Improving Access to Relevant Knowledge in Large Ontologies through Best Excerpts from Text

Ontology Summarization via Machine Learning Techniques

Martin Tri Vien Lam

Thesis submitted for the degree of
Master in Programming and System Architecture
60 credits

Institute of Informatics
Matnat

UNIVERSITY OF OSLO

Spring 2023

Improving Access to Relevant Knowledge in Large Ontologies through Best Excerpts from Text

Ontology Summarization via Machine Learning Techniques

Martin Tri Vien Lam

Improving Access to Relevant Knowledge in Large Ontologies through Best Excerpts from Text

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo

Acknowledgements

I would like to thank my supervisor Jieying Chen for her guidance throughout this project. She has taught me a lot, and her quick and concise feedback, input and ideas have been invaluable to me. I thank Jiaoyan Chen and Ernesto J. Ruiz, for their help and consultations regarding Owl2Vec*. I would also like to thank Hang Dong for providing data resources used in the evaluation of this project. Finally, I extend a heartfelt 'thank you' to my family and friends for their continued support throughout this project and academic career.

Martin Tri Vien Lam

Abstract

The use of ontologies in various fields has increased in recent years due to the need for structured and organized knowledge representation. However, the creation and maintenance of ontologies can be a time-consuming and complex task. In this paper we review the recent literature on the use of machine learning techniques such as natural language processing (NLP) and Owl2Vec* for ontology summarization. We highlight the strengths and limitations of what these techniques provide. We also provide a proof-of-concept program, OntoSum, which have been tested using the medical terminology ontology SNOMED CT and excerpts from medical texts. Overall the use of machine learning techniques for ontology summarization shows great promise.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In recent years, the proliferation of knowledge representation and management systems has led to the construction of large-scale ontologies encompassing various domains. These ontologies consist of two main components: the terminology component (TBox), which defines the vocabulary and hierarchical relationships within the application domain, and the assertion component (ABox), which contains specific facts about the world of interest. However, as ontologies continue to expand in size and complexity, efficient maintenance and utilization become significant challenges [9].

The need for ontology summarization has emerged as a crucial task in managing and comprehending these extensive knowledge structures. Summarization aims to distill essential information from ontologies, enabling users to grasp the key concepts and relationships without being overwhelmed by the sheer volume of data. Take the ontology SNOMED $CT_{3.2}$, the Systematized Nomenclature of Medicine - Clinical Terms, as an example. The 2017 version of the ontology contains more than 1.5 million axioms in total. A new user who is unfamiliar with the knowledge base has the daunting task in front of them of figuring out what already is in the knowledge base and what is not. In a field such as medicine, which is constantly updated with new information it is important to keep state-of-art knowledge up to date.

In this context, the utilization of machine learning techniques, specifically natural language processing (NLP) approaches such as keyword extraction and Sentence BERT [40], has gained considerable attention.

Sentence BERT, is a powerful NLP technique that utilizes pre-trained models to generate sentence embeddings. By mapping sentences into a continuous vector space, Sentence BERT captures the semantic meaning and contextual information of the text. This technique enables efficient similarity calculations between sentences and facilitates information retrieval

tasks. In the context of ontology summarization, Sentence BERT can assist in identifying relevant and informative sentences within the ontology, which can then be used to construct concise summaries.

Ontology embedding, provides a means to represent ontology elements in a continuous vector space. By mapping ontology terms, concepts, and relationships into low-dimensional vector representations, embedding techniques enable computational models to capture the rich semantic meaning and relational structure of the ontology. These learned embeddings facilitate various downstream tasks, including ontology summarization, by enabling efficient similarity calculations and information retrieval. To this end we employ the use of Owl2Vec* [6].

The primary motivation behind ontology summarization is to address the challenge of maintaining and utilizing large ontologies. As these knowledge structures grow in size and complexity, manual inspection and management become laborious and time-consuming. Summarization techniques offer a practical solution by condensing the ontology into concise and representative subsets, preserving the most relevant and informative aspects while less relevant details.

This thesis explores the application of machine learning techniques, specifically NLP and ontology embedding, for ontology summarization. By leveraging these approaches, we aim to develop automated methods to extract key information from ontologies, facilitating efficient comprehension and management of these vast knowledge structures. The ideas in this thesis is built upon ontology modularity and extraction techniques.

A module of an ontology is a subset of an ontology which retains all semantic knowledge on terms specified by a user. A drawback of this is that modules may be as large as the original ontology in the worst case. Ontology excerpts aim to redress this problem by adding a size limitation to the subset. However, due to this size limitation the semantic knowledge retained in an excerpt will most likely be incomplete. Through an experimental proof-of-concept program and user surveys, we intend to demonstrate the effectiveness and usefulness of our proposed approaches in tackling the challenges posed by large-scale ontologies.

In the following chapters, we will delve into the related literature, present the methodology employed, describe the implementation details, and showcase the experimental results.

Part I

Background

Chapter 2

Preliminary Knowledge

This chapter aims to present and define concepts that are essential to understanding this project. These concepts include ontologies and machine learning techniques.

2.1 Ontologies

An Ontology is an explicit specification of conceptualizations. [25] Ontologies themselves may have a very generalized definition of a finite set of first-order sentences belonging to an ontology language, like The Web Ontology Language. [15] It defines a set of representational primitives to model a domain of knowledge. The primitives are typically classes, properties and relationships. Ontologies are used in knowledge representation, reasoning and information retrieval. Ontologies provide a shared vocabulary which can be used to structure information and facilitate communication among people, software agents and databases. One of the most widely used ontology languages is the Web Ontology Language (OWL), which is based on the Resource Description Framework (RDF) and is used to represent knowledge on the World Wide Web. Ontologies and knowledge graphs [2] (KG) are quite similar. They are both used for knowledge represenation and management, but some differences between the two include:

- Scope: Ontologies typically define a conceptual model of a particular domain, while KGs may encapsulate many domains.
- Use: Ontologies are often used in areas such as semantic web [39], knowledge management and natural language processing. KGs on the other hand are used for search engines, recommendation systems and data analytics.

The ontology used for experiments in this thesis is the SNOMED CT ontology described further in section 3.2

Description Logic

Description Logic (DL), or Description Logics is a formal knowledge representation language based on a set-theoretical foundation [53]. DL is used to represent the knowledge of an application domain. It is particularly useful for providing logical formalism for ontologies and the Semantic Web: the Web Ontology Language (OWL) [24].

Axioms

An axiom is a statement that is taken to be true within a particular ontology. Axioms provide the foundation for reasoning within an ontology and are used to specify the relationships between different entities and concepts within an ontology. Axioms are used to define properties, classes and relationships that make up an ontology.

Example 2.1.1 *An example of an axiom in an ontology about animals could be:* $\mathit{Mammal} \sqsubseteq \mathit{Animal}$.

2.1.1 Ontology Modularity

Ontology modularization [14] is a research field of how ontologies can either be split into modules, built by gathering modules, or how a modules can be extracted. A module is a subset of an ontology that can act as a substitute for the ontology in certain contexts. The goal of ontology modularity is to improve the flexibility, scalability and maintainability of an ontology. Modularity allows different parts of an ontology to be addressed and appraised separately, facilitating easier development and comprehension of the ontology. An ontology module aims to keep as much semantic knowledge as possible while being as small as possible. As the modules aim to retain all semantic knowledge their size can not be predetermined. An important modularization principle to note is the concept of locality based modules. [7, 23] Locality based modules aim to preserve the entailed information over a signature. The idea behind a locality-based module is to isolate and encapsulate a coherent portion of an ontology that can be reused independently.

2.1.2 Ontology Excerpts

Ontology excerpts refers to the ability to decompose a large ontology into smaller, more manageable modules or sub-ontologies that can be developed and maintained independently. The paper Towards Extracting Ontology Excerpts [9] defines an ontology excerpt as:

> Definition 3 (Ontology Excerpt). Let $\mathcal{O}$ be an ontology and let $k > 0$ be a natural number. A k-excerpt of $\mathcal{O}$ is a subset $\mathcal{E} \subseteq \mathcal{O}$ consisting of k axioms, i.e. $|\mathcal{E}| = k$.

An ontology excerpt is a part or section of an ontology. An ontology excerpt contains a subset of the terms, concepts, and relationships that are defined in the larger ontology. The purpose of an ontology excerpt is to provide a more focused view of the ontology, highlighting the key concepts and relationships that are relevant to a particular application or task.

To quantify the meaning of an excerpt the authors introduce the metric μ. μ is an *incompleteness measure* where the lower the number, the more semantical meaning and knowledge is preserved for signature elements in Σ.

The authors introduce two extraction techniques. One is based on the simple intuition that axioms comprising more elements from Σ should be preferred to be included in an excerpt for Σ. The other approach was inspired by ideas from information retrieval [35]. In both techniques a similarity between axioms with regards to a signature Σ is measured such that the selected k most similar axioms result in a k-excerpt.

In their closing evaluation the technique based on information retrieval outperformed the other approach. Though it was flawed when $k < 19$.

2.2 Machine Learning Techniques

This section will describe machine learning techniques that are relevant to this project.

2.2.1 One-Hot Encoding

One-hot encoding is one of the simplest forms of encoding categorical data. Categorical data is data that can be separated into different groups. One-hot encoding's most relevant use-case is to encode ordinal and nominal categorical variables.

- Ordinal categorical variables are variables which follow a natural order such as levels of education such as: kindergarten, middle school, high school and university.
- Nominal categorical variables are the opposite of ordinal variables in that they do *not* follow a natural order for example categories like "dog" and "cat".

In this thesis one-hot encoding has been used to encode sentences and axioms, as nominal variables, into vectors on the form of an 1 x N matrix where N is the number of unique words in a vocabulary. Each index in matrix represents a word and if a sentence includes a word, the index in the matrix is a 1. If a word does not appear the value on the given index is a 0. [16] Advantages of one-hot encoding includes eliminating orders of importance between categories. At the start of this project this was an important attribute as it helped simplify the encoding of axioms. It was also helpful later to be able to refer to indecies backwards to find which words they referred to replace values at a later point.

2.2.2 K-Nearest-Neighbours

K-Nearest-Neighbours [27] is a supervised machine learning algorithm used for classification tasks. It is a simple algorithm which stores all available cases and classifies new cases using similarity measures, like Cosine Similarity.

Definition 2.2.1 *The formula for cosine-similarit is given by:*

$$cosine - similarity(v, u) = \frac{v \cdot u}{\|v\| \, \|u\|}$$

Where v and u are two vectors.

The KNN algorithm works as follows: given a new observation, the algorithm finds the k closest training examples in the feature space and assigns the class label of the majority of the k neighbors to the new observation. The choice of k determines the number of neighbors considered for classification.

The advantages of KNN comes with it's simplicity. It is an easy algorithm to interpret and implement, but this is also a disadvantage. This is because it stores all available cases and requires a large amount of memory to store them. KNN will also perform poorly if cases of different classes or categories are very intermingled in vector space and not easily separated.

2.2.3 Natural Language Processing

Natural Language Processing (NLP) is a vast subject. In this thesis we will focus on the tokenization and text summarization tasks NLP is used for. Tokenization [29] in the context of NLP is the process of breaking down text into smaller units, usually words, that are called tokens. Tokens are not only limited to words, they can also be sentences or sequences of special characters. The tokenization used in this thesis is word-based tokenization where words are separated on white space and punctuation. After tokenization, the resulting tokens are often normalized to reduce the number of unique tokens in a text. Normalization can involve converting all tokens to lowercase or removing stop words (i.e., common words such as "the" or "and" which provide no semantic value). Overall, tokenization is a crucial step in NLP as it transforms raw text into a format that can be easily processed by machine learning models and other NLP algorithms.

The other NLP task with relevance to this thesis is text summarization. For this task this project uses the tool YAKE! [5] for keyword extraction. YAKE! uses statistical techniques such as word frequency, word co-occurrence, and part-of-speech tagging to identify the most relevant candidates. Specifically, YAKE! identifies candidate keywords as single or multi-word phrases that consist of one or more content words, such as nouns or adjectives.

After candidate selection, YAKE! scores each candidate keyword based on its relevance to the text. YAKE! calculates a score for each candidate keyword using a statistical algorithm that considers the frequency, location, and context of the candidate keyword in the text. The scoring algorithm is designed to give higher scores to candidate keywords that are more relevant to the overall content of the text.

Overall, YAKE! is a simple and effective keyword extraction tool that can be used for a wide range of applications, including information retrieval, content analysis, and search engine optimization.

2.2.4 Gold and Silver Standards

A gold standard is a reference that is used to measure the performance of machine learning models. [12] It is often created by human experts who manually annotate and label data. A gold standard is the best possible performance a machine could achieve given a task. It is in important tool to ascertain the accuracy and reliability of machine learning model.

In NLP a gold corpus is a collection of manually annotated text that

serves as a reference for evaluating the performance an NLP system.

Conversely a silver corpus is a dynamically created reference used for evaluation. Silver corpora are typically less expensive and time-consuming to create than gold corpora, since they rely on automated annotation rather than manual annotation. However, their accuracy is typically lower than that of gold corpora, and they may require more manual review and correction to be useful for natural language processing tasks.

Chapter 3

State-of-Art

This chapter aims to describe some state-of-art methods that have been researched on the topics this thesis aims to discuss. These methods include embedding of: knowledge graphs, and by extension ontologies, the knowledge base SNOMED CT, entity linking and entity summarization.

3.1 Knowledge Graph Embedding

Though not a lot of research has been done on *ontology* embeddings. A lot has been done on knowledge graph and word embeddings [20–22]. Most of the ontology embeddings will be based on these. A knowledge base can be defined as a set of entities, a set of relations and a set of literals. An entity is anything a person can think of: Places, people, organizations or objects. A relation is a property that connect or relate two entities. Literals are immutable values integers or strings.

Embedding

The most basic embeddings map facts on the form $< s, p, o >$ in a way where $\vec{s} + \vec{p} \approx \vec{o}$. An example of this would be: if we know $<$ `Norway,hasCapital,Oslo` $>$ then the resulting vector of $\vec{Norway} + \vec{hasCapital}$ should be as close to $\vec{Oslo}$ as possible.
This approach has multiple benefits:

- Because all entities are vectors applying machine learning methods, like classification algorithms, become intuitive.
- The vectors themselves are typically low in dimension; because of this the run time when using or applying computations on them will not be too long.

- Similar entities are expectedly grouped together.
- Lastly we may compute additional information by for example computing $\vec{Norway} + \vec{hasCapital}$ and expect that the resulting vector will lead us to the capital of Norway, if the embedding is well made.

Knowledge bases are generally referred to as knowledge graphs because literals generally are abstracted away. Further they usually do not deal with classes or axioms. What is left are entities as nodes, connected together by relations as edges.

3.1.1 Geometric Embedding Models

TransE

TransE [4] is an embedding model that uses geometric distance to judge similarity between entities. It is implemented using a neural network and 1-hot encoding [16]. TransE is not able to model symmetry, anti-symmetry or reflexivity in relations. It will also struggle to embed 1 to n relations.

Example 3.1.1 *The two triples* $<$ `ElonMusk`, `founderOf`, `Tesla` $>$ *and* $<$ `ElonMusk`, `founderOf`, `SpaceX` $>$. *TransE would give similar embeddings to the entities* `Tesla` *and* `SpaceX`.

RotatE

RotatE [52] aims to be geared towards symmetric, anti symmetric, inverse and compositional relations. Symmetric relations include `marriedTo` as an example. Compositional relations may refer to familial relations like grandparent.

Example 3.1.2 *Example of a compositional relationship: If* $< x, \texttt{parentOf}, y >$, $< y, \texttt{parentOf}, z >$ *then* $< x, \texttt{grandparentOf}, z >$.

RotatE defines these relational patterns as rotations from the head- to the tail entity in a vector space. For a triple $< x, r, y >$ we want the vectors defined as $\vec{y} \approx x \circ r$ where $\circ$ is an element-wise product operation. A relation is symmetric if it belongs to $\{-1, +1\}^D$ *(A coordinate wise rotation of 0 or π radions). A relation r_3 is the composition of r_1 and r_2 if and only if $r_3 = r_1 \circ r_2$.

