Assignment 1

Joshua Cook

Georgia Institute of Technology

CS 6460

Research Log

Background

I am the cofounder of an educational technology startup currently in the pre-customer phase. As such, I am deeply invested in the intersection of education and technology. I have more than two decades combined experience in education and AI software engineering. When I discuss the work that I'm doing with my co-founders, we often talk not about product market fit, but about product founder fit as in the work we are doing aligns so well with the experiences that we have.

My background includes:

- 8 years teaching high school math in Los Angeles
- 2 years training student teachers for high school math instruction
- A master's degree in Education
- A second bachelor's degree in applied mathematics, focusing on computational software engineering
- Professional experience as an artificial intelligence engineer with a focus on natural language processing and numerous models in production
- Teaching data science, data engineering, and machine learning at UCLA and Caltech
- 3 years managing internal and external enablement at Databricks

As both a working engineer and student and teacher of applied mathematics, I have a particular interest in the development of best practices and rigor in software engineering. This is especially important working with extremely complicated systems, such as are required to deploy machine learning and artificial intelligence applications. I am also an early adopter of Jupyter notebooks.

Papers

A frustratingly easy approach for entity and relation extraction.

Citation: Zhong, Z., & Chen, D. (2021). A frustratingly easy approach for entity and relation extraction. In K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tur, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, & Y. Zhou (Eds.), Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 50-61). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.naacl-main.5

How Found: In references of another paper

My Summary: More canonical, large language model approach using tools derived from BERT architecture. Two models are used, one for entity recognition and one for relation extraction. Significant discussion on the trade-off between scalability and accuracy.

My Takeaway: While the paper focuses on using the BERT architecture, it may be useful to think about in terms of large language models, especially with different prompting strategies. While previously, this would've been too purpose trained models. We may simply think of the same large language model with two different prompts as two different models in a dag-centric transformation architecture. All in all very promising and is aligned with our DAG thinking for extraction.

A Simple Standard for Sharing Ontological Mappings (SSSOM)

Citation: Matentzoglu, N., Balhoff, J. P., Bello, S. M., Bizon, C., Brush, M., Callahan, T. J., Chute, C. G., Duncan, W. D., Evelo, C. T., Gabriel, D., Graybeal, J., Gray, A., Gyori, B. M., Haendel, M., Harmse, H., Harris, N. L., Harrow, I., Hegde, H. B., Hoyt, A. L., Hoyt, C. T., Jiao, D., Jiménez-Ruiz, E., Jupp, S., Kim, H., Koehler, S., Liener, T., Long, Q., Malone, J., McLaughlin, J. A., McMurry, J. A., Moxon, S., Munoz-Torres, M. C., Osumi-Sutherland, D., Overton, J. A., Peters, B., Putman, T., Queralt-Rosinach, N., Shefchek, K., Solbrig, H., Thessen, A., Tudorache, T., Vasilevsky, N., Wagner, A. H., & Mungall, C. J. (2022). A Simple Standard for Sharing Ontological Mappings (SSSOM). _Database_, 2022, baac035. https://doi.org/10.1093/database/baac035

How Found: Google Search

My Summary: A proposal for a standard for sharing ontological mappings. Useful for contexts where ontological mappings would be shared between labs working on a single subject area i.e. genomics. Quite a few contributors to the paper and the paper is probably very important to the field.

My Takeaway: Useful as we think about our own data model. We will not be sharing ontologies between organizations, even two different organizations that are both using our application, but we will need to think about our data model and it is useful to see something like this while we define that model. There seem to be a few other standards, so we'll need to do additional research looking at what others are using. Still need to get my head around concepts like the semantic web.

A Taxonomy for Human-LLM Interaction Modes: An Initial Exploration

Citation: Gao, J., Gebreegziabher, S. A., Choo, K. T. W., Li, T. J. J., Perrault, S. T., & Malone, T. W. (2024, May). A Taxonomy for Human-LLM Interaction Modes: An Initial Exploration. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (pp. 1-11).

How Found: Google Search

My Summary: Toward a classification of the different strategies that can be used for working with large language models. Toward describing their use in human-oriented applications. Four key phases in the human LLM interaction flow - planning, facilitating, iterating, and testing.

My Takeaway: Came up under a search for taxonomy, but this is really more of a collaboration-oriented paper. That is a concept area in which I have interest, but I will probably assign it mentally for review at a later time. Useful perhaps while developing our prompt engineering strategies.

The Artificial Intelligence Ontology: LLM-assisted construction of AI concept hierarchies

Citation: Joachimiak, M. P., Miller, M. A., Caufield, J. H., Ly, R., Harris, N. L., Tritt,A., Mungall, C. J., & Bouchard, K. E. (2024). The Artificial Intelligence Ontology: LLM-assisted construction of AI concept hierarchies. Lawrence Berkeley National Laboratory.

How Found: Google Search

My Summary: A paper discussing the ontology of AI concepts as a whole. Not a discussion for how to prepare ontology using AI but rather what is the ontology of AI work. Uses the Ontology Development Kit (ODK) for Curation.

My Takeaway: Paper is interesting in that it is using the ODK to prepare the ontology. We are not interested in this specific ontology other than our own education in working with artificial intelligence. But since we are building ontologies, it is useful to look at how the ODK was used here.

How much does curation cost?

Citation: Karp, P. D. (2016). How much does curation cost?. Database, 2016, baw110.

How Found: Google Search

My Summary: A discussion of the large costs associated with the curation of knowledge bases. Discusses alternative curation strategies but highlights how these may be less valuable in terms of knowledge quality. Applied to bio-informatics.

My Takeaway: Useful to think about in terms of making the case to organizations. Grounds of the conversation about knowledge based value in financial considerations. Maybe challenging to bridge the connection between an academic data store, which is significantly more fast moving and accompanies product data store. Will be useful to highlight how if large research institutions can do this then we should be able to do it for companies as well.

Mappergpt: Large language models for linking and mapping entities

Citation: Matentzoglu, N., Caufield, J. H., Hegde, H. B., Reese, J. T., Moxon, S., Kim, H., ... & Mungall, C. J. (2023). Mappergpt: Large language models for linking and mapping entities. arXiv preprint arXiv:2310.03666.

How Found: Google Search

My Summary: LLM-era paper on combining foundation models with conventional ontology preparation strategies. High recall ontology extraction method output is then passed to a large language model for refinement. Addresses a specific problem with automated ontological mapping in that they have high recall but low precision.

