

Creating a Fine Grained Entity Type Taxonomy Using LLMs

Michael Gunn

mwgunn2@illinois.edu

Dohyun Park

dohyunp2@illinois.edu

Nidhish Kamath

nkamath5@illinois.edu

Abstract

In this study, we investigate the potential of GPT-4 and its advanced iteration, GPT-4 Turbo, in autonomously developing a detailed entity type taxonomy. Our objective is to construct a comprehensive taxonomy, starting from a broad classification of entity types - including objects, time, locations, organizations, events, actions, and subjects - similar to existing manually curated taxonomies. This classification is then progressively refined through iterative prompting techniques, leveraging GPT-4’s internal knowledge base. The result is an extensive taxonomy comprising over 5000 nuanced entity types, which demonstrates remarkable quality upon subjective evaluation.

We employed a straightforward yet effective prompting strategy, enabling the taxonomy to be dynamically expanded. The practical applications of this detailed taxonomy are diverse and significant. It facilitates the creation of new, more intricate branches through pattern-based combinations and notably enhances information extraction tasks, such as relation extraction and event argument extraction. Our methodology not only introduces an innovative approach to taxonomy creation but also opens new avenues for applying such taxonomies in various computational linguistics and AI-related fields.

1 Introduction

Recent breakthroughs in large language models, notably GPT-4 (OpenAI, 2023), have ushered in an era of advanced natural language processing tools. These models exhibit extensive general language comprehension capabilities, resulting from their extensive training process. GPT-4, in particular, stands out for its deep understanding of language nuances. Utilizing their language understanding, these large language models have greatly improved the performance of various natural language tasks (Mao et al., 2023). However, despite

these advances, these models can struggle to maintain consistency in structured knowledge representation.

Addressing this gap motivates the development of a robust type taxonomy. Historically, type taxonomies, whether small-scale or human-crafted, have been pivotal in various domains, including entity typing, relation extraction, and event argument extraction. However, they often fall into two categories: either limited in scope or manually created, posing a significant challenge in being adapted for wide use. This project explores the possibility of leveraging GPT-4’s nuanced language understanding to construct a detailed, fine-grained entity type taxonomy.

Our project begins with an initial exploration, inspired by the granular entity types identified in Ultrafine (Choi et al., 2018). Next, we detail our comprehensive process of taxonomy development as well as the resulting taxonomy. We delve into the methodology, highlighting the iterative prompting techniques employed to refine and expand the taxonomy. Furthermore, we discuss the potential applications of this fine grained taxonomy, exploring its capability in augmenting existing methods in entity typing, relation extraction, and event argument extraction. This exploration not only showcases the capabilities of GPT-4 in structured knowledge organization but also illuminates new avenues for its application in computational linguistics and AI research.

2 Background

The evolution of Large Language Models (LLMs) such as GPT-4 has significantly impacted the field of natural language processing, particularly in the creation and utilization of type taxonomies. These taxonomies play a crucial role in structuring knowledge for complex NLP tasks. Among the pioneering works in this domain is the UltraFine Entity Typing Task (Choi et al., 2018). This endeavor fo-

All code and taxonomies created can be found at [github](#)

cuses on discerning fine-grained entity types, advancing the representation of structured knowledge in numerous NLP tasks like relation extraction and event argument extraction. Yet, despite its innovativeness, this method encountered scalability and consistency challenges, particularly in managing a wide spectrum of entity types.

Another significant contribution in this field is the OntoType model (Komarlu et al., 2023). OntoType introduces an ontology-guided, zero-shot fine-grained entity typing system with weak supervision from pre-trained language models. This method represents a substantial advancement in understanding and classifying fine-grained entity types, particularly in how it leverages the pre-existing knowledge embedded in language models.

Our work extends these efforts by exploring the potential of GPT-4 in autonomously creating a fine-grained entity type taxonomy. We build upon the foundational concepts established in the UltraFine Entity Typing Task and integrate them with the advanced capabilities of GPT-4 and GPT-4 Turbo. This integration aims to overcome the limitations faced in previous approaches, such as the UltraFine system’s reliance on crowdsourced results and the challenges in scaling up to thousands of entity types.