Other geometric embedding models include TransR [33], TransD [28] and HAKE [61]

3.1.2 Word Embedding

Like knowledge base embedding, word embedding aims to embed words into vectors which machines can operate on. Though knowledge base embeddings works with entities and relations, word embeddings, as the name suggests, aims to embed individual words based on semantic characteristics. Characteristics like definitions, context and relationships to other words. Word2Vec is a natural language processing technique which is very effective because of its ability to group together similar words.

Word2Vec

Word2Vec [36] creates a model using a corpus of text and a neural network network model to embed words into vectors. Once the model has been trained it should be able to detect synonyms and suggest words for an unfinished sentence. It trains the model by using techniques such as CBW (continuous bag of words) and the continuous skip gram model [37]. The continuous skip gram model, very simplified, tries to predict where certain words should end up in a sentence based on a corpus. It predicts the position of a word based on what words come before and after it.

BERT

BERT [18], which stands for Bidirectional Encoder Representations from Transformers, is a language representation model developed by Google. The model is based on transformers, but with an added change of bidirectional training. This change is significant in that previous language models only trained using left-to-right or right-to-left training. A transformer can be separated into two stages: an encoder that reads a text input and a decoder which outputs a prediction. Directional model's, previous models which used left-to-right or right-to-left training, encoders read one input word at a time sequentially. BERT's bidirectional encoder reads an entire sequence of words at once. This quirk allows the model to learn the context of words based on all of its surroundings. The training in BERT is done by applying two techniques: masked language modelling (MLM) and next sentence prediction (NSP). Before feeding the input to the encoder MLM masks one word of the sequence and the decoder will try to predict it, based on the other non-masked words in the sequence. Around 15% of the words all inputs are masked. Again during the input phase, BERT receives two sentences. 50% of the time the second sentence follows the first. The model will then predict whether the sentence actually follows or is a random sen-

tence from another sequence. The prediction is based on predicting the labels (isNext or notNext) placed on the inputs.

3.1.3 Ontology Embedding

Some early ontology embedding algorithms treated axioms as sentences. This meant they could not grasp the full context each axiom. Whether that was the correlation or logical relations between them. Other algorithms transformed logical relations into geometric relations like TransE. Encoding the logical and semantic constructors, but ignoring the additional lexical information the ontology provided. The algorithms EL embedding [30] and Quantum Embedding [20] could not fully explore the ontologies because they only considered `rdfs:subClassOf` and `rdf:type` as edges when traversing the ontology. Though these algorithms fail to capture *some* lexical information they do manage the embed ontologies into vector space.

Owl2Vec*

Owl2Vec* [6] is an ontology embedding algorithm which consists of two core strategies: corpus extraction and word embedding model training using the extracted corpus. The corpus is comprised of three "documents". They call the documents: Structure, Lexical and Combined documents. The first two documents collect the graph structure, logical constructors and the lexical information in the ontology. The third document aims to preserve the correlation between entities and lexical labels (`rdfs:label`). Because Owl2Vec* is a program written in python it was a suitable tool for this project's needs. Other python based ontology managers include Funowl [46, 47] and Owlready2 [32]. Owl2Vec* was chosen for it's embedding functionality. Owl2Vec* works by representing the ontology in the form of a KG, where entities are represented as nodes and relations as edges. The library then applies the skip-gram model from word2vec [11] to this knowledge graph, learning vector representations of each entity and relation in the ontology. These vector representations can then be used for various downstream tasks, such as ontology alignment, similarity-based reasoning, and link prediction.

The learning process in Owl2Vec* involves optimizing a cost function that aims to maximize the probability of observing the co-occurrence patterns of entities and relations in the knowledge graph. The optimization is performed using stochastic gradient descent, and the resulting vectors can be tuned for specific tasks by adjusting hyperparameters such as the

vector dimensionality, learning rate, and window size.

Overall, Owl2Vec* is a powerful tool for learning distributed representations of entities and relations in ontologies.

3.2 SNOMED CT

Also known as SNOMED CT stands for Systematized Nomenclature of Medicine and Clinical Terms. [50] SNOMED International is a not-for-profit organization that owns, administers and develops SNOMED CT, the world's most comprehensive clinical terminology. [45] SNOMED is maintained and developed by the International Health Terminology Standards Development Organisation (IHTSDO). It is used in electronic health records (EHRs), clinical decision support systems, and other healthcare information systems worldwide to support patient care, quality improvement, research, and public health reporting. SNOMED includes over 350,000 clinical concepts, organized into hierarchies and cross-referenced with other terminologies and classification systems.

SNOMED CT Ontology statistics	
Axiom	1,533,738
Logical axiom count	332,288
Declaration axioms count	332,293
Class count	332,209
Object property count	81
Annotation Property count	4
SubClassOf	250,131
EquivalentClasses	82,077
SubObjectPropertyOf	80
AnnotationAssertion	869,157

Table 3.1: A table showing the statistics of data in the SNOMED CT (2017) ontology.

3.3 Entity Linking

Entity linking with a knowledge base [43] is a type of entity linking that involves linking named entities in text to their corresponding entities in a knowledge base, such as a semantic web ontology, a database, or a KG. The entity linking process involves three main steps. First, the system generates a candidate entity set for each entity mention, which contains possible entities that the mention may refer to. This is done using

various techniques such as name dictionary-based methods or surface form expansion. Second, the system ranks the candidate entities to find the most likely link for each mention. This involves using different types of evidence such as supervised or unsupervised ranking methods. Finally, the system predicts unlinkable mentions by validating whether the top-ranked entity identified in the candidate entity ranking module is the correct target entity for the mention. If not, the system returns NIL for that mention. Current state of art entity linking methods include: LINDEN [44] is a framework to link named entities in text with YAGO [51], a knowledge base unifying Wikipedia and WordNet. LINDEN uses the rich semantic knowledge found in YAGO and Wikipedia to link entities and generate good results. ZenCrowd [17] is a probabilistic framework which uses both automatic techniques and punctual human intelligence feedback captured on a crowdsourcing platform. ZenCrowd provides a reliable approach to entity linking which exploits the trade-off between largescale automatic entity linking and high-quality human annotations.

3.4 Entity Summarization

Entity summarization [48] is the task of generating a concise summary of information related to a specific entity or set of entities. The goal of entity summarization is to provide a comprehensive overview of the key facts and attributes associated with an entity, while reducing redundancy and irrelevant information.

Entity summarization can be applied to various types of entities, such as people, organizations, products, or events. The summary can include various types of information, such as basic facts (e.g., name, date of birth, location), relationships to other entities (e.g., family members, business partners), accomplishments or contributions (e.g., awards, patents, publications), or opinions or sentiments expressed about the entity (e.g., in social media or news articles). Entity summarization has various applications in fields such as information retrieval, knowledge management, or data analytics. It can be used, for example, to provide a quick overview of a person or organization for a search engine result page, to extract insights from large datasets containing information about entities, or to facilitate decision-making in fields such as finance, marketing, or healthcare. Current state of art entity summarization methods include: RELIN [10], which aims to identify and extract key information about entities from large amounts of textual data. The

approach leverages relatedness and informativeness-based centrality to determine the importance of entities and their relationships within a given context. REMES [26], which summarize facts by analyzing the relatedness in a collection of entities based as opposed to summarizing each entity in isolation. Specifically, information is generated by inter-entity facts that are similar and intra-entity facts that are important and diverse. etc.

Chapter 4

Understanding Ontology Summarization

This chapter will describe the problem the project aims to address and the methodology behind it.

4.1 Describing the Project

As explained in the introduction the project entails researching methods and algorithms to determine an optimal way to summarize ontologies. Jieying Chen describes an ontology excerpt as a subset of axioms of a larger ontology[8]. (the subset may also be the whole ontology) More specifically an ontology excerpt is a subset of an ontology, which consisting of axioms that are related to a given set of terms. However, for a given set of terms there may exist an infinite number of summarizations. To determine which summarization best portrays an ontology some kind of measure must be used to rank summarizations. In this case Chen introduces two concepts: Logical Difference and Primitive Witnesses. Logical difference is defined as the difference between two TBoxes where T_1 entails all concepts that are included in the signature. On the contrary, T_2 does not entail any concepts that are found in the signature. If the two TBoxes are logically different the set consists of infinitely many concept inclusions. To only consider the concepts of interest they introduce the Primitive Witnesses theorem. Primitive witnesses are concepts that appear in the signature and occur on either the left or right side of concept inclusions in a set of logical difference. They use the set of primitive witnesses as a way to briefly and clearly express the logical difference. This method has been proven to quite efficient as they ran a test with an ontology consisting of more than 300 000

axioms where more than 3000 signatures were evaluated. The excerpts they computed were computed in a run time of around 150ms on average.

For the optimization part of the problem we want to use machine learning methods to create a model that can determine what "the best" ontology summarization is. To do this we will rely on previous knowledge of ontology embeddings. Jiaoyan Chen published an article in 2021 describing how his team embedded ontologies into vectors in a program they called Owl2Vec* [6]. Owl2Vec* used built upon previous research done on knowledge graph embeddings. Owl2Vec* uses two strategies to transform an Ontology O to a graph G.

1. Transform according to RDF graph mapping. Which means to exchange the owl ontology for RDF triples according to the W3 rules.

2. The second strategy being based on projection. Every axiom can be broken down to triples on the form $< X, r, Y >$.

Both ontology transformation strategies are compatible with OWL entailment reasoners which can compute the ABox realization and TBox classification before the transformation occurs. This ensures that no implicit knowledge is lost upon transforming the data. During Owl2Vec*'s training phase three "documents" are created. A structural, lexical and combined document. The structural document is used to capture the general structure of the RDF graph an the logical structure of the ontology. This is done by transforming the graph, G, into a single directed graph. For each triple on the form $< X, r, Y >$ in G, the subject X, the object Y and the relation r, they add a vertex between X and r and Y and r. The lexical document aims to capture non-logical information about the axioms; like names of entities. The document includes two kinds of sentences: The first are generated from IRI sentences in the structural document while the second are extracted from textual annotations from the ontology. The first sentences are found by replacing entities with the value of their rdfs:label value. Sometimes entities do not have an rdfs:label triple. In that case we use parse the name part of an IRI. En example of this could be vc:MilkAndYoghurt would be parsed into “milk”, “and” and “yogurt”. However, sometimes entities have more annotations that rdfs:label, like rdfs:comment or rdfs:seeAlso. The second sentence aims to extract this information to find out whether entities of different names may be interpreted as the same. For example the model should be able to tell that "edamame" is the same as "soybean". Lastly the combined document aims to capture the correlation between IRIs and words. Owl2Vec* then combines the three documents to train a lan-

guage model using word2vec. Word2vec is a natural language processing technique which embeds words into real numbers. The combined document lets us find the correlation between ontology entities in vector space.

4.2 Methodology

Figure 4.1 represents the overall framework of the program OntoSum. It can be broken into 3 parts: (i) Reading the ontology into Owl2Vec*, (ii) transforming the input text into a signature and (iii) using the signature to rank all axioms and extract the k most relevant.

4.2.1 The Inputs

The inputs of program are as follows:

- Text, natural language text that the user writes and is the prompt for which axioms the program will extract. The text may be single sentences like what one usually write as a prompt to a search engine, or short excerpts from papers that have some commonality with the ontology.
- Ontology, the ontology that one would want a summarization of.
- K, a natural number of how many axioms the summarization should consist of.

4.2.2 Reading the Ontology into Python

The first draft was created by recreating a program Jieying Chen wrote in Java, in python. The program Chen wrote computed the k most relevant axioms in an ontology based on a user inputted signature. As we want to expand the research of this topic by using machine learning techniques using a language like python, which has more support for data science and manipulation, is more beneficial.

The remade program used Owl2Vec* to read an ontology and output all* axioms in an easy to parse format, Manchester Syntax. The program then maps each entity into a vector. All axioms are initially a vector of zeros. If an entity appears in an axiom we assign the number 1 in the vector for a set index that matches the entity. Otherwise the zero is replaced by a value calculated by the average distance from each ontology entity to each signature entity. We can then compare all vectors to the signature vector using a cosine similarity function to select the k "best" axioms to be a part

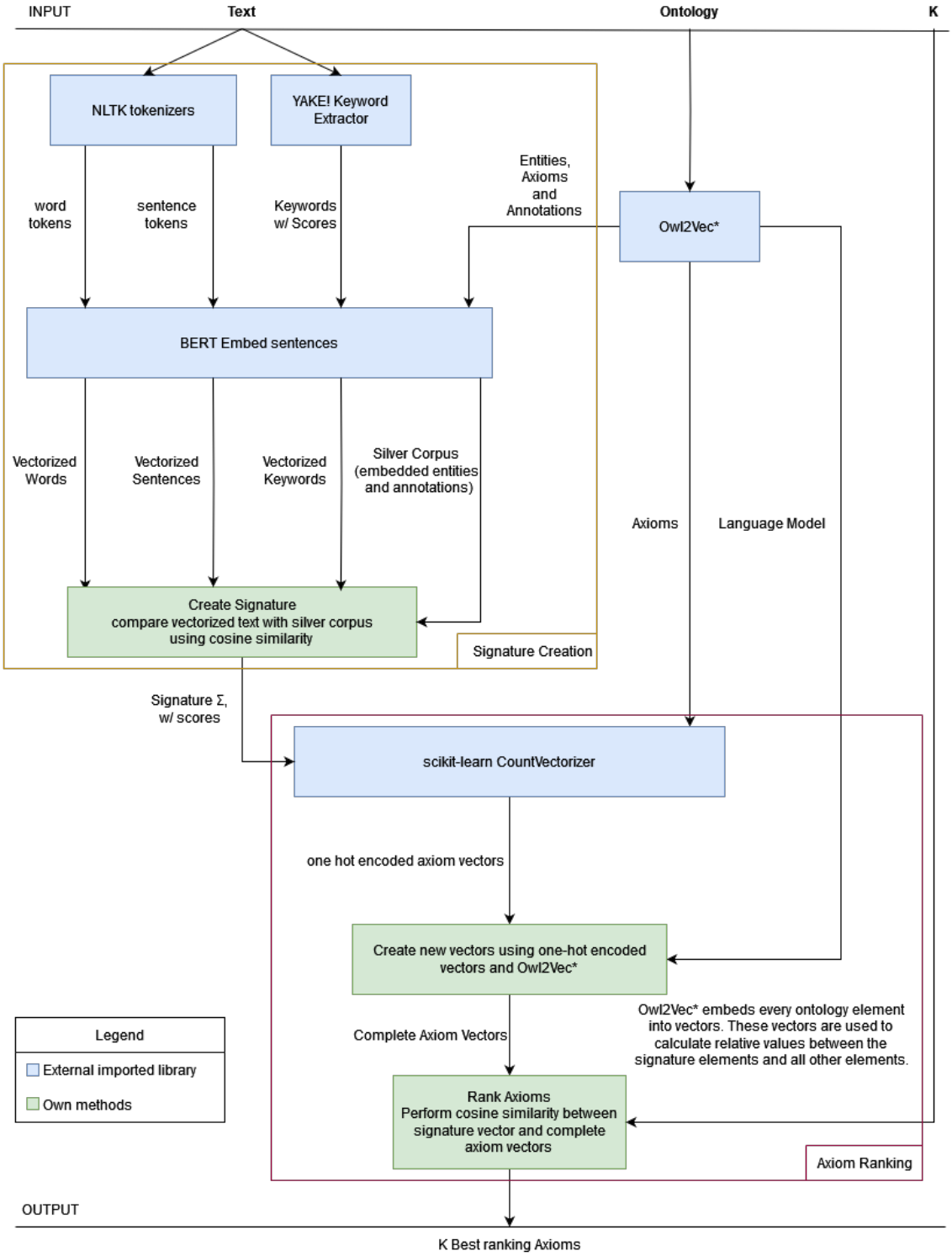

Figure 4.1: Pipeline of program

of the summarization. The problem with this approach is that it does not account for existential or universal restrictions. And it does not account for Intersection or Union. Before managing to create a functional program I tried using various ontology python libraries. FunOwl being one of them. Unfortunatley FunOwl was not a library suited for this kind of task as it formatted axioms in a way that was hard to parse and work with. Further it was mostly used to create and edit ontologies in python using a functional language. Using it for data science was not a good idea.

