric performance in fine-

tuning the encoder-decoder models.

3.2.1 Data preprocessing

For consumption by the models, a standardized format was chosen such that each of the four data sets was transformed into: [subject | predicate | object], [subject two | predicate two | object two], [third subject | predicate | final triple's example literal].

The idea is to make the triple structure as obvious as possible, but at the same time utilizing some common sentence patterns, namely commas between triples, and ending the sentences with punctuation. This seemed to help the model deliberately finish its output, instead of cutting off in the middle of a triple.

During training, when trained over multiple epochs, there would be a marked loss increase at the beginning of each epoch. The solution to this was to use `sample(frac)` to randomize the rows. The problem was that webNLG is structured by default in ascending order from one triple and short text to seven triples and long text, and so the model would continuously over-fit to the output length it was seeing at that point in the epoch, never being on its toes in regards to output length.

Task prefix, relevant in regards to T5 training, was not used. There was no difference between using one and not, as was known to be the case when T5 is fine-tuned for only one task.

3.3 Knowledge graph evaluation

3.3.1 Gold label alternative

There can be multiple sensible ways to represent information in a knowledge graph, it is therefore difficult to measure objectively how good a generative model is using a hard gold label of expected triples, as opposed to a classification model where outputs are pre-determined.

One could imagine that semantic similarity measures could do a good job, and for recall it could, "father" and "son" have closely related meanings, used in the same contexts and would therefore be given similar embeddings from the methods that are currently the

convention. Indeed, using sentence-bert to vectorize "James is John's father" and "John is James' son" yields a cosine similarity of 0.95. However, "James is John's father" and "James is John's son" yields a similarity of 0.96, even though the semantics are inversed, making it useless for precision.

Using a set of alternative gold label versions of each fact is challenging. Using ontology software to generate implicit information, such as inverse relations, super classes and aliases would reduce the job of making alternative gold labels to mostly creating aliases, but it is difficult to capture unexpected aliases. [John, familialTie, James], [John, hasFamilyRelation, James], [John, isRelatedTo, James], etc. And it would only be practical to measure precision, not recall which would always be low.

Relations and entities of different facts need to match each other to form a consistent graph. Even if accuracy is seemingly high, there can be a mismatch between aliases used in different triples. Resolution to standard vocabularies is a necessary additional step for the outputs to become real knowledge graphs, but if the graph is not internally consistent it is unfortunate, it destroys the applicability of the output as is, and would also make it difficult to resolve aliases with methods that rely on context or graph matching.

There is complexity in deciding which combinations of facts constitute a good representation of the relevant information, this is true already with one version of gold fact per piece of information in a natural language passage, since they can be more or less important. [Event, hasWeather, SunnyWeather] might not be an important fact.

As the challenges discussed above makes it difficult to measure the quality of the output quantitatively using a gold label correspondence approach, the evaluation will be done qualitatively; Instead of creating a gold label and comparing results to it, the results will be assessed in terms of how good or lacking they would be as gold labels. The creation of gold labels would introduce subjectivity in the evaluation, and therefore our methodology is not less objective. The more or less quantitative and objective metrics will serve as support for the goldlabelness assessment of outputs, which again will serve as support for a qualitative evaluation of RDF lifting transformers as a sequence to sequence generation task.

3.3.2 Metrics

syntax

Does it follow the triple syntax? Knowledge graph should be organized in structured triples.

conciseness

Triples should be concise. Are the URIs, or proto-URIs(before linking) self-descriptive? Does the subject, predicate and object/literal work as subject, predicate, and object respectively? Or is there packing of unnecessary information into semantically compound atom names or information that should instead be separated into their own atoms?

Consistency

Synonyms should be mapped and ambiguities should be eliminated to ensure reconcilable expressions. And beyond that, is the KG consistently represented in terms of disjoint of classes, inconsistent inverse functional property values, and schema restrictions?

precision

The percentage of correct and relevant triples relative to the total number of triples.

recall

The percentage estimation of information in the source input captured in the output triples. Subjectively relevant information will be weighted higher. "On Monday morning the sounds of exploding artillery shells could be heard in Kyiv even through a heavy snow storm, evidence that the war is nearing". [Explosions, inProximityTo, Kyiv], [War, inProximityTo, Kyiv] is more relevant than [Event_#ex_identifier, weather, Snow storm], [Event_#ex_identifier, location, Kyiv], [Event_#ex_identifier, day, Monday].

Density

How densely connected is the graph? There should be an appropriate amount of interlinking.

Size

How large is the graph? It must be noted that more is not necessarily better, the amount should be appropriate for the task.

Design Functionality

Do the design choices effectively serve the intended purpose, which in this context is to represent news articles? An example of a suitable design choice for representing information in news articles is using an event ontology, but just using sensible relations that adequately model the information can also be functional. Referring to the example presented in the recall section above: "On Monday morning, the sounds of exploding artillery shells could be heard in Kyiv even through a heavy snow storm, evidence that the war is nearing." The triples [Explosions, inProximityTo, Kyiv] and [War, inProximityTo, Kyiv] are more relevant in terms of recall than [Event_ID, a, <http://purl.org/NET/c4dm/event.owl#Event>], [Event_ID, weather, Snow storm], [Event_ID, location, Kyiv], and [Event_ID, day, Monday]. However, the latter set of triples is a better design choice for the task at hand, as it is event-centric.

Mapping existing evaluation framework to ours

In table 3.3.2, the relevant metrics ("goals") of *Huaman* (2022) are mapped to the corresponding metrics of this thesis in which the pertaining evaluation questions are taken into account or incorporated. In table 3.3.2, the same is done for the relevant metrics (criteria) of *Chen et al.* (2019a).

Table 3.1: Mapping metrics from existing framework to ours

<i>Huaman</i> (2022)'s Metrics	Our Corresponding Metrics
Accuracy	Syntax and Precision
Appropriate amount	Size
Completeness	Recall and Density
Concise representation	Design Functionality
Consistent representation	Consistency

Continued on next page

Continued from previous page

Their Metrics	Our Corresponding Metrics
Ease of understanding	Conciseness
Free of error	Precision
Interoperability	Design Functionality
Objectivity	Recall
Relevancy	Design Functionality

Table 3.2: Mapping metrics from existing framework to ours

Chen et al. (2019a)'s Metrics	Our Corresponding Metrics
"Triples should be concise."	Conciseness
"Contextual information of entities should be captured."	Recall
"Entities should be densely connected."	Recall and Density
"Relations among different types of entities should be included."	Recall
"The attributes of the entities should not be missed."	Recall
"Synonyms should be mapped and ambiguities should be eliminated."	Consistency
"Knowledge graph should be organized in structured triples."	Syntax

3.3.3 Process of evaluation

Custom metrics will be manually applied to the results as the motivated alternative to the standard gold label type of evaluation.

They all will be evaluated as if the other metrics are forgiven up to the point at which they affect each other, because the interaction and overlap of how the metrics affect each other can be redundant, and would potentially ruin the information of each metric. For example, precision and recall (accuracy) will be evaluated by the relevancy of the information which the model attempts to capture even if consistency is botched, because that error is captured in another metric. And low conciseness can be forgiven in terms of accuracy if it is not due to botched semantics.

Some of the custom metrics are objective, namely "syntax", "consistency", and "precision". These are calculated by counting hits and misses and calculating the proportion as a percentage. Recall is only objective as far as human ability goes, so uncertainty is introduced. It can be difficult to calculate the percentage of information from the source text that is captured. Some information can be more important than other information as well, which arguably should be accounted for even at the loss of objectivity. On the other end of the objectivity scale, there is conciseness and functionality. Loss of objectivity is a limitation.

3.3.4 Likert scale

The Likert scale will be needed to evaluate some of the more subjective metrics, such as density, size, conciseness, and design functionality. And for the more objective metrics, it will serve to normalize scores by mapping them onto the Likert value they correspond to based on interval. The Likert scale strikes a balance between allowing sufficiently nuanced evaluation of the knowledge graphs while being exceptionally easy and practical to use.

Syntax:

1. Triples do not follow the syntax in most cases (0-20% of the time).
2. Triples occasionally follow the syntax (21-40% of the time).

3. Triples moderately follow the syntax (41-60% of the time).
4. Triples mostly follow the syntax, with only a few exceptions (61-80% of the time).
5. Triples consistently follow the syntax (81-100% of the time).

Conciseness:

1. Triples are not self-descriptive or frequently pack unnecessary information into semantically compound atom names or contain information that should be separated into their own atoms or attributes of blank nodes.
2. Triples often are not self-descriptive, or include unnecessary information, leading to a less concise structure, though some parts are well-formed.
3. Triples are generally self-descriptive, with occasional instances of unnecessary information or less concise structures.
4. Triples are mostly self-descriptive and effectively convey information, with rare instances of unnecessary information or compound atom names.
5. Triples are highly self-descriptive, with subject, predicate, and object/literal consistently and effectively working as subject, predicate, and object, respectively, without unnecessary information or compound atom names.

Consistency:

1. The KG has numerous inconsistencies such as synonyms not being mapped, ambiguities, disjoint classes, inconsistent inverse functional property values, and schema restrictions, making it very difficult to reconcile.
2. The KG has several inconsistencies in the mentioned areas, requiring significant effort to reconcile and improve its representation.
3. The KG has some inconsistencies, but they can be made consistent with moderate effort by addressing issues in synonyms, ambiguities, and other areas.

4. The KG has minor inconsistencies, and they should be easily reconcilable by applying minor fixes to synonym mapping, ambiguity elimination, and other aspects.
5. The KG is consistent with proper synonym mapping, ambiguity elimination, and other aspects ensuring reconcilable expressions and a coherent representation.

Precision:

1. Low precision (0-20% of the triples are correct and relevant).
2. Below average precision (21-40% of the triples are correct and relevant).
3. Moderate precision (41-60% of the triples are correct and relevant).
4. Above average precision (61-80% of the triples are correct and relevant).
5. High precision (81-100% of the triples are correct and relevant).

Recall:

1. Low recall (0-20% of the information in the source input is captured).
2. Below average recall (21-40% of the information in the source input is captured).
3. Moderate recall (41-60% of the information in the source input is captured).
4. Above average recall (61-80% of the information in the source input is captured).
5. High recall (81-100% of the information in the source input is captured).

Note: essential information can be rated higher. Rating it in terms of proportion of the text length is meaningless if all that is captured is useless details of no importance.

Density:

1. Sparse graph: The graph has very few connections between nodes that are relevant and meaningful to the context.

2. Below average density: The graph has a limited number of connections between nodes that are relevant and meaningful to the context.
3. Moderate density: The graph has a balanced number of connections between nodes that are relevant and meaningful to the context.
4. Above average density: The graph has a high number of connections between nodes that are relevant and meaningful to the context.
5. Dense graph: The graph has a very high number of connections between nodes that are relevant and meaningful to the context.

Size:

1. Small graph.
2. Below average size.
3. Moderate size.
4. Above average size.
5. Large graph.

Design Functionality:

1. The graph is highly ineffective for representing news articles.
2. The graph is ineffective for representing news articles.
3. The graph can represent news articles.
4. The graph is effective for representing news articles.
5. The graph is highly effective for representing news articles.

3.3.5 Sample input data

"" In the quaint town of Maplewood, residents are rejoicing after local businessman Tom Smith donated \$100,000 to the community center. The center, which serves as a hub for social gatherings and after-school activities, has been struggling financially due to a lack of funding. Smith, who owns a chain of successful hardware stores in the area, has long been known for his philanthropy and commitment

to the community. He frequently donates to local charities and participates in volunteer efforts around town. However, some residents have raised eyebrows at Smith's generous donation, questioning his motives.

Rumors have circulated that Smith may be using his donations as a way to gain influence in the community and curry favor with local women. Despite the whispers, the community center's board of directors remains grateful for Smith's contribution and hopes that it will inspire others to give back as well.

The center has already begun planning new programs and activities thanks to Smith's generosity. Meanwhile, the town council, which includes several members with ties to Smith's hardware stores, has granted him permits to build three new stores in lucrative areas of the county. Among the council members is Rendley Scott, a prominent businessman and political figure in the area who has received donations from Smith in the past.

Scott's support for the new stores has drawn criticism from some residents who are concerned about the impact on local wildlife, including a threatened population of glowing frogs whose habitat is being encroached upon by one of the building projects. Only time will tell if Smith's donation and business ventures are truly altruistic or if there are hidden motives behind his actions. For now, the people of Maplewood are content to enjoy the benefits of his generosity and celebrate the positive impact it will have on their town, while others remain wary of the potential consequences. ""

All the articles can be found in the appendix for evaluation data B.

Chapter 4

Results

In this chapter, we present the results and discuss the performance of various text generation models based on our custom evaluation metrics.

4.1 Results

Table 4.1 shows the evaluation results of GPT-4, AMRBART with AMR-LD, T5 fine-tuned on DART, WebNLG, KELM, TEKGEN, and BART fine-tuned on DART.

Table 4.1: Comparative Analysis of Text Generation Models

Criteria	GPT-4	AMRBART with AMR-LD	T5 DART	T5 WebNLG	T5 KELM	T5 TEKGEN	BART DART
Syntax	5	5	5	3	3	5	5
Conciseness	4	1	3	3	2	2	2
Consistency	5	2	2	2	1	1	2
Precision	5	5	2	2	2	1	1
Recall	5	5	2	2	1	1	2
Density	3	5	1	1	1	1	2
Size	4	5	3	3	3	1	3
Functionality	4	2	2	2	1	1	1

4.2 Discussion

4.2.1 Limitations

The evaluation method is qualitative in nature, and the metrics are applied manually. This introduces subjectivity and potential inconsistencies in the assessment of the models. For example, what is

an appropriate graph size relative to the corresponding source text? Future research should focus on developing more objective and automated evaluation methods for knowledge graph extraction tasks.

KELM and TEKGEN was sampled to create bastardized versions due to their enormous sizes. It is most likely that they would have performed better if large enterprise-level computational resources were available to utilize them. Another problem is that long data instances were pruned to fit the model input capabilities. If they could have been useful in training a document-level extractor it would greatly enhance their usefulness.

Due to AMR-BART with AMR-LD's low conciseness, it was very difficult to evaluate its performance. It is highly possible that mistakes in its KG output have gone unnoticed. As opposed to the other methods, it was only evaluated on the "Smith's Donation Motivation" because of its enormous output size. These sources of uncertainty must be taken into consideration.

The consistency of models were measured internally for each knowledge graph, but especially for GPT-4 there is a lot of variety between outputs. This could stand to be tuned.

4.2.2 Syntax

All the models, except for T5_webNLG and T5_KELM, show high adherence to the triple syntax. The lower performance of T5_webNLG and T5_KELM in this metric can likely be attributed to the datasets since the same architecture was able to perform better when trained on DART and TEKGEN.

4.2.3 Conciseness

GPT-4 and T5_DART exhibit better conciseness compared to other models. The lower conciseness scores for AMRBART with AMR-LD is due to low human readability, specifically the excessive lifting of unnecessary syntactic information and use of unique identifiers for everything, while the lower conciseness of T5_webNLG, T5_KELM, and T5_TEKGEN is due to poorer quality of entity and relations, which are often unnecessarily compounded and some times does not make sense.

4.2.4 Consistency

GPT-4 outperforms other models in terms of consistency, as it has no trouble keeping track of the full text and the bits of information in it, resulting in a qualitative difference above what any of the other models are capable of. GPT-4 was also able to link the resources, and thereby completing full end-to-end parallel information extraction: from natural language, to finished knowledge graph that is in compliance with RDF standards such that the output could be parsed by a visualizer some times without error directly, or after being fed back the error message to correct itself. However, there was considerable variation from output to output.

AMRBART uses a unique identifier for every occurrence of the same expression but should be reconcilable after post-processing. All of the dataset fine-tuned models, particularly T5_KELM and T5_TEKGEN, struggle with ensuring consistent expressions due to not being able to handle the whole text in one go. Trying to feed the models more than one or two sentences at a time only results in nonsense output (after excessive inference time). This is because the training data is of limited length, and extending the training data results in enormous GPU memory and compute requirements.

4.2.5 Precision and Recall

GPT-4 and AMRBART with AMR-LD demonstrate the highest precision and recall scores, effectively capturing correct and relevant information from the input text. In contrast, the other models have lower scores in these metrics; they struggle to extract relevant information and generate accurate triples. These may be the most damning results for the data-set fine-tuning approach. T5 is considered a state-of-the-art sequence2sequence model, and as can be seen in preliminary assessments in appenixC on webNLG, BART performs slightly worse than T5 again, which is the reason T5 was selected to test the other datasets. Once the models have learned the syntax, they are already over-fitted to the sample distribution in the dataset, and are not able to generalize adequately, and there is a proneness to hallucinating entities and relations from the datasets.

4.2.6 Density and Size

GPT-4 and AMRBART produce relatively denser and larger graphs compared to the other models, especially AMRBART almost excessively so. These results suggest that the GPT-4 and AMRBART models are more effective at capturing the interconnectedness of information within the input text. However, the T5 models trained on various datasets (DART, webNLG, KELM, and TEKGEN) produce less dense and smaller graphs, with no high-level information like types for instance, which is expected given the datasets.

4.2.7 Functionality

GPT-4 achieves the highest score in terms of design utility, or functionality, creating graphs that are more suitable for representing news articles with a score of 4. The other methods perform badly with only 2s and 1s, which highlights how little they are tailored for the specific task of lifting new texts.

4.3 Overall Performance

4.3.1 GPT-4 is the winner

The other models perform poorly in this metric. Only GPT-4 seems to be actually useful as it stands, at the task of end-to-end knowledge graph extraction. The results indicate that GPT-4 outperforms other models in most evaluation metrics, making it a strong candidate for knowledge graph extraction tasks. The performance of other models is largely governed by the quality and focus of their training datasets, as well as their ability to generalize and adapt to the specific task of RDF lifting transformers.

4.3.2 The method of evaluation is useful

The evaluation methodology used in this study, which relies on custom metrics instead of a gold-label correspondence approach, has proven useful in assessing the performance of different models.

The development of the metrics took trial and error, but the resulting metrics were easy to use and the methodology allows for an

understanding of the models' strengths and weaknesses, and the resulting knowledge graphs' quality and applicability.

Chapter 5

Conclusion

5.1 Summary of the results

The results provide an extensive comparison between the different text generation models for knowledge graph extraction. GPT-4 outperforms other models in most of the evaluation criteria even without any specific fine-tuning for the task. The model demonstrates full end-to-end integrated knowledge graph extraction and strong performance in terms of syntax, consistency, precision, and recall, and was in terms of density and size taking the golden middle road indicating that GPT-4 is capable of generating high-quality knowledge graphs from natural language text. It is easy to imagine GPT-4 performing better than a novice human annotator.

AMRBART, when combined with AMR-LD, demonstrates remarkable performance in certain aspects such as syntax, precision, and recall. It can be viewed as the pinnacle of linguistic sentence-tree-based parsing. However, due to its inherent limitation as a sentence-level parser, it faces challenges in maintaining consistency and functionality, and it generates overly large graphs. These issues may potentially be addressed through the further development of tools that facilitate AMR-to-RDF translation. After all, the information is captured, and the focus should be on extracting relevant content.

On the other hand, extracting and presenting only useful information while discarding the unnecessary elements from the highly detailed semantic representation of each sentence may be a more difficult task than it initially appears. Given that current large language models (LLMs) can understand the semantics of natural language directly, the need for making semantics explicit to extract rel-

evant knowledge might no longer be essential. If making semantics explicit is no longer beneficial, then imposing semantics into a formal framework could lead to the loss of nuance and create additional sources of error without any advantages.

The supervised models had the lowest performance overall, but showed some differentiation between data sets, and models, with T5 fine-tuned on DART being the better-performing model versus BART fine-tuned on DART, while T5 fine-tuned on KELM and TEKGEN underperform in most criteria.

5.2 Implication and Recommendation

With regard to knowledge extraction, this study points to the benefit of moving from isolation to contextual integration; The main implication to highlight is how limiting it is for a lifter to look at each sentence of an article one at a time, without the chance of seeing the big picture of the full context.

Based on the comparative analysis, GPT-4 must be recommended for knowledge graph extraction from natural language text. The model demonstrates superior performance in all evaluation criteria, making it the most exciting candidate for extracting structured information from unstructured text. However, there is room for improvement; tailoring to specific domains and tasks to promote consistent design choices between outputs. By using existing human feedback based RL fine-tuning it is likely possible to tailor and perfect its performance at the domain and task of lifting news texts.

5.3 Future Research Directions

Some potential future research directions include:

- Developing more objective and automated evaluation methods for assessing the quality of generated knowledge graphs.
- Exploring the use of reinforcement learning of GPT-4 based on feedback from GPT-4 itself. Given specific instructions it is able to provide constructive feedback to itself, which may be the most promising direction of research.

- Moving from sentence-level approaches to document-level and beyond. Following the trend of more contextual integration is better, it might be a good idea to extrapolate this to the multi-document level in the context of news articles.