My Takeaway: Very applicable to our work. Is already a dag-centric approach. There are two different models being used here. Will be of interest to see which model is the first node on the dag that's being used to extract the ontologies. Also, the researchers on this paper are people that clearly "get it" in terms of the work that we are doing and are worth following.

GPT-4: A Stochastic Parrot or Ontological Craftsman? Discovering Implicit Knowledge Structures in Large Language Models

Citation: Procko, T. T., Elvira, T., & Ochoa, O. (2023, September). GPT-4: A Stochastic Parrot or Ontological Craftsman? Discovering Implicit Knowledge Structures in Large Language Models. In 2023 Fifth International Conference on Transdisciplinary AI (TransAI) (pp. 147-154). IEEE.

How Found: Google Search

My Summary: Looks at identifying ontologies via knowledge that foundation models have a priori. Foundation models have default knowledge generated during their training process. This paper looks at how this can be leveraged for the generation of ontological mappings.

My Takeaway: Extremely useful. This is something that I have thought about frequently. If you can define the problem that you're trying to solve in terms of a problem that the foundation model already knows how to solve, then you don't need to put effort into engineering a solution. Rather you're just encouraging the model to do what it already knows how to do. I haven't thought about this in terms of what exists within the large language model as a knowledge base and this is very profound and definitely worth exploring.

Hagrid: A human-llm collaborative dataset for generative information-seeking with attribution

Citation: Kamalloo, E., Jafari, A., Zhang, X., Thakur, N., & Lin, J. (2023). Hagrid: A human-llm collaborative dataset for generative information-seeking with attribution. arXiv preprint arXiv:2307.16883.

How Found: Google Search

My Summary: This paper presents a data set for use in the training of generative extraction models. Data set was prepared through a collaboration between humans and large language models. The data set is toward training models that are transparent, informative, and grounded in source material

My Takeaway: Likely not useful for my work as I will not be training my own models. Perhaps useful in reading about the researchers' take on source grounding. Perhaps useful in reading about the researchers' analysis as applied to existing models.

Leveraging passage retrieval with generative models for open domain question answering

Citation: Izacard, G., & Grave, E. (2021). Leveraging passage retrieval with generative models for open domain question answering. In P. Merlo, J. Tiedemann, & R. Tsarfaty (Eds.), Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume (pp. 874-880). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.eacl-main.74

How Found: In references of another paper

My Summary: Discusses the use of source grounding in generative extraction, though I'm concerned that it might not quite be making the leap to thinking about generative extraction. More focused on question answering trivia benchmarks, which for our purposes is pretty game oriented, "toy". Has performance metrics probably not quite useful for us.

My Takeaway: Is interesting in terms of using foundation models for applied work. Is interesting in terms of source grounding. Probably nothing novel in terms of our work.

Structured prompt interrogation and recursive extraction of semantics (SPIRES): A method for populating knowledge bases using zero-shot learning

Citation: Caufield, J. H., Hegde, H., Emonet, V., Harris, N. L., Joachimiak, M. P., Matentzoglu, N., ... & Mungall, C. J. (2024). Structured prompt interrogation and recursive extraction of semantics (SPIRES): A method for populating knowledge bases using zero-shot learning. Bioinformatics, 40(3), btae104.

How Found: Google Search

My Summary: A novel method and Python tool set for populating knowledge bases, using zero shot learning with foundation models. Introduces the idea of user defined schemes while working with these ontological mappings. Subject area is genomics, but specific attention is paid to the idea that this can be applied beyond genomics.

My Takeaway: State of the art work. The team working on this is definitely one to be followed. Very promising strategy in terms of working with our own knowledge extraction, perhaps combined with other research around identifying which ideologies exist a priori within the foundation model.

Retrieval-augmented generation for knowledge-intensive nlp task

Citation: Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. Advances in Neural Information Processing Systems, 33, 9459-9474.

How Found: In references of another paper

My Summary: An early paper discussing retrieval augmented generation (RAG). Models trained by researchers and are not foundation models as are typically used with RAG systems now. The source grounding system is Wikipedia. Probably very novel at the time.

My Takeaway: A little dated, probably not directly applicable to our work. Most of the concepts of strategies discussed here are already being used in our production systems. Of interest in that it is a watershed paper for the rag approach.

Semantic similarity metrics for evaluating source code summarization

Citation: Haque, S., Eberhart, Z., Bansal, A., & McMillan, C. (2022, May). Semantic similarity metrics for evaluating source code summarization. In Proceedings of the 30th IEEE/ACM International Conference on Program Comprehension (pp. 36-47).

How Found: In references of another paper

My Summary: A paper that focuses on different semantics similarity metrics. A sort of meta analysis on the viability of different metrics for different applications here with an emphasis on code summarization. The models and question are not foundation models, but are more from the BERT architecture.

My Takeaway: Also a little dated. In our experience, we will get more results just using default similarity metrics. The ANNOY algorithm gets very good results and is probably not something that we need to iterate on. Probably very groundbreaking at the time.

Text generation from knowledge graphs with graph transformers

Citation: Koncel-Kedziorski, R., Bekal, D., Luan, Y., Lapata, M., & Hajishirzi, H. (2019). Text generation from knowledge graphs with graph transformers. In J. Burstein, C. Doran, & T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human

Language Technologies, Volume 1 (Long and Short Papers) (pp. 2284-2293). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1238

How Found: In references of another paper

My Summary: An approach to generation from knowledge graphs using a researcher-trained transformer. An interesting combination of graph theory and transformer models. Assessment strategies for measuring generation coherence.

My Takeaway: Actually, probably extremely useful in the long run. There is probably significant low-hanging fruit to be achieved by combining foundation models with different extraction dags, but it is worth revisiting work like this at some point. Prior to leveraging this, we would need a method for extracting knowledge into knowledge graphs

Understanding By Design

Citation: Wiggins, G. (2005). Understanding by design. Association for Supervision and Curriculum Development.

How Found: Studied during my Master's Program in Education

My Summary: An educational planning framework favored in secondary education, instructional design. Emphasizes "backward planning":

1. Identify desired results
2. Determine acceptable evidence
3. Plan learning experiences and instruction

My Takeaway: A very strong framework for designing the components of our product. Identifying desired results might be mentor users of our application, defining the learning objectives of a guide. Determining acceptable evidence would be the development of summative assessment as applied to the learning acquisition of a guide. Planning learning experiences and instruction would be the substance of a guide and helping learner users to acquire the concepts presented in the guide.