Another library that was considered was OwlReady2. OwlReady2, like FunOwl, had the primary function of creating and managing ontologies using python. Where they differ is OwlReady2's focus on classes. Most likely through the principle of object oriented programming. Every interface in this library goes through the owl classes and it is not ideal to work with. So I ended up using Owl2Vec*. Owl2Vec* tries to imitate the functionality the java OWL API offered using OwlReady2. Where Owl2Vec* succeeds is providing an interface to access axioms.

*When referring to *all* axioms, Owl2Vec* does not actually include every single available axiom in it's output. This is because the authors decided that most atomic axioms would be implied by the included composite axioms.

4.2.3 Creating the Signature

First two definitions of what a signature is:

Definition 4.2.1 *$sig(\mathcal{O})$ is a function which returns a set of all elements. Elements are concepts (Classes), relations and attributes. The algorithm may use $sig(\Sigma)$ to get all the elements from a signature too.*

Definition 4.2.2 *A signature Σ is a set of elements created from an input described by a user. The signature is then the most relevant elements found in an ontology based on the input. $sig(\Sigma) \subseteq sig(\mathcal{O})$*

In Chen's initial Java program, the input signature was a file containing IRIs of concepts (OWL classes) and roles (OWL object properties).

Example 4.2.1 *An example of how the signature looked:*

```
Concept[
http://snomed.info/id/882784691000119100
http://snomed.info/id/880529761000119102
http://snomed.info/id/674814021000119106
```

```
]
Roles[
http://snomed.info/id/246090004
]
```

As we decided to improve upon this program, the input should be more user friendly and not require knowledge of IRIs or entity IDs. That is why we decided the input should be text in natural language.

The problem with a natural language input is that the program now needs to relate the text to ontology elements. Ontologies are used to represent real world objects or concepts. If the IRI of an ontology element does not describe the real world concept it is supposed to describe, it's annotations and axioms will. Using sentence-BERT [41] to encode the entities' class names and annotations, we have created a silver corpus. The reason we use a silver corpus instead of a gold corpus is because we lack the ideal manually annotated gold corpora for what ontology elements should or should not be selected for every available ontology. It is therefore more reasonable to dynamically create a silver corpus that we use to find the most relevant ontology elements. After using sentence-BERT to encode a silver corpus, we begin analyzing the input text. When analzying the input and finding the most relevant ontology elements two problems arose: (i) The generated signature was too bloated from having words with little relevance in input be matched with ontology elements. (ii) being too picky when selecting matching words resulted in an empty signature.

Problem (i) arose when doing a naïve approach where the input was split into word and sentence tokens using nltk tokenizers [34]. Each token would be encoded using sentence-BERT and compared to the silver corpus using cosine similarity as a measure. A more narrow approach employed the use of Yet Another Keyword Extractor, also known as "YAKE!" [5]. Yake would analyze the input and extract the most relevant keywords and phrases from the input and give them a term score. The same similarity measure would then be applied: encode the keywords and phrases using sentence-BERT, then perform cosine similarity between the encoded vector and silver corpus to find the most relevant ontology element. This set of keywords and phrases would be smaller than the set of tokens discussed previously and would give a more relevant signature. Due to the smaller set size problem (ii) arose.

Knowing this the function looks like this:

Algorithm 1 Create a signature Σ from an input i

Input: a text i, an ontology $\mathcal{O}$ and a threshold μ
Output: a set of elements Σ which contains the most relevant elements in the ontology $\mathcal{O}$ based on the input i

```
1:  function CREATE-SIGNATURE(i,O ,μ)
2:      Σ := {}
3:      keywords := EXTRACT-KEYWORDS(i)
4:      for word ∈ keywords do
5:          V_word := CONVERT-WORD-TO-VEC(word)
6:          B := GET-BEST-ONTOLOGY-ELEMENT(O ,V_word)
7:          if B.score > μ then
8:              B.weight := 1 - GET-WORD-SCORE(word)
9:              Σ := Σ ∪ {B}
10:         end if
11:     end for
12:     if Σ = {} then
13:         U := GET-WORDS-AND-SENTENCES(i)
14:         for u ∈ U do
15:             V_u := CONVERT-WORD-TO-VEC(u)
16:             B := GET-BEST-ONTOLOGY-ELEMENT(O ,V_u)
17:             if B.score > μ then
18:                 B.weight := 0
19:                 Σ := Σ ∪ {B}
20:             end if
21:         end for
22:     end if
23:     Σ := EXTEND-SIGNATURE(Σ ,O)
24:     return Σ
25: end function
```

Algorithm 2 Extend a signature Σ with an ontology $\mathcal{O}$

Input: a signature Σ and an ontology $\mathcal{O}$
Output: Σ extended with ontology elements found on the right side of axioms

```
1:  function EXTEND-SIGNATURE(Σ ,O)
2:      for e ∈ Σ do
3:          for A ∈ O.Axioms do
4:              if e appears on the left side of A then
5:                  for a ∈ right side of A do
6:                      B_a.weight := -0.05
7:                      Σ := Σ ∪ a
8:                  end for
9:              end if
10:         end for
11:     end for
12:     return Σ
13: end function
```

The function first extracts the keywords and phrases of the input using a keyword extractor, in this case YAKE! [5] as shown in in line 3 of algorithm 1. Each element in keywords is embedded into a vector using sentence-BERT [41] (line 5 of algorithm 1) so the algorithm can perform a cosine similarity comparison to the embedded $\mathcal{O}$ to find the most relevant elements in the ontology. This is abstracted by the function CONVERT-WORD-TO-VEC. The embedded $\mathcal{O}$ is the embedded silver corpus described in section 4.2.3. The highest ranking element is then added to the signature *if* it has a cosine-similarity score above μ. (lines 7-9 in algorithm 1) The threshold μ is implemented as 0.8. It is a number that was chosen because 0.7 resulted in too many empty signatures when testing on various inputs and ontologies. The algorithm then adds scoring modifiers, or weights, to the signature elements. The weight of the signature elements selected from the set of keywords is calculated by: 1 subtracted by a number generated by YAKE!. YAKE! gives each keyword a word score between 0 and 1. The closer to 0 it is the more relevant to the input text the word is. To keep a consistent value interpretation I subtract it from 1 so the values closer to 1 means they are more important.

In some cases the extracted keywords do not match to any ontology elements. In these cases the algorithm will continue a more granular search using nltk-tokenizers [34]. (lines 12-22 in algorithm 1) The input is divided into a set, $\mathcal{U}$, with every individual word and every sentence. The encoding

and comparisons of vectors are the same as with the keywords. However, the elements added by this granular extraction have a weight of 0. In the rest of this thesis, *granular signature*, will be used when referring to this part of the signature creation.

Lastly, to improve the relevancy of axioms to the signature the algorithm goes through each signature element it has compiled and sees if the element appears on the left hand side of an axiom (Algorithm 2). If it does, every other ontology element that appears in the axiom is added to the signature. These will have a weight of -0.05 to represent that axioms that include these elements, though relevant to the signature will have a scoring penalty of 5% deducted in the final ranking. In the rest of this thesis, *extended signature*, will be used when referring to this part of the signature creation.

This extended signature can be defined as:

Definition 4.2.3 *Given the creation of a signature Σ , for every element $e \in sig(\Sigma)$ at the moment. Check if e appears on the left hand side of any axioms $\mathcal{A}$ in the ontology $\mathcal{A}$. If e appears, add every element that appears on the right hand side in the axiom to Σ . If e does not appear on the left hand side, go to the next axiom.*

4.2.4 Ranking the Axioms

Having created a signature Σ we now had which ontology elements were important. Now we needed to encode the axioms and signature into data-structures that a machine could run comparisons on. To do this we employ the technique of one-hot encoding [16]. To do this, every ontology $\mathcal{O}$ has a strict order for each element in $sig(\mathcal{O})$. This means that when the program wants to access elements in the ontology it may refer to elements with the index i. When the algorithm vectorizes an axiom it starts with a vector of 0s, then for every ontology element that appears in the axiom, sets the appropriate index to 1.

Example 4.2.2 *An example of how the axioms are initially one-hot encoded. Consider the following:* Σ *= {A,B,D} and the Axioms = {* $A \sqcap B \sqsubseteq C$, $D \sqsubseteq A \sqcap E$ *}*
After running one-hot encoding on the axioms with the ontology as a vocabulary we have the vectors:
$\mathcal{V}_\Sigma$ *= [1,1,0,1,0]* $\mathcal{V}_1$ *= [1,1,1,0,0]* $\mathcal{V}_2$ *= [1,0,0,1,1] Where the first index refers to the ontology element A, the second index to B, the thrid to C and so on.*

With a signature- and axiom-vectors, we can perform cosine similarity between them and find which axiom is more relevant for the signature. In the case of example 4.2.2, the cosine similarity between $\mathcal{V}_\Sigma$ and $\mathcal{V}_1$ is 0.667 while $\mathcal{V}_\Sigma$ and $\mathcal{V}_2$ is also 0.667. This tells us the axioms are equally relevant to the signature, as both axioms only have 1 ontology element different that is not included in the signature itself. To differentiate ontology elements more from each other, we employ Owl2Vec*'s ontology element embeddings. We want the value 1 to represent signature elements as they are the most important. In Owl2Vec* each ontology element has it's own associated vector. We use these vectors to create values that are between 0 and 1 for the non-signature ontology elements that appear in an axiom. For the remainder of this thesis we will refer to this value as a *relative value*, $\mathcal{R}_\urcorner$, where e is the respective ontology element the relative value is associated with. The value is calculated using the average cosine distance of the signature elements' and ontology element's Owl2Vec* vectors. For the remainder of this thesis we will refer to these vectors by OV_e where e, where e is the respective ontology element the vector is associated with.

Example 4.2.3 *An example of how a relative value is calculated can be given by: Consider that* Σ *= {A,B,D} and the following Axioms = {* $A \sqcap B \sqsubseteq C$ *} The relative value is then calculated this formula*

$$\mathcal{R}_\mathcal{B} = \frac{\sum_{i=1}^{sig(\Sigma)} cosine - sim(OV_{\Sigma_i}, OV_B)}{\|sig(\Sigma)\|}$$

Having calculated $\mathcal{R}$ and replaced instances of non-signature element 1 values. The program can perform a cosine similarity calculation on the edited vectors to get a *cosine rank* for each axiom. The last step in the ranking process is to select the k highest ranking axioms. Those being the k axioms with a cosine rank closest to 1.

As stated at the beginning of the section, the summarization of the ontology starts with creating a signature. The algorithm then converts every axiom in the ontology into a vector and collects it in the set $\mathcal{V}_\mathcal{A}$. Before the ranking the algorithm also converts Σ into the vector $\mathcal{V}_\Sigma$.

The final ranking of axioms is once again done using cosine-similarity. However recall when we created the signature in Algorithm 1, each if an axiom included signature elements, the score would be weighted. This is applied through multiplying the cosine-similarity with the weights we get from GET-WEIGHT. GET-WEIGHT will go through each element in the axiom and fetch the weight. If the element is a keyword signature element $\mathcal{B}$.weight is multiplied with the score. If the element is not a

Algorithm 3 Rank axiom vectors, $\mathcal{V}_{\mathcal{A}}$, based on similarity to a signature vector,$\mathcal{V}_{\Sigma}$, then select the top k

Input: a set of axiom vectors $\mathcal{V}_{\mathcal{A}}$, a signature vector $\mathcal{V}_{\Sigma}$ and an integer k where k > 0
Output: a set of axioms $\mathcal{M}$ where $\mathcal{M} \in \mathcal{O}$ and $|\mathcal{M}| = k$

```
1:  function RANK-AXIOMS(V_A,V_Σ,k)
2:      S := []
3:      i := 0
4:      for V_a ∈ V_A do
5:          score := COSINE-SIMILARITY(V_Σ,V_a)
6:          score := score * GET-WEIGHT(a)
7:          S[i] := (s,a)
8:          i := i + 1
9:      end for
10:     ranking = SORT(S)
11:     M = SELECT-K-BEST(ranking,k)
12:     return M
13: end function
```

keyword signature, the weight is 1, which means the score is unaffected by this element. If the element was a right side entity that was included by a keyword element, the score is penalized by -5%. The score is then associated with the proper axiom in a tuple and collected in the list $\mathcal{S}$. The collection $\mathcal{S}$ is then sorted by the scores. the algorithm then selects the k-highest ranking axioms through SELECT-K-BEST.

Describing Vectors

The program works under the assumption that every ontology $\mathcal{O}$ has a strict order for each element in $sig(\mathcal{O})$. This means that when the program wants to access elements in the ontology it may refer elements with index i. Every element e in an ontology also has a corresponding vector called V_e. The reason for describing each axiom as a vector is to perform a similarity calculation to rank them.

Definition 4.2.4 *Given a signature $\Sigma \subseteq \mathcal{O}$, the algorithm can create a corresponding signature vector $\vec{\Sigma} = [v_1, v_2, ..., v_n]$ where $v_i = 1$ if the corresponding i-th element $e_i \in \Sigma$ otherwise $v_i = 0$.*

Definition 4.2.5 *Similarly, for every axiom $\alpha \in \mathcal{O}$, there is a corresponding $\mathcal{V}_\alpha$. In an axiom vector if $e_i \in \Sigma$ and $e_i \in \alpha$ then $v_i = 1$ otherwise if only $e_i \in \alpha$ then the algorithm calculates a relative value by calculating the average cosine distance between the vectors of the elements found in the signature and V_e. If the element does not appear in the axiom at all $v_i = 0$.*

Algorithm 4 Generate a Vector $\mathcal{V}_\alpha$ from an axiom α based on a signature Σ and an ontology $\mathcal{O}$

Input: an axiom α , a signature Σ , and an ontology $\mathcal{O}$
Output: a vector $\mathcal{V}_\alpha$ where $\|\mathcal{V}_\alpha\| = \|sig(\mathcal{O})\|$ and each element v_i is between [0,1]

```
1:  function CONVERT-AXIOM-TO-VECTOR(α ,Σ ,O)
2:      V_α := [0, 0, ..., 0]
3:      n := 0
4:      for e ∈ O do                          ▷ for every element in ontology
5:          v_e := 0
6:          if e ∈ sig(α) then                              ▷ Element in axiom
7:              if e ∈ Σ then                ▷ element is a signature element
8:                  v_e := 1
9:              else
10:                 v_e := CALCULATE-AVERAGE-COSINE-DISTANCE(Σ ,V_e)
11:             end if
12:         end if
13:         V_α[n] := v_e
14:         n := n + 1
15:     end for
16: end function
```

The Function takes an axiom, the signature Σ created using the previously described function and an ontology as inputs. The function first allocates a vector $\mathcal{V}_\alpha$ with a length of $\|sig(\mathcal{O})\|$. The function then checks if elements in the ontology appears in the axiom. If an element does appear in it, the value is replaced with a non-zero number. If the element appears in Σ the value is 1. If it does not appear in Σ the non-zero value is calculated using the average cosine distance of the signature elements' and ontology element's Owl2Vec* vectors. Example 4.2.3 shows the formula for this calculation.

The encapsulating function that calls on all the previously defined functions would is described in algorithm 5.