Appendix A

Code for repeatability

A.1 Preprocessing of datasets for fine-tuning

A.1.1 WebNLG

```
webnlg = load_dataset("web_nlg", "release_v3.0_en")
```

```
def prepare_webnlg_dataset(webnlg):
```

```
    input_texts = []
```

```
    target_texts = []
```

```
    for example in webnlg:
```

```
        input_text = example['modified_triple_sets']['mtriple_set']
```

```
        target_text = example['lex']['text'][0]
```

```
        input_texts.append(input_text)
```

```
        target_texts.append(target_text)
```

```
# create a DataFrame with the input and target texts
```

```
df = pd.DataFrame(columns=["input_text", "target_text"])
```

```
df = pd.DataFrame({'input_text': target_texts, 'target_text': inp
```

```
def clean(list):
```

```
    list[0] = list[0].replace("_", "␣")
```

```
    list[1] = underscore(list[1]).replace("_", "␣").replace("\\", "")
```

```
    list[2] = list[2].replace("\\", "").replace("_", "␣")
```

```
    return list
```

```
def transform(list):
```

```

list = [triple.split("_|_") for triple in list]
list = ["_|_".join(clean(elem)) for elem in list]
list = "["+str("),_|_".join(list))+"]."
return list
df['target_text'] = df['target_text'].apply(lambda x: ("_|_").join(x))
df = df.sample(frac=1).reset_index(drop=True)

```

```

return df
train = prepare_webnlg_dataset(webnlg['train'])
val = prepare_webnlg_dataset(webnlg['dev'])

```

A.1.2 DART

```

dart = load_dataset("dart")
def prepare_dart_dataset(dart):
    df = pd.DataFrame(dart)
    df = df[['annotations', 'tripleaset']].rename(columns={'annotations': 'tripleaset'})
    df['input_text'] = df['input_text'].apply(lambda x: '_|_'.join(x['tripleaset']))

    def transformations(list):
        list[0] = list[0].replace("_", "_|_")
        list[1] = underscore(list[1]).replace("_", "_|_");
        list[2] = list[2].replace("\\\"", "\"").replace("_", "_|_")
        return list

    # stringify the list of triples, separate triples by punctuation marks
    df["target_text"] = df["target_text"].apply(lambda x: "["+str("),_|_".join(x['tripleaset']))+"]")
    df = df.sample(frac=1).reset_index(drop=True)
    return df

```

```

dart_train = prepare_dart_dataset(dart['train'])
dart_val = prepare_dart_dataset(dart['validation'])

```

A.1.3 TEKGEN

```

tekgen_train_url = 'https://storage.googleapis.com/gresearch/kelm-corpora/tekgen_train.txt'
tekgen_val_url = 'https://storage.googleapis.com/gresearch/kelm-corpora/tekgen_val.txt'

```



```

tekgen_train_file = 'tekgen_train.tsv'
tekgen_val_file = 'tekgen_val.tsv'

# Download the file from the URL and save it to the local file system
urllib.request.urlretrieve(tekgen_train_url, tekgen_train_file)
urllib.request.urlretrieve(tekgen_val_url, tekgen_val_file)

# Load the TSV file into a Pandas DataFrame
with open(tekgen_train_file, 'r') as f:
    lines = f.readlines()
    data = [json.loads(line.strip()) for line in lines]
tekgen_train_df = pd.DataFrame(data, columns=['triples', 'serialized'])
tekgen_train_df = tekgen_train_df.sample(n=80000, random_state=1)

with open(tekgen_val_file, 'r') as f:
    lines = f.readlines()
    data = [json.loads(line.strip()) for line in lines]
tekgen_val_df = pd.DataFrame(data, columns=['triples', 'serialized'])
tekgen_val_df = tekgen_val_df.sample(n=8000, random_state=1)

def prepare_tekgen_dataset(tekgen):
    df = pd.DataFrame(columns=["input_text", "target_text"])
    df["input_text"] = tekgen["sentence"]
    df["target_text"] = tekgen["triples"]

    def transformations(list):
        #list[0] = list[0].replace("_", " ")
        #list[1] = underscore(list[1]).replace("_", " ");
        #list[2] = list[2].replace("\\\"", "\"").replace("_", " ")
        return list
    # stringify the list of triples, separate triples by punctuation
    df["target_text"] = df["target_text"].apply(lambda x: "["+str(x),
    #df = df.sample(frac=1).reset_index(drop=True)
    return df

train = prepare_tekgen_dataset(tekgen_train_df)

```

```
val = prepare_tekgen_dataset(tekgen_val_df)
```

A.1.4 KELM

```
kelm_url = 'https://storage.googleapis.com/gresearch/kelm-corpus/upo
kelm_file = 'kelm.jsonl'
urllib.request.urlretrieve(kelm_url, kelm_file)
```

```
# Initialize a list to store the dictionaries
```

```
kelm_dict_list = []
```

```
# Read in each line of the JSON file as a dictionary and append to t
```

```
with open(kelm_file, 'r') as f:
```

```
    for line in f:
```

```
        kelm_dict = json.loads(line)
```

```
        kelm_dict_list.append(kelm_dict)
```

```
# Create the dataframe from the list of dictionaries
```

```
kelm_df = pd.DataFrame(kelm_dict_list)
```

```
def prepare_kelm_dataset(kelm):
```

```
    df = pd.DataFrame(columns=["input_text", "target_text"])
```

```
    df["input_text"] = kelm["gen_sentence"]
```

```
    df["target_text"] = kelm["triples"]
```

```
def transformations(list):
```

```
    #list[0] = list[0].replace("_", " ")
```

```
    #list[1] = underscore(list[1]).replace("_", " ");
```

```
    #list[2] = list[2].replace("\'", "'").replace("_", " ")
```

```
    return list
```

```
df = df[df['target_text'].notna() & ~df['target_text'].apply(lambda
```

```
# stringify the list of triples, separate triples by punctuation m
```

```
df["target_text"] = df["target_text"].apply(lambda x: "["+str("),_
```

```
#df = df.sample(frac=1).reset_index(drop=True)
```

```
return df
```

```
kelm_df = kelm_df.sample(n=100000, random_state=1).reset_index(drop
```

```
val = prepare_kelm_dataset(kelm_df[:7999])
```

```
train = prepare_kelm_dataset(kelm_df[8000:88000])
```

A.2 Models

A.2.1 GPT-4

Hey! I am going to test you on the task of knowledge graph extraction

```
"""
```

```
\section{Knowledge Graph Evaluation}
```

```
\subsection{Gold label alternative}
```

There can be multiple sensible ways to represent information in a k

One could imagine that semantic similarity measures could do a better job than word embeddings. For example, it might be better to vectorize "James is John's father" and "John is James' son"

Using a set of alternative gold label versions of each fact is challenging

Relations and entities of different facts need to match each other

There is complexity in deciding which combinations of facts constitute a single fact

As the challenges discussed above makes it difficult to measure the quality of a knowledge graph

```
\subsection{Metrics}
```

They all will be evaluated as if the other metrics were perfect, because

```
\subsection{syntax}
```

Does it follow the triple syntax? Knowledge graph should be organized as a set of

```
\subsection{conciseness}
```

Triples should be concise. Does the subject, predicate and object/li

`\subsection{Consistency}`

Synonyms should be mapped and ambiguities should be eliminated to en
reconcilable expressions.

`\subsection{precision}`

The percentage of correct and relevant triples relative to the total

`\subsection{recall}`

The percentage estimation of information in the source input capture

`\subsection{Density}`

How densely connected is the graph? The higher the better.

`\subsection{Size}`

How large is the graph? This is an important bit of information that

`\subsection{Design utility}`

Is the structure of the graph functional with regard to its purpose,
centric. In the example introduced in the recall section above: "On
centric.

`\subsection{Test Data}`

In the quaint town of Maplewood, residents are rejoicing after local
The center, which serves as a hub for social gatherings and after-
school activities, has been struggling financially due to a lack of
Smith, who owns a chain of successful hardware stores in the area, h
He frequently donates to local charities and participates in volunte
However, some residents have raised eyebrows at Smith's generous don
Rumors have circulated that Smith may be using his donations as a wa
Despite the whispers, the community center's board of directors rema
The center has already begun planning new programs and activities th
Meanwhile, the town council, which includes several members with tie
Among the council members is Rendley Scott, a prominent businessman
Scott's support for the new stores has drawn criticism from some res
Only time will tell if Smith's donation and business ventures are tr

For now, the people of Maplewood are content to enjoy the benefits
 ""

Task: Please lift the news article into an RDF knowledge graph with

A.2.2 AMRBART

```
from google.colab import drive
drive.mount('/content/drive')
```

```
## Create the file to do inference on
```

```
import json
import os
```

```
# Define the data and file paths
```

```
data=[{"sent": "In_the_quaint_town_of_Maplewood,_residents_are_rejo",
{"sent": "The_center,_which_serves_as_a_hub_for_social_gatherings_a",
{"sent": "Smith,_who_owns_a_chain_of_successful_hardware_stores_in",
{"sent": "He_frequently_donates_to_local_charities_and_participates",
{"sent": "However,_some_residents_have_raised_eyebrows_at_Smith's_g",
{"sent": "Rumors_have_circulated_that_Smith_may_be_using_his_donati",
{"sent": "Despite_the_whispers,_the_community_center's_board_of_dir",
{"sent": "The_center_has_already_begun_planning_new_programs_and_a_c",
{"sent": "Meanwhile,_the_town_council,_which_includes_several_memb",
{"sent": "Among_the_council_members_is_Rendley_Scott,_a_prominent_b",
{"sent": "Scott's_support_for_the_new_stores_has_drawn_criticism_fro",
{"sent": "Only_time_will_tell_if_Smith's_donation_and_business_vent",
{"sent": "For_now,_the_people_of_Maplewood_are_content_to_enjoy_the",
}]
```

```
text_data_file = '/content/drive/MyDrive/datasets/data4parsing.json'
amr_data_file = '/content/drive/MyDrive/datasets/smith_donations.amr'
rdf_data_file = '/content/drive/MyDrive/datasets/smith_donations.rdf'
```

```
# Write the text data to the file
```

```
with open(text_data_file, 'w') as f:
```

```

    for line in data:
        json.dump(line, f)
        f.write('\n')

    ## Use AMRBART to do inference
    # This step should read from 'text_data_file' and write the output to

%cd /content
!git clone https://github.com/goodbai-nlp/AMRBART.git

inference_script_path = "/content/AMRBART/fine-tune/inference-amr.sh"
model_path = "/content/drive/MyDrive/models/amrbart"

# Set the data path in the inference script
with open(inference_script_path, 'r') as f:
    script_content = f.read()

script_content = script_content.replace("DataPath=$RootDir/../../Data")
script_content = script_content.replace("python_u_main.py", "cd_/co

with open(inference_script_path, 'w') as f:
    f.write(script_content)

# Get recommended virtual env manager: Conda
!wget https://repo.anaconda.com/miniconda/Miniconda3-py37_4.8.2-Linux
!chmod +x Miniconda3-py37_4.8.2-Linux-x86_64.sh
!bash ./Miniconda3-py37_4.8.2-Linux-x86_64.sh -b -f -p /usr/local
!conda init
!source ~/.bashrc
!conda env update —file /content/AMRBART/requirements.yml

# Run the inference scripty
!source activate base && bash {inference_script_path} {model_path}

import shutil

source_file = "/content/AMRBART/fine-tune/outputs/Infer-examples-AMR

```

```

destination_file = "/content/drive/MyDrive/datasets/smith_donations_
shutil.copyfile(source_file, destination_file)

## Translate the AMR to RDF
%cd /content
!apt-get install python2
!curl https://bootstrap.pypa.io/pip/2.7/get-pip.py --output get-pip
!python2 get-pip.py
!pip2 install argparse_config rdflib requests numpy
!git clone https://github.com/BMKEG/amr-ld.git
import sys
sys.path.append('/content/amr-ld')
%cd amr-ld
# Remove line 21 of amr_to_rdf.py because it's an import that cause
!sed -i '21d' amr_to_rdf.py

# Use the 'amr_data_file' as input and 'rdf_data_file' as output
!python2 amr_to_rdf.py -i {amr_data_file} -o {rdf_data_file} -f ttl

```

A.2.3 T5 and BART

```

wandb.login()
config = {
    "num_epochs": 10,
    "batch_size": 8,
    "learning_rate": 5e-5,
    "warmup_steps_as_fraction_of_total_steps": 0.1,
    "patience": 2,
    "scheduler": "linear_schedule_with_warmup",
    "model_name": "t5-large",
    "dataset": "dart"
}
wandb.init(project="MasterOppgaveDefinitive", config=config, reinit
# Saving path
BEST_MODEL_PATH = f"/content/drive/MyDrive/models/definitive/{wandb
wandb.config.update({"save_path": BEST_MODEL_PATH})

```

```

# Torch seed for reproducibility
torch.manual_seed(0)

# Set up device
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

# Load model and tokenizer
tokenizer = AutoTokenizer.from_pretrained(wandb.config.model_name)
#model = AutoModelForSeq2SeqLM.from_pretrained(wandb.config.model_name)

# Tokenize
def tokenize(example):
    model_inputs = tokenizer(example["input_text"], return_tensors="pt")
    with tokenizer.as_target_tokenizer():
        labels = tokenizer(example["target_text"], return_tensors="pt")

    model_inputs["labels"] = labels.squeeze()
    return model_inputs

# Split dataset into train and validation sets
train_dataset = Dataset.from_pandas(dart_train)
val_dataset = Dataset.from_pandas(dart_val)
train_dataset = train_dataset.map(tokenize, batched=True, batch_size=100)
val_dataset = val_dataset.map(tokenize, batched=True, batch_size=100)

def data_collator(batch):
    input_ids = [torch.tensor(item["input_ids"]) for item in batch]
    labels = [torch.tensor(item["labels"]) for item in batch]
    attention_mask = [torch.tensor(item["attention_mask"]) for item in batch]

    encodings = {
        "input_ids": torch.stack(input_ids),
        "attention_mask": torch.stack(attention_mask),
        "labels": torch.stack(labels)
    }
    return encodings

```



```
# Define dataloaders
train_dataloader = DataLoader(train_dataset, batch_size=wandb.config.batch_size)
val_dataloader = DataLoader(val_dataset, batch_size=wandb.config.batch_size)

#wandb.init(project="MasterOppgaveDefinitive", config=config, reinit=True)

model = AutoModelForSeq2SeqLM.from_pretrained(os.path.dirname(wandb.config.run_dir))

# Define optimizer and scheduler
optimizer = torch.optim.AdamW(model.parameters(), lr=wandb.config.lr)
total_steps = len(train_dataloader) * wandb.config.num_epochs
scheduler = get_linear_schedule_with_warmup(optimizer, num_warmup_steps=wandb.config.warmup_steps)

# Define training variables
patience = wandb.config.patience
best_val_loss = float('inf')
epochs_no_improve = 0

# Fine-tune the model
for epoch in range(wandb.config.num_epochs):
    # Train the model for one epoch
    model.train()
    for batch in train_dataloader:
        input_ids = batch["input_ids"].to(device)
        attention_mask = batch["attention_mask"].to(device)
        labels = batch["labels"].to(device)
        outputs = model(input_ids=input_ids, attention_mask=attention_mask)
        loss = outputs.loss
        loss.backward()
        optimizer.step()
        scheduler.step()
        optimizer.zero_grad()
        wandb.log({"train_loss": loss.item(), "scheduler": scheduler.get_lr()})

    # Evaluate the model on the validation set
```

```

model.eval()
val_loss = 0
with torch.no_grad():
    for batch in val_dataloader:
        input_ids = batch["input_ids"].to(device)
        attention_mask = batch["attention_mask"].to(device)
        labels = batch["labels"].to(device)
        outputs = model(input_ids=input_ids, attention_mask=attention_mask)
        val_loss += outputs.loss.item()
    val_loss /= len(val_dataloader)
    wandb.log({"val_loss": val_loss})

# Save the best model
if val_loss < best_val_loss:
    print(f'Validation_Loss_Decreased({best_val_loss}→{val_loss})')
    best_val_loss = val_loss
    epochs_no_improve = 0
    if not os.path.isdir(os.path.dirname(wandb.config.save_path)):
        os.makedirs(os.path.dirname(wandb.config.save_path))
    model.save_pretrained(os.path.dirname(wandb.config.save_path))

else:
    print(f"validation_loss_only_{val_loss},_not_saving")
    epochs_no_improve += 1
    if epochs_no_improve == patience:
        print("Early_stopping_-_the_model_has_not_improved_for_{patience} epochs")
        break

# Unassign after finishing to not waste compute points while sleeping
from google.colab import runtime
runtime.unassign()

```

Appendix B

Evaluation data

B.1 Russian official makes new claim

Russian official makes new claim of Ukrainian attacks on Bryansk region From CNN's Sebastian Shukla and Olga Voitovych

A Russian official on Tuesday said Ukrainian forces had shelled the border town of Kurkovichy overnight the third time in three days he has accused Ukraine of attacking the southwest Russian region of Bryansk.

Bryansk Gov. Alexander Bogomaz, whose region shares a border to its south with Ukraine and to its west with Belarus, made the latest allegations in a Twitter post, adding there were no casualties.

It comes after he claimed on Monday that an IED exploded along the Bryansk-Unecha railway, which borders Ukraine, causing a freight train to derail.

Video footage taken at the scene of the railway explosion shows dozens of destroyed and burned out rail cars smoldering as emergency crews put out several fires. There were no casualties in that incident, he said.

On Sunday, the governor said Ukrainian shelling killed four civilians after strikes hit residential buildings in the village of Suzemka, in Syzemsky district.

In all instances, Bogomaz has not provided any evidence and CNN cannot independently verify his claims. Ukraine has not commented on the alleged incidents.

B.2 AI pioneer quits Google to warn about the technology's dangers

AI pioneer quits Google to warn about the technology's dangers

New York CNN Geoffrey Hinton, who has been called the Godfather of AI, confirmed Monday that he left his role at Google last week to speak out about the dangers of the technology he helped to develop.

Hinton's pioneering work on neural networks shaped artificial intelligence systems powering many of today's products. He worked part-time at Google for a decade on the tech giant's AI development efforts, but he has since come to have concerns about the technology and his role in advancing it.

I console myself with the normal excuse: If I hadn't done it, somebody else would have, Hinton told the New York Times, which was first to report his decision.

In a tweet Monday, Hinton said he left Google so he could speak freely about the risks of AI, rather than because of a desire to criticize Google specifically.

I left so that I could talk about the dangers of AI without considering how this impacts Google, Hinton said in a tweet. Google has acted very responsibly.

Jeff Dean, chief scientist at Google, said Hinton has made foundational breakthroughs in AI and expressed appreciation for Hinton's decade of contributions at Google.

We remain committed to a responsible approach to AI, Dean said in a statement provided to CNN. We're continually learning to understand emerging risks while also innovating boldly.

Hinton's decision to step back from the company and speak out on the technology comes as a growing number of lawmakers, advocacy groups and tech insiders have raised alarms about the potential for a new crop of AI-powered chatbots to spread misinformation and displace jobs.

The wave of attention around ChatGPT late last year helped renew an arms race among tech companies to develop and deploy similar AI tools in their products. OpenAI, Microsoft and Google are at the forefront of this trend, but IBM, Amazon, Baidu and Tencent are

working on similar technologies.

In March, some prominent figures in tech signed a letter calling for artificial intelligence labs to stop the training of the most powerful AI systems for at least six months, citing profound risks to society and humanity. The letter, published by the Future of Life Institute, a nonprofit backed by Elon Musk, came just two weeks after OpenAI announced GPT-4, an even more powerful version of the technology that powers ChatGPT. In early tests and a company demo, GPT-4 was used to draft lawsuits, pass standardized exams and build a working website from a hand-drawn sketch.

In the interview with the Times, Hinton echoed concerns about AIs potential to eliminate jobs and create a world where many will not be able to know what is true anymore. He also pointed to the stunning pace of advancement, far beyond what he and others had anticipated.

The idea that this stuff could actually get smarter than people a few people believed that, Hinton said in the interview. But most people thought it was way off. And I thought it was way off. I thought it was 30 to 50 years or even longer away. Obviously, I no longer think that.

Even before stepping aside from Google, Hinton had spoken publicly about AIs potential to do harm as well as good.

I believe that the rapid progress of AI is going to transform society in ways we do not fully understand and not all of the effects are going to be good, Hinton said in a 2021 commencement address at the Indian Institute of Technology Bombay in Mumbai. He noted how AI will boost healthcare while also creating opportunities for lethal autonomous weapons. I find this prospect much more immediate and much more terrifying than the prospect of robots taking over, which I think is a very long way off.

Hinton isn't the first Google employee to raise a red flag on AI. In July, the company fired an engineer who claimed an unreleased AI system had become sentient, saying he violated employment and data security policies. Many in the AI community pushed back strongly on the engineer's assertion.

CNN's Samantha Murphy Kelly and Ramishah Maruf contributed to this report.

B.3 Palestinian Khader Adnan dies in Israel jail after 86 days on hunger strike

Palestinian Khader Adnan dies in Israel jail after 86 days on hunger strike

By David Gritten BBC News A prominent Palestinian prisoner has died in an Israeli jail after 86 days on hunger strike.

Khader Adnan was a senior figure in the militant group Islamic Jihad who Israel had charged with terrorism offences.

The Israel Prison Service said he had refused medical treatment before he was found unconscious in his cell on Tuesday.

Hours after his death, militants fired three rockets from the Gaza Strip towards Israel. No-one was hurt.

The Palestinian prime minister accused Israel of "deliberately assassinating" Adnan, while Islamic Jihad warned Israel that it would pay "the highest price".

Adnan, who was 44 and from the north of the occupied West Bank, had been in and out of detention by Israel over the past two decades.

He had been on hunger strike four times before in protest, helping to make his name well known to Palestinians.

It has become common for Palestinian prisoners to refuse food for extended periods while in Israeli jails, but in recent years such deaths have been rare due to medical intervention.

Adnan began a fifth hunger strike immediately after being detained by Israeli forces at his home in Arraba, near the city of Jenin, on 5 February.

Israeli authorities accused him of supporting terrorism, affiliation with a terrorist group and incitement, and he was due to go on trial this month.

But the Palestinian prisoners' rights group Addameer said he was being held on "spurious charges intended to further suppress Palestinian activists".

Last week, the Palestinian Prisoners' Club, another advocacy group, warned that Adnan's health situation was "very serious". It said he was being held in the clinic at Ramla prison in central Israel, and that he was refusing nutritional supplements and medical examinations.

Adnan's wife, Randa Mousa, said he was doing that because Israeli authorities had "refused to transfer him to a civilian hospital [and] refused to allow his lawyer a visit".

On Tuesday, the Israel Prison Service announced that Adnan was "found early this morning in his cell unconscious", and that was taken to a hospital where he was declared dead after efforts to revive him failed.

A senior Israeli official told AFP news agency that Adnan had risked his life by refusing medical attention, adding: "In recent days, the military appeal court decided against releasing him from detention solely on the merit of his medical condition."

However, Palestinian Authority Prime Minister Mohammed Shtayyeh said Israel had "carried out a deliberate assassination against the prisoner Khader Adnan by rejecting his request for his release, neglecting him medically, and keeping him in his cell despite the seriousness of his health condition."

Islamic Jihad, which is the second most powerful militant group in the Hamas-controlled Gaza Strip, said: "Our fight is continuing and the enemy will realise once again that its crimes will not pass without a response."

Randa Mousa said she did not want people to grieve her husband's death.

"We will only receive well-wishers, because this martyrdom is [like] a wedding, a [moment of] pride for us and a crown on our heads," she told journalists in Arraba, according to AFP.

She also insisted she did not want "a drop of blood to be shed" in retaliation.

"We don't want anyone to respond to the martyrdom. We don't want someone to launch rockets and then [Israel] strikes Gaza," she added.

There was no immediate claim for the three rockets that were fired from Gaza towards Israel soon after Adnan's death.

The Israeli military said the rockets set off sirens but that no interceptors were launched because they fell in open areas.

The fate of their prisoners in Israel is a top issue for Palestinians, who hold Israel responsible for their well-being.

There are some 4,900 Palestinians in Israeli prisons, according to

Addameer.

Most are serving sentences after being convicted by Israeli courts or are being held for questioning, have been charged, or are awaiting or standing trial. It says another 1,016 are in "administrative detention", a controversial measure under which suspects are held indefinitely without charge or trial for renewable six-month periods. Palestinians regard all those held by Israel as political prisoners.

Addameer says the deportation of Palestinians from the occupied West Bank to prisons in Israel is illegal under international law. Palestinians also protest that it makes prison visits difficult because of strict conditions on Palestinians entering Israel from the West Bank.