In our methodology, we utilize GPT-4’s language understanding to develop a comprehensive taxonomy, starting from a broad classification and progressively refining it through iterative prompting techniques. This process results in an extensive taxonomy comprising over 5000 nuanced entity types. Our approach not only introduces an innovative way of creating taxonomies, but also demonstrates the practical applications of such a detailed taxonomy in enhancing information extraction tasks. For instance, the taxonomy facilitates the creation of more specific event patterns using a hierarchical entity taxonomy, which is particularly beneficial in event argument extraction and relation extraction tasks.

In summary, our project draws on the foundations laid by previous work in type taxonomies, enriching it through the application of GPT-4’s capabilities. We propose a methodology that refines the concept of structured knowledge representation in NLP. While it marks progress in the field, we also acknowledge the complexities and ongoing challenges in creating and applying extensive taxonomies. Our work, therefore, should be seen

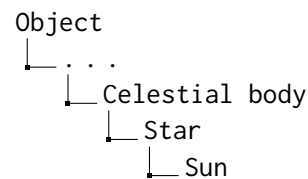
as a step towards enhancing information extraction tasks, offering insights into how advanced language models can contribute to the development of more nuanced and effective taxonomies in computational linguistics.

3 UltraFine Entity Typing Task

We motivate our work by highlighting the entity typing task (Choi et al., 2018) and contrasting how humans and LLMs might fare on it. The task is to predict natural-language phrases that describe an entity e contained in a given sentence. Consider the following sentence and the entity ‘it’s’ as an example:

“For starters, it’s not an ordinary sun but a Cepheid variable – a giant, pulsating star shining with the light of at least a thousand suns.”

What are the applicable types for ‘it’s’ in this sentence? Crowdsourced results from UltraFine (Choi et al., 2018) labeled it as ‘object’, ‘celestial body’, ‘sun’, ‘star’. An LLM such as GPT-4 can easily provide answers like ‘sun’ or ‘star’ directly, but would miss out on stating the other perfectly relevant entity types. This is where we feel that having an ontology of entity types would greatly improve results. A natural fine grained ontology would contain ‘sun’ under:



Any method that can infer ‘it’ as a type of ‘sun’ could then, by utilizing the structured knowledge of the ontology, infer all parent types as being applicable too.

From this example, we believe the role of a well structured fine grained type taxonomy in creating extensive multi-labeling of entities is clear. Given a well structured taxonomy like the one we create in this work, the task of providing a thorough list of applicable labels for an entity becomes much easier. In particular, existing simple prompting methods can already output accurate results of more specific type labels, which can then be traced using a type taxonomy to find all applicable labels.

4 Methodology

4.1 Initial Approach

In our initial approach, we explored the potential of leveraging a GPT-4 to structure a comprehensive ontology from a provided list of types. We use the list of 10,331 fine grained entity types provided by the Ultra Fine Entity typing task (Choi et al., 2018) and GPT-4 as our LLM. The fundamental strategy involved feeding this exhaustive list to a sufficiently advanced LLM, coupled with an aptly crafted prompt. The objective was to harness GPT-4’s extensive world knowledge to systematically organize these entity types into a cohesive ontology.