Unified named entity recognition as word-word relation classification

Citation: Li, J., Fei, H., Liu, J., Wu, S., Zhang, M., Teng, C., Ji, D., & Li, F. (2021). Unified named entity recognition as word-word relation classification. In Proceedings of the AAAI Conference on Artificial Intelligence. https://api.semanticscholar.org/CorpusID:245335089

How Found: In references of another paper

My Summary: Uses a researcher developed BERT architecture to perform NER. Includes discussion on measuring efficacy of the name identity recognition. Proposes a unified framework for named entity recognition.

My Takeaway: A little dated. Anyone anyone proposing a unified framework... Good luck with that. Again, we will not be training on models, possibly useful in terms of looking at measurement but probably not.

Unifying large language models and knowledge graphs: A roadmap

Citation: Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., & Wu, X. (2023). Unifying large language models and knowledge graphs: A roadmap. IEEE Transactions on Knowledge and Data Engineering, 36, 3580-3599. https://api.semanticscholar.org/CorpusID:259165563

How Found: In references of another paper

My Summary: An analysis of how knowledge graphs and foundation models can be leveraged to work together. Presents three frameworks for this kind of work - KG-enhanced LLMs, LLM-augmented KGs, and synergized LLMs + KGs. Especially focused on this work in terms of source grounding.

My Takeaway: Six months ago, I would have said, "this is the way". Source grounding is an important technique, but I think that ontological mapping is more important first. This paper is worth revisiting once the ontological mapping work is more mature.

Synthesis

Ontological Mappings

The creation and creation of knowledge bases is a particularly difficult challenge for most organizations, and in particular something that our team is looking to solve. Ontological mapping and knowledge extraction are very promising approaches to building these knowledge bases from raw source documents provided by organizations. We refer to this as "writing the textbook".

When writing this textbook, Based this research analysis and internal discussions, it seems as though establishing an ontology is a fundamental first step (Matentzoglu et al, 2023). This ontology provides a structured framework for organizing and concepts extracted from raw source documents.

Furthermore, we see these textbooks as foundational to developing the end-to-end educational experience that defines our product. We favor the approach outlined in *Understanding By Design*:

1. Stage 1—Identify Desired Results
2. Stage 2—Determine Assessment Evidence
3. Stage 3—Plan Learning Experiences and Instruction

An important distinction for the work that we are doing is that, while UBD is a fantastic framework, it is geared towards secondary instruction where an appropriate textbook is nearly always available i.e. an algebra book or a geometry book. For our mentor users, no such textbook will be available, rather the knowledge that they are charged with instructing their learners might be completely disorganized or even unavailable. As such, we might even design a stage 0, "write the textbook". This is why knowledge extraction of which ontological ming seems like the most promising approach is so critical to our work.

We can look at Matentzoglu et al. (2023): "Aligning terminological resources, including ontologies, controlled vocabularies, taxonomies, and value sets is a critical part of data integration in many domains such as healthcare, chemistry, and biomedical research". The approach described in this paper, that is combining a high recall, canonical NLP strategy with an LLM to improve precision that's improving overall accuracy is promising. The team behind this paper is now working on "BoomerGPT, a hybrid neuro-symbolic mapping tool that integrates probabilistic inference, description logic reasoning, lexical methods, rule-based methods, and LLMs for the problem of merging diverse ontologies".

Key Questions for Ontological Mappings

- How can large language models be effectively used for ontology mapping and knowledge extraction?
- What canonical techniques can be combined with LLM?
- How can we assess the efficacy?
- What is a useful way to encode the extraction DAG?

Canonical NLP & Large Language Models

There has been a sea change in natural language, processing and artificial intelligence with the advent of large language models. In my own work, this has shifted the focus from developing individual models to leveraging the inherent capabilities of LLMs and more recently combining this with traditional NLP methods. LLMs, especially the foundation models developed by large corporations at great expense, are so powerful that they can accomplish tasks that previously required significant experimentation and training. LLMs are extremely capable as "zero-shot learners". As noted in Caufield et al., (2024), "SPIRES is a new approach to information extraction that leverages recent advances in large language models to populate complex knowledge schemas from unstructured text. It uses ZSL to identify and extract relevant information from query text, which is then normalized and grounded using existing ontologies and vocabularies." As a result, the primary challenge has evolved from model development to the assessment of the quality of results.

As I was reading papers, I found myself almost immediately flipping to see when the paper had been written and if that date was prior to late 2022 when the release of GPT 3 to the public changed everything. Papers using researcher-trained models based on architecture like BERT are interesting, but likely less relevant than those utilizing foundation models like GPT. That said, significantly older techniques, such as rule-based methods and heuristics are suddenly interesting again as glue in between prompted foundation models. In my professional experience, foundation models developed by large organizations nearly always outperform custom trained models. It is also worth noting that as Caufield et al., (2024) states, "Current SPIRES accuracy is comparable to the mid-range of existing Relation Extraction methods, but greatly surpasses an LLM's native capability of grounding entities with unique identifiers. In other words, we must be careful not to chase what is new and exciting.

For our work, the approaches most likely to their fruit would combine heuristics, canonical techniques and specific prompt-LLM pairs organized in decision tree-like structures. This would be similar to "BoomerGPT, a hybrid neuro-symbolic mapping tool that integrates probabilistic inference, description logic reasoning, lexical methods, rule-based methods, and LLMs for the problem of merging diverse ontologies" (Matentzoglu et al. 2023). The challenges are not training models, but effective encoding strategies for designing Directed Acyclic Graphs (DAGs) And developing strategies to measure their efficacy. The assessment of these approaches is the biggest open research area for me and where I should put significant attention in the next few weeks.

Key Questions for Canonical NLP & Large Language Models

1. When was each paper written?
2. Are the researchers working with models they trained, probably using the BERT architecture or, are they working with foundation models like GPT?
3. How can we assess the efficacy?

4. What is a useful way to encode the extraction DAG?

Reflection

This research process has been very exciting for the future of our product and the R&D aspects of the development of that product. I used mostly google search and following citation trails to achieve what I think is a great leap in my own understanding of the work that we are doing. I learned about ontological mappings and interesting work being done in combining canonical NLP and LLMs when constructing knowledge bases and extracting information from raw documents. I am very excited about the MapperGPT and SPIRES approaches as they seem to be interesting applications at the intersection of canonical and contemporary NLP.

I also feel like my background in education is a tremendous asset. It allowed me to immediately leverage my experience with UBD and apply it to the work that we are doing. Combining education theory with AI and NLP strategies has the potential to be very transformative. I'm looking forward to continuing to combine my professional background with both AI and education on our specific project especially with regard to ontological mappings.