Algorithm 5 Extract the k most relevant axioms from an ontology $\mathcal{O}$, w.r.t. a user's input i

Input: an integer k where k > 0, an ontology $\mathcal{O}$ and input i
Output: a set of axioms $\mathcal{M}$ where $\mathcal{M} \in \mathcal{O}$ and $|\mathcal{M}| = k$

```
1: function SUMMARIZE-ONTOLOGY(k,O ,i)
2:     Σ := CREATE-SIGNATURE(O,i)
3:     V_A := {}
4:     for α ∈ O do
5:         V_α := CONVERT-AXIOM-TO-VECTOR(α ,Σ ,O )
6:         V_A := V_A ∪ {V_α}
7:     end for
8:     V_Σ := CONVERT-AXIOM-TO-VECTOR(Σ ,Σ ,O )
9:     M := RANK-AXIOMS(k,V_Σ,V_α)
10:    return M
11: end function
```

4.2.5 Modularity

When running this program on a large ontology the run time is quite long (25-30 min when Owl2Vec*'s output includes more than 360'000 axioms). To circumvent this we employ the use of ontology excerpts discussed in section 2.1.2. As discussed, an ontology excerpt is a subset of axioms of an ontology that aim to keep the semantic knowledge of elements given a signature Σ . Using the java program created by Jieying Chen discussed in the paper Towards Extracting Ontology Excerpts [9], the program can create an ontology excerpt which is a fraction of the size of the original. For example when running this on the SNOMED CT ontology with the signature:

> Nerve structure (body structure), Description type, Cochlear duct structure, Structure of accessory nerve tree (body structure), Body part structure, Entire cell (cell), Trunk structure (body structure), Organ of Corti structure (body structure), Spinal nerve structure, Peripheral nerve structure of upper limb (body structure), Trees and shrubs, Structure of median nerve (body structure), Unapproved attribute, Branch of, Extensiveness, Cranial nerve structure, Entire median nerve (body structure), Anatomical structure (body structure), Neuron (cell), Structure of spinal nerve proper, Body organ structure (body structure), Entire (qualifier value), Structure of nervous system, Cochlear canal structure (body structure), Tree (organism),

Definition, Branch of brachial plexus

The original number of axiom Owl2Vec* extracts is: 360714. When number of axioms Owl2Vec* extracts from the excerpt is: 40. That means 99.98% of the axioms have been removed. This speeds up the algorithm significantly which will be discussed later. Do note that this signature uses the labels of the SNOMED CT entities.

Chapter 5

Implementation

This chapter decribes the practical implementation of the proof-of-concept program OntoSum.

5.1 Description

In this chapter I present my proof of concept of the program OntoSum based on the previously described project description. This chapter will present abstracted psuedocode to describe the whole program. The source code can be found at https://github.uio.no/mtlam/OwlTest [31].

5.2 Tools

Dependencies

- Owl2Vec* [6] and all it's dependencies
- numpy [38]
- BERT sentence_transformers [41]
- YAKE! [5]
- nltk [34]
- scikit-learn CountVectorizer [42]
- h5py [13]
- tabulate [1]
- tqdm [54]

The implementation was written in Python specifically to make use of Owl2Vec*'s ontology embedding and numpy's efficient number manipulation. BERT's sentence transformers, nltk and YAKE! are used to create the signature Σ. scikit-learn's CountVectorizer is used to encode axioms and the signature to run similarity comparisons. This program uses the pre-trained all-MiniLM-L6-v2 BERT model which is suitable for English data and general use. h5py is used to save integer arrays to disk to avoid having to do the same calculations for the same ontology multiple times. tabulate and tqdm are libraries used for quality of life. Tabulate parses the resulting k best axioms into a formatted table for ease of readability. tqdm creates a progress bar in the terminal during the ranking of axioms.

5.3 Execution

The program is executed like any other python program.

```
python OntoSum.py [-h] -k K -ont ONTOLOGY -i INPUT
          [-s] [-l] [-im] [-ie] [-o OUTPUT]
```

The program has three mandatory arguments:

- -k, a natural number of how many axioms should be included in the final ranking.
- -ont, ontology, the filepath to the ontology one would want to summarize
- -i, the filepath to the user's input. Must be a file with readable text. Preferably somewhat related to the ontology for the result to be impactful.

The rest of the argument flags are optional:

- -h, prints the help message and details for what each flag's purpose is.
- -s, skip, if a user already has pre-computed an ontology this will use the stored information in the cache as data.
- -l, labels, choose whether the final result should display IRIs or annotation labels.

- -im, ignore modularity, if this flag is selected the program will not execute modularity for larger ontologies. This will significantly impact runtime, but guarantee that the ranked axioms have considered every available axiom.
- -ie, ignore extension, if this flag is selected the program will not use the signature extension method and only use granular method regardless of ontology size.
- -o, file path to the output, if nothing is specified the default will be "./outputs/o.txt"

Example 5.3.1

```
python OntoSum.py -k 5 -ont SNOMED.owl -i covid19.txt -l -o output.txt
```

This is an example of what the inputs may look like. Where k is `5`, the ontology is `SNOMED.owl`, the input text is a text file named `covid19.txt` and the output's given filename is `output.txt`. The `-l` flag denotes that the output will use the ontology entities' labels instead of IRIs.

5.4 Structure

The program itself is one class. This class' functions can be seen in the UML diagram figure 5.1

Initialization

When initializing the class the ontology has to be loaded. This is done entirely through Owl2Vec* [6]. Owl2Vec* takes an Ontology file and outputs 3 relevant files to this algorithm, axioms.txt, annotations.txt and output. axioms.txt includes every* axiom in the ontology and is the main body interest the program wants to summarize, the collection of these axioms will be referred to as the corpus. annotations.txt are comments used to describe ontology elements giving them more semantic value. These are used in the silver corpus. Lastly output is gensim model file which is used to fetch ontology elements on the form of embedded vectors this is also what was referred to as a language model in figure 4.1.

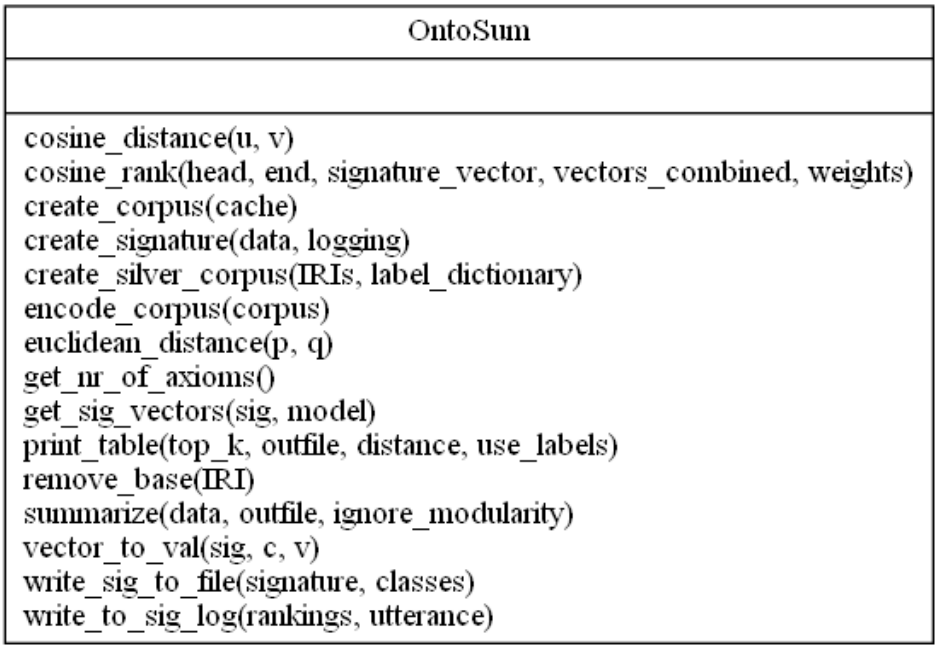

Figure 5.1: UML box of class

After loading the model, the program creates and embeds a silver corpus. The silver corpus created from every available IRI in the ontology and annotation. Each annotation is fetched from annotations.txt and is mapped to a corresponding IRI. All the IRIs and annotations are then encoded using sentence-BERT to have a mapping of IRIs and encoded matching text.

Part II

Evaluation

Chapter 6

Evaluation

This chapter begins by reiterating some of the central questions mentioned in the introduction. It continues by examining the results of the program, both qualitatively and quantitatively, before evaluating the runtime of the program in detail. Finally the discoveries are summarized.

6.1 Goals

To determine whether a method of summarization is good, we pose the following questions:

1. Do the selected axioms in the summarization provide an overview of what the ontology as a whole has to offer?
2. Are the selected axioms relevant to the user's input?
3. Is this the best summarization possible?
4. Are the summaries extracted in a reasonable amount of time?

All of the testing and creation of results were done on a computer with the following hardware:

Hardware

- Windows 64bit operating system
- 32GB installed RAM
- Intel(R) Core(TM) i7-4900MQ CPU @ 2.80GHz, 2794 Mhz, 4 Core(s), 8 Logical Processor(s)
- Python 3.9.13

6.2 Result Evaluation

For this thesis I conducted a survey on the program's summarization results as a qualitative evaluation. For the quantitative evaluation I evaluated the non-modular methods of OntoSum against the method IRExcerpts detailed in the paper Towards Extracting Ontology Excerpts [9].

6.2.1 Qualitiative Survey Evaluation

The survey's demographic was medical experts like doctors and specialists. The survey was divided into 5 medical topics: paracetamol, Tylenol, symptoms of heart disease, covid-19 and pneumonia. In the survey the participant was presented with 1 input text for a topic, 2 lists of ontology elements OntoSum had generated as signatures and 4 collections of ranked axioms. The input for each of the topics were the first paragraph of their respective Wikipedia pages [55, 58–60], except heart disease which used the list of symptoms [56] as it's input instead. The 2 signatures were generated using the granular signature creation and the extended signature creation methods. The 4 collections of ranked axioms were extracted using both a non-modular and modular function on both of the signatures. Collectively for each survey's list of ranked axioms is categorized by which methods have been used:

(A) No modularity and granular signature (GranSig)

(B) Modularity and granular signature (ModGranSig)

(C) No modularity and extended signature (ExtSig)

(D) Modularity and extended signature (ModExtSig)

Recall that modularity refers to the process of creating an ontology excerpt discussed in section 2.1.2. Granular signature refers to the signature creation process which matches all word and sentence tokens to find relevant ontology entities. Extended signature refers the signature creation process which extends a signature comprised by the ontology elements matched with keywords. It is extended by finding axioms where signature entities appear on the left hand side, then adding the right hand elements of these axioms to the signature. Both the signature creation methods were discussed in section 4.2.3. The alphabetical labels correspond to the the sections presented in the surveys.

1	2	3	4	5
Strongly Disagree	Disagree	Indifferent	Agree	Strongly Agree

Table 6.1: The evaluation scale users were asked to consider during the surveys. Where 1 means **strongly disagrees** and 5 means **strongly agrees**

To keep the demographic large some wording was translated to a simpler informal language so the medical experts did not need to have any understand of ontologies or description logic. The lists of ontology elements were translated to lists of keywords. The axioms were translated to natural language.

Example 6.2.1 *The axiom:*

> *"Paracetamol (substance)" SubClassOf "Non-opioid analgesic (substance)" and "Acetamide (substance)" and "Para-aminophenol derivative (substance)" and "Substance with phenol structure"*

would be translated to:

> *Paracetamol is a type of substance that belongs to the class of non-opioid analgesics. Paracetamol is also a type of acetamide. Additionally, paracetamol is a type of para-aminophenol derivative. Finally, paracetamol is a type of substance that has a phenol structure.*

The collections of ranked axioms were translated as enumerated lists of ranked sentences.

After presenting the input text, list of keywords (signature) and list of ranked sentences, the participant was asked to determine on a scale how much they disagreed or agreed with a couple statements. A visualisation of the scale is in table 6.1.

Lastly they were asked to rank the 4 solutions.
These were the presented statements:

- The selected keywords describe the text well.
- The selected sentences make sense according to the keywords.
- The selected sentences provide an overview of what the entire knowledge base has to offer.
- The selected sentences are informative
- With regards to what you know about [TOPIC], do you agree or disagree with the statement: The ranking of the sentences are good.

The selected keywords describe the text well.						
	Paracetamol	Tylenol	Heart Disease	Covid-19	Pneumonia	Complete Average
GranSig	4.00	4.33	4.60	4.25	4.50	4.34
ExtSig	4.56	4.33	4.60	4.00	4.00	4.30

Table 6.2: Survey table. Ratings of whether people believe the signature suits the input well.

The selected sentences make sense according to the keywords.						
	Paracetamol	Tylenol	Heart Disease	Covid-19	Pneumonia	Complete Average
GranSig	4.33	4.33	3.80	4.00	4.25	4.14
ModGranSig	4.33	4.17	4.40	4.00	4.25	4.23
ExtSig	4.56	3.50	4.20	4.00	3.75	4.00
ModExtSig	4.78	4.17	4.40	4.25	4.00	4.32

Table 6.3: Survey table. Ratings of whether people believe the summarization suits the signature.

[TOPIC] was replaced with the topic of the survey (i.e. Paracetamol, Tylenol or Covid-19 etc.).

Discussion of Results

The six tables compiling the average rating of each statement and method ranking will be discussed in this section. As mentioned earlier, the scale is from 1 - 5 where, the higher the number the more the participant agreed with the statement as shown in table 6.1. The rows describe how well a summarization method is received. The columns describe the average perception of the 5 different surveys, with the last column being the average of the preceding columns. Each survey was answered by at least 4 participants.

Table 6.2 shows the results of whether participants agreed with the statement: "The selected keywords describe the text well.". We asked participants this to judge which signature creation method were more liked. If there were to be an overwhelming majority that preferred the granular signature creation method there would be a good argument to use it instead of the extended method. We can see that the participants generally believe the signature describes the text well, as the lowest rating is 4. On average it seems like the granular signature method is preferred, though not by a large margin.

Table 6.3 shows the results of whether participants agreed with the statement: "The selected selected sentences makes sense according to the keywords.". The reasons for asking the opinions on this statement was to judge whether the program's summarizations held any merit to users. As there is no gold standard benchmark to compare our final results to,

The selected sentences are informative.						
	Paracetamol	Tylenol	Heart Disease	Covid-19	Pneumonia	Complete Average
GranSig	4.22	3.83	4.40	4.50	4.25	4.24
ModGranSig	4.00	4.00	4.40	4.50	4.25	4.23
ExtSig	4.44	3.67	4.60	4.25	3.75	4.14
ModExtSig	4.56	3.67	4.40	4.75	4.00	4.27

Table 6.4: Survey Table. Ratings of whether people find the axioms informative

The selected sentences provide an overview of what the entire knowledge base has to offer.						
	Paracetamol	Tylenol	Heart Disease	Covid-19	Pneumonia	Complete Average
GranSig	4.11	3.50	4.00	4.50	4.25	4.07
ModGranSig	3.67	3.83	4.20	4.50	4.25	4.09
ExtSig	4.56	3.67	4.40	4.25	3.75	4.12
ModExtSig	4.33	4.00	4.40	4.50	4.00	4.25

Table 6.5: Survey Table. Ratings of whether people believe the summarization represents the ontologies scope of the topic well.

a public evaluation is a suitable substitute. From table 6.3 we see some opinions dwindle a bit from 'Strongly Agree' and lean more to a neutral stance. This is notable in the survey with the topic Tylenol. My hypothesis for this is because Tylenol itself is not a term in SNOMED CTs ontology. Rather OntoSum latched onto Tylenol being associated with the mitigation of pain and diseases. Because of this the resulting summarizations for Tylenol were not as specifically topical as in the other surveys.

From table 6.4 and table 6.5 we see that generally the participants find the majority of summarizations informative and provide a good overview of the SNOMED CT ontology. Again like mentioned earlier, the less topical Tylenol summarization loses public opinion. Like the previous table 6.3, the reason for these statements were to judge results of program. If the program selects axioms that do not mean anything, then it has no value to users. If the program fails to provide an overview of what the knowledge base has to offer, the user may waste time as they could be testing other knowledge bases. For example if a summarization of the program returns a summarization that is not topical; it tells the user that the topic they are after is outside the scope of the knowledge base's domain.