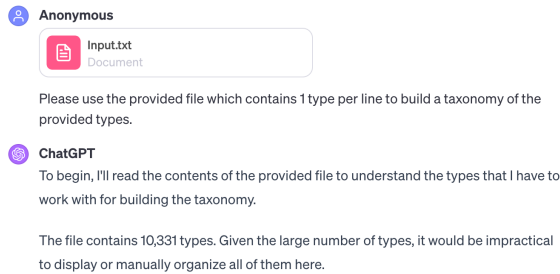


Figure 1: Initial Approach Prompt

4.1.1 Limitations of Initial Approach

This approach initially showed promise, as GPT-4 was capable of understanding the requirements and generating coherent ontology structures. However, certain limitations hindered further adoption. One significant issue was the occurrence of hallucinations. Despite being prompted to adhere to the provided UltraFine types, GPT-4 sometimes introduced unrelated terms into the ontology. Another challenge involved the method of matching terms based on **string similarity**. While we aimed to leverage GPT-4’s general language understanding to build an ontology, despite our prompts, we found that GPT-4 would tend to group phrases from the UltraFine types if they shared common words, regardless of their semantic context. An example of this was grouping 'Family_Business', 'Family_History', and 'Family_Name' under 'Family' simply due to the shared word. This approach often led to inaccurate grouping as can be seen in figure 2. Additionally, the unordered nature of the UltraFine types list itself was a limitation. An ideal scenario would involve feeding GPT-4 types that fit into an existing hierarchy. However, the lack of order in the list complicated the construction of a

semantically accurate hierarchy, often leading to subpar outputs.

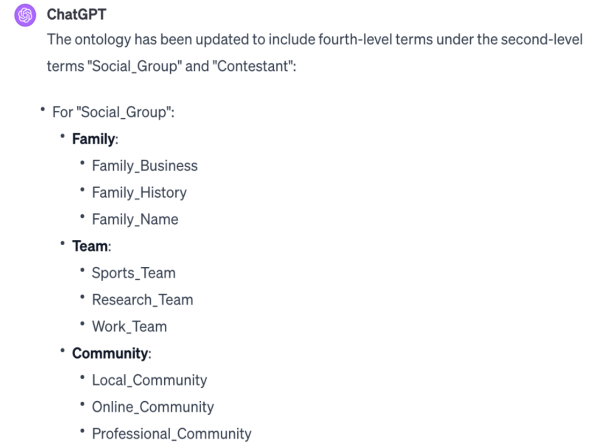


Figure 2: GPT-4 matching phrases on String Similarity

4.1.2 Initial Solutions

To address the limitations from Section 4.1.1, we opted for a more structured approach. Given that the Ultrafine types list doesn’t provide a clear starting point, we opted to develop a basic, broad level ontology using a selection of key terms. This initial framework serves as a *seed* ontology, which provides a foundational structure for GPT-4 to build upon.

Additionally, we formulated specific prompts aimed at encouraging GPT-4 to generate ontology relations grounded in a deeper comprehension of the meanings behind the terms instead of simple string similarity.

Contrary to our expectations, GPT-4 predominantly employed basic string matching rather than demonstrating its general language understanding while building ontologies in this manner. This tendency was particularly evident when the model was tasked with integrating *parent* entities into an existing structure where *child* entities had already been positioned. In such instances, GPT-4 exhibited a tendency to simply append the *parent* entities as a standalone node at the first level, rather than reorganizing the structure for optimal coherence.

This showed a critical limitation in the model’s capacity to dynamically restructure ontologies in response to the sequential presentation of entity types, a constraint notably evident in the context of the unordered UltraFine entity type list.

These observations suggest that our initial methodology may not be effectively scalable or applicable to extensive and complex entity lists,

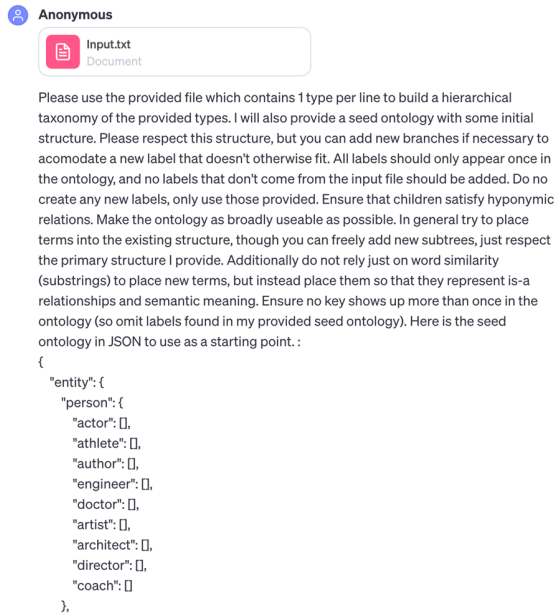


Figure 3: Revised prompt with seed ontology and other explicit instructions

such as the 10,331 types featured in the UltraFine list or other similar compilations. Consequently, this necessitates the development of a more robust approach for ontology creation with large language models like GPT-4, focusing on customization and scalability.