Planning

In the coming week I will dig deeper into the areas that emerged from my research this week. My primary research will be into investigating ontological mapping and methods for measuring the success of entity relation and knowledge extraction. I would like to explore BoomerGPT And other approaches to the use of heuristics and decision trees in extraction processes. I'd also like to learn more about non-LLM best practices in entity relation extraction.

Likely my strategy will be to follow research trails, looking to follow the work that went into the works into which I've already looked.

A significant portion of my time will be dedicated to investigating metrics and strategies for measuring the effectiveness of entity relations, and knowledge extraction. I also plan to investigate further some of the tools that I looked at this week – SPIRES, BoomerGPT, MapperGPT, Loom.

Another area to investigate is to try to find specific approaches to ontological mapping as applied to automated textbook, creation, and technical enablement.

Activity

A Simple Standard for Sharing Ontological Mappings (SSSOM)

Citation: Matentzoglu, N., Balhoff, J. P., Bello, S. M., Bizon, C., Brush, M., Callahan, T. J., Chute, C. G., Duncan, W. D., Evelo, C. T., Gabriel, D., Graybeal, J., Gray, A., Gyori, B. M., Haendel, M., Harmse, H., Harris, N. L., Harrow, I., Hegde, H. B., Hoyt, A. L., Hoyt, C. T., Jiao, D., Jiménez-Ruiz, E., Jupp, S., Kim, H., Koehler, S., Liener, T., Long, Q., Malone, J., McLaughlin, J. A., McMurry, J. A., Moxon, S., Munoz-Torres, M. C., Osumi-Sutherland, D., Overton, J. A., Peters, B., Putman, T., Queralt-Rosinach, N., Shefchek, K., Solbrig, H., Thessen, A., Tudorache, T., Vasilevsky, N., Wagner, A. H., & Mungall, C. J. (2022). A Simple Standard for Sharing Ontological Mappings (SSSOM). _Database_, 2022, baac035. https://doi.org/10.1093/database/baac035

Need: The paper addresses the problem of the lack of easy-to-use standards for sharing and interpreting ontological mappings across different databases and vocabularies, of particular interest when ontologies to find genetic relationships being studied by multiple different labs. It investigates the phenomenon of data integration and interoperability issues caused by imprecise, inaccurate, or incomplete mappings, which is a significant barrier in fields requiring high precision, such as diagnostics or risk prediction.

Method: The researchers developed the Simple Standard for Sharing Ontological Mappings (SSSOM), which introduces a machine-readable vocabulary to describe mapping metadata, defines a simple table-based format for integration into data science pipelines, implements collaborative workflows, and provides reference tools and software libraries. They conducted a detailed survey of existing mapping standards and specifications, and described several use cases to illustrate the utility of SSSOM.

Example:

```
# curie_map:
#   ENVO: http://purl.obolibrary.org/obo/ENVO_
#   RO: http://purl.obolibrary.org/obo/RO_
#   PATO: http://purl.obolibrary.org/obo/PATO_
# mapping_set_id: https://github.com/EnvironmentOntology/environmental-exposure-ontology/sssom/ecto.sssom.tsv
# mapping_set_version: 2021-08-18
# license: https://creativecommons.org/licenses/by/4.0/

subject_id    subject_label  predicate_id     object_id      object_label  mapping_justification author_id
mapping_date
ENVO:01001023 arctic sea ice skos:exactMatch  ENVO:01000348  sea ice       lexical               <AUTHOR_ID>
ENVO:01000338 fresh water    skos:exactMatch  ENVO:00002006  freshwater    lexical               <AUTHOR_ID>
ENVO:01000339 groundwater    skos:exactMatch  ENVO:01001004  groundwater   lexical               <AUTHOR_ID>
```

This figure demonstrates a concrete example of how SSSOM represents ontological mappings in a simple, tabular format, which is a key feature of the standard described in the paper (Matentzoglu et al., 2022, Fig. 3).

Tools

sssom-py: This is a Python library and command-line toolkit specifically designed for working with SSSOM. Its functionality includes:

- Importing files from different formats (e.g., OBO Graphs JSON, RDF Alignment API)
- Exporting SSSOM tables to various formats (RDF, OWL, JSON-LD)
- Merging and querying SSSOM tables
- Validating SSSOM tables

rdf-matcher: This is a matcher for RDF vocabularies or OWL ontologies that can export mapping sets as SSSOM tables. It includes metadata such as mapping tool, confidence, match fields, and match string.

LinkML (Linked Data Modeling Language): While not specific to SSSOM, this tool is used to manage the SSSOM schema. It allows for automatic conversion of the schema into various representations (JSON Schema, ShEx, SHACL, OWL) and generation of utility classes for data validation and conversion.

Boomer: This tool implements the k-BOOM algorithm for ontology merging and can read SSSOM files as mapping candidates.

Databases

Schema.org: While not strictly a mapping database, it provides vocabularies for structured data on the internet, which often involves mapping between different data schemas.

DBpedia: This project extracts structured content from Wikipedia and includes mappings between Wikipedia infobox templates and the DBpedia ontology.

PARIS (Probabilistic Alignment of Relations, Instances, and Schema): This is a system for automatically aligning ontologies of different knowledge bases, which has been used with general-purpose knowledge bases like YAGO and DBpedia.

Cultural Heritage Abstract Reference Model (CHARM): This involves mappings between different cultural heritage data models.

Linked Open Vocabularies (LOV): This is a catalog of RDFS vocabularies and OWL ontologies, which includes some mapping information between vocabularies.

Audience: The primary participants or subjects of the study are the various ontological mappings and databases in fields like biomedical research, healthcare, and data science. The intended audience includes researchers, data scientists, and developers who work on data integration and interoperability, especially those involved in ontology and database management.

Results: The key findings of the study are that SSSOM provides a rich, extensible vocabulary for mapping metadata, supports a simple tabular format for easy integration into data pipelines, and promotes community-driven collaborative workflows. The researchers concluded that SSSOM makes mappings more Findable, Accessible, Interoperable, and Reusable (FAIR), thereby addressing the identified problems in data integration.

Critique:

The paper presents intself as a broad solution for challenges in shared ontological mapping and collaborative ontologies, but the focus is on bioinformatics. All four use cases cited (Mondo Disease Ontology, O×O, National Microbiome Data Collaborative, EOSC-Life) are firmly rooted in biomedical or life sciences domains. The paper aims at broader applicability but there is little work pointing in that direction.