B.4 Erling Haaland scored his 50th goal of the season in all competitions

Manchester City beat Fulham to return to the top of the Premier League as Erling Haaland scored his 50th goal of the season in all competitions.

Haaland converted a third-minute penalty, after Julian Alvarez had been fouled by Tim Ream, to become the first top-flight player to reach a half-century of goals in one campaign since Tom 'Pongo' Waring did so for Aston Villa in 1931.

The Norwegian striker also moved level for most goals in a Premier League season - matching the 34 goals Andy Cole scored for Newcastle in 1993-94 and the same amount Alan Shearer registered for Blackburn a year later.

However, Fulham grabbed a 15th-minute equaliser with their first attempt of the match as Carlos Vinicius powerfully shot past Ederson after Harry Wilson had headed the ball into his path.

But the hosts could not hang on to record what might have been a remarkable result.

Jack Grealish had an effort pushed on to the crossbar for Manchester City, who regained their lead in the 36th minute through Alvarez's superb strike from 25 yards out.

City's win moves them above - and one point clear of - Arsenal, with Pep Guardiola's side having a game in hand.

This is the first time City have been above Arsenal since mid-February as they aim to win the Premier League, Champions League and FA Cup treble.

B.5 Smith's Donation Motivation

In the quaint town of Maplewood, residents are rejoicing after local businessman Tom Smith donated \$100,000 to the community center. The center, which serves as a hub for social gatherings and after-school activities, has been struggling financially due to a lack of funding. Smith, who owns a chain of successful hardware stores in the area, has long been known for his philanthropy and commitment to the community. He frequently donates to local charities and participates in volunteer efforts around town. However, some residents have raised eyebrows at Smith's generous donation, questioning his motives. Rumors have circulated that Smith may be using his donations as a way to gain influence in the community and curry favor with local women. Despite the whispers, the community center's board of directors remains grateful for Smith's contribution and hopes that it will inspire others to give back as well. The center has already begun planning new programs and activities thanks to Smith's generosity. Meanwhile, the town council, which includes several members with ties to Smith's hardware stores, has granted him permits to build three new stores in lucrative areas of the county. Among the council members is Rendley Scott, a prominent businessman and political figure in the area who has received donations from Smith in the past. Scott's support for the new stores has drawn criticism from some residents who are concerned about the impact on local wildlife, including a threatened population of glowing frogs whose habitat is being encroached upon by one of the building projects. Only time will tell if Smith's donation and business ventures are truly altruistic or if there are hidden motives behind his actions. For now, the people of Maplewood are content to enjoy the benefits of his generosity and celebrate the positive impact it will have on their town, while others remain wary of the potential consequences.

Appendix C

Output of GPT-4

C.1 First output

The syntax is not perfect, but nearly.

```
@prefix ex: http://example.org/ .
@prefix rdfs: http://www.w3.org/2000/01/rdf-schema# .
@prefix foaf: http://xmlns.com/foaf/0.1/ .
@prefix schema: http://schema.org/ .
@prefix xsd: http://www.w3.org/2001/XMLSchema# .
```

```
ex:TomSmith a foaf:Person ;
foaf:name "Tom Smith" ;
ex:owns ex:HardwareStoreChain ;
ex:donatedTo ex:CommunityCenter ;
schema:memberOf ex:TownCouncil ;
ex:donationAmount "100000"^^xsd:integer .
```

```
ex:HardwareStoreChain a ex:Business ;
schema:numberOfItems 3 ;
ex:locatedIn ex:MaplewoodArea .
```

```
ex:MaplewoodArea a ex:GeographicalArea ;
rdfs:label "Maplewood" .
```

```
ex:CommunityCenter a schema:CivicStructure ;
schema:name "Maplewood Community Center" ;
ex:receivesFundingFrom ex:TomSmith ;
ex:plans ex:NewProgramsAndActivities .
```

```
ex:TownCouncil a schema:Organization ;
ex:hasMember ex:TomSmith ;
ex:hasMember ex:RendleyScott ;
ex:grantedPermit ex:BuildingProjects .
```

```
ex:RendleyScott a foaf:Person ;
foaf:name "Rendley Scott" ;
schema:memberOf ex:TownCouncil ;
ex:receivedDonationsFrom ex:TomSmith ;
ex:supports ex:BuildingProjects .
```

```
ex:BuildingProjects a ex:Construction ;
schema:numberOfItems 3 ;
ex:approvedBy ex:TownCouncil ;
ex:impacts ex:GlowingFrogHabitat .
```

```
ex:GlowingFrogHabitat a ex:Habitat ;
ex:hasSpecies ex:GlowingFrogs ;
ex:threatenedBy ex:BuildingProjects .
```

```
ex:GlowingFrogs a ex:Species ;
rdfs:label "Glowing Frogs" ;
ex:populationStatus ex:Threatened .
```

```
ex:NewProgramsAndActivities a ex:Programs ;
ex:fundedBy ex:TomSmith ;
ex:offeredAt ex:CommunityCenter .
```

```
ex:Threatened a ex:ConservationStatus ;
rdfs:label "Threatened" .
```

C.2 First regenerate output

```
@prefix ex: http://example.org/ .
```

```
ex:TomSmith a ex:Person ;
```

ex:hasOccupation ex:Businessman ;
ex:owns ex:HardwareStores ;
ex:donatedTo ex:MaplewoodCommunityCenter ;
ex:amountDonated 100000 ;
ex:knownFor ex:Philanthropy, ex:CommitmentToCommunity ;
ex:participatesIn ex:VolunteerEfforts ;
ex:hasRumoredMotive ex:GainingInfluence, ex:CurryingFavorWithLocalW

ex:MaplewoodCommunityCenter a ex:CommunityCenter ;
ex:locatedIn ex:Maplewood ;
ex:hosts ex:SocialGatherings, ex:AfterSchoolActivities ;
ex:strugglesWith ex:LackOfFunding ;
ex:receivesDonationFrom ex:TomSmith ;
ex:boardOfDirectors ex:GratefulToTomSmith ;
ex:plans ex:NewPrograms, ex:NewActivities .

ex:HardwareStores a ex:ChainOfStores ;
ex:ownedBy ex:TomSmith ;
ex:locatedIn ex:MaplewoodArea .

ex:TownCouncil a ex:Council ;
ex:locatedIn ex:Maplewood ;
ex:hasMembers ex:RendleyScott, ex:MembersWithTiesToHardwareStores ;
ex:grantsPermitTo ex:TomSmith ;
ex:permits ex:ThreeNewStores .

ex:RendleyScott a ex:Person ;
ex:hasRole ex:Businessman, ex:PoliticalFigure ;
ex:locatedIn ex:Maplewood ;
ex:memberOf ex:TownCouncil ;
ex:receivesDonationsFrom ex:TomSmith ;
ex:supports ex:NewStores .

ex:NewStores a ex:BuildingProjects ;
ex:owner ex:TomSmith ;
ex:grantedBy ex:TownCouncil ;
ex:locatedIn ex:LucrativeAreasOfCounty ;

ex:impactOn ex:LocalWildlife ;
ex:encroachesUpon ex:GlowingFrogsHabitat .

ex:GlowingFrogs a ex:ThreatenedSpecies ;
ex:habitat ex:AffectedByNewStores .

C.3 Second generate output

@prefix : http://example.org/ .
@prefix dbo: http://dbpedia.org/ontology/ .
@prefix foaf: http://xmlns.com/foaf/0.1/ .
@prefix xsd: http://www.w3.org/2001/XMLSchema# .

:TomSmith a foaf:Person ;
foaf:name "Tom Smith" ;
dbo:occupation :HardwareStoreOwner ;
dbo:philanthropy :DonationToCommunityCenter ;
dbo:associatedWith :TownCouncil ;
dbo:relatedTo :RendleyScott .

:DonationToCommunityCenter a dbo:Donation ;
dbo:amount "100000"^^xsd:integer ;
dbo:recipient :MaplewoodCommunityCenter .

:MaplewoodCommunityCenter a dbo:CommunityCenter ;
dbo:location :Maplewood ;
dbo:fundingStatus :StrugglingFinancially ;
dbo:newPrograms :NewProgramsAndActivities .

:HardwareStoreOwner a dbo:Occupation ;
dbo:businessName "Tom Smith's Hardware Store Chain" ;
dbo:location :MaplewoodArea .

:TownCouncil a dbo:Organization ;
dbo:member :TomSmith, :RendleyScott ;
dbo:decision :GrantBuildingPermits .

```
:RendleyScott a foaf:Person ;
foaf:name "Rendley Scott" ;
dbo:occupation :BusinessmanAndPoliticalFigure ;
dbo:location :MaplewoodArea ;
dbo:donationFrom :TomSmith .
```

```
:GrantBuildingPermits a dbo:Decision ;
dbo:permitCount "3"^^xsd:integer ;
dbo:recipient :TomSmith ;
dbo:impact :LocalWildlifeImpact .
```

```
:LocalWildlifeImpact a dbo:Impact ;
dbo:affectedSpecies :GlowingFrogs .
```

```
:GlowingFrogs a dbo:Species ;
dbo:populationStatus :Threatened ;
dbo:habitat :BuildingProjectArea .
```

```
:Maplewood a dbo:Place ;
dbo:name "Maplewood" .
```

```
:StrugglingFinancially a dbo:FinancialStatus ;
dbo:description "Lack of funding" .
```

```
:NewProgramsAndActivities a dbo:Program ;
dbo:description "New programs and activities enabled by Tom Smith's
```

```
:BusinessmanAndPoliticalFigure a dbo:Occupation ;
dbo:location :MaplewoodArea .
```

```
:BuildingProjectArea a dbo:Place ;
dbo:affectedBy :TomSmithsBuildingProjects .
```

```
:TomSmithsBuildingProjects a dbo:Project ;
dbo:location :LucrativeAreasOfCounty .
```

```
:LucrativeAreasOfCounty a dbo:Place ;
```

dbo:partOf :MaplewoodCounty .

:MaplewoodCounty a dbo:Place ;
dbo:name "Maplewood County" .

C.4 GPT-4 evaluating itself

PROMPT:

\subsection{Metrics}

\subsubsection{syntax}

Does it follow the triple syntax? Knowledge graph should be organized

\subsubsection{conciseness}

Triples should be concise. Are the URIs, or proto-URIs (before linking) descriptive? Does the subject, predicate and object/literal work as

\subsubsection{Consistency}

Synonyms should be mapped and ambiguities should be eliminated to ensure reconcilable expressions. And beyond that, is the KG consistently re

\subsubsection{precision}

The percentage of correct and relevant triples relative to the total

\subsubsection{recall}

The percentage estimation of information in the source input captured

\subsubsection{Density}

How densely connected is the graph? There should be an appropriate a

\subsubsection{Size}

How large is the graph? It must be noted that more is not necessarily

\subsubsection{Design Functionality}

Is the structure of the graph functional with regard to its purpose,

centric. In the example introduced in the recall section above: "On
centric.

\subsection{Process of evaluation}

They all will be evaluated as if the other metrics are forgiven up

Custom metrics will be manually applied to the results as the motiv

Some of the custom metrics are objective, namely "syntax", "consist

\subsection{Likert scale}

The Likert scale will be needed to evaluate some of the more subjec

\textbf{Syntax:}

1: Triples do not follow the syntax in most cases (0-20\% of the ti

2: Triples occasionally follow the syntax (21-40\% of the time).

3: Triples moderately follow the syntax (41-60\% of the time).

4: Triples mostly follow the syntax, with only a few exceptions (61
80\% of the time).

5: Triples consistently follow the syntax (81-100\% of the time).

\textbf{Conciseness:}

1: Triples are not self-descriptive or frequently pack unnecessary

2: Triples often are not self-descriptive, or include unnecessary in
formed.

3: Triples are generally self-descriptive, with occasional instance

4: Triples are mostly self-descriptive and effectively convey inform

5: Triples are highly self-descriptive, with subject, predicate, an

\textbf{Consistency:}\

1: The KG has numerous inconsistencies such as synonyms not being m

2: The KG has several inconsistencies in the mentioned areas, requi

3: The KG has some inconsistencies, but they can be made consistent

4: The KG has minor inconsistencies, and they should be easily reco

5: The KG is consistent with proper synonym mapping, ambiguity elim

\textbf{Precision:}

- 1: Low precision (0-20\% of the triples are correct and relevant).
- 2: Below average precision (21-40\% of the triples are correct and relevant).
- 3: Moderate precision (41-60\% of the triples are correct and relevant).
- 4: Above average precision (61-80\% of the triples are correct and relevant).
- 5: High precision (81-100\% of the triples are correct and relevant).

\textbf{Recall:}

- 1: Low recall (0-20\% of the information in the source input is captured).
 - 2: Below average recall (21-40\% of the information in the source input is captured).
 - 3: Moderate recall (41-60\% of the information in the source input is captured).
 - 4: Above average recall (61-80\% of the information in the source input is captured).
 - 5: High recall (81-100\% of the information in the source input is captured).
- Note; essential information can be rated higher. Rating it in terms of

\textbf{Density:}

- 1: Sparse graph: The graph has very few connections between nodes throughout.
- 2: Below average density: The graph has a limited number of connections between nodes.
- 3: Moderate density: The graph has a balanced number of connections between nodes.
- 4: Above average density: The graph has a high number of connections between nodes.
- 5: Dense graph: The graph has a very high number of connections between nodes.

\textbf{Size:}

- 1: Small graph.
- 2: Below average size.
- 3: Moderate size.
- 4: Above average size.
- 5: Large graph.

\textbf{Design Functionality:}

- 1: The graph is highly ineffective for representing news articles.
- 2: The graph is ineffective for representing news articles.
- 3: The graph can represent news articles.
- 4: The graph is effective for representing news articles.
- 5: The graph is highly effective for representing news articles.

\subsection{Test Data}

In the quaint town of Maplewood, residents are rejoicing after local school activities, has been struggling financially due to a lack of

Rumors have circulated that Smith may be using his donations as a w

The center has already begun planning new programs and activities t

Scott's support for the new stores has drawn criticism from some re

RESULTING OUTPUT FROM GPT-4:

\subsection{Test Results}

The following triples were generated based on the input text:

- 1: [Tom Smith, donates, \$100,000]
- 2: [Tom Smith, donatesTo, community center]
- 3: [community center, function, social gatherings and after-school activities]
- 4: [community center, financialStatus, struggling]
- 5: [Tom Smith, owns, hardware stores]
- 6: [Tom Smith, knownFor, philanthropy]
- 7: [Tom Smith, participatesIn, volunteer efforts]
- 8: [community center, boardOfDirectors, grateful]
- 9: [community center, plans, new programs and activities]
- 10: [town council, grants, building permits]
- 11: [Tom Smith, builds, new stores]
- 12: [Rendley Scott, councilMember, town council]
- 13: [Rendley Scott, relationship, Tom Smith]
- 14: [Rendley Scott, supports, new stores]
- 15: [new stores, impact, local wildlife]
- 16: [local wildlife, includes, glowing frogs]
- 17: [glowing frogs, status, threatened]

\textbf{Syntax:} 5

\textbf{Conciseness:} 4

\textbf{Consistency:} 4

\textbf{Precision:} 4

\textbf{Recall:} 4

\textbf{Density:} 3

\textbf{Size:} 3

\textbf{Design Functionality:} 4

Appendix D

Output of AMR-BART with AMR-LD

```
@prefix amr-core: <http://amr.isi.edu/rdf/core-amr#> .
@prefix amr-terms: <http://amr.isi.edu/rdf/amr-terms#> .
@prefix entity-types: <http://amr.isi.edu/entity-types#> .
@prefix propbank: <http://amr.isi.edu/frames/ld/v1.2.2/> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
```

```
amr-core:Concept a rdfs:Class ;
  rdfs:label "AMR-Concept" .
```

```
amr-core:Role a rdfs:Class ;
  rdfs:label "AMR-Role" .
```

```
<http://amr.isi.edu/amr_data/0#root01> a amr-core:AMR ;
  amr-core:has-id "0" ;
  amr-core:has-sentence "In the quaint town of Maplewood, residen
  amr-core:root <http://amr.isi.edu/amr_data/0#z0> .
```

```
<http://amr.isi.edu/amr_data/0#z10> a propbank:local-02 ;
  propbank:local-02.ARG1 <http://amr.isi.edu/amr_data/0#z8> .
```

```
<http://amr.isi.edu/amr_data/0#z2> a propbank:reside-01 ;
  propbank:reside-01.ARG0 <http://amr.isi.edu/amr_data/0#z1> ;
  propbank:reside-01.ARG1 <http://amr.isi.edu/amr_data/0#z3> .
```

```
<http://amr.isi.edu/amr_data/1#root01> a amr-core:AMR ;
  amr-core:has-id "1" ;
  amr-core:has-sentence "The center, which serves as a hub for so
```

school activities, has been struggling financially due to a lack of
 amr-core:root <http://amr.isi.edu/amr_data/1#z0> .

<http://amr.isi.edu/amr_data/1#z11> a propbank:cause-01 ;
 propbank:cause-01.ARG0 <http://amr.isi.edu/amr_data/1#z12> ;
 propbank:cause-01.ARG1 <http://amr.isi.edu/amr_data/1#z0> .

<http://amr.isi.edu/amr_data/1#z2> a propbank:serve-01 ;
 propbank:serve-01.ARG0 <http://amr.isi.edu/amr_data/1#z1> ;
 propbank:serve-01.ARG1 <http://amr.isi.edu/amr_data/1#z3> .

<http://amr.isi.edu/amr_data/10#root01> a amr-core:AMR ;
 amr-core:has-id "10" ;
 amr-core:has-sentence "Among the council members is Rendley Scott
 amr-core:root <http://amr.isi.edu/amr_data/10#z0> .

<http://amr.isi.edu/amr_data/10#z10> a propbank:donate-01 ;
 propbank:donate-01.ARG0 <http://amr.isi.edu/amr_data/10#z11> ;
 propbank:donate-01.ARG1 <http://amr.isi.edu/amr_data/10#z9> .

<http://amr.isi.edu/amr_data/10#z15> a propbank:have-org-role-
 91 ;
 propbank:have-org-role-91.ARG0 <http://amr.isi.edu/amr_data/10#z
 propbank:have-org-role-91.ARG1 <http://amr.isi.edu/amr_data/10#z
 propbank:have-org-role-91.ARG2 <http://amr.isi.edu/amr_data/10#z

<http://amr.isi.edu/amr_data/10#z8> a propbank:receive-01 ;
 propbank:receive-01.ARG0 <http://amr.isi.edu/amr_data/10#z1> ;
 propbank:receive-01.ARG1 <http://amr.isi.edu/amr_data/10#z9> ;
 amr-terms:time <http://amr.isi.edu/amr_data/10#z13> .

<http://amr.isi.edu/amr_data/11#root01> a amr-core:AMR ;
 amr-core:has-id "11" ;
 amr-core:has-sentence "Scott's support for the new stores has d
 amr-core:root <http://amr.isi.edu/amr_data/11#z0> .

<http://amr.isi.edu/amr_data/11#z10> a propbank:concern-01 ;

propbank:concern-01.ARG0 <http://amr.isi.edu/amr_data/11#z7> ;
propbank:concern-01.ARG1 <http://amr.isi.edu/amr_data/11#z11> .

<http://amr.isi.edu/amr_data/11#z13> a propbank:local-02 ;
propbank:local-02.ARG1 <http://amr.isi.edu/amr_data/11#z12> .

<http://amr.isi.edu/amr_data/11#z14> a propbank:include-01 ;
propbank:include-01.ARG1 <http://amr.isi.edu/amr_data/11#z15> ;
propbank:include-01.ARG2 <http://amr.isi.edu/amr_data/11#z12> .

<http://amr.isi.edu/amr_data/11#z17> a propbank:glow-01 ;
propbank:glow-01.ARG0 <http://amr.isi.edu/amr_data/11#z16> .

<http://amr.isi.edu/amr_data/11#z18> a propbank:threaten-01 ;
propbank:threaten-01.ARG1 <http://amr.isi.edu/amr_data/11#z15>

<http://amr.isi.edu/amr_data/11#z20> a propbank:encroach-01 ;
propbank:encroach-01.ARG0 <http://amr.isi.edu/amr_data/11#z21>
propbank:encroach-01.ARG1 <http://amr.isi.edu/amr_data/11#z19>

<http://amr.isi.edu/amr_data/11#z22> a propbank:include-91 ;
propbank:include-91.ARG1 <http://amr.isi.edu/amr_data/11#z21> ;
propbank:include-91.ARG2 <http://amr.isi.edu/amr_data/11#z23> .

<http://amr.isi.edu/amr_data/11#z5> a propbank:new-01 ;
propbank:new-01.ARG1 <http://amr.isi.edu/amr_data/11#z4> .

<http://amr.isi.edu/amr_data/11#z8> a propbank:reside-01 ;
propbank:reside-01.ARG0 <http://amr.isi.edu/amr_data/11#z7> .

<http://amr.isi.edu/amr_data/12#root01> a amr-core:AMR ;
amr-core:has-id "12" ;
amr-core:has-sentence "Only time will tell if Smith's donation
amr-core:root <http://amr.isi.edu/amr_data/12#z0> .

<http://amr.isi.edu/amr_data/12#z14> a propbank:hide-01 ;
propbank:hide-01.ARG1 <http://amr.isi.edu/amr_data/12#z13> .

<http://amr.isi.edu/amr_data/13#root01> a amr-core:AMR ;
amr-core:has-id "13" ;
amr-core:has-sentence "For now, the people of Maplewood are cont
amr-core:root <http://amr.isi.edu/amr_data/13#z0> .

<http://amr.isi.edu/amr_data/2#root01> a amr-core:AMR ;
amr-core:has-id "2" ;
amr-core:has-sentence "Smith, who owns a chain of successful ha
amr-core:root <http://amr.isi.edu/amr_data/2#z0> .

<http://amr.isi.edu/amr_data/2#z13> a propbank:long-03 ;
propbank:long-03.ARG1 <http://amr.isi.edu/amr_data/2#z0> .

<http://amr.isi.edu/amr_data/2#z3> a propbank:own-01 ;
propbank:own-01.ARG0 <http://amr.isi.edu/amr_data/2#z1> ;
propbank:own-01.ARG1 <http://amr.isi.edu/amr_data/2#z4> .

<http://amr.isi.edu/amr_data/2#z7> a propbank:succeed-01 ;
propbank:succeed-01.ARG1 <http://amr.isi.edu/amr_data/2#z5> .

<http://amr.isi.edu/amr_data/3#root01> a amr-core:AMR ;
amr-core:has-id "3" ;
amr-core:has-sentence "He frequently donates to local charities
amr-core:root <http://amr.isi.edu/amr_data/3#z0> .

<http://amr.isi.edu/amr_data/3#z4> a propbank:local-02 ;
propbank:local-02.ARG1 <http://amr.isi.edu/amr_data/3#z3> .