4.2 Final Approach

Despite our initial efforts, as outlined in section 4.1.2, we found the constraints of UltraFine Types more limiting than beneficial for inference. Consequently, we shifted from our initial strategy to a more unrestricted process in our final approach, in which we leveraged the capabilities of GPT-4 without the constraints of the UltraFine types. In this approach, we do not limit types in our ontology to UltraFine types list, and we allow the same type to be able to be inserted in the ontology multiple times as well. We do not provide a Seed Ontology, allowing GPT-4 to structure the ontology as needed. We develop an iterative process to expand an ontology, allowing us to fully utilize GPT-4's language understanding to add new types to our ontology.

5 Iterative Process

GPT-4's ability to utilize file storage for ontology maintenance simplifies the process of keeping a previous ontology in context. This feature was pivotal in designing an iterative process for building our ontology, enabling us to add new types seam-

lessly. Each iteration involved presenting GPT-4 with an existing ontology, either from a prior generation or uploaded as a file. With the previous ontology, we now prompt GPT-4 to expand our ontology. We employed two main methods for prompting: one was requesting the expansion of a specific subtree, and the other involved suggesting a new type for GPT-4 to integrate into the most fitting subtree. This approach ensured a dynamic evolution of our ontology, achieving both broad coverage and detailed categorization. Focusing on expanding one subtree (or multiple subtrees of the same type) at a time, as illustrated in figure 4, resulted in faster, more accurate expansions. Another significant advantage of our approach is the ability for multiple team members to concurrently work on different subtrees, enhancing the efficiency of the ontology expansion. This allows each person to focus on a specific area, diversifying the development process and contributing unique perspectives to various parts of the ontology. Additionally, the flexibility to add any desired subtree into the ontology furthers its comprehensiveness, ensuring that our ontology is not only expansive but also tailored to our specific project needs.

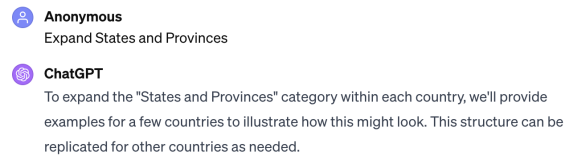


Figure 4: Prompt expanding specific subtree

5.1 Extending to Domain Specific Taxonomies

While the iterative process of building ontologies leverages GPT-4's general language understanding capabilities, it can be further enhanced to incorporate domain-specific knowledge when constructing relevant branches. A recent feature of GPT-4, the *GPT Builder*, enables users to directly feed domain-specific knowledge sources into GPT-4. While untested in this work, utilizing this feature could allow for the creation of more accurate and

relevant ontologies within specific fields. Even without a GPT-4-specific feature, our iterative process permits the expansion of specific branches through prompts that can be enhanced with additional knowledge. For example, in the medical context, medical research papers and journals can be input either through the GPT-4 knowledge source or within each prompt. This tailored approach, where domain knowledge is provided with each prompt, can lead to the creation of high-quality and accurate domain-specific ontologies.

6 Evaluation

Utilizing the iterative process, we were able to generate an ontology with over 5000 types, reaching depths up to 10. This method proved to be highly scalable, allowing for the addition of any necessary subtree branches. As illustrated in figure 5, our system efficiently categorizes a diverse range of types. Particularly noteworthy is the 'organization' subtree, showcasing our method's ability to delve into detailed categorizations. This illustrates the model's strength in grasping general language concepts and generating intricate, fine-grained ontological structures.