The Artificial Intelligence Ontology: LLM-assisted construction of AI concept hierarchies

Citation: Joachimiak, M. P., Miller, M. A., Caufield, J. H., Ly, R., Harris, N. L., Tritt,A., Mungall, C. J., & Bouchard, K. E. (2024). The Artificial Intelligence Ontology: LLM-assisted construction of AI concept hierarchies. Lawrence Berkeley National Laboratory.

Need: The paper addresses the challenge of standardizing AI concepts, methodologies, and their interrelations in the nascent field of post-LLM advent artificial intelligence and natural language processing. It investigates the phenomenon of concept hierarchies in AI, emphasizing both technical and ethical aspects. The field is quite new and such a taxonomy is not yet largely agreed upon.

Methods: The researchers developed the Artificial Intelligence Ontology (AIO) using a multi-faceted approach:

- Ontology Development Kit (ODK): Used to organize the AIO and set up a GitHub repository containing source components and workflows for building the ontology. Note that this work was largely developed by Matentzoglu, whose work shows up frequently in this analysis.
- Large Language Models (LLMs): Employed for content suggestion and curation support, particularly for developing the LLM and Preprocessing branches of the ontology. The models were intentionally foundation models in particular largely GPT-4 via the ChatGPT web interface.
- ROBOT templates: Created for each main ontology branch, allowing for easy authoring and contribution by domain experts. https://github.com/ontodev/robot
- Manual curation: Used various sources including publications, the Asimov Institute Neural Network Zoo, Wikipedia, and documentation from PyTorch and TensorFlow.
- AI-assisted ontology development: Leveraged LLMs to streamline the process of ontology expansion and refinement, especially few shot learning with examples from the AIO worksheet.
- Evaluation: Conducted an NLP evaluation using the Papers with Code dataset to assess coverage and applicability of AIO terms in practical AI research.

The ontology was structured around six top-level branches: Networks, Layers, Functions, LLMs, Preprocessing, and Bias. This structure was designed to support modular composition of AI methods and facilitate a deeper understanding of deep learning architectures and ethical considerations in AI.

Audience: The primary audience for the AIO includes:

1. AI researchers: Who can use the standardized terminology for more effective communication and collaboration.
2. Developers: Who can leverage the ontology for more consistent documentation and implementation of AI systems.
3. Educators: Who can utilize the standardized concepts for teaching AI methodologies.
4. Data scientists: Who can benefit from the standardized terminology in data integration and interoperability tasks.
5. Ethicists and policymakers: Who can use the Bias branch to address ethical considerations in AI development and deployment.

The study also indirectly involves the broader AI community, including those who contribute to and use resources like Papers with Code, Hugging Face, and GitHub.

Results: The key findings of the study include:

1. Creation of a comprehensive ontology: AIO contains 417 classes, 360 synonyms, and 414 is_a relationships, structured to support modular composition of AI methods.
2. High coverage of current AI research: The NLP evaluation demonstrated significant coverage of AIO terms in the Papers with Code dataset, with 205 out of 417 AIO terms found in paper titles and method classification fields.
3. Flexible and extensible framework: The use of ROBOT templates and LLM-assisted curation allows for easy updates and extensions to keep pace with the rapidly evolving field of AI.
4. Ethical considerations: The inclusion of a dedicated Bias branch addresses the growing need for ethical considerations in AI development.
5. Integration with existing resources: AIO was successfully integrated into BioPortal, demonstrating its potential for cross-disciplinary research and interoperability with other ontologies.
6. Enhanced model cards: The researchers proposed using AIO to enhance Model Cards, improving transparency and understanding of AI models.

Critique: While the paper presents a significant advancement in AI concept standardization, there are some limitations to consider:

1. Scope limitations: The ontology does not delve into specifics of individual model implementations or parameter values, which may limit its utility for highly specialized applications.
2. Simplification of network layers: The representation of Network Layers as a list may not fully capture the complexity of nonlinear architectures with loops.
3. Bias towards biomedical applications: Although intended for broader use, the integration with BioPortal and the authors' affiliations suggest a potential bias towards biomedical applications of AI.
4. Rapid evolution of AI: The fast-paced nature of AI advancements may challenge the ontology's ability to stay current, despite the implemented update mechanisms.
5. Adoption challenges: The paper does not extensively address potential barriers to widespread adoption of the AIO in the AI community, which could impact its long-term success.

Mappergpt: Large language models for linking and mapping entities

Citation: Matentzoglu, N., Caufield, J. H., Hegde, H. B., Reese, J. T., Moxon, S., Kim, H., ... & Mungall, C. J. (2023). Mappergpt: Large language models for linking and mapping entities. arXiv preprint arXiv:2310.03666.

Need: The paper explores a novel usage of LLMs. LLMs have an established practice of using SOA relation extraction models but these models are high recall. This paper looks at using LLMs to balance this by providing a filter on entities returned by these high recall processes in order to also achieve high precision.

Methods: The researchers developed MapperGPT, an innovative approach that combines existing high-recall methods for generating candidate mappings with LLMs to review and refine these mappings as a post-processing step. The methodology involves several key components:

1. Input Generation: MapperGPT takes two ontologies and a set of candidate mappings as input. These mappings are typically generated from high-recall methods like LOOM.
2. Prompt Generation: For each candidate mapping, a prompt is generated that includes:
    - Descriptions of the two concepts being mapped
    - Examples of different mapping categories
    - A request for the LLM to categorize the relationship between the concepts
3. LLM Processing: The prompt is passed to a GPT model (either GPT-3.5-turbo or GPT-4) via the OpenAI API.
4. Response Parsing: The LLM's response is parsed to extract key elements such as the mapping category, confidence level, similarities, and differences between the concepts.
5. Output Generation: The results are formatted in the Simple Standard for Sharing Ontological Mapping (SSSOM) format.

Audience: The primary audience for this research includes:

1. Biomedical researchers and data scientists working on data integration challenges.
2. Ontology developers and curators in fields such as healthcare, biology, and chemistry.
3. AI and machine learning researchers interested in applications of LLMs to structured data problems.

4. Database administrators and information architects dealing with heterogeneous data sources.
5. Software developers working on tools for data harmonization and interoperability.
6. Policy makers and standardization bodies in the biomedical and healthcare domains.

The study also indirectly involves the broader AI and data science community, as it demonstrates a novel application of LLMs to a long-standing problem in data integration.