The ranking of the sentences are good.						
	Paracetamol	Tylenol	Heart Disease	Covid-19	Pneumonia	Complete Average
GranSig	3.67	4.17	4.00	4.50	4.25	4.12
ModGranSig	4.11	3.83	4.20	4.50	4.00	4.13
ExtSig	4.44	3.67	4.20	4.50	3.50	4.06
ModExtSig	4.33	4.17	4.40	4.50	4.00	4.28

Table 6.6: Survey Table. Ratings of whether people agree with the ranking of the selected axioms

Rank the 4 summarization methods.					
	Best	2nd Best	3rd Best	4th Best	AVG RANK
GranSig	8.00	10.00	6.00	4.00	2.79
ModGranSig	6.00	6.00	8.00	9.00	2.31
ExtSig	8.00	6.00	10.00	4.00	2.64
ModExtSig	6.00	6.00	4.00	11.00	2.26

Table 6.7: Survey Table. Table shows how many times a method was ranked according to the other methods. For the sake of consistency the higher the number, the better it is in the AVG RANK column.

From table 6.6 we find that people generally believe the ranking of the axioms make sense. The ranking of the axioms' importance and relevancy to the input is generally accepted which means the weighted signatures and scoring methods are effective.

Overall from the tables we see that participants on average believe the summarization program does a good job at producing signatures and summarizations. Though it is a bit lacking when the input is not directly associated to the ontologies domain.

Lastly we discuss how the participants rated each summarization method against each other. This is shown in table 6.7 we see many times each method was selected in a ranking spot against each other. When looking at the opinions in the survey we see that the results created by ModExtSig methods were ranked higher on average than the others. The granular method of signature creation is therefore, not as good as the extended signature method. This is not reflected in Table 6.7 where ModExtSig score the lowest. This is because in the survey about Tylenol both extended signatures gave the same results. This caused participants to put D (ModExtSig) lower simply because it appeared later, which is fair.

6.2.2 Quantitative Benchmark Evaluation

As of May 2023 there exist no gold standard benchmark data which compiles the most relevant axioms in an ontology with regards to a short text or signature. Because there are no manually annotated datasets for in the field of ontology summarization, I manually picked 10 axioms for 9 different, arbitrarily chosen,topics related to the ontology FoodOn [19]. FoodOn was selected because of it smaller size to the SNOMED CT ontology while still being a well known large ontology in it's own right. (Though the ontology was not large enough after Owl2Vec*'s embedding to warrant the use of the modularity methods.) To further evaluate the

methods I compared the summaries OntoSum creates against the method IRExcerpts. My selection process of which axioms are important was to scan through all the axioms in the ontology using standfords' ontology editor, Protégé [49], to manually find the ones I found to be the most relevant. Due to this thorough approach not a lot of topics were considered. After collecting the 10 axioms of a given topic I compared them to what my own methods and IRExcerpts found to be the top 10 axioms in their respects and scored them based on the evaluation metrics mean reciprocal rank (MRR) [57] and hits@k [3]. MRR is a widely used evaluation metric for ranking algorithms in information retrieval and natural language processing tasks. It is defined as the average of the reciprocal ranks of the correct answers. In other words, it measures how well a ranking algorithm places the correct answer at the top of the list.

Definition 6.2.1 *MRR is given by the formula*

$$MRR = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{rank_i} \tag{6.1}$$

where:

$$N : \textit{Total number of queries}$$
$$rank_i : \textit{Rank of the first relevant item for query } i$$

The MRR value ranges from 0 to 1, where a higher MRR indicates better ranking performance, with 1 representing perfect ranking (the first relevant item is always ranked first).

In addition to a metric that highly values the most correct answer we use hits@k as a metric too, as it measures the percentage of correct answers in the top k results of a ranked list.

Definition 6.2.2 *Hits@k is given by the formula*

$$Hits@K = \frac{1}{N}\sum_{i=1}^{N} hit_i \tag{6.2}$$

where:

$$N : \textit{Total number of queries or users}$$
$$hit_i : \begin{cases} 1 & \textit{if at least one relevant item appears in the top K positions for query or user } i \\ 0 & \textit{otherwise} \end{cases}$$

The Hits@K metric ranges from 0 to 1, with a higher value indicating a better performance in retrieving relevant items within the top K positions.

This makes it easy to understand how well a ranking algorithm is performing.

The results can be viewed in Table 6.8. From the table we see that the GranSig method is generally better at finding the *best* axiom and ranking it as the best found axiom. However, ExtSig is generally more consistent at finding the best axiom at all as indicated by the higher MRR rating. Overall GranSig loses to ExtSig and IRExcerpt when more axioms are selected. Lastly we see that ExtSig scores higher than IRExcerpt in all categories, which helps support that this method has merits over existing methods.

Method	MRR	Hits@1	Hits@5	Hits@10
GranSig	0.27	0.24	0.39	0.41
ExtSig	0.35	0.19	0.60	0.60
IRExcerpt	0.28	0.14	0.45	0.49

Table 6.8: Benchmark evaluation on the ontology FoodOnt. Hits@k denotes the test axioms that have been ranked among the top k axioms.

6.3 Runtime Evaluation

This section will discuss the runtime of the program OntoSum and how the various methods affect runtime. First I will describe how the experiment was conducted. As stated in the implementation section 5.3, the unchanging inputs of the program is as follows:

- k = 5
- SNOMED CT ontology
- The input text is a short excerpt about nerves from the SNOMED 2017 editorial guide (Appendix 8.1)
- The flags -s, skip, and -l, labels

The file path for output irrelevant to the discussion.

In this runtime experiment we run the 4 different methods of summarization dicussed in the previous section: GranSig, ModGranSig, ExtSig and ModExtSig.

Each method is ran 30 times to ensure the average time is consistent.

6.3.1 Preprocessed Data

In this runtime experiment I have pre-processed 2 parts of the algorithm.

1. Owl2Vec*'s encoding of the SNOMED CT ontology
2. Sentence-BERT's encoding of the silver corpus

Both of these parts will always be the same for an ontology. Owl2Vec*s encoding of the SNOMED CT ontology takes around 8 hours. The embedding of all the classes and annotations of the SNOMED CT ontology takes around 6265 seconds or 104 minutes. To save time I have pre-processed both of these processes and loaded the necessary data in the experiment. Owl2Vec* caches all it's results in a defined cache folder, which makes fetching encoded data easy. Sentence-BERT does not have this caching function. To emulate this I used h5py [13] to save the data to be loaded later.

6.3.2 Runtime Results

Time spent on:	GranSig	ModGranSig	ExtSig	ModExtSig
Corpus construction	13.43 / 14.69 / 13.76	13.34 / 14.24 / 13.63	13.47 / 14.96 / 13.93	13.35 / 14.50 / 13.73
Signature creation	398.69 / 412.61 / 403.16	398.80 / 411.92 / 402.47	42.52 / 45.10 / 43.19	42.36 / 44.23 / 42.94
Ranking of vectors	1204.473 / 1244.19 / 1222.15	0.0023 / 0.0265 / 0.0015	1153.97 / 1195.29 / 1171.65	0.0053 / 0.0089 / 0.0058
Total Time	1848.32 / 1905.59 / 1863.34	475.19 / 505.11 / 478.83	1422.31 / 1497.65 / 1457.15	114.67 / 118.61 / 115.87

Table 6.9: Summarized time breakdown. Min / Max / Mean runtime (s). Granular, Extended signatures and Modular.

From the table 6.9 it is clear to see that modularity leads to a speedup of 3.9x in the case of granular signatures and 12.67x in the case of extended signatures when looking at the mean runtimes. In looking at the tables it is clear to see where one could try to improve the program.

1. The creation of the signature
2. The ranking of axioms which is mitigated through modularity

Time Complextity of the Signature

The Algorithm 1 has a big O-notation of O(wm + sm$|\Sigma|$m$|A|$).
Where `EXTRACT-KEYWORDS()` takes O($|i|$) time. The first loop takes O(w) time, where "w" is the number of keywords extracted in line 3. Within the loop `CONVERT-WORD-TO-VEC()` takes O(m) time, where "m" is the number of elements in the ontology. Likewise `GET-BEST-ONTOLOGY-ELEMENT()` takes O(m) time. The if statement takes O(1) time, and `SET-WEIGHT()` takes O(1)

time. The union operation takes O($|\Sigma|$) time, where $|\Sigma|$ is the current size of the signature.

The granular part of the signature creation (line 13-22) takes O(s) time where "s" is the number of words and sentences in the input. Within the loop `CONVERT-WORD-TO-VEC()` and `GET-BEST-ONTOLOGY-ELEMENT()` both takes O(m) time like described earlier. The if statement takes O(1) time, and `SET-WEIGHT()` takes O(1) time. The union operation takes O($|\Sigma|$) time

The extended signature part of the signature creation takes O($|\Sigma|$ * $|A|$) where $|A|$ is the number of axioms in the ontology. The innermost loop takes O($|a|$) time where $|a|$ is the size of the axiom signature. `SET-WEIGHT()` takes O(1) time. The union operation takes O($|\Sigma|$) time.

The runtime of this function scales with the number of elements in the ontology, m and the size of the input. We see from the tables 8.1 and 8.2 that the granular input matching slows down the execution significantly. From the table we see that the extended signature method is around 9x faster than the granular method in the signature creation.

6.3.3 The Effects of Modularity

We see clearly the effects of modularity by the difference in runtime when ranking vectors in table 6.9. The ranking process of ranking the vectors themselves has been sped up by around 800x. This is because the ontology excerpt removes around 99% of the axioms the algorithm has to search through.

The finding of 1-values

The finding 1-values refers to section 4.2.4 where I aim to retain the semantic knowledge of ontology elements. This only takes a long time in the methods where modularity is not used as the runtime scales with the size of the ontology. The axiom vectors created by one-hot encoding are sparse in the case where the input ontology is large. Like the test case with the SNOMED CT ontology. An axiom may at it's smallest be comprised of 3 ontology elements in this program. For example `Moderate Pain SubClassOf Pain`. This would result in a vector containing 2 instances of 1 in a vector where $|\mathcal{V}| = 360'000$. Filtering through indecies of sparse vectors are time-consuming because of the large amount of zeros. This was mitigated with the use of Numpy [38], however it is still very time-consuming. A solution to speed up this process could be pre-processing using a data structure that saves the indecies of the ontology elements that

appear in the axiom, then create a sparse vector for the cosine-similarity comparison.

Cosine-Similarity Comparison

Like the problem in the previous subsection, the cosine-similarity comparison takes a long time because the signature vector, $\mathcal{V}_\Sigma$, is compared against every other axiom vector in the ontology. From the table we see that each iteration of comparisons takes around 3 seconds. Over the span of 180 iterations the time we see for the row "Ranking of vectors" is the total time, including sorting the vectors. The sorting of the vectors barely take any time comparatively though. A way to speed this up in python would be to compare every single vector at the same time. However due to the memory constraints of scikit-learn's CountVectorizer [42]. The class allows the creation of vectors that retains what index refers to which ontology element. Due to this constraint of index to element maintenance a more memory efficient class like scikit-learn's HashVectorizer could not be used instead.

Chapter 7

Closing Thoughts and Future Work

In this project we have presented an approach to ontology summarization using unsupervised learning techniques, specifically KNN. Through user surveys we have found that participants' general consensus towards our approach positive, with users finding our generated summaries to be informative and relevant to the inputs. However, our approach suffers from the drawback of slow runtime. This is due to the unsupervised nature of our learning techniques, which require extensive computation and optimization. In order to improve the runtime of our approach, we suggest the use of supervised learning techniques, which have the potential to greatly reduce the computation required. Though to create a supervised learning model we would need gold standard benchmark data which currently does not exist for this case.

Despite this limitation, we believe that our approach has significant potential for real-world applications, such as aiding in the navigation and maintenance of large ontologies. Further due to the lack of available gold standard benchmarks evaluation is tough. An attempt was made to create an evaluation of the quality of the program's abilities' in entity linking using some benchmarks created using Unified Medical Language System (UMLS). Though due to conflicting IDs this did not lead anywhere.

In conclusion, our results demonstrates that ontology summarization using unsupervised learning techniques is a promising approach that has potential. However further optimizations are necessary to overcome the boundaries of slow runtime. We hope that our work will motivate future research in this area and contribute to the development of more efficient and effective ontology summarization techniques.

Chapter 8

Appendix

8.1 Nerve Input

Nerve There are different interpretations of nerve in the current modeling and the Editorial Guide. In the 2017 July version Editorial Guide, The FMA has defined three meanings for nerve: 1. a nerve trunk 2. the entire neural organ including nuclei, ganglia, roots, etc. 3. a nerve trunk plus all its branches (excluding nuclei, ganglia, and roots) It can create significant confusion we recognize that nerve is commonly used as a homonym for all three meanings. The FMA assigns the third meaning as the one that they adopt for the class labelled nerve. The trouble in this approach to resolving the problem of what "nerve" means is that when we call the first meaning "nerve trunk" and the second meaning "neural tree", it is difficult to decide how to refer to the third meaning. One solution is to use the phrase "neural organ" for the second meaning, since it is not really just a tree structure (at least an above-ground tree: the ganglia are not in the trunk or the branches); then the phrase "nerve tree" can be used for the third meaning. This would give us the trio of nerve trunk, neural organ, and nerve tree. These terms have better transparency than the trio of "nerve trunk", "nerve tree organ" and "nerve". Unlike clinical usage for arteries and veins, the clinical usage of the word "nerve" does not reliably refer to one of the three possible meanings, but instead varies much more between the different interpretations, based on context. If one severs the facial nerve, the meaning refers to the trunk. But if one has facial nerve palsy, the meaning refers to the entire distribution of the nerve and the functions served by it. The latest draft version: The word "nerve" potentially has multiple meanings: 'nerve tree' - encapsulating a nerve's entire structure as an organ which may include a nucleus, root, ganglion,

main trunk and branches; Note, this is different from the above because it includes the nucleus, root and ganglion. 'nerve trunk and branch' - including a nerve's main trunk and all its branches of distribution (but excluding all proximal constituent structures such as its nucleus) 'nerve trunk' - referring exclusively to the core structure of a nerve, excluding its proximal origin and its distal branches FMA defines peripheral nerve as "Organ component cluster which has as its direct parts components from two or more spinal nerves": This approach utilises the 'tree' definition, thus the median nerve in FMA is superordinate to its roots (C6-8, T1), central segments and its trunk and peripheral branches. SNOMED CT, in contrast to FMA, uses the term median nerve to mean the trunk of the median nerve representing its association with its origin by it being subordinate to Branch of brachial plexus and its regional location by an additional superordinate Peripheral nerve structure of upper limb. Thus the 'trunk' of peripheral and cranial nerves are arranged separately from their branches so that a specified nerve is not subordinate to the nerve from which it originates. The consequence of these semantics is that 'nerve concepts' are arranged in the following hierarchical relationships in general: X nerve tree X neural organ Nucleus of X neural organ Ganglion of X neural organ nerve root of X neural organ nerve trunk of X neural organ X nerve trunk and branch X nerve trunk Branch of X nerve trunk The agreement of this generic model needs to be achieved with the consultation to distinguish between these three meanings. The specific instances of 'tree' and 'trunk and branch', and "neural organ" concepts are not currently explicitly identified in SCT. Entire nerve The notion of "entire nerve" relates to the meaning of the applicable nerve - so Entire median nerve means the entire structure of the median nerve 'trunk' (rather than the entire median nerve tree which would include nucleus, roots, trunk and branches).