In our proof of concept, we did not expand all possible branches to as deep as possible, but the technique we employed has no limitation for scalability that we observed. GPT-4 is able to read JSON files directly and reference their contents for later expansions. This means that the approach we employed in principle should extend into the many tens of thousands of entity types with no problems. Additionally, we believe the recursive expansion technique we utilized even allows for branch expansion to be performed past the entire taxonomy being able to fit inside the 128k context window of GPT-4 Turbo. It would be possible to pass only the relevant subset of the entire taxonomy into the model in order to allow branch expansion to be performed essentially without limit.

Overall the quality of the resulting taxonomy that we observed from our approach subjectively was extremely high. All of the hierarchical divisions created made logical sense and were accurate to real world relationships between entities. We were not able to manually check every single entity type in our resulting taxonomy, but all that we did examine were of high quality.

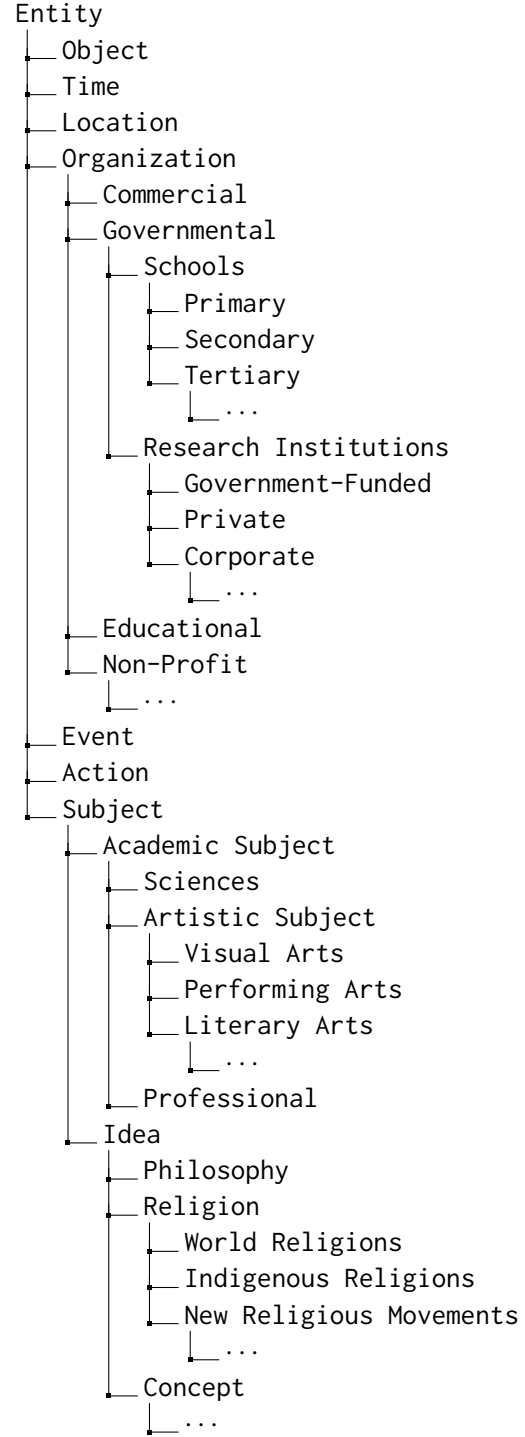


Figure 5: GPT-4 Generated Ontology subtrees

6.1 Limitations

The iterative prompting approach that we developed and employed to create our taxonomy using GPT-4 was very effective, but we did observe and theorize some potential limitations and shortcomings of the method. One such limitation comes from the manner in which we perform branch expansion. We decided to construct the taxonomy

using a hierarchical iterative deepening approach, which involved first creating a high level structure of the most general categories for entities, then recursively expanding each branch to obtain more fine grained types and hierarchical structures. In practice this meant that in order to expand a given branch of the taxonomy, we would prompt it to generate a more detailed version of that branch. The resulting output branch then replaced the previous branch entirely. This branch replacement strategy allowed simple expansion of any given branch of the taxonomy, but it means that the selection of which branches should be expanded next is left up to the user.