Results: The key findings of the study include:

1. Overall Performance: MapperGPT with GPT-4 achieved an overall accuracy (F1 score) of 0.672 across all tasks, representing a 24% improvement over the state-of-the-art LogMap system (0.527).
2. Precision vs. Recall: MapperGPT with GPT-4 consistently achieved higher precision than other methods while maintaining competitive recall rates.
3. Model Comparison: GPT-4 significantly outperformed GPT-3.5-turbo in all tasks, highlighting the importance of using more advanced language models.
4. Domain Adaptability: MapperGPT demonstrated strong performance across different domains (anatomy, developmental biology, disease), suggesting good generalizability.

Critique: MapperGPT represents a significant advancement in the field of ontology and entity mapping, showcasing the potential of LLMs to enhance precision in complex data integration tasks. The approach's ability to leverage both lexical information and semantic understanding positions it as a promising tool for improving interoperability across diverse terminological resources.

GPT-4: A Stochastic Parrot or Ontological Craftsman? Discovering Implicit Knowledge Structures in Large Language Models

Citation: Procko, T. T., Elvira, T., & Ochoa, O. (2023, September). GPT-4: A Stochastic Parrot or Ontological Craftsman? Discovering Implicit Knowledge Structures in Large Language Models. In 2023 Fifth International Conference on Transdisciplinary AI (TransAI) (pp. 147-154). IEEE.

Need: The paper addresses the fundamental challenge of understanding the implicit knowledge structures that Large Language Models (LLMs), specifically

GPT-4, can create when prompted about a novel domain without explicit instructions to form an ontology. This investigation is crucial for several reasons:

1. Ontological understanding: There's a need to explore whether LLMs can generate unique or novel ontological forms, potentially offering new insights into knowledge representation.
2. Default knowledge organization: Understanding the default structure LLMs use to organize knowledge can provide insights into their internal representations and biases.
3. Creativity in AI: Assessing the ability of LLMs to generate creative or novel knowledge structures beyond their training data.
4. Comparison with human ontology creation: Exploring how LLM-generated knowledge structures compare to those created by human ontologists.
5. Implications for AI-assisted ontology development: Insights from this study could inform the development of more sophisticated AI-assisted ontology creation tools.

The research is motivated by the philosophical question of whether LLMs truly "understand" the world in a way comparable to human ontologists, or if they are merely "stochastic parrots" recapitulating patterns from their training data.

Methods: The researchers employed a multi-faceted approach to investigate GPT-4's implicit knowledge structures:

1. Baseline Ontology Creation: A fiat ontology for the domain of Machine Learning was created, drawing inspiration from existing AI ontologies (AIO and ANNETT-O). This served as a control for comparison with GPT-4's outputs.
2. Term Selection: A set of "potentially particular" terms (mostly leaf nodes) from the baseline ontology was compiled into a comma-separated list.
3. Initial Prompting: GPT-4 was prompted with a general request to organize the provided terms, without explicitly mentioning ontologies or taxonomies.
4. Iterative Prompting: Follow-up prompts were used to encourage GPT-4 to extend and refine its initial organization.
5. Creative Prompting: GPT-4 was explicitly prompted to be creative and unique in its organization, moving beyond traditional knowledge structures.
6. Domain Variation: The experiment was repeated with terms from disparate domains to test GPT-4's versatility.
7. No-Lexicon Test: GPT-4 was prompted to organize the entire domain of Machine Learning without being provided specific terms.

8. API Testing: The gpt-4-0613 model was tested through OpenAI's API playground with varying temperature settings to explore more creative outputs.
9. Analysis: The researchers analyzed GPT-4's responses across multiple trials, focusing on the structure, depth, and novelty of the generated organizations.

Audience: The primary audience for this research includes:

1. AI Researchers: Particularly those working on language models, knowledge representation, and machine learning.
2. Ontology Experts: Professionals and academics involved in the development and study of ontologies.
3. Cognitive Scientists: Those interested in comparing AI knowledge structures to human cognition.
4. AI Ethicists: Individuals concerned with the implications of AI's understanding and representation of knowledge.
5. Software Developers: Those working on AI-assisted tools for knowledge organization and ontology creation.
6. Philosophers of AI: Scholars exploring the nature of machine intelligence and knowledge.
7. Data Scientists: Professionals interested in how AI models organize and represent complex information.

The study also indirectly involves the broader AI and cognitive science communities, as it touches on fundamental questions about machine intelligence and knowledge representation.

Results: The key findings of the study include:

1. Default to Taxonomy: GPT-4 consistently defaulted to creating taxonomic structures when organizing knowledge, even when explicitly prompted to be creative or unique.
2. Shallow Hierarchies: The generated taxonomies were often shallow, rarely exceeding two levels of depth.
3. Domain Consistency: GPT-4's tendency to create taxonomies persisted across different domains, from Machine Learning to disparate concepts like philosophy and biology.
4. Categorical Binning: When not creating explicit hierarchies, GPT-4 tended to bin terms into categories.

5. Limited Creativity: Even with explicit prompts for creativity, GPT-4's outputs remained fundamentally taxonomic, though sometimes with novel perspectives (e.g., musical or spatial metaphors).
6. Natural Language Influence: The researchers concluded that GPT-4's reliance on taxonomies is likely due to the inherently taxonomic nature of natural language, on which it was trained.
7. Iterative Improvement: The quality and depth of GPT-4's organizations improved with iterative prompting, suggesting the importance of interaction in extracting more sophisticated structures.
8. Consistency with Established Knowledge: When organizing the entire ML domain without a provided lexicon, GPT-4 produced a taxonomy largely consistent with established domain knowledge.

The researchers concluded that GPT-4's organization of knowledge is heavily influenced by the taxonomic nature of natural language and its training data. This leads them to characterize GPT-4 as more of a "stochastic parrot" than a creator of novel ontological forms, at least in the context of unprompted knowledge organization.

Critique: While the study provides valuable insights, there are several limitations and areas for future research:

1. Limited Domain Exploration: The primary focus on Machine Learning may not fully represent GPT-4's capabilities across all domains.
2. Lack of Quantitative Metrics: The analysis is largely qualitative, lacking numerical metrics to quantify the similarity or novelty of generated structures.
3. Potential Prompt Bias: The phrasing of prompts, even when attempting to be neutral, may inadvertently bias GPT-4 towards certain types of responses.
4. Single Model Focus: The study focuses solely on GPT-4, without comparison to other LLMs or earlier versions of GPT.
5. Limited Exploration of Multimodal Capabilities: The paper mentions GPT-4's potential for multimodal inputs but does not explore this aspect in depth.
6. Absence of Human Baseline: The study could benefit from a comparison with how humans organize the same terms when given similar prompts.