8.2 Detailed Time Breakdown of Runtime

	GranSig			ModGranSig		
Time spent on:	Min	Max	Mean	Min	Max	Mean
Constructing corpus	1.23	1.45	1.32	1.22	1.37	1.29
Embedding axioms	4.73	5.38	4.83	4.67	4.89	4.78
Silver corpus	7.48	7.87	7.61	7.46	7.98	7.56
Keyword matching	38.03	40.06	38.65	38.07	39.46	38.52
Granular input matching	360.29	372.07	364.08	360.73	372.45	363.95
Encoding signature	0.36	0.49	0.42	0.0003	0.0062	0.0006
Owl2Vec* averages	29.70	31.86	31.36	0.0013	0.0186	0.0001
Find 1 values in vectors	575.46	590.98	585.27	0.0001	0.0002	0.0001
Replace non-sig 1s	6.68	6.77	6.71	0.0006	0.0008	0.0007
Cosine-Sim comparison	578.93	766.08	597.98	0.0003	0.0069	0.0006
Sort axioms vectors	0.02	0.25	0.05	0.0001	0.0001	0.0001
Total time	1848.32	1905.59	1863.34	475.19	505.11	478.83

Table 8.1: Table for runs using granular signature methods. All values are in seconds

	ExtSig			ModExtSig		
Time spent on:	Min	Max	Mean	Min	Max	Mean
Constructing corpus	1.22	1.61	1.33	1.22	1.37	1.23
Embedding axioms	4.77	5.34	4.95	4.67	5.08	4.87
Constructing silver corpus	7.48	8.00	7.65	7.46	8.05	7.56
Keyword matching	38.04	40.36	38.64	38.20	39.82	38.71
Extend Signature	4.11	4.34	4.17	4.16	4.41	4.23
Encoding signature	0.37	0.40	0.38	0.0003	0.0008	0.0004
Owl2Vec* averages	16.91	17.41	17.11	0.0044	0.0069	0.0048
Find 1 values in vectors	555.48	572.63	564.06	0.0001	0.0002	0.0001
Replace non-sig 1s	6.48	6.64	6.55	0.0005	0.0006	0.0005
Cosine-Sim comparison	561.60	856.08	583.20	0.0002	0.0003	0.0002
Sort axioms vectors	0.0001	0.0013	0.036	0.0001	0.0001	0.0001
Total time	1422.31	1497.65	1457.15	114.67	118.61	115.87

Table 8.2: Runtime breakdown (s) Table for runs using extended signature methods.

8.3 Surveys

Links to each Google Forms survey

- Paracetamol: https://forms.gle/7A6oPa2wigUNNm616
- Tylenol: https://forms.gle/d8F7h8rYe2jpGdyH7
- Heart Disease: https://forms.gle/1r9cwGftgFSqAXiv9
- Covid-19: https://forms.gle/v6ef8p6CYYS3Vzqr9
- Pneumonia: https://forms.gle/qJzomkQDVz8Xtk2t6

8.3.1 Survey: Paracetamol

Survey for Ontology Summarization: Paracetamol https://docs.google.com/forms/u/0/d/1wNvzvNrNk6Y-usp1fl71G1ity...

Survey for Ontology Summarization: Paracetamol

This is a collection of surveys made for the purpose of evaluating the results of Martin Lam's Master's project. The project is about summarizing **knowledge bases** (Ontologies).
To summarize knowledge bases a program has been created that ranks and extracts *axioms*. Axioms can be viewed as **sentences**. Do keep in mind that the selected sentences were selected from a pool of 360'000 sentences.
This is 1 of 5 available surveys.
This survey will cover the topic: **Paracetamol**

The survey is divided into **5** sections (A, B, C, D and ranking).

The **text** that the program evaluates is the same for **all** sections.

The first **4** sections each have their own enumerated list of ranked sentences. There may be shared sentences between them. The ranking in the lists are ordered in descending order, where rank 1 is the most relevant sentence and rank 5 is the least relevant sentence among those selected.

You will be presented with **2** different lists of keywords. The keyword list in section (A) and (B) is the same and the list in (C) and (D) is the same.

Lastly you will be asked to rank the 4 enumerated lists.

For the survey to have the best results, I kindly ask that you answer as honestly as you can.

* Indicates required question

General Information About Paracetamol (A)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Paracetamol)
Paracetamol (acetaminophen or para-hydroxyacetanilide) is a medication used to treat fever and mild to moderate pain. Common brand names include Tylenol and Panadol.

At a standard dose, paracetamol only slightly decreases body temperature; it is inferior to ibuprofen in that respect, and the benefits of its use for fever are unclear. Paracetamol may relieve pain in acute mild migraine but only slightly in episodic tension headache. However, the aspirin/paracetamol/caffeine combination helps with both conditions where the pain is mild and is recommended as a first-line treatment for them. Paracetamol is effective for post-surgical pain, but it is inferior to ibuprofen. The paracetamol/ibuprofen combination provides further increase in potency and is superior to either drug alone. The pain relief paracetamol provides in osteoarthritis is small and clinically insignificant. The evidence in its favor for the use in low back pain, cancer pain, and neuropathic pain is insufficient.

Keywords selected by program:

- Acetamide (substance)
- Acuteness (qualifier value)
- Administration of drug or medicament (procedure)
- Body part structure
- Combined (qualifier value)
- Common
- Condition
- Decreasing (qualifier value)
- Drug or medicament (substance)
- Effective
- Episodic
- Evidence of (contextual qualifier) (qualifier value)
- Feeling relief (finding)
- Fever (finding)
- Headache (finding)
- Ibuprofen
- Including (qualifier value)
- Increase (qualifier value)
- Inferior (qualifier value)
- Insufficient
- Lives alone (finding)
- Low (qualifier value)
- Malignant neoplasm (morphologic abnormality)
- Migraine (disorder)
- Mild (qualifier value)
- Moderate (severity modifier) (qualifier value)
- Moderate pain (finding)
- Name (property)
- Neuropathic pain
- Osteoarthritis (disorder)
- Pain
- Paracetamol (substance)
- Recommended
- Relieves
- Slightly (qualifier value)
- Small
- Structure of back of trunk (body structure)
- Superior (qualifier value)
- Temperature
- Tension
- Treatment given
- Treatment of fever (procedure)
- Unit dose
- Use - dosing instruction imperative (qualifier value)

53

1. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking:

1. Moderate pain is a subclass of Pain
2. Headache is a subclass of Pain located in the head
3. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
4. Backache is equivalent to Pain located in the Entire back structure excluding neck
5. A medicinal product that contains only ibuprofen and paracetamol is a type of product that is equivalent to a medicinal product in general. This type of product is associated with a group of roles that describe its active ingredients. Specifically, it has ibuprofen and paracetamol as its active ingredients, both of which are substances used to relieve pain and reduce fever. Additionally, the count of the base of the active ingredient is included in the description, which refers to the amount of ibuprofen and paracetamol present in the product.

2. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Survey for Ontology Summarization: Paracetamol https://docs.google.com/forms/u/0/d/1wNvzvNrNk6Y-usp1fl71G1ity...

3. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

4. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

5. With regards to what **you** know about **Paracetamol**, do you agree or disagree with the statement: * 54
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

General Information About Paracetamol (B)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Paracetamol)

Paracetamol (acetaminophen or para-hydroxyacetanilide) is a medication used to treat fever and mild to moderate pain. Common brand names include Tylenol and Panadol.

At a standard dose, paracetamol only slightly decreases body temperature; it is inferior to ibuprofen in that respect, and the benefits of its use for fever are unclear. Paracetamol may relieve pain in acute mild migraine but only slightly in episodic tension headache. However, the aspirin/paracetamol/caffeine combination helps with both conditions where the pain is mild and is recommended as a first-line treatment for them. Paracetamol is effective for post-surgical pain, but it is inferior to ibuprofen. The paracetamol/ibuprofen combination provides further increase in potency and is superior to either drug alone. The pain relief paracetamol provides in osteoarthritis is small and clinically insignificant. The evidence in its favor for the use in low back pain, cancer pain, and neuropathic pain is insufficient.

Keywords selected by program:

- Acetamide (substance)
- Acuteness (qualifier value)
- Administration of drug or medicament (procedure)
- Body part structure
- Combined (qualifier value)
- Common
- Condition
- Decreasing (qualifier value)
- Drug or medicament (substance)
- Effective
- Episodic
- Evidence of (contextual qualifier) (qualifier value)
- Feeling relief (finding)
- Fever (finding)
- Headache (finding)
- Ibuprofen
- Including (qualifier value)
- Increase (qualifier value)
- Inferior (qualifier value)
- Insufficient
- Lives alone (finding)
- Low (qualifier value)
- Malignant neoplasm (morphologic abnormality)
- Migraine (disorder)
- Mild (qualifier value)
- Moderate (severity modifier) (qualifier value)
- Moderate pain (finding)
- Name (property)
- Neuropathic pain
- Osteoarthritis (disorder)
- Pain
- Paracetamol (substance)
- Recommended
- Relieves
- Slightly (qualifier value)
- Small
- Structure of back of trunk (body structure)
- Superior (qualifier value)
- Temperature
- Tension
- Treatment given
- Treatment of fever (procedure)
- Unit dose
- Use - dosing instruction imperative (qualifier value)

55

Sentence ranking:

1. Moderate pain is a subclass of Pain
2. Headache is a subclass of Pain located in the head
3. Backache is equivalent to Pain located in the Entire back structure excluding neck
4. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
5. Pain is a subclass of Pain and sensation

6. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

7. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

8. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

9. With regards to what **you** know about **Paracetamol**, do you agree or disagree * with the statement:
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

56

General Information About Paracetamol (C)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Paracetamol)
Paracetamol (acetaminophen or para-hydroxyacetanilide) is a medication used to treat fever and mild to moderate pain. Common brand names include Tylenol and Panadol.

At a standard dose, paracetamol only slightly decreases body temperature; it is inferior to ibuprofen in that respect, and the benefits of its use for fever are unclear. Paracetamol may relieve pain in acute mild migraine but only slightly in episodic tension headache. However, the aspirin/paracetamol/caffeine combination helps with both conditions where the pain is mild and is recommended as a first-line treatment for them. Paracetamol is effective for post-surgical pain, but it is inferior to ibuprofen. The paracetamol/ibuprofen combination provides further increase in potency and is superior to either drug alone. The pain relief paracetamol provides in osteoarthritis is small and clinically insignificant. The evidence in its favor for the use in low back pain, cancer pain, and neuropathic pain is insufficient.

Keywords selected by program:

- Abnormal body temperature
- Above reference range
- Acetamide (substance)
- Acetamide and/or acetamide derivative (substance)
- Administration - action (qualifier value)
- Administration of drug or medicament (procedure)
- Body temperature (observable entity)
- Cooling the patient (procedure)
- Cyclooxygenase-1 inhibitor
- Cyclooxygenase-2 inhibitor
- Degree findings
- Direct substance (attribute)
- Drug or medicament (substance)
- Fever (finding)
- General site descriptor (qualifier value)
- Has disposition (attribute)
- Has focus (attribute)
- Has interpretation
- Ibuprofen
- Inferior (qualifier value)
- Interprets
- Method (attribute)
- Mild (qualifier value)
- Moderate (severity modifier) (qualifier value)
- Moderate pain (finding)
- Non-opioid analgesic (substance)
- Non-steroidal anti-inflammatory agent (substance)
- Pain
- Pain / sensation finding
- Para-aminophenol derivative (substance)
- Paracetamol (substance)
- Procedure (procedure)
- Propionic acid and/or propionic acid derivative
- Role group
- Severities
- Slightly (qualifier value)
- Substance with phenol structure
- Treatment of fever (procedure)

10. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking

1. Paracetamol is a type of substance that belongs to the class of non-opioid analgesics. Paracetamol is also a type of acetamide. Additionally, paracetamol is a type of para-aminophenol derivative. Finally, paracetamol is a type of substance that has a phenol structure.

2. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.

3. A medicinal product that contains only ibuprofen and paracetamol is a type of product that is equivalent to a medicinal product in general. This type of product is associated with a group of roles that describe its active ingredients. Specifically, it has ibuprofen and paracetamol as its active ingredients, both of which are substances used to relieve pain and reduce fever. Additionally, the count of the base of the active ingredient is included in the description, which refers to the amount of ibuprofen and paracetamol present in the product.

4. A medicinal product that contains both ibuprofen and paracetamol is a type of product that is equivalent to a medicinal product in general. This type of product is associated with a group of roles that describe its active ingredients. Specifically, it has ibuprofen and paracetamol as its active ingredients

5. Moderate pain is a subclass of Pain

11. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

12. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

13. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

14. With regards to what **you** know about **Paracetamol**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

General Information About Paracetamol (D)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Paracetamol)

Paracetamol (acetaminophen or para-hydroxyacetanilide) is a medication used to treat fever and mild to moderate pain. Common brand names include Tylenol and Panadol.

At a standard dose, paracetamol only slightly decreases body temperature; it is inferior to ibuprofen in that respect, and the benefits of its use for fever are unclear. Paracetamol may relieve pain in acute mild migraine but only slightly in episodic tension headache. However, the aspirin/paracetamol/caffeine combination helps with both conditions where the pain is mild and is recommended as a first-line treatment for them. Paracetamol is effective for post-surgical pain, but it is inferior to ibuprofen. The paracetamol/ibuprofen combination provides further increase in potency and is superior to either drug alone. The pain relief paracetamol provides in osteoarthritis is small and clinically insignificant. The evidence in its favor for the use in low back pain, cancer pain, and neuropathic pain is insufficient.

58

Keywords selected by program:

- Abnormal body temperature
- Above reference range
- Acetamide (substance)
- Acetamide and/or acetamide derivative (substance)
- Administration - action (qualifier value)
- Administration of drug or medicament (procedure)
- Body temperature (observable entity)
- Cooling the patient (procedure)
- Cyclooxygenase-1 inhibitor
- Cyclooxygenase-2 inhibitor
- Degree findings
- Direct substance (attribute)
- Drug or medicament (substance)
- Fever (finding)
- General site descriptor (qualifier value)
- Has disposition (attribute)
- Has focus (attribute)
- Has interpretation
- Ibuprofen
- Inferior (qualifier value)
- Interprets
- Method (attribute)
- Mild (qualifier value)
- Moderate (severity modifier) (qualifier value)
- Moderate pain (finding)
- Non-opioid analgesic (substance)
- Non-steroidal anti-inflammatory agent (substance)
- Pain
- Pain / sensation finding
- Para-aminophenol derivative (substance)
- Paracetamol (substance)
- Procedure (procedure)
- Propionic acid and/or propionic acid derivative
- Role group
- Severities
- Slightly (qualifier value)
- Substance with phenol structure
- Treatment of fever (procedure)

Sentence Ranking:

1. Paracetamol is a type of substance that belongs to the class of non-opioid analgesics. Paracetamol is also a type of acetamide. Additionally, paracetamol is a type of para-aminophenol derivative. Finally, paracetamol is a type of substance that has a phenol structure.
2. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
3. Moderate pain is a subclass of Pain
4. Symptoms of fever is a subclass of Abnormal body temperature.
5. Ibuprofen is an Non-steroidal anti-inflammatory agent and a Propionic acid and/or propionic acid derivative. It has the qualities of Cyclooxygenase-1 inhibitor and Cyclooxygenase-2 inhibitor.

15. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

16. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

17. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

18. With regards to what **you** know about **Paracetamol**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Ranking the tables: **Paracetamol**

In this section you will rank the 4 different collections on which one describes the text best.

Text:

Paracetamol (acetaminophen or para-hydroxyacetanilide) is a medication used to treat fever and mild to moderate pain. Common brand names include Tylenol and Panadol.

At a standard dose, paracetamol only slightly decreases body temperature; it is inferior to ibuprofen in that respect, and the benefits of its use for fever are unclear. Paracetamol may relieve pain in acute mild migraine but only slightly in episodic tension headache. However, the aspirin/paracetamol/caffeine combination helps with both conditions where the pain is mild and is recommended as a first-line treatment for them. Paracetamol is effective for post-surgical pain, but it is inferior to ibuprofen. The paracetamol/ibuprofen combination provides further increase in potency and is superior to either drug alone. The pain relief paracetamol provides in osteoarthritis is small and clinically insignificant. The evidence in its favor for the use in low back pain, cancer pain, and neuropathic pain is insufficient.

Survey for Ontology Summarization: Paracetamol https://docs.google.com/forms/u/0/d/1wNvzvNrNk6Y-usp1fl71G1ity...

Paracetemol (A)

1. Moderate pain is a subclass of Pain
2. Headache is a subclass of Pain located in the head
3. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
4. Backache is equivalent to Pain located in the Entire back structure excluding neck
5. A meicinal product that contains only ibuprofen and paracetamol is a type of product that is equivalent to a medicinal product in general. This type of product is associated with a group of roles that describe its active ingredients. Specifically, it has ibuprofen and paracetamol as its active ingredients, both of which are substances used to relieve pain and reduce fever. Additionally, the count of the base of the active ingredient is included in the description, which refers to the amount of ibuprofen and paracetamol present in the product.