Another limitation that we observed in our method comes from the allowance of repeated terms in different places in the taxonomy. From our first approach using the type list from UltraFine (Choi et al., 2018) we observed that GPT-4 could sometimes place the same term in multiple places in the taxonomy. We initially did not feel this was appropriate, but upon further consideration recognized that in a highly extensive type taxonomy, certain terms should occur in multiple sections as they have multiple uses across different areas. One example of this is the term "technology". In an ideal extensive taxonomy the term "technology" should likely occur under the categories of professional occupations, as a subset of objects, under infrastructure, and likely multiple other branches as well. As a result, in our final approach in which we constructed the taxonomy from scratch without restriction on the terms to be used, we relaxed the constraint on term uniqueness.

In general we feel that this decision did allow for a more useful and informative taxonomy, but it does introduce a complication for later applications of the taxonomy. The multiple occurrences of the same term in the taxonomy means that performing a task like entity typing using our taxonomy would necessitate a more detailed description of each type than just its individual label. This problem can be addressed in a number of ways, but the most obvious is to make the resulting type label not just be the final type label from the taxonomy, but instead the entire branch of the taxonomy that ends in that type. Thus instead of "technology" as a type label, one would obtain "Entity / Object / Physical Object / Non-living Beings / Artificial objects / Infrastructure / Technology". The handling of duplicate terms in taxonomy applications

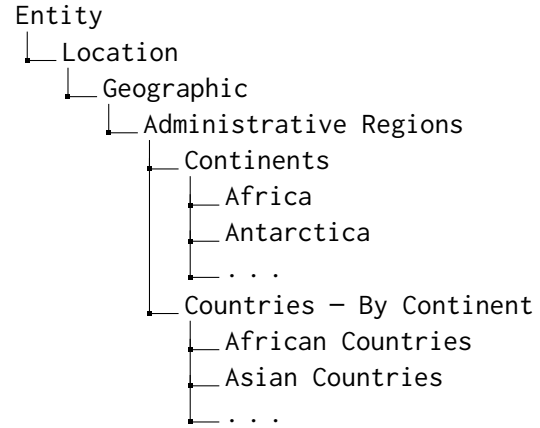


Figure 6: Repetition in Ontology

does introduce some challenges, but we believe that it is worth those costs in order to have a more informative and unrestricted type tree.

7 Ontology Combinations

The GPT-4 generated ontology, while detailed, suffered from excessive granularity, leading to heavy repetition within similar but different categories. For example, as can be seen in figure 6, the name of each country can be seen being repeated in both **Countries** and **Continents**. Furthermore, the name of each country can be seen in totally separate subtrees such as **mountains**, where categories such as **american mountains**, **canadian mountains** can be seen. Such repetition can hinder search within an ontology, and make it overwhelming for users. To overcome these deficiencies, we define new ontology combinations, such that new logical categories can be defined without having excessive repetition.

Following our example from figure 6, we can define a new combination : **Continent + Country** → **Country in Continent**. This leads to expressions like '**Countries in Europe**' or '**Asian Nations**', reducing repetition. Similarly, defining the combination **Country + fauna** leads to expressions such as '**Australian Mammals**' or '**Indonesian Tiger**'.

This approach of defining ontology combinations makes generating ontologies simpler while also making it easier for users to find relevant information and understand the relation between the different categories.

8 Practical Applications

There are many tangible benefits of employing a fine-grained ontology in NLP and AI tasks, and

given our high quality ontology, there are many practical applications where our ontology can be used. Using our ontology can significantly enhance event argument extraction, enabling the identification of more nuanced event patterns that traditional methods may overlook. For the task of entity typing, we believe that the approach introduced in OntoType(Komarlu et al., 2023) is extendable to use a fine grained type taxonomy. Their method of performing flat type classification at each level of the tree can be directly extended to work with a deeper and more extensive taxonomy like the one we created. The only modification we foresee in employing their method with our larger taxonomy is to allow multiple subbranches to be explored concurrently in finding the final type label using something like beam search. This modification allows the approach of OntoType to handle any challenging or opaque decisions, resulting from the broader structure of our taxonomy, which must be made to arrive at the most specific label in our taxonomy for a given entity.