Despite these limitations, the paper provides valuable insights into the knowledge organization capabilities of GPT-4 and raises important questions about the nature of knowledge representation in LLMs. It sets a foundation for future research into more sophisticated prompting techniques, multimodal inputs, and comparative studies across different AI models and human ontologists.

Structured prompt interrogation and recursive extraction of semantics (SPIRES): A method for populating knowledge bases using zero-shot learning

Citation: Caufield, J. H., Hegde, H., Emonet, V., Harris, N. L., Joachimiak, M. P., Matentzoglu, N., ... & Mungall, C. J. (2024). Structured prompt interrogation and recursive extraction of semantics (SPIRES): A method for populating knowledge bases using zero-shot learning. Bioinformatics, 40(3), btae104.

Need: The paper addresses the critical challenge of efficiently populating knowledge bases and ontologies, which traditionally rely on time-consuming manual curation. This challenge is particularly acute in fields such as biomedicine, where the rapid growth of scientific literature outpaces manual curation efforts. The research investigates the potential of using Large Language Models (LLMs) to automate and accelerate this process through zero-shot learning capabilities. Specific needs addressed include:

1. Reducing the time and effort required for manual curation of knowledge bases.
2. Developing methods to populate arbitrarily complex nested knowledge schemas without extensive training data.
3. Improving the accuracy and consistency of information extraction from unstructured text.
4. Bridging the gap between natural language understanding and structured knowledge representation.
5. Enabling cross-domain application of knowledge extraction techniques.

Methods: The researchers developed a novel approach called Structured Prompt Interrogation and Recursive Extraction of Semantics (SPIRES). The methodology involves several key components:

1. Knowledge Schema Definition: Users define detailed schemas using LinkML, a general-purpose data modeling framework.
2. Prompt Generation: SPIRES generates prompts based on the schema structure and input text.
3. LLM Querying: The generated prompts are sent to an LLM (e.g., GPT-3.5-turbo) for completion.
4. Recursive Extraction: SPIRES recursively processes the LLM's responses to handle nested structures in the schema.

5. Entity Grounding: Extracted entities are mapped to identifiers in existing vocabularies, ontologies, or databases using tools like the Ontology Access Kit (OAKlib).
6. OWL Translation: Optional conversion of results to Web Ontology Language (OWL) format for further reasoning and consistency checking.

The researchers evaluated SPIRES across multiple domains:

- Food recipes
- Multi-species cellular signaling pathways
- Disease treatments
- Multi-step drug mechanisms
- Chemical to disease relationships

They also conducted a specific evaluation on the BioCreative Chemical-Disease-Relation (BC5CDR) task corpus to benchmark SPIRES against existing relation extraction methods.

Audience: The primary audience for this research includes:

1. Bioinformaticians and computational biologists working on knowledge base construction and curation.
2. Data scientists and AI researchers interested in natural language processing and knowledge extraction.
3. Ontology developers and curators in life sciences and other domains.
4. Researchers in fields requiring large-scale information extraction from scientific literature.
5. Software developers working on tools for automated knowledge management.
6. Domain experts in fields such as biomedicine, chemistry, and nutrition who rely on curated knowledge bases.

The study also indirectly involves the broader scientific community, as improved knowledge base population could accelerate research across multiple disciplines.

Results: The key findings of the study include:

1. Zero-Shot Performance: SPIRES demonstrated the ability to populate complex knowledge schemas without task-specific training, leveraging the zero-shot learning capabilities of LLMs.

2. Accuracy Comparison: SPIRES achieved accuracy comparable to mid-range existing Relation Extraction methods on the BC5CDR task, with an F-score of 41.16 using GPT-3.5-turbo with chunking.
3. Entity Grounding Improvement: SPIRES significantly outperformed LLMs' native capabilities in grounding entities with unique identifiers. For example, it correctly identified 98 out of 100 Gene Ontology terms, compared to only 3 correct identifications by GPT-3.5-turbo alone.
4. Cross-Domain Applicability: The method successfully extracted structured information across various domains, from food recipes to biomedical pathways.
5. Flexibility and Customization: SPIRES demonstrated the ability to adapt to user-defined schemas without requiring retraining or extensive modifications.
6. Recursive Handling: The approach effectively managed nested structures in complex knowledge schemas, a capability not typically found in traditional relation extraction methods.
7. Integration with Existing Resources: SPIRES successfully leveraged existing ontologies and databases for entity grounding, enhancing the reliability of extracted information.

The researchers concluded that SPIRES offers a promising approach for assisting manual knowledge curation and acquisition, particularly in domains with complex, hierarchical knowledge structures. They emphasized its advantages in customization, flexibility, and the ability to perform new tasks without additional training data.

Critique: While the study presents significant advancements, there are several limitations and areas for future research:

1. Model Dependency: The current implementation relies heavily on proprietary models like GPT-3.5, which may limit accessibility and reproducibility.
2. Hallucination Risk: Despite efforts to mitigate it, the risk of LLM hallucination remains a concern, particularly for critical applications like clinical decision support.
3. Evaluation Scope: While the BC5CDR task provides a useful benchmark, evaluation on a broader range of tasks and domains would strengthen the generalizability claims.
4. Computational Resources: The recursive nature of SPIRES and the use of large language models may require significant computational resources, potentially limiting its application in resource-constrained settings.

5. Human-in-the-Loop Integration: The study focuses on automated extraction, but future work could explore how to optimally integrate SPIRES with human expert curation workflows.
6. Ontology Alignment: While SPIRES improves entity grounding, challenges in ontology alignment and integration across different knowledge bases remain to be fully addressed.

Despite these limitations, SPIRES represents a significant step forward in automating knowledge base population, offering a flexible and powerful tool for researchers and practitioners across various domains. Its ability to handle complex, nested knowledge structures while leveraging existing ontologies positions it as a valuable asset in the ongoing effort to organize and utilize the vast amounts of unstructured information in scientific literature.

Nicolas Matentzoglu

An emerging phenomenon in the work that was done. Here is the recurring appearance of a single name, Nicolas Matentzoglu. Below is a list of publications by this researcher of which three appeared in the activity portion of this assignment. These are just a sampling of his works better included here as potentially useful for ongoing research. I am definitely a believer in the principle of follow of the leader, especially when it comes to frontier research.