<http://amr.isi.edu/amr_data/3#z5> a propbank:frequent-02 ;
propbank:frequent-02.ARG1 <http://amr.isi.edu/amr_data/3#z1> .

<http://amr.isi.edu/amr_data/4#root01> a amr-core:AMR ;
amr-core:has-id "4" ;
amr-core:has-sentence "However, some residents have raised eyeb
amr-core:root <http://amr.isi.edu/amr_data/4#z0> .

<http://amr.isi.edu/amr_data/4#z10> a propbank:generous-01 ;
propbank:generous-01.ARG0 <http://amr.isi.edu/amr_data/4#z7> .

<http://amr.isi.edu/amr_data/4#z3> a propbank:reside-01 ;
propbank:reside-01.ARG0 <http://amr.isi.edu/amr_data/4#z2> .

<http://amr.isi.edu/amr_data/4#z6> a propbank:cause-01 ;
propbank:cause-01.ARG0 <http://amr.isi.edu/amr_data/4#z7> ;
propbank:cause-01.ARG1 <http://amr.isi.edu/amr_data/4#z1> .

<http://amr.isi.edu/amr_data/5#root01> a amr-core:AMR ;
amr-core:has-id "5" ;
amr-core:has-sentence "Rumors have circulated that Smith may be
amr-core:root <http://amr.isi.edu/amr_data/5#z0> .

<http://amr.isi.edu/amr_data/5#z15> a propbank:local-02 ;
propbank:local-02.ARG1 <http://amr.isi.edu/amr_data/5#z14> .

<http://amr.isi.edu/amr_data/5#z7> a propbank:donate-01 ;
propbank:donate-01.ARG0 <http://amr.isi.edu/amr_data/5#z4> ;
propbank:donate-01.ARG1 <http://amr.isi.edu/amr_data/5#z6> .

<http://amr.isi.edu/amr_data/6#root01> a amr-core:AMR ;
amr-core:has-id "6" ;
amr-core:has-sentence "Despite the whispers, the community cent
amr-core:root <http://amr.isi.edu/amr_data/6#z0> .

<http://amr.isi.edu/amr_data/6#z4> a amr-terms:center ;
amr-terms:mod <http://amr.isi.edu/amr_data/6#z5> ;
amr-terms:part <http://amr.isi.edu/amr_data/6#z2> .

<http://amr.isi.edu/amr_data/7#root01> a amr-core:AMR ;
amr-core:has-id "7" ;
amr-core:has-sentence "The center has already begun planning ne
amr-core:root <http://amr.isi.edu/amr_data/7#z0> .

<http://amr.isi.edu/amr_data/7#z6> a propbank:new-01 ;

propbank:new-01.ARG1 <http://amr.isi.edu/amr_data/7#z3> .

<http://amr.isi.edu/amr_data/7#z8> a propbank:cause-01 ;
propbank:cause-01.ARG0 <http://amr.isi.edu/amr_data/7#z9> ;
propbank:cause-01.ARG1 <http://amr.isi.edu/amr_data/7#z0> .

<http://amr.isi.edu/amr_data/8#root01> a amr-core:AMR ;
amr-core:has-id "8" ;
amr-core:has-sentence "Meanwhile, the town council, which includ
amr-core:root <http://amr.isi.edu/amr_data/8#z0> .

<http://amr.isi.edu/amr_data/8#z16> a propbank:new-01 ;
propbank:new-01.ARG1 <http://amr.isi.edu/amr_data/8#z15> .

<http://amr.isi.edu/amr_data/8#z19> a entity-types:county ;
amr-terms:part <http://amr.isi.edu/amr_data/8#z17> .

<http://amr.isi.edu/amr_data/8#z3> a propbank:include-01 ;
propbank:include-01.ARG1 <http://amr.isi.edu/amr_data/8#z4> ;
propbank:include-01.ARG2 <http://amr.isi.edu/amr_data/8#z1> .

<http://amr.isi.edu/amr_data/8#z5> a propbank:have-org-role-
91 ;
propbank:have-org-role-91.ARG0 <http://amr.isi.edu/amr_data/8#z4>
propbank:have-org-role-91.ARG1 <http://amr.isi.edu/amr_data/8#z1>
propbank:have-org-role-91.ARG2 <http://amr.isi.edu/amr_data/8#z6>

<http://amr.isi.edu/amr_data/8#z8> a propbank:tie-01 ;
propbank:tie-01.ARG1 <http://amr.isi.edu/amr_data/8#z4> ;
propbank:tie-01.ARG2 <http://amr.isi.edu/amr_data/8#z9> .

<http://amr.isi.edu/amr_data/9#root01> a amr-core:AMR ;
amr-core:has-id "9" ;
amr-core:has-sentence "Among the council members is Rendley Scot
amr-core:root <http://amr.isi.edu/amr_data/9#z0> .

<http://amr.isi.edu/amr_data/9#z10> a propbank:donate-01 ;

propbank:donate-01.ARG0 <http://amr.isi.edu/amr_data/9#z11> ;
propbank:donate-01.ARG1 <http://amr.isi.edu/amr_data/9#z9> .

<http://amr.isi.edu/amr_data/9#z15> a propbank:have-org-role-91 ;

propbank:have-org-role-91.ARG0 <http://amr.isi.edu/amr_data/9#z15> ;
propbank:have-org-role-91.ARG1 <http://amr.isi.edu/amr_data/9#z15> ;
propbank:have-org-role-91.ARG2 <http://amr.isi.edu/amr_data/9#z15> .

<http://amr.isi.edu/amr_data/9#z8> a propbank:receive-01 ;
propbank:receive-01.ARG0 <http://amr.isi.edu/amr_data/9#z1> ;
propbank:receive-01.ARG1 <http://amr.isi.edu/amr_data/9#z9> ;
amr-terms:time <http://amr.isi.edu/amr_data/9#z13> .

propbank:act-02.ARG0 a propbank:FrameRole .

propbank:appreciate-02.ARG0 a propbank:FrameRole .

propbank:appreciate-02.ARG1 a propbank:FrameRole .

propbank:begin-01.ARG0 a propbank:FrameRole .

propbank:begin-01.ARG1 a propbank:FrameRole .

propbank:benefit-01.ARG0 a propbank:FrameRole .

propbank:benefit-01.ARG1 a propbank:FrameRole .

propbank:build-01.ARG0 a propbank:FrameRole .

propbank:build-01.ARG1 a propbank:FrameRole .

propbank:cause-01.ARG0 a propbank:FrameRole .

propbank:cause-01.ARG1 a propbank:FrameRole .

propbank:celebrate-01.ARG0 a propbank:FrameRole .

propank:celebrate-01.ARG1 a propank:FrameRole .

propank:circulate-01.ARG1 a propank:FrameRole .

propank:commit-01.ARG1 a propank:FrameRole .

propank:commit-01.ARG2 a propank:FrameRole .

propank:concern-01.ARG0 a propank:FrameRole .

propank:concern-01.ARG1 a propank:FrameRole .

propank:consequence-03.ARG1 a propank:FrameRole .

propank:content-02.ARG0 a propank:FrameRole .

propank:content-02.ARG1 a propank:FrameRole .

propank:contrast-01.ARG1 a propank:FrameRole .

propank:contrast-01.ARG2 a propank:FrameRole .

propank:contribute-01.ARG0 a propank:FrameRole .

propank:criticize-01.ARG0 a propank:FrameRole .

propank:curry-01.ARG0 a propank:FrameRole .

propank:curry-01.ARG1 a propank:FrameRole .

propank:curry-01.ARG2 a propank:FrameRole .

propank:donate-01.ARG0 a propank:FrameRole .

propank:donate-01.ARG1 a propank:FrameRole .

propbank:donate-01.ARG2 a propbank:FrameRole .
propbank:draw-02.ARG0 a propbank:FrameRole .
propbank:draw-02.ARG1 a propbank:FrameRole .
propbank:effort-01.ARG1 a propbank:FrameRole .
propbank:encroach-01.ARG0 a propbank:FrameRole .
propbank:encroach-01.ARG1 a propbank:FrameRole .
propbank:enjoy-01.ARG0 a propbank:FrameRole .
propbank:enjoy-01.ARG1 a propbank:FrameRole .
propbank:frequent-02.ARG1 a propbank:FrameRole .
propbank:fund-01.ARG1 a propbank:FrameRole .
propbank:gain-02.ARG0 a propbank:FrameRole .
propbank:gain-02.ARG1 a propbank:FrameRole .
propbank:generous-01.ARG0 a propbank:FrameRole .
propbank:give-back-03.ARG0 a propbank:FrameRole .
propbank:glow-01.ARG0 a propbank:FrameRole .
propbank:have-org-role-91.ARG0 a propbank:FrameRole .
propbank:have-org-role-91.ARG1 a propbank:FrameRole .
propbank:have-org-role-91.ARG2 a propbank:FrameRole .
propbank:hide-01.ARG1 a propbank:FrameRole .

propbank:hope-01.ARG0 a propbank:FrameRole .
propbank:hope-01.ARG1 a propbank:FrameRole .
propbank:impact-01.ARG0 a propbank:FrameRole .
propbank:impact-01.ARG1 a propbank:FrameRole .
propbank:include-01.ARG1 a propbank:FrameRole .
propbank:include-01.ARG2 a propbank:FrameRole .
propbank:include-91.ARG1 a propbank:FrameRole .
propbank:include-91.ARG2 a propbank:FrameRole .
propbank:influence-01.ARG0 a propbank:FrameRole .
propbank:influence-01.ARG1 a propbank:FrameRole .
propbank:inspire-01.ARG0 a propbank:FrameRole .
propbank:inspire-01.ARG1 a propbank:FrameRole .
propbank:inspire-01.ARG2 a propbank:FrameRole .
propbank:know-02.ARG1 a propbank:FrameRole .
propbank:know-02.ARG3 a propbank:FrameRole .
propbank:lack-01.ARG0 a propbank:FrameRole .
propbank:lack-01.ARG1 a propbank:FrameRole .
propbank:local-02.ARG1 a propbank:FrameRole .

propbank:long-03.ARG1 a propbank:FrameRole .
propbank:new-01.ARG1 a propbank:FrameRole .
propbank:own-01.ARG0 a propbank:FrameRole .
propbank:own-01.ARG1 a propbank:FrameRole .
propbank:participate-01.ARG0 a propbank:FrameRole .
propbank:participate-01.ARG1 a propbank:FrameRole .
propbank:permit-01.ARG0 a propbank:FrameRole .
propbank:permit-01.ARG1 a propbank:FrameRole .
propbank:permit-01.ARG2 a propbank:FrameRole .
propbank:plan-01.ARG0 a propbank:FrameRole .
propbank:plan-01.ARG1 a propbank:FrameRole .
propbank:possible-01.ARG1 a propbank:FrameRole .
propbank:question-03.ARG0 a propbank:FrameRole .
propbank:question-03.ARG1 a propbank:FrameRole .
propbank:raise-01.ARG0 a propbank:FrameRole .
propbank:raise-01.ARG1 a propbank:FrameRole .
propbank:receive-01.ARG0 a propbank:FrameRole .
propbank:receive-01.ARG1 a propbank:FrameRole .
propbank:rejoice-01.ARG0 a propbank:FrameRole .

propank:remain-01.ARG1 a propank:FrameRole .
propank:remain-01.ARG3 a propank:FrameRole .
propank:reside-01.ARG0 a propank:FrameRole .
propank:reside-01.ARG1 a propank:FrameRole .
propank:rumor-01.ARG1 a propank:FrameRole .
propank:serve-01.ARG0 a propank:FrameRole .
propank:serve-01.ARG1 a propank:FrameRole .
propank:struggle-02.ARG0 a propank:FrameRole .
propank:struggle-02.ARG1 a propank:FrameRole .
propank:succeed-01.ARG1 a propank:FrameRole .
propank:support-01.ARG0 a propank:FrameRole .
propank:support-01.ARG1 a propank:FrameRole .
propank:tell-01.ARG0 a propank:FrameRole .
propank:tell-01.ARG1 a propank:FrameRole .
propank:threaten-01.ARG1 a propank:FrameRole .
propank:tie-01.ARG1 a propank:FrameRole .
propank:tie-01.ARG2 a propank:FrameRole .
propank:use-01.ARG0 a propank:FrameRole .

propbank:use-01.ARG1 a propbank:FrameRole .
propbank:use-01.ARG2 a propbank:FrameRole .
amr-terms:concession a amr-core:Role .
amr-terms:consist a amr-core:Role .
amr-terms:degree a amr-core:Role .
amr-terms:domain a amr-core:Role .
amr-terms:location a amr-core:Role .
amr-terms:manner a amr-core:Role .
amr-terms:mod a amr-core:Role .
amr-terms:op1 a amr-core:Role .
amr-terms:op2 a amr-core:Role .
amr-terms:part a amr-core:Role .
amr-terms:poss a amr-core:Role .
amr-terms:purpose a amr-core:Role .
amr-terms:quant a amr-core:Role .
amr-terms:time a amr-core:Role .
amr-terms:topic a amr-core:Role .
amr-terms:unit a amr-core:Role .
<http://amr.isi.edu/amr_data/0#z0> a propbank:rejoice-01 ;

propbank:rejoice-01.ARG0 <http://amr.isi.edu/amr_data/0#z1> ;
amr-terms:time <http://amr.isi.edu/amr_data/0#z6> .

<http://amr.isi.edu/amr_data/0#z11> a amr-terms:monetary-quantity ;
amr-terms:quant "100000" ;
amr-terms:unit <http://amr.isi.edu/amr_data/0#z12> .

<http://amr.isi.edu/amr_data/0#z12> a amr-terms:dollar .

<http://amr.isi.edu/amr_data/0#z13> a amr-terms:center ;
amr-terms:mod <http://amr.isi.edu/amr_data/0#z14> .

<http://amr.isi.edu/amr_data/0#z14> a amr-terms:community .

<http://amr.isi.edu/amr_data/0#z3> a amr-terms:town ;
rdfs:label "Maplewood" ;
amr-terms:mod <http://amr.isi.edu/amr_data/0#z5> ;
amr-terms:wiki "Maplewood,_New_Jersey" .

<http://amr.isi.edu/amr_data/0#z5> a amr-terms:quaint .

<http://amr.isi.edu/amr_data/0#z6> a amr-core:after ;
amr-terms:op1 <http://amr.isi.edu/amr_data/0#z7> .

<http://amr.isi.edu/amr_data/0#z7> a propbank:donate-01 ;
propbank:donate-01.ARG0 <http://amr.isi.edu/amr_data/0#z8> ;
propbank:donate-01.ARG1 <http://amr.isi.edu/amr_data/0#z11> ;
propbank:donate-01.ARG2 <http://amr.isi.edu/amr_data/0#z13> .

<http://amr.isi.edu/amr_data/1#z10> a amr-terms:finance .

<http://amr.isi.edu/amr_data/1#z12> a propbank:lack-01 ;
propbank:lack-01.ARG0 <http://amr.isi.edu/amr_data/1#z1> ;
propbank:lack-01.ARG1 <http://amr.isi.edu/amr_data/1#z13> .

<http://amr.isi.edu/amr_data/1#z13> a propbank:fund-01 ;
propbank:fund-01.ARG1 <http://amr.isi.edu/amr_data/1#z1> .

<http://amr.isi.edu/amr_data/1#z3> a amr-terms:hub ;
 amr-terms:purpose <http://amr.isi.edu/amr_data/1#z4> .

<http://amr.isi.edu/amr_data/1#z4> a amr-core:and ;
 amr-terms:op1 <http://amr.isi.edu/amr_data/1#z5> ;
 amr-terms:op2 <http://amr.isi.edu/amr_data/1#z7> .

<http://amr.isi.edu/amr_data/1#z5> a propbank:gather-03 ;
 amr-terms:mod <http://amr.isi.edu/amr_data/1#z6> .

<http://amr.isi.edu/amr_data/1#z6> a amr-terms:social .

<http://amr.isi.edu/amr_data/1#z7> a propbank:activity-06 ;
 amr-terms:time <http://amr.isi.edu/amr_data/1#z8> .

<http://amr.isi.edu/amr_data/1#z8> a amr-core:after ;
 amr-terms:op1 <http://amr.isi.edu/amr_data/1#z9> .

<http://amr.isi.edu/amr_data/1#z9> a entity-types:school .

<http://amr.isi.edu/amr_data/10#z0> a propbank:include-91 ;
 propbank:include-91.ARG1 <http://amr.isi.edu/amr_data/10#z1> ;
 propbank:include-91.ARG2 <http://amr.isi.edu/amr_data/10#z14> .

<http://amr.isi.edu/amr_data/10#z11> a entity-types:person ;
 rdfs:label "Smith" ;
 amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/10#z13> a amr-terms:past .

<http://amr.isi.edu/amr_data/10#z16> a amr-terms:council .

<http://amr.isi.edu/amr_data/10#z17> a amr-terms:member .

<http://amr.isi.edu/amr_data/10#z3> a amr-terms:businessman ;
 amr-terms:mod <http://amr.isi.edu/amr_data/10#z4> .

<http://amr.isi.edu/amr_data/10#z4> a amr-terms:prominent .

<http://amr.isi.edu/amr_data/10#z5> a amr-terms:figure ;
amr-terms:mod <http://amr.isi.edu/amr_data/10#z6> .

<http://amr.isi.edu/amr_data/10#z6> a amr-terms:politics .

<http://amr.isi.edu/amr_data/10#z7> a amr-terms:area .

<http://amr.isi.edu/amr_data/11#z0> a propbank:draw-02 ;
propbank:draw-02.ARG0 <http://amr.isi.edu/amr_data/11#z1> ;
propbank:draw-02.ARG1 <http://amr.isi.edu/amr_data/11#z6> .

<http://amr.isi.edu/amr_data/11#z1> a propbank:support-01 ;
propbank:support-01.ARG0 <http://amr.isi.edu/amr_data/11#z2> ;
propbank:support-01.ARG1 <http://amr.isi.edu/amr_data/11#z4> .

<http://amr.isi.edu/amr_data/11#z11> a propbank:impact-01 ;
propbank:impact-01.ARG0 <http://amr.isi.edu/amr_data/11#z4> ;
propbank:impact-01.ARG1 <http://amr.isi.edu/amr_data/11#z12> .

<http://amr.isi.edu/amr_data/11#z16> a amr-terms:frog ;
amr-terms:consist <http://amr.isi.edu/amr_data/11#z15> .

<http://amr.isi.edu/amr_data/11#z19> a amr-terms:habitat ;
amr-terms:poss <http://amr.isi.edu/amr_data/11#z15> .

<http://amr.isi.edu/amr_data/11#z2> a entity-types:person ;
rdfs:label "Scott" ;
amr-terms:wiki "Rick_Scott" .

<http://amr.isi.edu/amr_data/11#z23> a amr-terms:project ;
amr-terms:mod <http://amr.isi.edu/amr_data/11#z24> .

<http://amr.isi.edu/amr_data/11#z24> a propbank:build-01 .

<http://amr.isi.edu/amr_data/11#z6> a propbank:criticize-01 ;
propbank:criticize-01.ARG0 <http://amr.isi.edu/amr_data/11#z7>

<http://amr.isi.edu/amr_data/11#z9> a amr-terms:some .

<http://amr.isi.edu/amr_data/12#z0> a propbank:tell-01 ;
propbank:tell-01.ARG0 <http://amr.isi.edu/amr_data/12#z1> ;
propbank:tell-01.ARG1 <http://amr.isi.edu/amr_data/12#z3> .

<http://amr.isi.edu/amr_data/12#z1> a amr-core:time ;
amr-terms:mod <http://amr.isi.edu/amr_data/12#z2> .

<http://amr.isi.edu/amr_data/12#z10> a amr-terms:business .

<http://amr.isi.edu/amr_data/12#z11> a amr-terms:truly .

<http://amr.isi.edu/amr_data/12#z12> a propbank:cause-01 ;
propbank:cause-01.ARG0 <http://amr.isi.edu/amr_data/12#z13> ;
propbank:cause-01.ARG1 <http://amr.isi.edu/amr_data/12#z15> .

<http://amr.isi.edu/amr_data/12#z15> a propbank:act-02 ;
propbank:act-02.ARG0 <http://amr.isi.edu/amr_data/12#z7> .

<http://amr.isi.edu/amr_data/12#z2> a amr-terms:only .

<http://amr.isi.edu/amr_data/12#z3> a amr-core:or ;
amr-terms:op1 <http://amr.isi.edu/amr_data/12#z4> ;
amr-terms:op2 <http://amr.isi.edu/amr_data/12#z12> .

<http://amr.isi.edu/amr_data/12#z4> a amr-terms:altruistic ;
amr-terms:degree <http://amr.isi.edu/amr_data/12#z11> ;
amr-terms:domain <http://amr.isi.edu/amr_data/12#z5> .

<http://amr.isi.edu/amr_data/12#z5> a amr-core:and ;
amr-terms:op1 <http://amr.isi.edu/amr_data/12#z6> ;
amr-terms:op2 <http://amr.isi.edu/amr_data/12#z9> .

<http://amr.isi.edu/amr_data/12#z6> a propbank:donate-01 ;
propbank:donate-01.ARG0 <http://amr.isi.edu/amr_data/12#z7> .

<http://amr.isi.edu/amr_data/12#z9> a amr-terms:venture ;
amr-terms:mod <http://amr.isi.edu/amr_data/12#z10> ;
amr-terms:poss <http://amr.isi.edu/amr_data/12#z7> .

<http://amr.isi.edu/amr_data/13#z0> a propbank:contrast-01 ;
propbank:contrast-01.ARG1 <http://amr.isi.edu/amr_data/13#z1> ;
propbank:contrast-01.ARG2 <http://amr.isi.edu/amr_data/13#z14> .

<http://amr.isi.edu/amr_data/13#z1> a propbank:content-02 ;
propbank:content-02.ARG0 <http://amr.isi.edu/amr_data/13#z2> ;
propbank:content-02.ARG1 <http://amr.isi.edu/amr_data/13#z4> ;
amr-terms:time <http://amr.isi.edu/amr_data/13#z13> .

<http://amr.isi.edu/amr_data/13#z10> a propbank:celebrate-
01 ;
propbank:celebrate-01.ARG0 <http://amr.isi.edu/amr_data/13#z4> ;
propbank:celebrate-01.ARG1 <http://amr.isi.edu/amr_data/13#z11>

<http://amr.isi.edu/amr_data/13#z11> a propbank:impact-01 ;
propbank:impact-01.ARG0 <http://amr.isi.edu/amr_data/13#z8> ;
propbank:impact-01.ARG1 <http://amr.isi.edu/amr_data/13#z5> ;
amr-terms:mod <http://amr.isi.edu/amr_data/13#z12> .

<http://amr.isi.edu/amr_data/13#z12> a amr-terms:positive .