In addition to the task of entity typing, an extensive taxonomy proves invaluable in pattern-based relation extraction and exploring meta-patterns in data. Using our ontology with fine-grained entity types would make relation extraction significantly simpler as type-level patterns describe relations well. For example, the ability to discern complex relations like **\$President, leader, \$Country** illustrates the depth of understanding a quality ontology can provide.

9 Future Work

We believe the methods explored in this work can serve as a foundation and motivation to explore the general approach of utilizing LLMs like GPT-4 to build structured knowledge representations, specifically in the form of extensive type taxonomies. This work has demonstrated that from its training, GPT-4 and GPT-4 Turbo have obtained a deep understanding of natural language and the structured hierarchical relationships between entities. We believe that our iterative prompting approach to taxonomy creation can easily be extended to create fine grained type taxonomies for specific domains. If successful, this process could make it tractable to build structured hierarchical tree representations for topics and entities in both existing and emerging domains in an automated fashion.

Despite the difficulty of manual creation, this

task has been seen as important enough to motivate the creation of large specialized domain type taxonomies such as the PubMed taxonomy (PubMed). Taxonomies like that one allow more informative structuring of academic literature and entities within the field. We believe that iterative prompting to build taxonomies should be explored in the context of the construction of specialized branches of a larger general taxonomy like the one we created, or the construction of entirely new taxonomies for specific fields. This process could involve the same zero shot prompting style we utilized, or potentially utilize providing relevant domain specific documents as grounding for the model to utilize when creating a new specialized taxonomy. Retrieval augmented generation (RAG) (Lewis et al., 2021) has shown success in many other contexts, and the application of relevant context using domain documents has become even easier with the new capability of GPT-4 to ingest existing document formats directly and fit them within its massive context window.

In addition to creating taxonomies, the area of utilizing fine grained entity taxonomies is also ripe for further exploration. Extensive taxonomies such as the one created in this work can be used in existing ontology guided approaches (Komarlu et al., 2023) (Qu et al., 2017) to increase granularity, as well as applied to new tasks in information extraction that have yet to employ structured taxonomies.

This work surely only scratches the surface of exploring the capability of LLMs such as GPT-4 in creating and utilizing structured knowledge. We feel the technique we developed and the high quality output we were able to produce using it can be extended and utilized in future work.

10 Conclusion

In this work we aimed to explore the potential of using GPT-4 to create a fine grained type taxonomy. We believe that structured knowledge representation and consistency remains a shortcoming of existing LLMs, and extensive taxonomies are one way to address that. We demonstrated that GPT-4 has obtained significant understanding of the structural and hierarchical relations between real world entities and concepts. Utilizing the cutting edge file input and structured output capability of GPT-4, we were able to construct an extensive type taxonomy using iterative prompting that is broadly applicable, high quality, and highly exten-

sible. In the process we also developed an approach that can be applied to construct or even augment existing domain specific type trees. Additionally we explored some of the many potential applications for an extensive type taxonomy. These range from extensions of existing methods that already utilize ontologies, to methods of pattern and branch combination approaches to extending taxonomies. Lastly we consider the potential application of structured type trees to existing information extraction tasks. The exploration of the capabilities of massive and highly capable LLMs like GPT-4 and future models is in its early days, and we hope that this work can serve as inspiration for the further utilization of these new powerful tools.

References

- Eunsol Choi, Omer Levy, Yejin Choi, and Luke Zettlemoyer. 2018. [Ultra-fine entity typing](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 87–96, Melbourne, Australia. Association for Computational Linguistics.
- Tanay Komarlu, Minhao Jiang, Xuan Wang, and Jiawei Han. 2023. [Ontotype: Ontology-guided zero-shot fine-grained entity typing with weak supervision from pre-trained language models](#).
- Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2021. [Retrieval-augmented generation for knowledge-intensive nlp tasks](#).
- Rui Mao, Guanyi Chen, Xulang Zhang, Frank Guerin, and Erik Cambria. 2023. [GpTEval: A survey on assessments of chatgpt and gpt-4](#).
- OpenAI. 2023. [Gpt-4 technical report](#).
- PubMed. 2023. Pubmed taxonomy. <https://www.ncbi.nlm.nih.gov/taxonomy>. Accessed: 2023-12-08.
- Meng Qu, Xiang Ren, Yu Zhang, and Jiawei Han. 2017. [Weakly-supervised relation extraction by pattern-enhanced embedding learning](#).