09

Paracetemol (B)

1. Moderate pain is a subclass of Pain
2. Headache is a subclass of Pain located in the head
3. Backache is equivalent to Pain located in the Entire back structure excluding neck
4. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
5. Pain is a subclass of Pain and sensation

Paracetemol (C)

1. Paracetamol is a type of substance that belongs to the class of non-opioid analgesics. Paracetamol is also a type of acetamide. Additionally, paracetamol is a type of para-aminophenol derivative. Finally, paracetamol is a type of substance that has a phenol structure.
2. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
3. A medicinal product that contains only ibuprofen and paracetamol is a type of product that is equivalent to a medicinal product in general. This type of product is associated with a group of roles that describe its active ingredients. Specifically, it has ibuprofen and paracetamol as its active ingredients, both of which are substances used to relieve pain and reduce fever. Additionally, the count of the base of the active ingredient is included in the description, which refers to the amount of ibuprofen and paracetamol present in the product.
4. A medicinal product that contains both ibuprofen and paracetamol is a type of product that is equivalent to a medicinal product in general. This type of product is associated with a group of roles that describe its active ingredients. Specifically, it has ibuprofen and paracetamol as its active ingredients
5. Moderate pain is a subclass of Pain

Paracetemol (D)

1. Paracetamol is a type of substance that belongs to the class of non-opioid analgesics. Paracetamol is also a type of acetamide. Additionally, paracetamol is a type of para-aminophenol derivative. Finally, paracetamol is a type of substance that has a phenol structure.

2. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.

3. Moderate pain is a subclass of Pain

4. Symptoms of fever is a subclass of Abnormal body temperature.

5. Ibuprofen is an Non-steroidal anti-inflammatory agent and a Propionic acid and/or propionic acid derivative. It has the qualities of Cyclooxygenase-1 inhibitor and Cyclooxygenase-2 inhibitor.

19. Rank the collections of sentences from best to worst *

61

Mark only one oval per row.

	A	B	C	D
Best				
2nd Best				
3rd Best				
4th Best				

Closing comments

Thank you for taking the survey!

20. If you have any comments to the survey please leave them here!

8.3.2 Survey: Tylenol

Survey for Ontology Summarization: Tylenol https://docs.google.com/forms/u/1/d/1wEiybVT9nJTXwgihV_gA1zGi...

Survey for Ontology Summarization: Tylenol

This is a collection of surveys made for the purpose of evaluating the results of Martin Lam's Master's project. The project is about summarizing **knowledge bases** (Ontologies).
To summarize knowledge bases a program has been created that ranks and extracts *axioms*. Axioms can be viewed as **sentences**. Do keep in mind that the selected sentences were selected from a pool of 360'000 sentences.
This is 1 of 5 available surveys.
This survey will cover the topic: **Tylenol**

The survey is divided into **5** sections (A, B, C, D and ranking).

The **text** that the program evaluates is the same for **all** sections.

The first **4** sections each have their own enumerated list of ranked sentences. There may be shared sentences between them. The ranking in the lists are ordered in descending order, where rank 1 is the most relevant sentence and rank 5 is the least relevant sentence among those selected.

63

You will be presented with **2** different lists of keywords. The keyword list in section (A) and (B) is the same and the list in (C) and (D) is the same.

Lastly you will be asked to rank the 4 enumerated lists.

For the survey to have the best results, I kindly ask that you answer as honestly as you can.

* Indicates required question

General Information About Tylenol (A)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Tylenol_(brand))
Tylenol is a brand of medication, advertised for reducing pain, reducing fever, and relieving the symptoms of allergies, cold, cough, headache, and influenza. The active ingredient of its original flagship product is paracetamol (known in the United States, Canada, and various other countries as acetaminophen), an analgesic and antipyretic. Like the words paracetamol and acetaminophen, the brand name Tylenol is derived from a chemical name for the compound, N-acetyl-para-aminophenol (APAP). The brand name is owned by McNeil Consumer Healthcare, a subsidiary of Johnson & Johnson.

Keywords selected by program:

- Acetamide (substance)
- Active (qualifier value)
- Administration of antipyretic (procedure)
- Administration of drug or medicament (procedure)
- Allergic disposition (finding)
- Analgesic (substance)
- Canada (geographic location)
- Chemical (substance)
- Common cold
- Compound
- Consumer product (product)
- Cough
- Country
- Fever (finding)
- Headache (finding)
- Influenza (disorder)
- Known (qualifier value)
- Medical care
- Name (property)
- Original (qualifier value)
- Pain
- Paracetamol (substance)
- Preventing pain
- Product containing paracetamol (medicinal product)
- Reduction plasty (qualifier value)
- State (environment)
- Treatment of fever (procedure)
- United States of America (geographic location)

Survey for Ontology Summarization: Tylenol https://docs.google.com/forms/u/1/d/1wEiybVT9nJTXwgihV_gA1zGi...

1. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking:

1. Administration of antipyretic is a subclass of Administration of drug or medicament
2. Cough with fever is a subclass of Fever and Cough
3. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
4. Headache (finding) is a subclass of Pain located in the head
5. Painful cough is equivalent to Pain and Cough found in Respiratory tract structure

64

2. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

3. With regards to what **you** know about **Tylenol**, do you agree or disagree with the * statement:
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

4. The selected sentences provide an overview of what the entire knowledge base * has to offer.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

5. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

General Information About Tylenol (B)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Tylenol_(brand))
Tylenol is a brand of medication, advertised for reducing pain, reducing fever, and relieving the symptoms of allergies, cold, cough, headache, and influenza. The active ingredient of its original flagship product is paracetamol (known in the United States, Canada, and various other countries as acetaminophen), an analgesic and antipyretic. Like the words paracetamol and acetaminophen, the brand name Tylenol is derived from a chemical name for the compound, N-acetyl-para-aminophenol (APAP). The brand name is owned by McNeil Consumer Healthcare, a subsidiary of Johnson & Johnson.

Keywords selected by program:

- Acetamide (substance)
- Active (qualifier value)
- Administration of antipyretic (procedure)
- Administration of drug or medicament (procedure)
- Allergic disposition (finding)
- Analgesic (substance)
- Canada (geographic location)
- Chemical (substance)
- Common cold
- Compound
- Consumer product (product)
- Cough
- Country
- Fever (finding)
- Headache (finding)
- Influenza (disorder)
- Known (qualifier value)
- Medical care
- Name (property)
- Original (qualifier value)
- Pain
- Paracetamol (substance)
- Preventing pain
- Product containing paracetamol (medicinal product)
- Reduction plasty (qualifier value)
- State (environment)
- Treatment of fever (procedure)
- United States of America (geographic location)

65

Sentence Ranking:

1. Administration of antipyretic is a subclass of Administration of drug or medicament
2. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
3. Headache (finding) is a subclass of Pain located in the head
4. Painful cough is equivalent to Pain and Cough found in Respiratory tract structure
5. Administration of analgesic is equivalent to Administration of a drug or medicament and is associated with a group of roles that describe the method of administration and the substance being administered. Specifically, it is associated with a method role that describes the action of administration, and a direct substance role that describes the substance being administered as an analgesic.

6. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

7. With regards to what **you** know about **Tylenol**, do you agree or disagree with the * statement:
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

8. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

9. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

General Information About Tylenol (C)

99

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Tylenol_(brand))
Tylenol is a brand of medication, advertised for reducing pain, reducing fever, and relieving the symptoms of allergies, cold, cough, headache, and influenza. The active ingredient of its original flagship product is paracetamol (known in the United States, Canada, and various other countries as acetaminophen), an analgesic and antipyretic. Like the words paracetamol and acetaminophen, the brand name Tylenol is derived from a chemical name for the compound, N-acetyl-para-aminophenol (APAP). The brand name is owned by McNeil Consumer Healthcare, a subsidiary of Johnson & Johnson.

Keywords selected by program:

- Administration - action (qualifier value)
- Administration of drug or medicament (procedure)
- Causative agent
- Common cold
- Cooling the patient (procedure)
- Cough
- Direct substance (attribute)
- Disease
- Drug or medicament (substance)
- Fever (finding)
- Finding site
- Has focus (attribute)
- Head structure
- Headache (finding)
- Infectious process (qualifier value)
- Method (attribute)
- North American country
- Pain
- Pain control
- Pathological process
- Plastic operation (qualifier value)
- Preventing pain
- Preventive procedure (procedure)
- Procedure (procedure)
- Reduction plasty (qualifier value)
- Respiratory function finding (finding)
- Respiratory tract structure (body structure)
- Role group
- Treatment of fever (procedure)
- United States of America (geographic location)
- Upper respiratory tract structure (body structure)
- Virus

10. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking:

1. The common cold is a type of disease that is associated with a group of roles. The first role describes the causative agent of the disease, which is a virus. Another role describes the finding site of the disease, which is the upper respiratory tract structure of the body. Additionally, the disease is associated with a pathological process, which is an infectious process.

2. The administration of a drug or medicament is a type of procedure that is equivalent to a general procedure. This type of procedure is associated with a group of roles that describe how it is carried out. Specifically, it is associated with a method role that describes the action of administration, and a direct substance role that describes the drug or medicament that is being administered.

3. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.

4. Headache caused by oral contraceptive pill is a subclass of headaches. This specific type of headache finding is associated with a group of roles that describe its cause and location. Specifically, it is associated with a causative agent role that describes the drug or medicament that causes the headache, and a finding site role that describes the location of the headache as the head structure.

5. Cough is a subclass of respiratory functions and is associated with findings in the Respiratory tract structure

11. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

12. With regards to what **you** know about **Tylenol**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

13. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

14. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

67

General Information About Tylenol (D)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Tylenol_(brand))

Tylenol is a brand of medication, advertised for reducing pain, reducing fever, and relieving the symptoms of allergies, cold, cough, headache, and influenza. The active ingredient of its original flagship product is paracetamol (known in the United States, Canada, and various other countries as acetaminophen), an analgesic and antipyretic. Like the words paracetamol and acetaminophen, the brand name Tylenol is derived from a chemical name for the compound, N-acetyl-para-aminophenol (APAP). The brand name is owned by McNeil Consumer Healthcare, a subsidiary of Johnson & Johnson.

89

Keywords selected by program:

- Administration - action (qualifier value)
- Administration of drug or medicament (procedure)
- Causative agent
- Common cold
- Cooling the patient (procedure)
- Cough
- Direct substance (attribute)
- Disease
- Drug or medicament (substance)
- Fever (finding)
- Finding site
- Has focus (attribute)
- Head structure
- Headache (finding)
- Infectious process (qualifier value)
- Method (attribute)
- North American country
- Pain
- Pain control
- Pathological process
- Plastic operation (qualifier value)
- Preventing pain
- Preventive procedure (procedure)
- Procedure (procedure)
- Reduction plasty (qualifier value)
- Respiratory function finding (finding)
- Respiratory tract structure (body structure)
- Role group
- Treatment of fever (procedure)
- United States of America (geographic location)
- Upper respiratory tract structure (body structure)
- Virus

Sentence Ranking:

1. The common cold is a type of disease that is associated with a group of roles. The first role describes the causative agent of the disease, which is a virus. Another role describes the finding site of the disease, which is the upper respiratory tract structure of the body. Additionally, the disease is associated with a pathological process, which is an infectious process.

2. The administration of a drug or medicament is a type of procedure that is equivalent to a general procedure. This type of procedure is associated with a group of roles that describe how it is carried out. Specifically, it is associated with a method role that describes the action of administration, and a direct substance role that describes the drug or medicament that is being administered.

3. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.

4. Headache caused by oral contraceptive pill is a subclass of headaches. This specific type of headache finding is associated with a group of roles that describe its cause and location. Specifically, it is associated with a causative agent role that describes the drug or medicament that causes the headache, and a finding site role that describes the location of the headache as the head structure.

5. Cough is a subclass of respiratory functions and is associated with findings in the Respiratory tract structure

69

15. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

16. With regards to what **you** know about **Tylenol**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

17. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

18. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Survey for Ontology Summarization: Tylenol https://docs.google.com/forms/u/1/d/1wEiybVT9nJTXwgihV_gA1zGi...

Ranking the tables: **Tylenol**

In this section you will rank the 4 different tables on how which one describes the text best.

Text:

Tylenol is a brand of medication, advertised for reducing pain, reducing fever, and relieving the symptoms of allergies, cold, cough, headache, and influenza. The active ingredient of its original flagship product is paracetamol (known in the United States, Canada, and various other countries as acetaminophen), an analgesic and antipyretic. Like the words paracetamol and acetaminophen, the brand name Tylenol is derived from a chemical name for the compound, N-acetyl-para-aminophenol (APAP). The brand name is owned by McNeil Consumer Healthcare, a subsidiary of Johnson & Johnson.

Tylenol (A)

70

1. Administration of antipyretic is a subclass of Administration of drug or medicament
2. Cough with fever is a subclass of Fever and Cough
3. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
4. Headache (finding) is a subclass of Pain located in the head
5. Painful cough is equivalent to Pain and Cough found in Respiratory tract structure

Tylenol (B)

1. Administration of antipyretic is a subclass of Administration of drug or medicament
2. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.
3. Headache (finding) is a subclass of Pain located in the head
4. Painful cough is equivalent to Pain and Cough found in Respiratory tract structure
5. Administration of analgesic is equivalent to Administration of a drug or medicament and is associated with a group of roles that describe the method of administration and the substance being administered. Specifically, it is associated with a method role that describes the action of administration, and a direct substance role that describes the substance being administered as an analgesic.

Tylenol (C)

1. The common cold is a type of disease that is associated with a group of roles. The first role describes the causative agent of the disease, which is a virus. Another role describes the finding site of the disease, which is the upper respiratory tract structure of the body. Additionally, the disease is associated with a pathological process, which is an infectious process.

2. The administration of a drug or medicament is a type of procedure that is equivalent to a general procedure. This type of procedure is associated with a group of roles that describe how it is carried out. Specifically, it is associated with a method role that describes the action of administration, and a direct substance role that describes the drug or medicament that is being administered.

3. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.

4. Headache caused by oral contraceptive pill is a subclass of headaches. This specific type of headache finding is associated with a group of roles that describe its cause and location. Specifically, it is associated with a causative agent role that describes the drug or medicament that causes the headache, and a finding site role that describes the location of the headache as the head structure.

5. Cough is a subclass of respiratory functions and is associated with findings in the Respiratory tract structure

Tylenol (D)

1. The common cold is a type of disease that is associated with a group of roles. The first role describes the causative agent of the disease, which is a virus. Another role describes the finding site of the disease, which is the upper respiratory tract structure of the body. Additionally, the disease is associated with a pathological process, which is an infectious process.

2. The administration of a drug or medicament is a type of procedure that is equivalent to a general procedure. This type of procedure is associated with a group of roles that describe how it is carried out. Specifically, it is associated with a method role that describes the action of administration, and a direct substance role that describes the drug or medicament that is being administered.

3. Treating a fever is a type of medical procedure that belongs to the class of cooling the patient and has a focus on fever symptoms.

4. Headache caused by oral contraceptive pill is a subclass of headaches. This specific type of headache finding is associated with a group of roles that describe its cause and location. Specifically, it is associated with a causative agent role that describes the drug or medicament that causes the headache, and a finding site role that describes the location of the headache as the head structure.

5. Cough is a subclass of respiratory functions and is associated with findings in the Respiratory tract structure

19. Rank the collections of sentences from best to worst *

Mark only one oval per row.

	A	B	C	D
Best				
2nd Best				
3rd Best				
4th Best				

Closing comments

Thank you for taking the survey!

20. If you have any comments to the survey please leave them here!

8.3.3 Survey: Heart Disease

Survey for Ontology Summarization: Heart Disease https://docs.google.com/forms/u/0/d/1Zg_tyoTKfZWj0Wz_3k1nXEc...

Survey for Ontology Summarization: Heart Disease

This is a collection of surveys made for the purpose of evaluating the results of Martin Lam's Master's project. The project is about summarizing **knowledge bases** (Ontologies).