- Caufield, J. H., Hegde, H., Emonet, V., Harris, N. L., Joachimiak, M. P., Matentzoglu, N., ... & Mungall, C. J. (2024). Structured prompt interrogation and recursive extraction of semantics (SPIRES): A method for populating knowledge bases using zero-shot learning. Bioinformatics, 40(3), btae104.
- Matentzoglu, N., Caufield, J. H., Hegde, H. B., Reese, J. T., Moxon, S., Kim, H., ... & Mungall, C. J. (2023). Mappergpt: Large language models for linking and mapping entities. arXiv preprint arXiv:2310.03666.
- Matentzoglu, N., Balhoff, J. P., Bello, S. M., Bizon, C., Brush, M., Callahan, T. J., Chute, C. G., Duncan, W. D., Evelo, C. T., Gabriel, D., Graybeal, J., Gray, A., Gyori, B. M., Haendel, M., Harmse, H., Harris, N. L., Harrow, I., Hegde, H. B., Hoyt, A. L., Hoyt, C. T., Jiao, D., Jiménez-Ruiz, E., Jupp, S., Kim, H., Koehler, S., Liener, T., Long, Q., Malone, J., McLaughlin, J. A., McMurry, J. A., Moxon, S., Munoz-Torres, M. C., Osumi-Sutherland, D., Overton, J. A., Peters, B., Putman, T., Queralt-Rosinach, N., Shefchek, K., Solbrig, H., Thessen, A., Tudorache, T., Vasilevsky, N., Wagner, A. H., & Mungall, C. J. (2022). A Simple Standard for Sharing Ontological Mappings (SSSOM). *Database*, 2022, baac035. https://doi.org/10.1093/database/baac035

- Matentzoglu, N., Braun, I., Caron, A. R., Goutte-Gattat, D., Gyori, B. M., Harris, N. L., ... & Mungall, C. J. (2023). A simple standard for ontological mappings 2023: updates on data model, collaborations and tooling. In OM@ ISWC (pp. 73-78).
- Vasilevsky, N., Overton, J., Jackson, R., Toro, S., Tan, S., Varner, B., ... & Matentzoglu, N. (2022). OBO Academy: Training materials for bio-ontologists. ISMB Bio-Ontologies Community, 273(10.5281).
- Goldbeck, G., Simperler, A., Bull, L., Gao, D., Ghedini, E., Karray, M., ... & Waaler, A. (2022). The Translator in Knowledge Management for Innovation–towards Industry Commons.
- Trojahn, C., & Matentzoglu, N. (2022). What should be the minimum requirements for making FAIR ontology alignments?. In OM@ ISWC (pp. 220-222).
- Matentzoglu, N., Goutte-Gattat, D., Tan, S. Z. K., Balhoff, J. P., Carbon, S., Caron, A. R., ... & Osumi-Sutherland, D. (2022). Ontology Development Kit: a toolkit for building, maintaining and standardizing biomedical ontologies. Database, 2022, baac087.
- Chan, Lauren E., et al. "A Semantic Model Leveraging Pattern-based Ontology Terms to Bridge Environmental Exposures and Health Outcomes." ICBO. 2021.
- Dhombres, F., & Charlet, J. (2019). Formal medical knowledge representation supports deep learning algorithms, bioinformatics pipelines, genomics data analysis, and big data processes. Yearbook of medical informatics, 28(01), 152-155.
- Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology?. Handbook on ontologies, 1-17.

As I dig deeper into his work, I find a couple of resources that are particularly valuable. First is the OBO Academy (https://oboacademy.github.io/obook/). This has promise as a tool for my own learning to become an expert in ontologies. And looking at the OBO Academy, it is stated that the course was developed using the Diataxis Framework (https://diataxis.fr/). This framework has promise as a fundamental framework for our work in that. It is a framework for building courses. Our product is essentially a course building and delivering as a service product and it seems that this may be a great foundational framework for each course that we build and deliver for our users.

Appendix

The activity portion of this assignment was developed with assistance from openAI GPT4o using the following prompt:

```
def logic_chart_prompt(RESEARCH_PAPER_TEXT, RESEARCH_LOG_NOTES):
    return f"""
    You are tasked with analyzing a research paper and creating a logic chart based
on the paper's content and your research log notes.
    Follow these instructions carefully to complete the task:

    First, carefully read through the following research paper:

    <research_paper>
    {RESEARCH_PAPER_TEXT}
    </research_paper>

    Now, review your research log notes related to this paper:

    <research_log>
    {RESEARCH_LOG_NOTES}
    </research_log>

    Using the information from both the research paper and your notes, complete the
following steps:


    1. Create a logic chart by answering the following questions in 2-3 sentences
each. Use the specified tags for each section:

    <need>
    a) Need:
   - What problem is the paper trying to solve?
    - What phenomenon is it investigating?
    </need>

    <method>
   b) Method:
   - What did the researchers build or do to address the need?
   - How did they investigate the phenomenon or solve the problem?
   </method>

   <audience>
   c) Audience:
   - Who are the participants or subjects of the study?
   - Who is the intended audience for the results?
```

```
   </audience>

   <results>
   d) Results:
   - What were the key findings of the study?
   - What conclusions did the researchers draw?
   </results>

    2. Critique the alignment between need, audience, method, and results. Address
the following points in your critique:
    - Are the results justified by the method used?
    - Does the method effectively address the identified need?
    - Is the audience appropriate for the need and method?
    - Overall, how strong is the alignment between these four elements?

    Place your critique in <alignment_critique> tags.

    3. Reflect on any challenges you encountered in fitting the paper into this
structure. If the paper doesn't fit neatly into these categories,
    explain why and how you interpreted it. Place your reflection in <reflection>
tags.

    5. Summarize your analysis in 3-4 paragraphs, covering all the above points.
Place your summary in <summary> tags.

    Throughout your analysis, incorporate insights from your research log notes to
provide additional context or perspectives you may have
    gained while reading the paper.

    Present your complete analysis within <research_paper_analysis> tags, following
this structure:

    <artifact name="logic_chart">

    <need>
    [Need content here]
    </need>

    <method>
    [Method content here]
    </method>

    <audience>
    [Audience content here]
    </audience>

    <results>
```

```
[Results content here]
</results>
</artifact>

<artifact name="alignment_critique">
[Alignment critique here]
</artifact>

<artifact name="reflection">
[Reflection on challenges here]
</artifact>

<artifact name="summary">
[Summary paragraphs here]
</artifact>
"""
```

Results were iterated upon via discussion with Anthropic's Claude. The most common strategy used there has been to present the text e.g. the returned logic chart and the original pdf, then to ask Claude to ask my questions about the work so that Claude can understand its importance to me. Claude will then help to organize my thoughts which are then resynthesized by me for the assignment.