<http://amr.isi.edu/amr_data/13#z13> a amr-terms:now .

<http://amr.isi.edu/amr_data/13#z14> a propbank:remain-01 ;
propbank:remain-01.ARG1 <http://amr.isi.edu/amr_data/13#z15> ;
propbank:remain-01.ARG3 <http://amr.isi.edu/amr_data/13#z17> .

<http://amr.isi.edu/amr_data/13#z16> a amr-terms:other .

<http://amr.isi.edu/amr_data/13#z17> a amr-terms:wary ;

amr-terms:domain <http://amr.isi.edu/amr_data/13#z15> ;
amr-terms:topic <http://amr.isi.edu/amr_data/13#z18> .

<http://amr.isi.edu/amr_data/13#z18> a propbank:consequence-03 ;

propbank:consequence-03.ARG1 <http://amr.isi.edu/amr_data/13#z8> ;
amr-terms:mod <http://amr.isi.edu/amr_data/13#z19> .

<http://amr.isi.edu/amr_data/13#z19> a amr-terms:potential .

<http://amr.isi.edu/amr_data/13#z2> a amr-core:and ;
amr-terms:op1 <http://amr.isi.edu/amr_data/13#z3> ;
amr-terms:op2 <http://amr.isi.edu/amr_data/13#z10> .

<http://amr.isi.edu/amr_data/13#z3> a propbank:enjoy-01 ;
propbank:enjoy-01.ARG0 <http://amr.isi.edu/amr_data/13#z4> ;
propbank:enjoy-01.ARG1 <http://amr.isi.edu/amr_data/13#z7> .

<http://amr.isi.edu/amr_data/13#z7> a propbank:benefit-01 ;
propbank:benefit-01.ARG0 <http://amr.isi.edu/amr_data/13#z8> ;
propbank:benefit-01.ARG1 <http://amr.isi.edu/amr_data/13#z4> .

<http://amr.isi.edu/amr_data/13#z9> a amr-terms:he .

<http://amr.isi.edu/amr_data/2#z10> a amr-terms:philanthropy ;
amr-terms:poss <http://amr.isi.edu/amr_data/2#z1> .

<http://amr.isi.edu/amr_data/2#z11> a propbank:commit-01 ;
propbank:commit-01.ARG1 <http://amr.isi.edu/amr_data/2#z1> ;
propbank:commit-01.ARG2 <http://amr.isi.edu/amr_data/2#z12> .

<http://amr.isi.edu/amr_data/2#z12> a amr-terms:community .

<http://amr.isi.edu/amr_data/2#z5> a amr-terms:store ;
amr-terms:consist <http://amr.isi.edu/amr_data/2#z4> ;
amr-terms:location <http://amr.isi.edu/amr_data/2#z8> ;
amr-terms:mod <http://amr.isi.edu/amr_data/2#z6> .

<http://amr.isi.edu/amr_data/2#z6> a amr-terms:hardware .

<http://amr.isi.edu/amr_data/2#z8> a amr-terms:area .

<http://amr.isi.edu/amr_data/2#z9> a amr-core:and ;
amr-terms:op1 <http://amr.isi.edu/amr_data/2#z10> ;
amr-terms:op2 <http://amr.isi.edu/amr_data/2#z11> .

<http://amr.isi.edu/amr_data/3#z0> a amr-core:and ;
amr-terms:op1 <http://amr.isi.edu/amr_data/3#z1> ;
amr-terms:op2 <http://amr.isi.edu/amr_data/3#z6> .

<http://amr.isi.edu/amr_data/3#z10> a amr-terms:town .

<http://amr.isi.edu/amr_data/3#z6> a propbank:participate-01 ;

propbank:participate-01.ARG0 <http://amr.isi.edu/amr_data/3#z2>
propbank:participate-01.ARG1 <http://amr.isi.edu/amr_data/3#z7>

<http://amr.isi.edu/amr_data/3#z7> a propbank:effort-01 ;
propbank:effort-01.ARG1 <http://amr.isi.edu/amr_data/3#z8> ;
amr-terms:location <http://amr.isi.edu/amr_data/3#z9> .

<http://amr.isi.edu/amr_data/3#z8> a propbank:volunteer-01 .

<http://amr.isi.edu/amr_data/3#z9> a amr-terms:around ;
amr-terms:op1 <http://amr.isi.edu/amr_data/3#z10> .

<http://amr.isi.edu/amr_data/4#z0> a propbank:contrast-01 ;
propbank:contrast-01.ARG2 <http://amr.isi.edu/amr_data/4#z1> .

<http://amr.isi.edu/amr_data/4#z11> a propbank:question-03 ;
propbank:question-03.ARG0 <http://amr.isi.edu/amr_data/4#z2> ;
propbank:question-03.ARG1 <http://amr.isi.edu/amr_data/4#z12> .

<http://amr.isi.edu/amr_data/4#z12> a amr-terms:motive ;

amr-terms:poss <http://amr.isi.edu/amr_data/4#z8> .

<http://amr.isi.edu/amr_data/4#z4> a amr-terms:some .

<http://amr.isi.edu/amr_data/4#z5> a amr-terms:eyebrow .

<http://amr.isi.edu/amr_data/5#z0> a propbank:circulate-01 ;
propbank:circulate-01.ARG1 <http://amr.isi.edu/amr_data/5#z1> .

<http://amr.isi.edu/amr_data/5#z1> a propbank:rumor-01 ;
propbank:rumor-01.ARG1 <http://amr.isi.edu/amr_data/5#z2> .

<http://amr.isi.edu/amr_data/5#z10> a propbank:influence-01 ;
propbank:influence-01.ARG0 <http://amr.isi.edu/amr_data/5#z4> ;
propbank:influence-01.ARG1 <http://amr.isi.edu/amr_data/5#z11>

<http://amr.isi.edu/amr_data/5#z11> a amr-terms:community .

<http://amr.isi.edu/amr_data/5#z12> a propbank:curry-01 ;
propbank:curry-01.ARG0 <http://amr.isi.edu/amr_data/5#z4> ;
propbank:curry-01.ARG1 <http://amr.isi.edu/amr_data/5#z13> ;
propbank:curry-01.ARG2 <http://amr.isi.edu/amr_data/5#z14> .

<http://amr.isi.edu/amr_data/5#z13> a amr-terms:favor .

<http://amr.isi.edu/amr_data/5#z2> a propbank:possible-01 ;
propbank:possible-01.ARG1 <http://amr.isi.edu/amr_data/5#z3> .

<http://amr.isi.edu/amr_data/5#z3> a propbank:use-01 ;
propbank:use-01.ARG0 <http://amr.isi.edu/amr_data/5#z4> ;
propbank:use-01.ARG1 <http://amr.isi.edu/amr_data/5#z6> ;
propbank:use-01.ARG2 <http://amr.isi.edu/amr_data/5#z8> .

<http://amr.isi.edu/amr_data/5#z8> a amr-core:and ;
amr-terms:op1 <http://amr.isi.edu/amr_data/5#z9> ;
amr-terms:op2 <http://amr.isi.edu/amr_data/5#z12> .

<http://amr.isi.edu/amr_data/5#z9> a propbank:gain-02 ;
propbank:gain-02.ARG0 <http://amr.isi.edu/amr_data/5#z4> ;
propbank:gain-02.ARG1 <http://amr.isi.edu/amr_data/5#z10> .

<http://amr.isi.edu/amr_data/6#z0> a amr-core:and ;
amr-terms:concession <http://amr.isi.edu/amr_data/6#z16> ;
amr-terms:op1 <http://amr.isi.edu/amr_data/6#z1> ;
amr-terms:op2 <http://amr.isi.edu/amr_data/6#z10> .

<http://amr.isi.edu/amr_data/6#z1> a propbank:remain-01 ;
propbank:remain-01.ARG1 <http://amr.isi.edu/amr_data/6#z2> ;
propbank:remain-01.ARG3 <http://amr.isi.edu/amr_data/6#z6> .

<http://amr.isi.edu/amr_data/6#z10> a propbank:hope-01 ;
propbank:hope-01.ARG0 <http://amr.isi.edu/amr_data/6#z2> ;
propbank:hope-01.ARG1 <http://amr.isi.edu/amr_data/6#z11> .

<http://amr.isi.edu/amr_data/6#z11> a propbank:inspire-01 ;
propbank:inspire-01.ARG0 <http://amr.isi.edu/amr_data/6#z7> ;
propbank:inspire-01.ARG1 <http://amr.isi.edu/amr_data/6#z12> ;
propbank:inspire-01.ARG2 <http://amr.isi.edu/amr_data/6#z14> .

<http://amr.isi.edu/amr_data/6#z13> a amr-terms:other .

<http://amr.isi.edu/amr_data/6#z14> a propbank:give-back-03 ;
propbank:give-back-03.ARG0 <http://amr.isi.edu/amr_data/6#z12> ;
amr-terms:mod <http://amr.isi.edu/amr_data/6#z15> .

<http://amr.isi.edu/amr_data/6#z15> a amr-terms:as-well .

<http://amr.isi.edu/amr_data/6#z16> a propbank:whisper-01 .

<http://amr.isi.edu/amr_data/6#z5> a amr-terms:community .

<http://amr.isi.edu/amr_data/6#z6> a propbank:appreciate-02 ;
propbank:appreciate-02.ARG0 <http://amr.isi.edu/amr_data/6#z2> ;
propbank:appreciate-02.ARG1 <http://amr.isi.edu/amr_data/6#z7> .

<http://amr.isi.edu/amr_data/6#z8> a entity-types:person ;
 rdfs:label "Smith" ;
 amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/7#z10> a entity-types:person ;
 rdfs:label "Smith" ;
 amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/7#z2> a propbank:plan-01 ;
 propbank:plan-01.ARG0 <http://amr.isi.edu/amr_data/7#z1> ;
 propbank:plan-01.ARG1 <http://amr.isi.edu/amr_data/7#z3> .

<http://amr.isi.edu/amr_data/7#z4> a amr-terms:program .

<http://amr.isi.edu/amr_data/7#z5> a propbank:activity-06 .

<http://amr.isi.edu/amr_data/7#z7> a amr-terms:already .

<http://amr.isi.edu/amr_data/7#z9> a propbank:generous-01 ;
 propbank:generous-01.ARG0 <http://amr.isi.edu/amr_data/7#z10> .

<http://amr.isi.edu/amr_data/8#z0> a propbank:permit-01 ;
 propbank:permit-01.ARG0 <http://amr.isi.edu/amr_data/8#z1> ;
 propbank:permit-01.ARG1 <http://amr.isi.edu/amr_data/8#z13> ;
 propbank:permit-01.ARG2 <http://amr.isi.edu/amr_data/8#z14> ;
 amr-terms:time <http://amr.isi.edu/amr_data/8#z20> .

<http://amr.isi.edu/amr_data/8#z10> a amr-terms:hardware .

<http://amr.isi.edu/amr_data/8#z11> a entity-types:person ;
 rdfs:label "Smith" ;
 amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/8#z13> a propbank:build-01 ;
 propbank:build-01.ARG0 <http://amr.isi.edu/amr_data/8#z14> ;
 propbank:build-01.ARG1 <http://amr.isi.edu/amr_data/8#z15> ;

amr-terms:location <http://amr.isi.edu/amr_data/8#z17> .
<http://amr.isi.edu/amr_data/8#z18> a amr-terms:lucrative .
<http://amr.isi.edu/amr_data/8#z2> a amr-terms:town .
<http://amr.isi.edu/amr_data/8#z20> a amr-terms:meanwhile .
<http://amr.isi.edu/amr_data/8#z6> a amr-terms:member .
<http://amr.isi.edu/amr_data/8#z7> a amr-terms:several .
<http://amr.isi.edu/amr_data/8#z9> a amr-terms:store ;
 amr-terms:mod <http://amr.isi.edu/amr_data/8#z10> ;
 amr-terms:poss <http://amr.isi.edu/amr_data/8#z11> .
<http://amr.isi.edu/amr_data/9#z0> a propbank:include-91 ;
 propbank:include-91.ARG1 <http://amr.isi.edu/amr_data/9#z1> ;
 propbank:include-91.ARG2 <http://amr.isi.edu/amr_data/9#z14> .
<http://amr.isi.edu/amr_data/9#z11> a entity-types:person ;
 rdfs:label "Smith" ;
 amr-terms:wiki "-" .
<http://amr.isi.edu/amr_data/9#z13> a amr-terms:past .
<http://amr.isi.edu/amr_data/9#z16> a amr-terms:council .
<http://amr.isi.edu/amr_data/9#z17> a amr-terms:member .
<http://amr.isi.edu/amr_data/9#z3> a amr-terms:businessman ;
 amr-terms:mod <http://amr.isi.edu/amr_data/9#z4> .
<http://amr.isi.edu/amr_data/9#z4> a amr-terms:prominent .
<http://amr.isi.edu/amr_data/9#z5> a amr-terms:figure ;
 amr-terms:mod <http://amr.isi.edu/amr_data/9#z6> .

<http://amr.isi.edu/amr_data/9#z6> a amr-terms:politics .

<http://amr.isi.edu/amr_data/9#z7> a amr-terms:area .

entity-types:city a amr-core:NamedEntity .

entity-types:county a amr-core:NamedEntity .

entity-types:organization a amr-core:NamedEntity .

entity-types:school a amr-core:NamedEntity .

propbank:act-02 a amr-core:Frame .

propbank:appreciate-02 a amr-core:Frame .

propbank:begin-01 a amr-core:Frame .

propbank:benefit-01 a amr-core:Frame .

propbank:celebrate-01 a amr-core:Frame .

propbank:circulate-01 a amr-core:Frame .

propbank:commit-01 a amr-core:Frame .

propbank:concern-01 a amr-core:Frame .

propbank:consequence-03 a amr-core:Frame .

propbank:content-02 a amr-core:Frame .

propbank:contribute-01 a amr-core:Frame .

propbank:criticize-01 a amr-core:Frame .

propbank:curry-01 a amr-core:Frame .
propbank:draw-02 a amr-core:Frame .
propbank:effort-01 a amr-core:Frame .
propbank:encroach-01 a amr-core:Frame .
propbank:enjoy-01 a amr-core:Frame .
propbank:frequent-02 a amr-core:Frame .
propbank:fund-01 a amr-core:Frame .
propbank:gain-02 a amr-core:Frame .
propbank:gather-03 a amr-core:Frame .
propbank:give-back-03 a amr-core:Frame .
propbank:glow-01 a amr-core:Frame .
propbank:hide-01 a amr-core:Frame .
propbank:hope-01 a amr-core:Frame .
propbank:influence-01 a amr-core:Frame .
propbank:inspire-01 a amr-core:Frame .
propbank:know-02 a amr-core:Frame .
propbank:lack-01 a amr-core:Frame .
propbank:long-03 a amr-core:Frame .
propbank:own-01 a amr-core:Frame .

propbank:participate-01 a amr-core:Frame .

propbank:permit-01 a amr-core:Frame .

propbank:plan-01 a amr-core:Frame .

propbank:possible-01 a amr-core:Frame .

propbank:question-03 a amr-core:Frame .

propbank:raise-01 a amr-core:Frame .

propbank:rejoice-01 a amr-core:Frame .

propbank:rumor-01 a amr-core:Frame .

propbank:serve-01 a amr-core:Frame .

propbank:struggle-02 a amr-core:Frame .

propbank:succeed-01 a amr-core:Frame .

propbank:support-01 a amr-core:Frame .

propbank:tell-01 a amr-core:Frame .

propbank:threaten-01 a amr-core:Frame .

propbank:tie-01 a amr-core:Frame .

propbank:use-01 a amr-core:Frame .

propbank:volunteer-01 a amr-core:Frame .

propbank:whisper-01 a amr-core:Frame .

amr-terms:already a amr-core:Concept .

amr-terms:altruistic a amr-core:Concept .

amr-terms:around a amr-core:Concept .

amr-terms:as-well a amr-core:Concept .

amr-terms:business a amr-core:Concept .

amr-terms:chain a amr-core:Concept .

amr-terms:charity a amr-core:Concept .

amr-terms:dollar a amr-core:Concept .

amr-terms:eyebrow a amr-core:Concept .

amr-terms:favor a amr-core:Concept .

amr-terms:finance a amr-core:Concept .

amr-terms:frog a amr-core:Concept .

amr-terms:habitat a amr-core:Concept .

amr-terms:hub a amr-core:Concept .

amr-terms:lucrative a amr-core:Concept .

amr-terms:meanwhile a amr-core:Concept .

amr-terms:monetary-quantity a amr-core:Concept .

amr-terms:now a amr-core:Concept .

amr-terms:only a amr-core:Concept .

```
amr-terms:philanthropy a amr-core:Concept .
amr-terms:population a amr-core:Concept .
amr-terms:positive a amr-core:Concept .
amr-terms:potential a amr-core:Concept .
amr-terms:program a amr-core:Concept .
amr-terms:quaint a amr-core:Concept .
amr-terms:several a amr-core:Concept .
amr-terms:social a amr-core:Concept .
amr-terms:truly a amr-core:Concept .
amr-terms:venture a amr-core:Concept .
amr-terms:wary a amr-core:Concept .
amr-terms:wildlife a amr-core:Concept .
amr-terms:woman a amr-core:Concept .
amr-core:or a amr-core:Concept .
amr-core:time a amr-core:Concept .
<http://amr.isi.edu/amr_data/0#z1> a entity-types:person .
<http://amr.isi.edu/amr_data/0#z8> a amr-terms:businessman ;
    rdfs:label "Tom Smith" ;
    amr-terms:wiki "-" .
```

<http://amr.isi.edu/amr_data/1#z0> a propbank:struggle-02 ;
propbank:struggle-02.ARG0 <http://amr.isi.edu/amr_data/1#z1> ;
propbank:struggle-02.ARG1 <http://amr.isi.edu/amr_data/1#z10> .

<http://amr.isi.edu/amr_data/10#z1> a entity-types:person ;
rdfs:label "Rendley Scott" ;
amr-terms:location <http://amr.isi.edu/amr_data/10#z7> ;
amr-terms:mod <http://amr.isi.edu/amr_data/10#z3>,
<http://amr.isi.edu/amr_data/10#z5> ;
amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/10#z14> a entity-types:person .

<http://amr.isi.edu/amr_data/10#z9> a entity-types:thing .

<http://amr.isi.edu/amr_data/11#z21> a amr-terms:project .

<http://amr.isi.edu/amr_data/12#z13> a amr-terms:motive .

<http://amr.isi.edu/amr_data/13#z15> a entity-types:person ;
amr-terms:mod <http://amr.isi.edu/amr_data/13#z16> .

<http://amr.isi.edu/amr_data/13#z5> a entity-types:city ;
rdfs:label "Maplewood" ;
amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/2#z0> a propbank:know-02 ;
propbank:know-02.ARG1 <http://amr.isi.edu/amr_data/2#z1> ;
propbank:know-02.ARG3 <http://amr.isi.edu/amr_data/2#z9> .

<http://amr.isi.edu/amr_data/2#z4> a amr-terms:chain .

<http://amr.isi.edu/amr_data/3#z1> a propbank:donate-01 ;
propbank:donate-01.ARG0 <http://amr.isi.edu/amr_data/3#z2> ;
propbank:donate-01.ARG2 <http://amr.isi.edu/amr_data/3#z3> .

<http://amr.isi.edu/amr_data/3#z2> a amr-terms:he .

<http://amr.isi.edu/amr_data/3#z3> a amr-terms:charity .

<http://amr.isi.edu/amr_data/4#z1> a propbank:raise-01 ;
propbank:raise-01.ARG0 <http://amr.isi.edu/amr_data/4#z2> ;
propbank:raise-01.ARG1 <http://amr.isi.edu/amr_data/4#z5> ;
amr-terms:manner <http://amr.isi.edu/amr_data/4#z11> .

<http://amr.isi.edu/amr_data/4#z7> a propbank:donate-01 ;
propbank:donate-01.ARG0 <http://amr.isi.edu/amr_data/4#z8> .

<http://amr.isi.edu/amr_data/4#z8> a entity-types:person ;
rdfs:label "Smith" ;
amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/5#z14> a amr-terms:woman .

<http://amr.isi.edu/amr_data/5#z6> a entity-types:thing .

<http://amr.isi.edu/amr_data/6#z12> a entity-types:person ;
amr-terms:mod <http://amr.isi.edu/amr_data/6#z13> .

<http://amr.isi.edu/amr_data/6#z7> a propbank:contribute-01 ;
propbank:contribute-01.ARG0 <http://amr.isi.edu/amr_data/6#z8>

<http://amr.isi.edu/amr_data/7#z0> a propbank:begin-01 ;
propbank:begin-01.ARG0 <http://amr.isi.edu/amr_data/7#z1> ;
propbank:begin-01.ARG1 <http://amr.isi.edu/amr_data/7#z2> ;
amr-terms:time <http://amr.isi.edu/amr_data/7#z7> .

<http://amr.isi.edu/amr_data/7#z1> a amr-terms:center .

<http://amr.isi.edu/amr_data/7#z3> a amr-core:and ;
amr-terms:op1 <http://amr.isi.edu/amr_data/7#z4> ;
amr-terms:op2 <http://amr.isi.edu/amr_data/7#z5> .

<http://amr.isi.edu/amr_data/8#z14> a amr-terms:he .

<http://amr.isi.edu/amr_data/8#z15> a amr-terms:store ;
 amr-terms:quant "3" .

<http://amr.isi.edu/amr_data/8#z17> a amr-terms:area ;
 amr-terms:mod <http://amr.isi.edu/amr_data/8#z18> .

<http://amr.isi.edu/amr_data/9#z1> a entity-types:person ;
 rdfs:label "Rendley Scott" ;
 amr-terms:location <http://amr.isi.edu/amr_data/9#z7> ;
 amr-terms:mod <http://amr.isi.edu/amr_data/9#z3>,
 <http://amr.isi.edu/amr_data/9#z5> ;
 amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/9#z14> a entity-types:person .

<http://amr.isi.edu/amr_data/9#z9> a entity-types:thing .

propbank:activity-06 a amr-core:Frame .

propbank:build-01 a amr-core:Frame .

propbank:contrast-01 a amr-core:Frame .

propbank:impact-01 a amr-core:Frame .

propbank:include-01 a amr-core:Frame .

propbank:receive-01 a amr-core:Frame .

propbank:remain-01 a amr-core:Frame .

amr-terms:figure a amr-core:Concept .

amr-terms:hardware a amr-core:Concept .

amr-terms:motive a amr-core:Concept .

amr-terms:other a amr-core:Concept .

amr-terms:past a amr-core:Concept .

amr-terms:politics a amr-core:Concept .

amr-terms:project a amr-core:Concept .

amr-terms:prominent a amr-core:Concept .

amr-terms:some a amr-core:Concept .

amr-core:after a amr-core:Concept .