11 Appendix

11.1 Ontology Depth

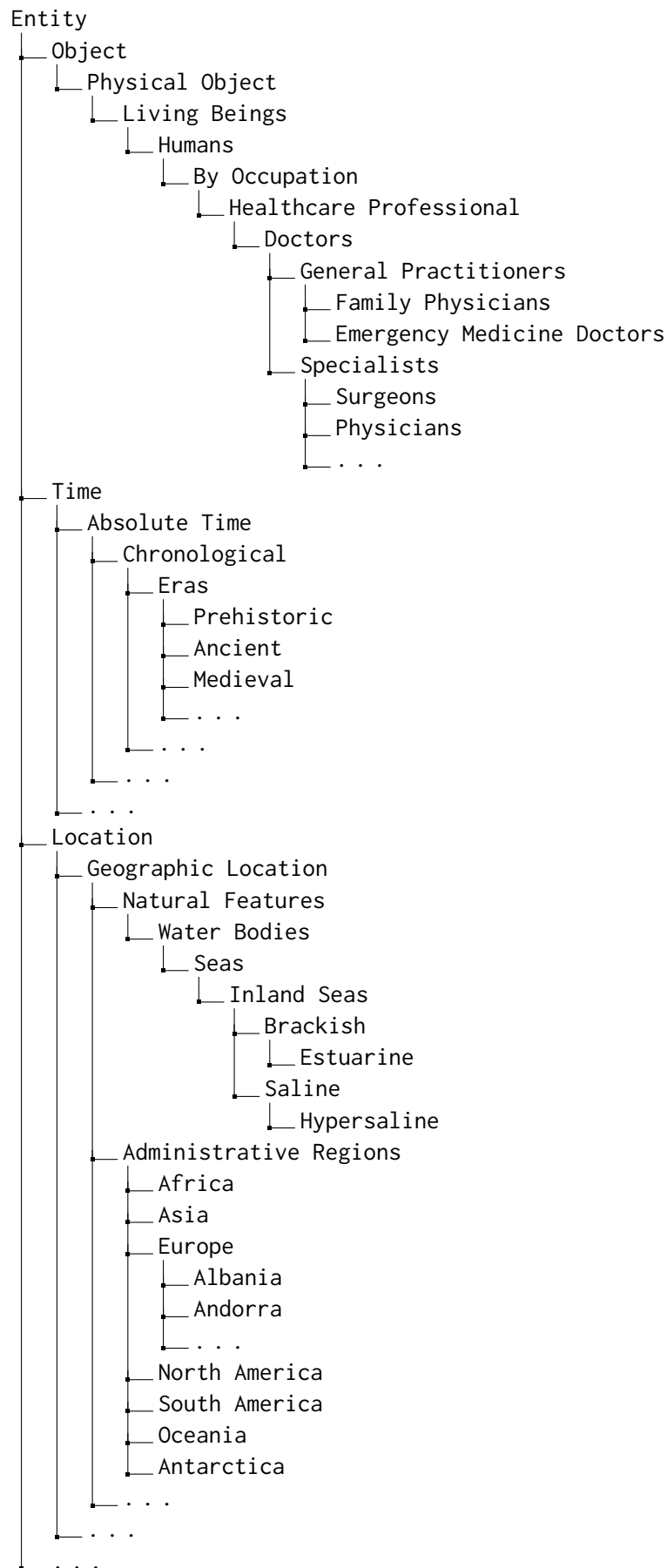


Figure 7: More examples from our ontology covering more depth and breadth