<http://amr.isi.edu/amr_data/11#z12> a amr-terms:wildlife .

<http://amr.isi.edu/amr_data/11#z4> a amr-terms:store .

<http://amr.isi.edu/amr_data/11#z7> a entity-types:person ;
amr-terms:quant <http://amr.isi.edu/amr_data/11#z9> .

<http://amr.isi.edu/amr_data/12#z7> a entity-types:person ;
rdfs:label "Smith" ;
amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/13#z8> a propbank:generous-01 ;
propbank:generous-01.ARG0 <http://amr.isi.edu/amr_data/13#z9> .

<http://amr.isi.edu/amr_data/4#z2> a entity-types:person ;
amr-terms:quant <http://amr.isi.edu/amr_data/4#z4> .

<http://amr.isi.edu/amr_data/8#z1> a amr-terms:council ;
amr-terms:mod <http://amr.isi.edu/amr_data/8#z2> .

<http://amr.isi.edu/amr_data/8#z4> a entity-types:person ;
amr-terms:quant <http://amr.isi.edu/amr_data/8#z7> .

entity-types:thing a amr-core:NamedEntity .

propbank:generous-01 a amr-core:Frame .

propbank:have-org-role-91 a amr-core:Frame .

propbank:include-91 a amr-core:Frame .

propbank:new-01 a amr-core:Frame .

propbank:reside-01 a amr-core:Frame .

amr-terms:businessman a amr-core:Concept .

amr-terms:council a amr-core:Concept .

amr-terms:he a amr-core:Concept .

amr-terms:member a amr-core:Concept .

amr-terms:town a amr-core:Concept .

<http://amr.isi.edu/amr_data/1#z1> a amr-terms:center .

<http://amr.isi.edu/amr_data/11#z15> a amr-terms:population .

<http://amr.isi.edu/amr_data/13#z4> a entity-types:person ;
amr-terms:mod <http://amr.isi.edu/amr_data/13#z5> .

<http://amr.isi.edu/amr_data/2#z1> a entity-types:person ;
rdfs:label "Smith" ;
amr-terms:wiki "-" .

<http://amr.isi.edu/amr_data/6#z2> a entity-types:organization ;
rdfs:label "Board of Directors" ;
amr-terms:wiki "-" .

```
propbank:cause-01 a amr-core:Frame .

propbank:local-02 a amr-core:Frame .

amr-terms:area a amr-core:Concept .

amr-terms:center a amr-core:Concept .

amr-terms:community a amr-core:Concept .

amr-terms:store a amr-core:Concept .

<http://amr.isi.edu/amr_data/5#z4> a entity-types:person ;
    rdfs:label "Smith" ;
    amr-terms:wiki "-" .

amr-core:NamedEntity a amr-core:Concept ;
    rdfs:label "AMR-EntityType",
        "AMR-Term" .

propbank:donate-01 a amr-core:Frame .

amr-core:and a amr-core:Concept .

entity-types:person a amr-core:NamedEntity .

amr-core:Frame a amr-core:Concept ;
    rdfs:label "AMR-PropBank-Frame" .

propbank:FrameRole a amr-core:Role ;
    rdfs:label "AMR-PropBank-Role" .
```


Appendix E

Outputs of T5 fine tuned on different data sets

E.1 T5-DART

[Maplewood Community Center | owner | Tom Smith],
[Tom Smith | donation amount | \$100,000].

[Centre for Community Development | age | 18-24],
[Centre for Community Development | location | South End],
[Centre for Community Development | budget | \$0].

[Jim Smith | business | Hardware],
[Jim Smith | philanthropy | A. J. Smith & Co.],
[Jim Smith | area | North Carolina],
[Jim Smith | family name | Smith].

[Jimmy Johnson | community involvement |
Volunteering, philanthropy, and other activities].

[St. George's Episcopal Church | donation | \$150,000],
[St. George's Episcopal Church | residents | Questioned by some].

[Sarah Smith | rumor | using his donations as a way to gain influence]

[Community Center | thanks | Smith].

[St. Joseph's Episcopal Church | activities | Various],

[St. Joseph's Episcopal Church | gifts | Smith's generosity].

[Smith's Hardware | county | Douglas County],
 [Desert County | town council | Smith's Hardware],
 [Smith's Hardware | area | lucrative],
 [Desert County | town council | Several members with ties to Smith's]

[Rendley Scott | past donors | Smith],
 [Rendley Scott | political affiliation | Rendley Scott],
 [Rendley Scott | area | North Carolina].

[Glowing Frog | location | Northampton],
 [Northampton | wildlife | Glowing Frog],
 [Northampton | building projects | Greenfield],
 [Greenfield | residents' concerns | Concerns about the impact on loc

[William Smith | motives | Altruistic].

[Malewood | residents | Content to enjoy the benefits of his generosity],
 [Malewood | population | 67,020].

E.2 T5-WebNLG

[Maplewood, California | is part of | Maplewood, California],
 [Maplewood, California | population total | 164],
 [Tom Smith (politician) | donation | \$100,000],
 [Tom Smith (politician) | hometown | Maplewood, California].

[Center for Social and Cultural Learning | shortfall | \$400,000],
 [Center for Social and Cultural Learning | activity | After-school programs],
 [Center for Social and Cultural Learning | location | Parkview Park],
 [Center for Social and Cultural Learning | financial strain | inadequate]

[Smith (entrepreneur) | longtime community activist],
 [Smith (entrepreneur) | chain of hardware stores],
 [Smith (entrepreneur) | charitable giving | longtime commitment to t

[Davey (politician) | charity | Local charities],
[Davey (politician) | volunteerism | active years spent in the comm

[Smith (politician) | donation | €100],
[Smith (politician) | motive | Charity],
[Antrimothy Smith | donation | €100].

[Smith (politician) | rumor | curry favor with local women],
[Smith (politician) | donations | countless].

[Community Center | board of directors | members | Smith family],
[Smith family | legacy | Community Center],
[Community Center | fundraising goal | \$10,000],
[Smith family | status | Quiet],
[Smith family | inspiration | others].

[Center for Cultural Diversity | new programs and activities | plan
[Smith Family Foundation | gift | \$1,000].

[Smith's Hardware | locations | County of Orange],
[Smith's Hardware | store | Town Council],
[Town Council | member | Smith's Hardware],
[Smith's Hardware | building permits | three new stores].

[Rendley Scott | businessman | Rendley Scott],
[Rendley Scott | party | Council of Economic Advisors],
[Rendley Scott | political party | Democratic Party (United States)]

[Glen frog | habitat | Encroaching Building],
[Scott's support for building has drawn criticism from some residen
[Glen frog | population | threatened],
[Scott's support for building | impacts on wildlife].

[Smith (entrepreneur) | charitable giving | unknown motive],
[Smith (entrepreneur) | business | Donation],
[Smith (entrepreneur) | motive | Altruism].

[Maplewood, New Jersey | leader | David McKinley],
 [Dave McKinley | generosity | many],
 [Manhattan | leader | John McKinley],
 [Manhattan | residents of Maplewood, New Jersey],
 [Alternative title | Maplewood, New Jersey],
 [Alternative title | Controversy].

E.3 T5-KELM

[Malewood, quaint town | Tom Smith (local businessman) follows | L

[Center (social gatherings and after-school) | instance of | Cente
 [Center (social gatherings and after-school) | use | Social gather
 [Center (social gatherings and after-school) | dissolved, abolishe
 [Center (social gatherings and after-school) | use | after-
 school].

[Smith (retailer) | owner of | Chain of successful hardware stores
 [Smith (retailer) | Chain of successful hardware stores in the are

[He often donates to local charities and volunteers in town | instar

[Smith's generous donation | raising eyebrows | questioning his moti

[Rumors | circulated that Smith | sex or gender | female],
 [Rumors | Smith (community and curry favor) | object has role | do

[Despite the whispers | board of directors (community center) obje

[Smith's generosity | planning and programming of events in a specif

[While, the Town Council (county) | applies to jurisdiction | Cour
 [While, the Town Council (county) | County of Smith | ties to | Sm
 [While, the Town Council (county) | County of Smith | granting aut

[Rendley Scott (council member) | instance of | Human],
 [Rendley Scott (council member) | sex or gender | male],
 [Rendley Scott (council member) | position held | Member of the C
 [Rendley Scott (council member) | given name | Rendley (given na
 [Rendley Scott (council member) | position held | Member of the C

[Scott's support for the new stores | draws criticism from some res
 [Scott's support for the new stores | instance of | drawing], [Scot
 [Scott's support for the new stores | instance of | local wildlife]
 [Scott's support for the new stores | species kept | Gullivar frogs
 [Scott's support for the new stores | endangered species].

[Smith's donation and business ventures | time of discovery or inve
 [Smith's donation and business ventures | Time of discovery or inve
 [Smith's donation and business ventures | Time of discovery or inve
 [Smith's donation and business ventures | Time of discovery or inve
 [Smith's donation and business ventures | Time of discovery or inve
 [Smith's donation and business ventures | True altruistic or if the

[Maplewood | people's generosity | positive impact],
 [Maplewood | people's generosity | end cause | Potential consequenc

E.4 T5-TekGen

[Maplewood, California | instance of | Town].
 [Greenwood Academy | instance of | School].
 [Smith | instance of | Neighbourhood].
 [Doncaster | place of birth | Town].
 [Earls at Smith | cause of death | Erectiles].
 [Rumors | subclass of | Don].
 [The Community Center | instance of | Community center].
 [Pennsylvania State Forest | instance of | Center].
 [Borough of Smith | instance of | Town].
 [Midland (UK Parliament constituency) | named after | Rendley Sco
 [Borough of Scott, New South Wales | instance of | Building].
 [Leifif Smith | occupation | business].
 [Maplewood | instance of | Town].

Appendix F

Preliminary outputs of T5 and BART

Results comparing T5 and BART, testing various settings, and testing metrics.

F.1 T5

When applying [an earlier version of the evaluation metrics] to the initial paragraphs of a CNN news article, the following results were obtained. The metric 'article_level_avg' represents the average evaluation results for all outputs generated by the various checkpoints of the T5 model. In essence, it is the mean of the checkpoint averages. The checkpoint averages are presented below, starting with 'run_one-1200'. It should be noted that 'run_one' does not refer to the first run in the strictest sense, but rather the first run after finalizing numerous settings and choices. The differences between runs lie in the learning rate and weight decay settings. Subsequently, the outputs of all checkpoints for each input paragraph or sentence are listed.

Interestingly, in journalistic writing, paragraphs and sentences are often interchangeable, resulting in paragraph splitting that also separates sentences.

```
1 {"https://edition.cnn.com/2022/10/13/europe/ukraine-
   russia-war-next-stages-intl/index.html": {
2 {
3   "https://edition.cnn.com/2022/10/13/europe/ukraine-
   russia-war-next-stages-intl/index.html": {
4     "Article_level_avg": {
5       "Consistency": 0.7281481481481482,
```

```
6       "Precision": 0.32962962962962966,  
7       "Hallucination": 0.23296296296296296,  
8       "Recall": 0.27333333333333333,  
9       "Quality": 0.21851851851851853,  
10      "Syntax": 0.9125925925925927  
11    },  
12    "run_one/checkpoint-1200": {  
13      "Consistency": 0.66666666666666666,  
14      "Precision": 0.38666666666666667,  
15      "Hallucination": 0.33,  
16      "Recall": 0.26666666666666666,  
17      "Quality": 0.20000000000000004,  
18      "Syntax": 0.77333333333333334  
19    },  
20    "run_two/checkpoint-800": {  
21      "Consistency": 0.83333333333333334,  
22      "Precision": 0.5,  
23      "Hallucination": 0.0,  
24      "Recall": 0.33333333333333333,  
25      "Quality": 0.16666666666666666,  
26      "Syntax": 0.83333333333333334  
27    },  
28    "run_two/checkpoint-1000": {  
29      "Consistency": 0.61,  
30      "Precision": 0.31,  
31      "Hallucination": 0.37000000000000005,  
32      "Recall": 0.23333333333333333,  
33      "Quality": 0.20000000000000004,  
34      "Syntax": 0.94333333333333334  
35    },  
36    "run_two/checkpoint-1400": {  
37      "Consistency": 0.83333333333333334,  
38      "Precision": 0.33333333333333333,  
39      "Hallucination": 0.16666666666666666,  
40      "Recall": 0.18333333333333335,  
41      "Quality": 0.20000000000000004,  
42      "Syntax": 1.0
```



```
43     },
44     "run_two/checkpoint-1600": {
45         "Consistency": 0.61,
46         "Precision": 0.370000000000000005,
47         "Hallucination": 0.3,
48         "Recall": 0.2333333333333333,
49         "Quality": 0.200000000000000004,
50         "Syntax": 1.0
51     },
52     "run_two/checkpoint-2200": {
53         "Consistency": 0.8333333333333334,
54         "Precision": 0.3333333333333333,
55         "Hallucination": 0.2866666666666667,
56         "Recall": 0.1766666666666667,
57         "Quality": 0.3,
58         "Syntax": 1.0
59     },
60     "run_two/checkpoint-3200": {
61         "Consistency": 0.8333333333333334,
62         "Precision": 0.2333333333333333,
63         "Hallucination": 0.3,
64         "Recall": 0.2666666666666667,
65         "Quality": 0.2666666666666667,
66         "Syntax": 0.9433333333333334
67     },
68     "run_two/checkpoint-3800": {
69         "Consistency": 0.6666666666666667,
70         "Precision": 0.2166666666666667,
71         "Hallucination": 0.21,
72         "Recall": 0.3833333333333333,
73         "Quality": 0.2333333333333333,
74         "Syntax": 0.87
75     },
76     "run_two/checkpoint-4200": {
77         "Consistency": 0.6666666666666667,
78         "Precision": 0.2833333333333333,
79         "Hallucination": 0.1333333333333333,
```

```
80     "Recall": 0.38333333333333333,
81     "Quality": 0.200000000000000004,
82     "Syntax": 0.85
83 },
84 "Paragraph 1": {
85     "text": "The relative calm in Ukrainian
            cities far from the country\u2019s
            battlefields was shattered by two painfully
            familiar sounds this week: the ominous
            ring of the air raid sirens, and the
            eruptions of Russian attacks.",
86     "run_one/checkpoint-1200": {
87         "triples": "
88             <Ukrainian_city | relative_calm |
            eruptions of Russian attacks>,
89             <Russian_attack | eruption | Russian
            attacks>.
90         ",
91         "Consistency": 0.5,
92         "Precision": 0.5,
93         "Hallucination": 0,
94         "Recall": 0.7,
95         "Quality": 0.2,
96         "Syntax": 1,
97         "comment": "I give this a high recall
            because the first triple tries to
            capture the essence of the passage,\n
            even though there are details
            not captured. And a high precision
            because what's wrong with the triples
            is the quality, not that the
            information is not relevant"
98     },
99     "run_two/checkpoint-800": {
100         "triples": "
101             [Ukraine | battlefield | Ukrainian
            cities],
```

```
102         [Ukrainian_city | relative calm |
103             eruption of Russian attacks].
104     ",
105     "Consistency": 0.5,
106     "Precision": 0.5,
107     "Hallucination": 0,
108     "Recall": 0.8,
109     "Quality": 0.3,
110     "Syntax": 1
111 },
112 "run_two/checkpoint-1000": {
113     "triples": "
114         [Ukraine | battlefield | Ukrainian],
115         [Ukrainian | relative_calm |
116             eruptedBy | Russian attacks],
117         [Ukrainian | battlefield | Ukrainian
118             military],
119         [Ukrainian | militaryBranch |
120             Ukrainian military],
121         [Ukrainian | militaryBranch |
122             Ukrainian military],
123         [Ukrainian | militaryBranch |
124             Ukrainian military battlefield].
125     ",
126     "Consistency": 1,
127     "Precision": 0.33,
128     "Hallucination": 0.66,
129     "Recall": 0.5,
130     "Quality": 0.2,
131     "Syntax": 0.83
132 },
133 "run_two/checkpoint-1400": {
134     "triples": "
135         [Ukraine | battle | Ukrainian
136             military],
137         [Ukrainian_military | eruption |
138             Russian attacks],
```

```
131         [Air_raid_sirens | ring | Ukrainian
132             military],
133         [Ukrainian_military | significantly
134             changed | the relative calm in
135             Ukrainian cities far from the
136             battlefield].
137     ",
138     "Consistency": 1,
139     "Precision": 0,
140     "Hallucination": 0,
141     "Recall": 0,
142     "Quality": 0.2,
143     "Syntax": 1,
144     "Comment": "Objects are not connected
145         with '_' by data design choice because
146         it could be a literal"
147 },
148 "run_two/checkpoint-1600": {
149     "triples": "
150         [Ukraine | battlefield | Ukrainian
151             military],
152         [Ukrainian_military_force | eruption
153             | Russian attacks],
154         [Air_raid_sirens | ringing |
155             significantly higher in Ukrainian
156             cities],
157         [Ukraine | militaryBranch | Military
158             Defense Force].
159     ",
160     "Consistency": 0.33,
161     "Precision": 0.25,
162     "Hallucination": 0.25,
163     "Recall": 0.25,
164     "Quality": 0.2,
165     "Syntax": 1
166 },
167 "run_two/checkpoint-2200": {
```

```
157     "triples": "  
158         [Ukraine | battle | Russian invasion]  
159         ,  
160         [Air_raid_sirens | ringing |  
161             Ukrainian military],  
162         [Ukraine | ethnicGroup | Ukrainian  
163             people].  
164     ",  
165     "Consistency": 1,  
166     "Precision": 0.5,  
167     "Hallucination": 0.33,  
168     "Recall": 0.33,  
169     "Quality": 0.5,  
170     "Syntax": 1  
171 },  
172 "run_two/checkpoint-3200": {  
173     "triples": "  
174         [Ukraine | battlefield | Ukrainian  
175             military],  
176         [Ukrainian_military | eruption |  
177             Russian attacks],  
178         [Air_raid_sirens | ringing |  
179             significantly more significant than  
180             the ring of the air raid sirens].  
181     ",  
182     "Consistency": 1,  
183     "Precision": 0.5,  
184     "Hallucination": 0,  
185     "Recall": 0.6,  
186     "Quality": 0.4,  
187     "Syntax": 1  
188 },  
189 "run_two/checkpoint-3800": {  
190     "triples": "  
191         [Ukraine | battle | Ukrainian Army],  
192         [Air_raid_sirens | ringing |  
193             significantly more than the ring of
```

```
        the Air raid siren],
186      [Ukraine | militaryBranch | Russian
        Army],
187      [Ukraine | relativeCalm | stunnedBy |
        Russian eruption of Russian
        aircraft].
188      ",
189      "Consistency": 1,
190      "Precision": 0.25,
191      "Hallucination": 0,
192      "Recall": 0.85,
193      "Quality": 0.3,
194      "Syntax": 0.75
195    },
196    "run_two/checkpoint-4200": {
197      "triples": "[Ukraine | battle | Ukrainian
198        battlefield],
199      [Air_raid_sirens | ringing | Ukrainian
200        military],
201      [Ukraine | relative_calm | stunnedBy |
202        Russian eruptions],
203      [Ukraine | militaryBranch | Battlefield
204        Ukrainian].
205      ",
206      "Consistency": 1,
207      "Precision": 0.25,
208      "Hallucination": 0,
209      "Recall": 0.85,
210      "Quality": 0.2,
211      "Syntax": 0.75
    }
  },
  "Paragraph 2": {
    "text": "A wave of missiles, rockets and
    drones has struck dozens of locations
    across Ukraine since Monday, according to
    officials, targeting civilian
```

```
infrastructure in several major cities,  
including Kyiv, located hundreds of miles  
from the front lines in the east and south  
."  
212 "run_one/checkpoint-1200": {  
213   "triples": "  
214     <Ukraine | military | rockets,  
        rockets and drones>,  
215     <Ukraine | largestCity | Kyiv>,  
216     <Ukraine | military | targets |  
        Civilian infrastructure>.  
217   ",  
218   "Consistency": 1,  
219   "Precision": 0.33,  
220   "Hallucination": 0.33,  
221   "Recall": 0,  
222   "Quality": 0.2,  
223   "Syntax": 0.66  
224 },  
225 "run_two/checkpoint-800": {  
226   "triples": "  
227     [Ukraine | military infrastructure |  
        hundreds of miles from the  
        frontline],  
228     [Ukraine | military infrastructure |  
        hundreds of miles from the  
        frontline],  
229     [Ukraine | military infrastructure |  
        hundreds of miles from the  
        frontline].  
230   ",  
231   "Consistency": 1,  
232   "Precision": 0,  
233   "Hallucination": 0,  
234   "Recall": 0,  
235   "Quality": 0.1,  
236   "Syntax": 1
```

```
237     },
238     "run_two/checkpoint-1000": {
239         "triples": "
240             [Ukraine | military targets | rockets
241               , rockets, drones],
242             [Ukraine | largestCity | Kyiv\n
243               ],
244             [Ukraine | military targets | rockets
245               , rockets, drones and rockets],
246             [Ukraine | military targets | rockets
247               , rockets, drones and drones]
248         "
249         "Consistency": 0.33,
250         "Precision": 0,
251         "Hallucination": 0.25,
252         "Recall": 0,
253         "Quality": 0.2,
254         "Syntax": 1
255     },
256     "run_two/checkpoint-1400": {
257         "triples": "
258             [Ukraine | military attack | rockets,
259               rockets and drones],
260             [Ukraine | largestCity | Kyiv],
261             [Ukraine | military attack | rockets,
262               rockets and drones],
263             [Ukraine | largestCity | Kyiv],
264             [Ukraine | military attack | rockets,
265               rockets and drones],
266             [Ukraine | military | attacking
267               infrastructure in the city of Kyiv]
268         .
269         "
270         "Consistency": 1,
271         "Precision": 0.25,
272         "Hallucination": 0.25,
273         "Recall": 0.25,
```



```
265         "Quality": 0.2,
266         "Syntax": 1
267     },
268     "run_two/checkpoint-1600": {
269         "triples": "
270             [Ukraine | military attack | rockets,
271               rockets and drones],
272             [Ukraine | largestCity | Kyiv],
273             [Kyiv | isPartOf | Front lines in the
274               east and south],
275             [Ukraine | military attack | rockets,
276               rockets and drones].
277         ",
278         "Consistency": 1,
279         "Precision": 0.66,
280         "Hallucination": 0.25,
281         "Recall": 0.25,
282         "Quality": 0.2,
283         "Syntax": 1
284     },
285     "run_two/checkpoint-2200": {
286         "triples": "
287             [Ukraine | military | attackAircraft
288               | rockets, rockets and drones],
289             [Ukraine | largestCity | Kyiv],
290             [Kyiv | location | South-West],
291             [Ukraine | military | rockets,
292               rockets and drones].
293         ",
294         "Consistency": 1,
295         "Precision": 0,
296         "Hallucination": 0.33,
297         "Recall": 0,
298         "Quality": 0.2,
299         "Syntax": 1
300     },
301     "run_two/checkpoint-3200": {
```

```
297     "triples": "  
298         [Ukraine | military | attackAircraft  
299           | rockets, rockets and drones],  
300         [Ukraine | largestCity | Kyiv],  
301         [Kyiv | isPartOf | Front lines in the  
302           east and South],  
303         [Ukraine | leader | Igor Kyiv],  
304         [Ukraine | runwayLength | 100.0],[  
305           Ukraine | leader | Ukrainian  
306           military].  
307     ",  
308     "Consistency": 1,  
309     "Precision": 0,  
310     "Hallucination": 0.5,  
311     "Recall": 0,  
312     "Quality": 0.2,  
313     "Syntax": 0.83  
314 },  
315 "run_two/checkpoint-3800": {  
316     "triples": "[Ukraine | military |  
317       attackAircraft | rockets],  
318     [Ukraine | largestCity | Kyiv],  
319     [Ukraine | runwayLength | 99.0],  
320     [Ukraine | military | rockets, rockets  
321       and drones],  
322     [Ukraine | runwayName | Ukrainian  
323       military aircraft],  
324     [Ukraine | hasToItsWest | South],  
325     [Ukraine | runwayName | Ukrainian  
326       military aircraft].  
327   ",  
328   "Consistency": 0.5,  
329   "Precision": 0,  
330   "Hallucination": 0.43,  
331   "Recall": 0,  
332   "Quality": 0.2,  
333   "Syntax": 0.86
```

```
326     },
327     "run_two/checkpoint-4200": {
328         "triples": "[Ukraine | military |
329             aircraft rockets],
330         [Ukraine | largestCity | Kyiv],
331         [Kyiv | isPartOf | Front lines in the
332             East and South],
333         [Ukraine | military | rockets, rockets
334             and drones],
335         [Kyiv | location | country | Ukraine].
336     ",
337     "Consistency": 0.5,
338     "Precision": 0,
339     "Hallucination": 0.2,
340     "Recall": 0,
341     "Quality": 0.2,
342     "Syntax": 0.8
343 }
344 },
345 "Paragraph 3": {
346     "text": "The wide bombardment echoed the
347         early days of Russia\u2019s scattershot
348         initial invasion in February, but also
349         underlined that the conflict in Ukraine,
350         which for months appeared to be descending
351         into a slow and painful grind in the Donbas
352         , has erupted once again as winter nears.",
353     "run_one/checkpoint-1200": {
354         "triples": "<Ukraine | invasion |
355             February>,
356         <Ukraine | scattershot_invasion | epoch |
357             2014>,
358         <Ukraine | epoch | 2014>.
359     ",
360     "Consistency": 0.5,
361     "Precision": 0.33,
362     "Hallucination": 0.66,
```

```
352         "Recall": 0.1,  
353         "Quality": 0.2,  
354         "Syntax": 0.66  
355     },  
356     "run_two/checkpoint-800": {  
357         "triples": "[Ukraine | invasion |  
358             February],  
359         [Ukraine | invasion | February],  
360         [Ukraine | eruption | once again | winter  
361             nears].",  
362         "Consistency": 1,  
363         "Precision": 1,  
364         "Hallucination": 0,  
365         "Recall": 0.2,  
366         "Quality": 0.1,  
367         "Syntax": 0.5  
368     },  
369     "run_two/checkpoint-1000": {  
370         "triples": "[Ukraine | invasion |  
371             February 2],  
372         [Ukraine | scattershot | Russia],  
373         [Ukraine | invasion | February],  
374         [Ukraine | crisis | erupted once again],  
375         [Ukraine | militaryConflict | erupted  
376             once again].",  
377         "Consistency": 0.5,  
378         "Precision": 0.6,  
379         "Hallucination": 0.2,  
380         "Recall": 0.2,  
381         "Quality": 0.2,  
382         "Syntax": 1  
383     },  
384     "run_two/checkpoint-1400": {  
385         "triples": "[Ukraine | invasion |  
386             February],  
387         [Ukraine | militaryConflict | erupted
```

```
        once again],
384     [Ukraine | ethnicGroup | Ukrainian people
        ],
385     [Ukraine | invasion | scattershot
        invasion in February].",
386     "Consistency": 0.5,
387     "Precision": 0.75,
388     "Hallucination": 0.25,
389     "Recall": 0.3,
390     "Quality": 0.2,
391     "Syntax": 1
392 },
393 "run_two/checkpoint-1600": {
394     "triples": "[Ukraine | invasion | Russia]
        ,
395     [Ukraine | ethnicGroup | Ukrainian people
        ],
396     [Ukraine | invasion | February],
397     [Ukraine | epoch | 2014],
398     [Ukraine | ethnicGroup | Donbas].",
399     "Consistency": 0.5,
400     "Precision": 0.2,
401     "Hallucination": 0.4,
402     "Recall": 0.2,
403     "Quality": 0.2,
404     "Syntax": 1
405 },
406 "run_two/checkpoint-2200": {
407     "triples": "[Ukraine | invasion | Russia]
        ,
408     [Russia | scattershot | February invasion
        ],
409     [Ukraine | ethnicConflict | Ukraine],
410     [Ukraine | epoch | 2014\u20132015].",
411     "Consistency": 0.5,
412     "Precision": 0.5,
413     "Hallucination": 0.2,
```

```
414     "Recall": 0.2,  
415     "Quality": 0.2,  
416     "Syntax": 1,  
417     "Comment": "The first and second tripels  
                don't work on their own, and not really  
                together either in terms of quality,  
                but in each other's consistency-context  
                they contribute to precision"  
418 },  
419 "run_two/checkpoint-3200": {  
420     "triples": "[Ukraine | invasion | Russia]  
                ,  
421     [Ukraine | epoch | February 2],  
422     [Ukraine | militaryConflict | eased in  
                February],  
423     [Ukraine | battle | Donbas],  
424     [Ukraine | epoch | 2014].",  
425     "Consistency": 0.5,  
426     "Precision": 0.2,  
427     "Hallucination": 0.4,  
428     "Recall": 0.2,  
429     "Quality": 0.2,  
430     "Syntax": 1  
431 },  
432 "run_two/checkpoint-3800": {  
433     "triples": "[Ukraine | invasion | Russia]  
                ,  
434     [Russia | scattershot | February invasion  
                ],  
435     [Ukraine | ethnicConflict | Ukraine],  
436     [Ukraine | epoch | 2014],  
437     [Ukraine | battle | Donbas].",  
438     "Consistency": 0.5,  
439     "Precision": 0.4,  
440     "Hallucination": 0.2,  
441     "Recall": 0.3,  
442     "Quality": 0.2,
```

```

443     "Syntax": 1
444   },
445   "run_two/checkpoint-4200": {
446     "triples": "[Ukraine | invasion | Russia]
447     ,
448     [Russia | scattershot | February invasion
449     ],
450     [Ukraine | ethnicConflict | Ukraine],
451     [Ukraine | epoch | 2014],
452     [Ukraine | battle | Donbas].",
453     "Consistency": 0.5,
454     "Precision": 0.6,
455     "Hallucination": 0.2,
456     "Recall": 0.3,
457     "Quality": 0.2,
458     "Syntax": 1
459   }
460 }

```

F.2 BART

```

1 {
2   "https://edition.cnn.com/2022/10/13/europe/ukraine-
3   russia-war-next-stages-intl/index.html": {
4     "Article_level_avg": {
5       "Consistency": 0.8008333333333333,
6       "Precision": 0.1952777777777778,
7       "Hallucination": 0.6069444444444446,
8       "Recall": 0.21527777777777776,
9       "Quality": 0.18194444444444446,
10      "Syntax": 0.8972222222222221
11    },
12    "First run/checkpoint-??/pytorch_model.bin": {
13      "Consistency": 0.8333333333333334,
14      "Precision": 0.3066666666666667,

```

```
14     "Hallucination": 0.4366666666666667,  
15     "Recall": 0.45,  
16     "Quality": 0.2333333333333333,  
17     "Syntax": 0.9166666666666666  
18 },  
19 "second_run/checkpoint-600/pytorch_model.bin": {  
20     "Consistency": 0.6666666666666666,  
21     "Precision": 0.21333333333333335,  
22     "Hallucination": 0.7033333333333333,  
23     "Recall": 0.3333333333333333,  
24     "Quality": 0.16666666666666666,  
25     "Syntax": 1.0  
26 },  
27 "third_run/checkpoint-1600/pytorch_model.bin": {  
28     "Consistency": 0.8333333333333334,  
29     "Precision": 0.06666666666666667,  
30     "Hallucination": 0.5333333333333333,  
31     "Recall": 0.08333333333333333,  
32     "Quality": 0.20000000000000004,  
33     "Syntax": 1.0  
34 },  
35 "fourth_run/checkpoint-1800/pytorch_model.bin": {  
36     "Consistency": 0.8333333333333334,  
37     "Precision": 0.0,  
38     "Hallucination": 0.9166666666666666,  
39     "Recall": 0.0,  
40     "Quality": 0.18333333333333335,  
41     "Syntax": 0.6933333333333334  
42 },  
43 "fourth_run/checkpoint-2000/pytorch_model.bin": {  
44     "Consistency": 0.8333333333333334,  
45     "Precision": 0.0,  
46     "Hallucination": 0.75,  
47     "Recall": 0.0,  
48     "Quality": 0.15,  
49     "Syntax": 0.9433333333333334  
50 },
```



```
51 "fourth_run/checkpoint-2200/pytorch_model.bin": {
52     "Consistency": 0.8333333333333334,
53     "Precision": 0.08333333333333333,
54     "Hallucination": 0.5833333333333334,
55     "Recall": 0.06666666666666667,
56     "Quality": 0.25,
57     "Syntax": 0.9166666666666666
58 },
59 "fourth_run/checkpoint-3600/pytorch_model.bin": {
60     "Consistency": 0.5,
61     "Precision": 0.25,
62     "Hallucination": 0.75,
63     "Recall": 0.26666666666666666,
64     "Quality": 0.16666666666666666,
65     "Syntax": 0.64
66 },
67 "fifth_run/checkpoint-200/pytorch_model.bin": {
68     "Consistency": 0.7766666666666667,
69     "Precision": 0.3333333333333333,
70     "Hallucination": 0.5533333333333333,
71     "Recall": 0.3,
72     "Quality": 0.16666666666666666,
73     "Syntax": 0.8333333333333334
74 },
75 "sixth_run/checkpoint-400/pytorch_model.bin": {
76     "Consistency": 0.8333333333333334,
77     "Precision": 0.0,
78     "Hallucination": 0.5533333333333333,
79     "Recall": 0.0,
80     "Quality": 0.13333333333333333,
81     "Syntax": 1.0
82 },
83 "sixth_run/checkpoint-800/pytorch_model.bin": {
84     "Consistency": 0.6666666666666666,
85     "Precision": 0.14,
86     "Hallucination": 0.5366666666666667,
87     "Recall": 0.15,
```

```
88     "Quality": 0.15,  
89     "Syntax": 0.9633333333333334  
90 },  
91 "seventh_run/checkpoint-2400/pytorch_model.bin":  
92   {  
93     "Consistency": 1.0,  
94     "Precision": 0.56,  
95     "Hallucination": 0.5033333333333333,  
96     "Recall": 0.5666666666666665,  
97     "Quality": 0.18333333333333335,  
98     "Syntax": 0.9166666666666666  
99   },  
100 "seventh_run/checkpoint-2600/pytorch_model.bin":  
101   {  
102     "Consistency": 1.0,  
103     "Precision": 0.38999999999999996,  
104     "Hallucination": 0.46333333333333334,  
105     "Recall": 0.36666666666666667,  
106     "Quality": 0.20000000000000004,  
107     "Syntax": 0.9433333333333334  
108   },  
109 "Paragraph 1": {  
110   "text": "The relative calm in Ukrainian  
111     cities far from the country\u2019s  
112     battlefields was shattered by two painfully  
113     familiar sounds this week: the ominous  
114     ring of the air raid sirens, and the  
115     eruptions of Russian attacks.",  
116   "First great run/pytorch_model.bin": {  
117     "triples": "<Ukrain | demonym | calm>,  
118     <Ukrain | battlefield | Ukrainian cities  
119     far from the battlefields>,  
120     <Ukrain | attacker | Russian attacks>,  
121     <Ukrain | air_raid_Sister | ominous ring  
122     | the air raid sirens>.",  
123     "Consistency": 1,  
124     "Precision": 0.75,
```

```
116         "Hallucination": 0.25,
117         "Recall": 0.85,
118         "Quality": 0.2,
119         "Syntax": 0.75
120     },
121     "checkpoint-600/pytorch_model.bin": {
122         "triples": "<Ukrain | battle | Battle of
123             the air raids>,
124             <Ukrain | demonym | Ukrainians>,
125             <Ukrain | demonym | Ukrainians>,
126             <Ukrain | painfully familiar_sounds |
127                 Ominous ring of the air raid sirens>,
128             <Ukrain | country | Russia>.",
129         "Consistency": 1,
130         "Precision": 0.5,
131         "Hallucination": 0.25,
132         "Recall": 0.5,
133         "Quality": 0.2,
134         "Syntax": 1
135     },
136     "checkpoint-1600/pytorch_model.bin": {
137         "triples": "<Ukrain | battle | Ukrainian
138             cities far from the country\u2019s
139             battlefields>,
140             <Ukrain | leaderTitle | Air raid>,
141             <Ukrain | leader | Russia>,
142             <Ukrain | language | Ominous ring of the
143                 air raid sirens>.",
144         "Consistency": 1,
145         "Precision": 0.2,
146         "Hallucination": 0.25,
147         "Recall": 0.25,
148         "Quality": 0.2,
149         "Syntax": 1
150     },
151     "checkpoint-1800/pytorch_model.bin": {
```

```
148     "triples": "<U.K. | air_raid_sirens |  
149         country | Ukraine>,  
150     <U.K. | air_raid_sirens | alternativeName  
151         | Ominous ring of the air raid sirens>  
152     ,  
153     <U.K. | battle | Ukraine>.",  
154     "Consistency": 1,  
155     "Precision": 0,  
156     "Hallucination": 1,  
157     "Recall": 0,  
158     "Quality": 0.15,  
159     "Syntax": 0.33  
160 },  
161 "checkpoint-2000/pytorch_model.bin": {  
162     "triples": "<U.K. | leader | Ukrainians>,  
163     <U.K. | battle | Ukrainian cities>,  
164     <U.K. | leader | Vladimir Putin>,  
165     <U.K. | leader | Ukrainians>,  
166     <U.K. | air raid sirens>,  
167     <U.K. | demonym | Ominous ring>.",  
168     "Consistency": 1,  
169     "Precision": 0,  
170     "Hallucination": 1,  
171     "Recall": 0,  
172     "Quality": 0.15,  
173     "Syntax": 0.83  
174 },  
175 "checkpoint-2200/pytorch_model.bin": {  
176     "triples": "  
177         <Ukrain | leaderTitle | Fighter of  
178             the Ukrainians>,  
179         <Air_Raid_Sirens | country | Ukraine>  
180     ,  
181         <Ukrain | leaderTitle | Fighter of  
182             the Ukrainians>,  
183         <Air_Raid_Sirens | alternativeName |  
184             Ominous ring of the air raid sirens
```

```
178         >,
179         <Ukrain | country | battlefield |
180         Ukraine>.",
179         "Consistency": 0.5,
180         "Precision": 0.25,
181         "Hallucination": 0.25,
182         "Recall": 0.2,
183         "Quality": 0.3,
184         "Syntax": 0.75
185     },
186     "checkpoint-3600/pytorch_model.bin": {
187         "triples": "<Ukrain | air_raid_sirens |
188         country | Ukraine>,
189         <Ukrain | demonym | calm>,
190         <Ukrain | battle | Ukrainian cities far
191         from the battlefields>,
192         <Ukrain | attacks | Russian air raids>.",
193         "Consistency": 0.5,
194         "Precision": 0.75,
195         "Hallucination": 0.25,
196         "Recall": 0.8,
197         "Quality": 0.2,
198         "Syntax": 0.75
199     },
200     "checkpoint-200/pytorch_model.bin": {
201         "triples": "<Ukraine | demonym |
202         Ukrainians>,
203         <Ukraine | battle | Ukrainian cities far
204         from the country's battlefields>,
205         <Ukraine | air raid sirens>,
206         <Ukraine | attacker | Russian attacks>.",
207         "Consistency": 1,
208         "Precision": 0.75,
209         "Hallucination": 0.25,
210         "Recall": 0.7,
211         "Quality": 0.2,
212         "Syntax": 0.75
```

```
209     },
210     "checkpoint-400/pytorch_model.bin": {
211         "triples": "<Ucrain | battle | Battle of
212             Ukrainians>,
213             <Air_Rauts_Sirens | region | Ukrainian
214             cities>,
215             <Air_Rauts_Sirens | country | Russian
216             attacks on Ukrainians>.",
217         "Consistency": 1,
218         "Precision": 0,
219         "Hallucination": 0.66,
220         "Recall": 0,
221         "Quality": 0.1,
222         "Syntax": 1
223     },
224     "checkpoint-800/pytorch_model.bin": {
225         "triples": "<Ukrain | stateOfOrigin |
226             Ukrainian people>,
227             <Ukrain | country | Russia>,
228             <Ukrain | stateOfOrigin | Ominous ring of
229             the air raid sirens>,
230             <Ukrain | battlefield | Ukraine>,
231             <Ukrain | country | Russia>.",
232         "Consistency": 0.5,
233         "Precision": 0,
234         "Hallucination": 0.5,
235         "Recall": 0,
236         "Quality": 0.15,
237         "Syntax": 1
238     },
239     "checkpoint-2400/pytorch_model.bin": {
240         "triples": "<Ukrain | battlefield |
241             Battlefields | Russian attacks>,
242             <Ukrain | demonym | calm>,
243             <Air_Rattle_Sirns | region | Ukrainian
244             cities far from the country's
245             battlefields>,"
```

```
238         <Ukrain | battle | Ominous ring of the
          air raid sirens>. ",
239         "Consistency": 1,
240         "Precision": 0.75,
241         "Hallucination": 0.25,
242         "Recall": 0.7,
243         "Quality": 0.2,
244         "Syntax": 0.75
245     },
246     "checkpoint-2600/pytorch_model.bin": {
247         "triples": "<Ukrain | battlefield |
          Russian attacks>,
248         <Ukrain | demonym | calm>,
249         <Air_Raise_Sirens | region | Ukrainian
          cities far from the battlefields>,
250         <Ukrain | battle | Ominous ring of the
          air raid sirens>.",
251         "Consistency": 1,
252         "Precision": 0.75,
253         "Hallucination": 0.25,
254         "Recall": 0.7,
255         "Quality": 0.2,
256         "Syntax": 1
257     }
258 },
259 "Paragraph 2": {
260     "text": "A wave of missiles, rockets and
          drones has struck dozens of locations
          across Ukraine since Monday, according to
          officials, targeting civilian
          infrastructure in several major cities,
          including Kyiv, located hundreds of miles
          from the front lines in the east and south
          .",
261     "First great run/pytorch_model.bin": {
262         "triples": "<Ukraine | capital | Kyiv>,
          <Ukraine | capital | Ukraine>,"
```

```
264     <Ukraine | hasToItsWest | east and
        southwest>,
265     <Ukraine | hasToItsNorth | Ukraine>,
266     <Ukraine | attackedBy | A wave of
        missiles, rockets, drones>,
267     <Ukraine | capital | Kiev>.
268     ",
269     "Consistency": 0.5,
270     "Precision": 0.17,
271     "Hallucination": 0.66,
272     "Recall": 0.5,
273     "Quality": 0.3,
274     "Syntax": 1
275 },
276 "checkpoint-600/pytorch_model.bin": {
277     "triples": "
278     <Ukraine | leader | Ukraine>,
279     <Ukraine | leader | Ukraine>,
280     <Ukraine | leader | Kyiv>,
281     <Ukraine | country | Ukraine>,
282     <Ukraine | attackedBy | A wave of
        missiles, rockets, drones>,
283     <Ukraine | country | Ukraine>,
284     <Ukraine | east | east>.
285     ",
286     "Consistency": 0.5,
287     "Precision": 0.14,
288     "Hallucination": 0.86,
289     "Recall": 0.5,
290     "Quality": 0.2,
291     "Syntax": 1
292 },
293 "checkpoint-1600/pytorch_model.bin": {
294     "triples": "<Ukraine | frontline |
        Ukraine>,
295     <Ukraine | military | Defense Department
        (Ukraine))>,"
```



```
296     <Ukraine | military | Defense Department
297         (Ukraine))>,
298     <Ukraine | city | Kyiv>,
299     <Ukraine | country | Ukraine>,
300     <Ukraine | military | missiles, rockets,
301         drones>,
302     <Ukraine | country | Ukraine>.
303     ",
304     "Consistency": 1,
305     "Precision": 0,
306     "Hallucination": 0.85,
307     "Recall": 0,
308     "Quality": 0.2,
309     "Syntax": 1
310 },
311 "checkpoint-1800/pytorch_model.bin": {
312     "triples": "
313     <U.S._Airport | location | Kyiv>,
314     <U.S._Airport | attackWave | 3>,
315     <U.S._Airport | location | Ukraine>,
316     <U.S._Airport | attackWave | 3>,
317     <U.S._Airport | location | Ukraine>,
318     <U.S._Airport | attackWave | 3>,
319     <U.S._Airport | city | Kyiv>.
320     ",
321     "Consistency": 1,
322     "Precision": 0,
323     "Hallucination": 1,
324     "Recall": 0,
325     "Quality": 0.2,
326     "Syntax": 1
327 },
328 "checkpoint-2000/pytorch_model.bin": {
329     "triples": "<Ukraine | frontline |
    Ukraine>,
    <Ukraine | leader | Vladimir Putin>,
    <Ukraine | leader | Vladimir Putin>,"
```

```
330     <Ukraine | military | attacks on civilian
        infrastructure in several major cities
        , including Kyiv, located hundreds of
        miles from the front lines in the east
        and south>,
331     <Ukraine | military | wave of missiles,
        rockets and drones>.",
332     "Consistency": 1,
333     "Precision": 0,
334     "Hallucination": 0.25,
335     "Recall": 0,
336     "Quality": 0.15,
337     "Syntax": 1
338 },
339 "checkpoint-2200/pytorch_model.bin": {
340     "triples": "<Ukraine | frontline |
        Ukraine>,
341     <Ukraine | frontline | Ukraine>,
342     <Ukraine | city | Kyiv>,
343     <Ukraine | leader | Vladimir Putin>,
344     <Ukraine | leader | Ukraine>,
345     <Ukraine | leader | Vladimir Putin>,
346     <Ukraine | military | attacks on missiles
        ,
347     rockets, drones>,
348     <Ukraine | country | United States>.",
349     "Consistency": 1,
350     "Precision": 0,
351     "Hallucination": 0.5,
352     "Recall": 0,
353     "Quality": 0.2,
354     "Syntax": 1
355 },
356 "checkpoint-3600/pytorch_model.bin": {
357     "triples": "
```

```
359     <U.S._Missile_(rocket) | location |
360         Ukraine>,
361     <U.S._Missile_(rocket) | attacker |
362         Ukraine>,
363     <U.S._Missile_(rocket) | attacker |
364         Ukraine>,
365     <U.S._Missile_(rocket) | location | Kyiv>
366     ,
367     <U.S._Missile_(rocket) | attacker |
368         Ukraine>,
369     <U.S._State | frontline | east and
370         southwest>.
371     "
372     ,
373     "Consistency": 0.5,
374     "Precision": 0,
375     "Hallucination": 1,
376     "Recall": 0,
377     "Quality": 0.1,
378     "Syntax": 1
379     },
380     "checkpoint-200/pytorch_model.bin": {
381         "triples": "
382         <Ukraine | frontline | Ukraine>,
383         <Ukraine | capital | Kyiv>,
384         <Ukraine | largestCity | Ukraine>,
385         <Ukraine | military | military personnel>
386         ,
387         <Ukraine | military | rockets and drones>
388         ,
389         <Ukraine | country | Ukraine>.
390     "
391     ,
392     "Consistency": 1,
393     "Precision": 0,
394     "Hallucination": 0.66,
395     "Recall": 0,
396     "Quality": 0.2,
397     "Syntax": 1
```

```
388     },
389     "checkpoint-400/pytorch_model.bin": {
390         "triples": "
391         <Kyiv | isPartOfMilitaryConstruction |
392         Ukraine>,
393         <Kyiv | isPartOfMilitaryConstruction |
394         Ukraine>,
395         <Kyiv | isPartOfMilitaryConstruction |
396         Kyiv>,
397         <Kyiv | isPartOfMilitaryConstruction |
398         Ukraine>.
399     ",
400     "Consistency": 1,
401     "Precision": 0,
402     "Hallucination": 0,
403     "Recall": 0,
404     "Quality": 0.15,
405     "Syntax": 1
406 },
407 "checkpoint-800/pytorch_model.bin": {
408     "triples": "
409     <Kyiv | cityServed | Kyiv>,
410     <Kyiv | country | Ukraine>,
411     <Kyiv | isPartOf | Ukraine>,
412     <Kyiv | isPartOf | Ukraine>,
413     <Kyiv | isPartOf | Ukraine>,
414     <Kyiv | cityServed | Kyiv>,
415     <Kyiv | location | Ukraine>,
416     <Kyiv | country | United States>,
417     <Kyiv | attackAircraft | missiles,
418     rockets and drones>.
419     ",
420     "Consistency": 1,
421     "Precision": 0.2,
422     "Hallucination": 0.33,
423     "Recall": 0.25,
424     "Quality": 0.2,
```

```
421     "Syntax": 1
422   },
423   "checkpoint-1000/pytorch_model.bin": {
424     "triples": "
425     <U.S._Wave_ofMissiles,_Rockets_and_Roads
426     | location | Ukraine>,
427     <U.S._Wave_ofMissiles,_Rockets_and_Roads
428     | struck | 1 (U.S.)>,
429     <U.S._Wave_ofMissiles,_Rockets_and_Roads
430     | targeting | Civilian infrastructure
431     in several major cities, including Kyiv
432     >,
433     <U.S._Wave_ofMissiles,_Rockets_and_Roads
434     | struck | 1 (U.S.)>.
435   ",
436   "Consistency": 1,
437   "Precision": 0,
438   "Hallucination": 1,
439   "Recall": 0,
440   "Quality": 0.15,
441   "Syntax": 1
442 },
443 "checkpoint-2400/pytorch_model.bin": {
444   "triples": "<Kyiv | isPartOf | Ukraine>,
445   <Ukraine | military | Department of
446   Defense (United States)>,
447   <Ukraine | battle | Fighting missiles,
448   rockets, drones>,
449   <Ukraine | frontline | east and southwest
450   >,
451   <Ukraine | country | United States>,
452   <Ukraine | largestCity | Kyiv>.
453   ",
454   "Consistency": 1,
455   "Precision": 0.6,
456   "Hallucination": 0.6,
457   "Recall": 0.6,
```

```

449         "Quality": 0.15,
450         "Syntax": 1
451     },
452     "checkpoint-2600/pytorch_model.bin": {
453         "triples": "<Ukraine | battle | Civilians
454             attack>,
455         <Ukraine | attackSite | Kyiv>,
456         <Kyiv | isPartOf | Ukraine>,
457         <Ukraine | hasToItsNortheast | Front
458             Lines in the east and southwest>,
459         <Ukraine | attackSite | Ukraine>,
460         <Ukraine | militaryInfrastructure |
461             rockets, drones>,
462         <Ukraine | largestCity | Kiev>.
463     ",
464     "Consistency": 1,
465     "Precision": 0.42,
466     "Hallucination": 0.14,
467     "Recall": 0.4,
468     "Quality": 0.2,
469     "Syntax": 1
470 }
471 },
472 "Paragraph 3": {
473     "text": "The wide bombardment echoed the
474         early days of Russia\u2019s scattershot
475         initial invasion in February, but also
476         underlined that the conflict in Ukraine,
477         which for months appeared to be descending
478         into a slow and painful grind in the Donbas
479         , has erupted once again as winter nears.",
480     "First great run/pytorch_model.bin": {
481         "triples": "
482         <Russia | scattershot_Identifier |
483             February>,
484         <Ukrain | leader | Ukraine>,
485         <Ukrain | hasToItsNorth | Donbass>,

```

```
476     <Ukrain | hasToItsEnd | December>,
477     <Ukrain | leader | Russia>.
478     ",
479     "Consistency": 1,
480     "Precision": 0,
481     "Hallucination": 0.4,
482     "Recall": 0,
483     "Quality": 0.2,
484     "Syntax": 1
485 },
486 "checkpoint-600/pytorch_model.bin": {
487     "triples": "
488     <Ukray | isPartOf | Donbas>,
489     <Ukray | isPartOf | Ukraine>,
490     <Ukray | isPartOf | Russia>,
491     <Ukray | country | Russia>,
492     <Ukray | isPartOf | Ukraine>,
493     <Ukray | isPartOf | Russia>,
494     <Ukray | isPartOf | Russia>.
495     ",
496     "Consistency": 0.5,
497     "Precision": 0,
498     "Hallucination": 1,
499     "Recall": 0,
500     "Quality": 0.1,
501     "Syntax": 1
502 },
503 "checkpoint-1600/pytorch_model.bin": {
504     "triples": "
505     <Russia | invasion | February 2014>,
506     <Ukrain | leader | Ukraine>,
507     <Ukrain | grinding | Ukraine>,
508     <Ukrain | leader | Russia>,
509     <Ukrain | leader | Ukraine>,
510     <Ukrain | ethnicGroup | Donbass>.
511     ",
512     "Consistency": 0.5,
```

```
513     "Precision": 0,  
514     "Hallucination": 0.5,  
515     "Recall": 0,  
516     "Quality": 0.2,  
517     "Syntax": 1  
518 },  
519 "checkpoint-1800/pytorch_model.bin": {  
520     "triples": "  
521     <U.K. | invasion | February 2014>,  
522     <U.K. | leader | Ukraine>,  
523     <U.K. | ground | Donbas>,  
524     <Russia | wide bombardment | beginning |  
525     February 2014>.  
526     ",  
527     "Consistency": 0.5,  
528     "Precision": 0,  
529     "Hallucination": 0.75,  
530     "Recall": 0,  
531     "Quality": 0.2,  
532     "Syntax": 0.75  
533 },  
534 "checkpoint-2000/pytorch_model.bin": {  
535     "triples": "  
536     <U.K. | invasion | February 2014>,  
537     <U.K. | leader | Russia>,  
538     <U.K. | leader | Ukraine>,  
539     <U.K. | leader | Ukraine>,  
540     <U.K. | ground | Donbas>,  
541     <U.K. | wide bombardment | February 2014  
542     >.  
543     ",  
544     "Consistency": 0.5,  
545     "Precision": 0,  
546     "Hallucination": 1,  
547     "Recall": 0,  
548     "Quality": 0.15,  
549     "Syntax": 1
```



```
548     },
549     "checkpoint-2200/pytorch_model.bin": {
550         "triples": "
551         <U.S._invasion | dateOfFear | February 20
552             14>,
553         <U.S._invasion | location | Ukraine>,
554         <U.S._invasion | impactType | Wide
555             bombardment>,
556         <U.S._invasion | ground | Donbas>,
557         <U.S._invasion | course | slow and
558             painful grind>.
559     ",
560     "Consistency": 1,
561     "Precision": 0,
562     "Hallucination": 1,
563     "Recall": 0,
564     "Quality": 0.25,
565     "Syntax": 1,
566     "Comment": "maybe it's a stretch, but U.S
567         ._invasion could be seen as a blank
568         event_identifier by ignoring the name."
569 },
570     "checkpoint-3600/pytorch_model.bin": {
571         "triples": "
572         <U.S._invasion | February 2015>,
573         <U.S._invasion | region | Ukraine>,
574         <U.S._invasion | impactType | Wide
575             bombardment>,
576         <U.S._invasion | date | February 2015>,
577         <U.S._invasion | region | Donbas>,
578         <U.S._invasion | intensity | slow and
579             painful grind>.
580     ",
581     "Consistency": 0.5,
582     "Precision": 0,
583     "Hallucination": 1,
584     "Recall": 0,
```

```
578         "Quality": 0.2,  
579         "Syntax": 0.17  
580     },  
581     "checkpoint-200/pytorch_model.bin": {  
582         "triples": "  
583         <U.K._(U.K._K._Donbas | isPartOf |  
584         Ukraine>,  
585         <U.K._Donbas | isPartOf | Donbas people>,  
586         <Russia | scattershot initial invasion in  
587         February>,  
588         <U.K._Donbas | isPartOf | Ukraine>.  
589         ",  
590         "Consistency": 0.33,  
591         "Precision": 0.25,  
592         "Hallucination": 0.75,  
593         "Recall": 0.2,  
594         "Quality": 0.1,  
595         "Syntax": 0.75  
596     },  
597     "checkpoint-400/pytorch_model.bin": {  
598         "triples": "<United_Kingdom | isPartOf |  
599         Donbas (Ucrania)>,  
600         <United_Kingdom | isPartOf | Ukraine>,  
601         <United_Kingdom | isPartOf | Ukraine>,  
602         <United_Kingdom | season | 2014\u20132015>.  
603         ",  
604         "Consistency": 0.5,  
605         "Precision": 0,  
606         "Hallucination": 1,  
607         "Recall": 0,  
608         "Quality": 0.15,  
609         "Syntax": 1  
610     },  
611     "checkpoint-800/pytorch_model.bin": {  
612         "triples": "  
613         <Russia <Russia <Russia | echoedBy |  
614         Russia>,"
```

```
611     <Russia | scattershotInvasion | February>
612     ,
613     <Russia | stateOfOrigin | Ukraine>,
614     <Ukrain | country | Russia>,
615     <Ukrain | stateOfOrigin | Ukraine>,
616     <Ukrain | regionServedBy | Ukraine>,
617     <Ukrain | country | Russia>,
618     <Ukrain | regionServedBy | Donbas>,
619     <Ukrain | regionServedBy | Ukraine>.
620     "Consistency": 0.5,
621     "Precision": 0.22,
622     "Hallucination": 0.78,
623     "Recall": 0.2,
624     "Quality": 0.1,
625     "Syntax": 0.89
626 },
627 "checkpoint-2400/pytorch_model.bin": {
628     "triples": "
629     <Russia | invasion | February 2014>,
630     <Ukraine | battle | Donbass>,
631     <Ukraine | leader | Ukraine>,
632     <Russia | attack | Wide bombardment>,
633     <Ukraine | country | Russia>,
634     <Ukraine | hasToItsNorth | Ukraine>.
635     "Consistency": 1,
636     "Precision": 0.33,
637     "Hallucination": 0.66,
638     "Recall": 0.4,
639     "Quality": 0.2,
640     "Syntax": 1
641 },
642 "checkpoint-2600/pytorch_model.bin": {
643     "triples": "
644     <U.S._invasion | February 2014>,"
```

```

646     <U.S._invasion | magnitude | Wide
        bombardment>,
647     <U.S._invasion | region | Donbass>,
648     <U.S._invasion | year | 2014>,
649     <U.S._invasion | precededBy | 2013>,
650     <U.S._invasion | state | Ukraine>.
651     ",
652     "Consistency": 1,
653     "Precision": 0,
654     "Hallucination": 1,
655     "Recall": 0,
656     "Quality": 0.2,
657     "Syntax": 0.83
658     }
659   }
660 }
661 }

```

F.3 Intermediate evaluation

T5 results: Consistency: 0.73, Precision: 0.33, Hallucination: 0.23, Recall: 0.27, Quality: 0.22, Syntax: 0.91. BART results: Consistency: 0.80, Precision: 0.20, Hallucination: 0.61, Recall: 0.22, Quality: 0.18, Syntax: 0.90.

The models use the same terms for the same concepts more often than not inside of a single output, but some times entities showed slight differences. For example, <Ukraine>, <Ukrain> could both be in the same output, one a subject and the other an object. Inside of the sets of subjects of a single output, consistency was good for all models and checkpoints, however between different outputs of even the same checkpoint it consistency was very low, take BART checkpoint-600, 'Ukrain', 'Ukraine' and 'Ukray' were each the subject for paragraph 1,2 and 3 respectively, and none of them correct.

The results reveals that some issues needs to be addressed before these state of the art models can be used for any adequate lifting of RDF graphs. What wasn't captured well in the consistency metric is that the models often were not consistent between paragraphs,

which highlights the limitation of sentence-level KG extraction versus document level.

The comparability of metrics was unfortunately greatly affected by the length of the output. One error would lead to .0 of the relevant metric if the output was only one triple, while one error out of 10 triples would be a .9 score.

Bibliography

- Al-Moslmi, T., M. Gallofre Ocana, A. L. Opdahl, and C. Veres (2020), Named Entity Extraction for Knowledge Graphs: A Literature Overview, *IEEE Access*, 8, 32,862–32,881, doi:10.1109/ACCESS.2020.2973928. 2.3
- Boettiger, C., J. Ooms, pdatascience, K. Leinweber, and J. Hester (2020), ropensci/rdfliib: v0.2.3, doi:10.5281/ZENODO.3604372. 2.3.5
- Burns, G., and J. L. Ambite (2014), BMKEG/amr-ld: A Python library for mapping AMRs to linked data formats (such as RDF and JSON-LD). This work is heavily derived from Naomi Saphra’s ‘AMERICA’ library (<https://github.com/nsaphra/AMERICA>) and SMATCH. 2.3.5
- Cai, S., and K. Knight (2013), Smatch: an Evaluation Metric for Semantic Feature Structures. 2.3.5
- Chen, H., G. Cao, J. Chen, and J. Ding (2019a), A Practical Framework for Evaluating the Quality of Knowledge Graph, *Communications in Computer and Information Science*, 1134 CCIS, 111–122, doi:10.1007/978-981-15-1956-7{_}10. 2.2, 3.3.2, 3.2
- Chen, Y., J. Kuang, D. Cheng, J. Zheng, M. Gao, and A. Zhou (2019b), AgriKG: An Agricultural Knowledge Graph and Its Applications, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11448 LNCS, 533–537, doi:10.1007/978-3-030-18590-9{_}81/COVER. 10
- Deng, S., J. Chen, N. Zhang, J. Z. Pan, W. Zhang, and H. Chen (2019), Knowledge-driven stock trend prediction and explanation via temporal convolutional network, *The Web Conference 2019 - Compan-*

- ion of the World Wide Web Conference, WWW 2019*, pp. 678–685, doi:10.1145/3308560.3317701. 1, 10, 1, 2
- Devlin, J., M. W. Chang, K. Lee, and K. Toutanova (2018), BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 1, 4171–4186. 2.3.1
- Dietz, L., A. Kotov, and E. Meij (2018), Utilizing knowledge graphs for text-centric information retrieval, *41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018*, pp. 1387–1390, doi:10.1145/3209978.3210187. 6
- Eberts, M., and A. Ulges (2019), Span-based Joint Entity and Relation Extraction with Transformer Pre-training, *Frontiers in Artificial Intelligence and Applications*, 325, 2006–2013, doi:10.3233/FAIA200321. 2.3
- Giorgi, J., G. D. Bader, and B. Wang (2022), A sequence-to-sequence approach for document-level relation extraction. 2.3
- Huaman, E. (2022), Steps to Knowledge Graphs Quality Assessment. 2.2, 3.3.2, 3.1
- Huang, X., J. Zhang, D. Li, and P. Li (2019), Knowledge graph embedding based question answering, *WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining*, pp. 105–113, doi:10.1145/3289600.3290956. 4, 9, 6
- Jo, J.-y., and S.-H. Myaeng (2020), Roles and Utilization of Attention Heads in Transformer-based Neural Language Models, pp. 3404–3417, Association for Computational Linguistics (ACL), doi:10.18653/v1/2020.acl-main.311. 2.3
- Jones, M., P. Slaughter, J. Ooms, C. Boettiger, and S. Chamberlain (2022), redland: RDF Library Bindings in R. 2.3.5
- Keller, R. M. (2019), Building a knowledge graph for the air traffic management community, *The Web Conference 2019 - Companion*

- of the World Wide Web Conference, WWW 2019*, pp. 700–704, doi:10.1145/3308560.3317706. 3
- Lewis, M., Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer (2019), BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 7871–7880, doi:10.18653/v1/2020.acl-main.703. 2.3.1, 2.3.3
- Liang, Y., N. Duan, Y. Gong, N. Wu, F. Guo, W. Qi, M. Gong, L. Shou, D. Jiang, G. Cao, X. Fan, R. Zhang, R. Agrawal, E. Cui, S. Wei, T. Bharti, Y. Qiao, J. H. Chen, W. Wu, S. Liu, F. Yang, D. Campos, R. Majumder, and M. Zhou (2020), XGLUE: A New Benchmark Dataset for Cross-lingual Pre-training, Understanding and Generation, *EMNLP 2020 - 2020 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, pp. 6008–6018, doi:10.48550/arxiv.2004.01401. 2.3
- Lukovnikov, D., A. Fischer, J. Lehmann, and S. Auer (2017), Neural network-based question answering over knowledge graphs on word and character level, *26th International World Wide Web Conference, WWW 2017*, pp. 1211–1220, doi:10.1145/3038912.3052675. 4, 6
- Miwa, M., and Y. Sasaki (2014), Modeling Joint Entity and Relation Extraction with Table Representation, *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, pp. 1858–1869, doi:10.3115/V1/D14-1200. 2.3
- Muppalla, R., S. Lalithsena, T. Banerjee, and A. Sheth (2017), A Knowledge Graph Framework for Detecting Traffic Events Using Stationary Cameras, pp. 431–436, doi:10.1145/3091478.3162384. 2
- OpenAI (2023), GPT-4 Technical Report. 2.3.1, 2.3.6
- Openai, A. R., K. N. Openai, T. S. Openai, and I. S. Openai (2018), Improving Language Understanding by Generative Pre-Training. 2.3.6

- Ostapuk, N., J. Yang, and P. Cudré-Mauroux (2019), ActiveLink: Deep active learning for link prediction in knowledge graphs, *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, pp. 1398–1408, doi:10.1145/3308558.3313620. 5
- Ouyang, L., J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Aspell, P. Welinder, P. Christiano, J. Leike, and R. Lowe (2022), Training language models to follow instructions with human feedback, doi:10.48550/arxiv.2203.02155. 2.3, 2.3.6
- Ping, P., K. Watson, J. Han, and A. Bui (2017), Individualized knowledge graph: a viable informatics path to precision medicine, *Circulation Research*, 120(7), 1078–1080, doi:10.1161/CIRCRESAHA.116.310024. 3
- Radhakrishnan, P., P. Talukdar, and V. Varma (2018), ELDEN: Improved Entity Linking Using Densified Knowledge Graphs, *NAACL HLT 2018 - 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 1*, 1844–1853, doi:10.18653/V1/N18-1167. 3
- Raffel, C., N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu (2019), Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer, *Journal of Machine Learning Research*, 21, 1–67. 2.3.1, 2.3.2
- Saphra, N. (2014), nsaphra/AMRICA: Visualize AMRs with inter-annotator disagreement and bitext alignment. 2.3.5
- Sui, D., Y. Chen, K. Liu, J. Zhao, X. Zeng, and S. Liu (2020), Joint Entity and Relation Extraction with Set Prediction Networks, doi:10.1109/TNNLS.2023.3264735. 2.3
- Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Kaiser, and I. Polosukhin (2017), Attention Is All You Need, *Advances in Neural Information Processing Systems, 2017-December*, 5999–6009. 2.3.1

- Wang, A., Y. Pruksachatkun, N. Nangia, A. Singh, J. Michael, F. Hill, O. Levy, and S. R. Bowman (2019a), SuperGLUE: A Stickier Benchmark for General-Purpose Language Understanding Systems, *Advances in Neural Information Processing Systems*, 32. 2.3
- Wang, C., X. Liu, and D. Song (2020), Language Models are Open Knowledge Graphs, doi:10.48550/arxiv.2010.11967. 2.3
- Wang, H., J. Leskovec, F. Zhang, M. Zhao, W. Li, M. Zhang, and Z. Wang (2019b), Knowledge-aware Graph Neural Networks with Label Smoothness Regularization for Recommender Systems, *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 968–977, doi:10.1145/3292500.3330836. 5, 9
- Wang, R., Y. Yan, J. Wang, Y. Jia, Y. Zhang, W. Zhang, and X. Wang (2018), AceKG: A Large-scale Knowledge Graph for Academic Data Mining, *International Conference on Information and Knowledge Management, Proceedings*, pp. 1487–1490, doi:10.1145/3269206.3269252. 3, 4, 4, 6, 7
- Zhao, Q., Q. Li, and J. Wen (2018), Construction and application research of knowledge graph in aviation risk field, *MATEC Web of Conferences*, 151, 05,003, doi:10.1051/MATECCONF/201815105003. 7, 8, 10, 1, 3
- Zheng, W., J. X. Yu, L. Zou, and H. Cheng (2018), Question answering over knowledge graphs, *Proceedings of the VLDB Endowment*, 11(11), 1373–1386, doi:10.14778/3236187.3236192. 11
- Ziegler, D. M., N. Stiennon, J. Wu, T. B. Brown, A. Radford, D. A. Paul, C. Geoffrey, and I. Openai (2019), Fine-Tuning Language Models from Human Preferences. 2.3, 2.3.6