To summarize knowledge bases a program has been created that ranks and extracts *axioms*. Axioms can be viewed as **sentences**. Do keep in mind that the selected sentences were selected from a pool of 360'000 sentences.
This is 1 of 5 available surveys.
This survey will cover the topic: **Heart Disease**

The survey is divided into **5** sections (A, B, C, D and ranking).

The **text** that the program evaluates is the same for **all** sections.

The first **4** sections each have their own enumerated list of ranked sentences. There may be shared sentences between them. The ranking in the lists are ordered in descending order, where rank 1 is the most relevant sentence and rank 5 is the least relevant sentence among those selected.

You will be presented with **2** different lists of keywords. The keyword list in section (A) and (B) is the same and the list in (C) and (D) is the same.

Lastly you will be asked to rank the 4 enumerated lists.

For the survey to have the best results, I kindly ask that you answer as honestly as you can.

* Indicates required question

Symptoms of Heart Disease (A)

Text the program evaluates
(First paragraph of the wikipedia article: https://simple.wikipedia.org/wiki/Heart_disease)
Symptoms

A person can have heart disease and not feel sick. Some people with heart disease have symptoms. This is when there are changes or pain in the body to show a disease is there. Some symptoms of heart disease are:
-Pain in the chest—the heart muscle is not getting enough flow to keep it going.
- Trouble breathing—blood may back up into the lungs.
- Palpitations (a feeling that the heart is beating too fast, too hard, or not regularly).
- Swelling of feet or legs—blood is backing up from the heart into the lower body.
- Feeling weak because the body and brain are not getting enough blood to supply them with oxygen.
- Cyanosis (skin turning a blue colour) means that too little oxygen is in the bloodstream to supply the cells in the body.

Keywords selected by program:

- Blood (substance)
- Blue color (finding)
- Body part structure
- Brain structure (body structure)
- Cell structure
- Changing
- Cyanosis
- Difficulty breathing (finding)
- Disease
- Does turn
- Emotion (observable entity)
- Fast
- Foot swelling
- Frequent (qualifier value)
- Hard (qualifier value)
- Heart disease
- Heart structure (body structure)
- Lower
- Lung structure
- Mean
- Muscle structure
- Myocardium structure
- Oxygen
- Pain
- Palpitations (finding)
- Person (person)
- Skin structure
- Small
- Structure of back of trunk (body structure)
- Swelling (finding)
- Weak (qualifier value)
- ft

75

Survey for Ontology Summarization: Heart Disease https://docs.google.com/forms/u/0/d/1Zg_tyoTKfZWj0Wz_3k1nXEc...

1. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking:

1. Heart disease is equivalent to a Disease found in the heart.

2. Angina is a subclass of pain and a disease. It is associated with structural changes due to ischemia (a morphological abnormality) and occurs in the heart structure (a body structure)

3. Medical condition where a mother has a heart disease that complicates her pregnancy, childbirth, and/or puerperium. This medical condition is a subtype of "heart disease" and "disease of the circulatory system complicating pregnancy childbirth and puerperium". It also has a role group specifying the finding site, which is the heart structure.

4. Fetal echocardiography screening is a subclass of heart disease screening and Fetal echocardiography procedure. The procedure involves ultrasound imaging of the heart structure, with the intent of screening for heart disease. The method attribute is Ultrasound imaging - action, and the direct procedure site attribute is Heart structure. Additionally, the procedure has a focus attribute of heart disease.

5. Positron emission tomography using fluorodeoxyglucose with computed tomography myocardial rest imaging is equivalent to The procedure is a combination of positron emission tomography myocardial rest imaging using fluorodeoxyglucose and computed tomography imaging. The procedure is performed on the heart structure and involves the use of the substance fluorodeoxyglucose (18-F) as a tracer. The method of computed tomography imaging and positron emission tomographic imaging are used to produce images of the myocardium structure.

2. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

3. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

4. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

5. With regards to what **you** know about **heart disease**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

76

Symptoms of Heart Disease (B)

Text the program evaluates
(First paragraph of the wikipedia article: https://simple.wikipedia.org/wiki/Heart_disease)
Symptoms

A person can have heart disease and not feel sick. Some people with heart disease have symptoms. This is when there are changes or pain in the body to show a disease is there. Some symptoms of heart disease are:
-Pain in the chest—the heart muscle is not getting enough flow to keep it going.
- Trouble breathing—blood may back up into the lungs.
- Palpitations (a feeling that the heart is beating too fast, too hard, or not regularly).
- Swelling of feet or legs—blood is backing up from the heart into the lower body.
- Feeling weak because the body and brain are not getting enough blood to supply them with oxygen.
- Cyanosis (skin turning a blue colour) means that too little oxygen is in the bloodstream to supply the cells in the body.

Keywords selected by program:

- Blood (substance)
- Blue color (finding)
- Body part structure
- Brain structure (body structure)
- Cell structure
- Changing
- Cyanosis
- Difficulty breathing (finding)
- Disease
- Does turn
- Emotion (observable entity)
- Fast
- Foot swelling
- Frequent (qualifier value)
- Hard (qualifier value)
- Heart disease
- Heart structure (body structure)
- Lower
- Lung structure
- Mean
- Muscle structure
- Myocardium structure
- Oxygen
- Pain
- Palpitations (finding)
- Person (person)
- Skin structure
- Small
- Structure of back of trunk (body structure)
- Swelling (finding)
- Weak (qualifier value)
- ft

77

Sentence Ranking:

1. Heart disease is equivalent to a Disease found in the heart
2. Backache is equivalent to Pain located in the Entire back structure excluding neck
3. Pain of skin is equivalent to Pain and is found in skin
4. A cardiac complication is a disease that affects the heart and is caused by another underlying disease.
5. Palpitations are a type of clinical finding, characterized by the perception of an abnormal heart beat. They are associated with the heart structure and interpreted as awareness of heart beat.

6. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

7. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

8. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

9. With regards to what **you** know about **heart disease**, do you agree or disagree * with the statement:
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

78

Symptoms of Heart Disease (C)

Text the program evaluates
(First paragraph of the wikipedia article: https://simple.wikipedia.org /wiki/Heart_disease)
Symptoms

A person can have heart disease and not feel sick. Some people with heart disease have symptoms. This is when there are changes or pain in the body to show a disease is there. Some symptoms of heart disease are:
-Pain in the chest—the heart muscle is not getting enough flow to keep it going.
- Trouble breathing—blood may back up into the lungs.
- Palpitations (a feeling that the heart is beating too fast, too hard, or not regularly).
- Swelling of feet or legs—blood is backing up from the heart into the lower body.
- Feeling weak because the body and brain are not getting enough blood to supply them with oxygen.
- Cyanosis (skin turning a blue colour) means that too little oxygen is in the bloodstream to supply the cells in the body.

Keywords selected by program:

- Associated morphology (attribute)
- Awareness of heart beat (observable entity)
- Blood (substance)
- Blood gas
- Blood material (substance)
- Body fluid
- Body part structure
- Body region structure (body structure)
- Cardiac wall structure
- Chemical element (substance)
- Clinical finding (finding)
- Difficulty breathing (finding)
- Disease
- Emotion (observable entity)
- Finding of ease of respiration (finding)
- Finding site
- Foot structure (body structure)
- Foot swelling
- Heart and/or pericardium structure
- Heart disease
- Heart structure (body structure)
- Interprets
- Mental state (observable entity)
- Myocardium structure
- Oxygen
- Pain
- Pain / sensation finding
- Palpitations (finding)
- Person (person)
- Role group
- Social context
- Structure of striated muscle
- Structure of thoracic viscus
- Swelling (morphologic abnormality)
- Swelling of lower limb (finding)

Survey for Ontology Summarization: Heart Disease https://docs.google.com/forms/u/0/d/1Zg_tyoTKfZWj0Wz_3k1nXEc...

10. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking:

1. Heart disease is equivalent to a Disease found in the heart
2. Angina is a subclass of pain and a disease. It is associated with structural changes due to ischemia (a morphological abnormality) and occurs in the heart structure (a body structure)
3. Repetitive motion disorder is equivalent to Pain and Disease and is found in Body parts and is caused by Repetitive routines.
4. Malignant tumor of heart (disorder) is equivalent to Disease and is associated with Malignant neoplasm found in the Heart structure.
5. Benign neoplasm of heart is equivalent to Disease and associated with Neoplasm, benign found in heart structure.

11. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

12. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

13. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

14. With regards to what **you** know about **heart disease**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Symptoms of Heart Disease (D)

12 of 21 08/05/2023, 12:09

Text the program evaluates

(First paragraph of the wikipedia article: https://simple.wikipedia.org/wiki/Heart_disease)

Symptoms

A person can have heart disease and not feel sick. Some people with heart disease have symptoms. This is when there are changes or pain in the body to show a disease is there. Some symptoms of heart disease are:
-Pain in the chest—the heart muscle is not getting enough flow to keep it going.
- Trouble breathing—blood may back up into the lungs.
- Palpitations (a feeling that the heart is beating too fast, too hard, or not regularly).
- Swelling of feet or legs—blood is backing up from the heart into the lower body.
- Feeling weak because the body and brain are not getting enough blood to supply them with oxygen.
- Cyanosis (skin turning a blue colour) means that too little oxygen is in the bloodstream to supply the cells in the body.

80

Keywords selected by program:

- Associated morphology (attribute)
- Awareness of heart beat (observable entity)
- Blood (substance)
- Blood gas
- Blood material (substance)
- Body fluid
- Body part structure
- Body region structure (body structure)
- Cardiac wall structure
- Chemical element (substance)
- Clinical finding (finding)
- Difficulty breathing (finding)
- Disease
- Emotion (observable entity)
- Finding of ease of respiration (finding)
- Finding site
- Foot structure (body structure)
- Foot swelling
- Heart and/or pericardium structure
- Heart disease
- Heart structure (body structure)
- Interprets
- Mental state (observable entity)
- Myocardium structure
- Oxygen
- Pain
- Pain / sensation finding
- Palpitations (finding)
- Person (person)
- Role group
- Social context
- Structure of striated muscle
- Structure of thoracic viscus
- Swelling (morphologic abnormality)
- Swelling of lower limb (finding)

Survey for Ontology Summarization: Heart Disease https://docs.google.com/forms/u/0/d/1Zg_tyoTKfZWj0Wz_3k1nXEc...

Sentence Ranking:

1. Heart disease is equivalent to a Disease found in the heart
2. Palpitations are a type of clinical finding, characterized by the perception of an abnormal heart beat. They are associated with the heart structure and interpreted as awareness of heart beat.
3. A "structural disorder of the heart" refers to a disease that affects the structure of the heart and is characterized by morphological abnormalities in the heart structure.
4. Myocardial disease is equivalent to disease found in myocardium structure.
5. Finding related to awareness of heartbeat is a type of clinical finding and is associated with the heart structure as its finding site. Additionally, it is interpreted by the awareness of a heartbeat.

15. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

16. With regards to what **you** know about **heart disease**, do you agree or disagree * with the statement:
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

17. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

18. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Ranking the tables: **Symptoms of Heart Disease**

In this section you will rank the 4 different tables on how which one describes the text best.

Text:

Symptoms

A person can have heart disease and not feel sick. Some people with heart disease have symptoms. This is when there are changes or pain in the body to show a disease is there. Some symptoms of heart disease are:
-Pain in the chest—the heart muscle is not getting enough flow to keep it going.
- Trouble breathing—blood may back up into the lungs.
- Palpitations (a feeling that the heart is beating too fast, too hard, or not regularly).
- Swelling of feet or legs—blood is backing up from the heart into the lower body.
- Feeling weak because the body and brain are not getting enough blood to supply them with oxygen.
- Cyanosis (skin turning a blue colour) means that too little oxygen is in the bloodstream to supply the cells in the body.

82

Survey for Ontology Summarization: Heart Disease https://docs.google.com/forms/u/0/d/1Zg_tyoTKfZWj0Wz_3k1nXEc...

Symptoms of Heart Disease (A)

1. Heart disease is equivalent to a Disease found in the heart.

2. Angina is a subclass of pain and a disease. It is associated with structural changes due to ischemia (a morphological abnormality) and occurs in the heart structure (a body structure)

3. Medical condition where a mother has a heart disease that complicates her pregnancy, childbirth, and/or puerperium. This medical condition is a subtype of "heart disease" and "disease of the circulatory system complicating pregnancy childbirth and puerperium". It also has a role group specifying the finding site, which is the heart structure.

4. Fetal echocardiography screening is a subclass of heart disease screening and Fetal echocardiography procedure. The procedure involves ultrasound imaging of the heart structure, with the intent of screening for heart disease. The method attribute is Ultrasound imaging - action, and the direct procedure site attribute is Heart structure. Additionally, the procedure has a focus attribute of heart disease.

5. Positron emission tomography using fluorodeoxyglucose with computed tomography myocardial rest imaging is equivalent to The procedure is a combination of positron emission tomography myocardial rest imaging using fluorodeoxyglucose and computed tomography imaging. The procedure is performed on the heart structure and involves the use of the substance fluorodeoxyglucose (18-F) as a tracer. The method of computed tomography imaging and positron emission tomographic imaging are used to produce images of the myocardium structure.

Symptoms of Heart Disease (B)

1. Heart disease is equivalent to a Disease found in the heart
2. Backache is equivalent to Pain located in the Entire back structure excluding neck
3. Pain of skin is equivalent to Pain and is found in skin
4. A cardiac complication is a disease that affects the heart and is caused by another underlying disease.
5. Palpitations are a type of clinical finding, characterized by the perception of an abnormal heart beat. They are associated with the heart structure and interpreted as awareness of heart beat.

Symptoms of Heart Disease (C)

1. Heart disease is equivalent to a Disease found in the heart
2. Angina is a subclass of pain and a disease. It is associated with structural changes due to ischemia (a morphological abnormality) and occurs in the heart structure (a body structure)
3. Repetitive motion disorder is equivalent to Pain and Disease and is found in Body parts and is caused by Repetitive routines.
4. Malignant tumor of heart (disorder) is equivalent to Disease and is associated with Malignant neoplasm found in the Heart structure.
5. Benign neoplasm of heart is equivalent to Disease and associated with Neoplasm, benign found in heart structure.

83

Symptoms of Heart Disease (D)

1. Heart disease is equivalent to a Disease found in the heart
2. Palpitations are a type of clinical finding, characterized by the perception of an abnormal heart beat. They are associated with the heart structure and interpreted as awareness of heart beat.
3. A "structural disorder of the heart" refers to a disease that affects the structure of the heart and is characterized by morphological abnormalities in the heart structure.
4. Myocardial disease is equivalent to disease found in myocardium structure.
5. Finding related to awareness of heartbeat is a type of clinical finding and is associated with the heart structure as its finding site. Additionally, it is interpreted by the awareness of a heartbeat.

19. Rank the collections of sentences from best to worst *

Mark only one oval per row.

	A	B	C	D
Best				
2nd Best				
3rd Best				
4th Best				

Closing comments

Thank you for taking the survey!

20. If you have any comments to the survey please leave them here!

8.3.4 Survey: Covid-19

Survey for Ontology Summarization: Covid-19

This is a collection of surveys made for the purpose of evaluating the results of Martin Lam's Master's project. The project is about summarizing **knowledge bases** (Ontologies).
To summarize knowledge bases a program has been created that ranks and extracts *axioms*. Axioms can be viewed as **sentences**. Do keep in mind that the selected sentences were selected from a pool of 360'000 sentences.
This is 1 of 5 available surveys.
This survey will cover the topic: **Covid-19**

The survey is divided into **5** sections (A, B, C, D and ranking).

The **text** that the program evaluates is the same for **all** sections.

The first **4** sections each have their own enumerated list of ranked sentences. There may be shared sentences between them. The ranking in the lists are ordered in descending order, where rank 1 is the most relevant sentence and rank 5 is the least relevant sentence among those selected.

You will be presented with **2** different lists of keywords. The keyword list in section (A) and (B) is the same and the list in (C) and (D) is the same.

Lastly you will be asked to rank the 4 enumerated lists.

For the survey to have the best results, I kindly ask that you answer as honestly as you can.

* Indicates required question

General information of Covid-19 (A)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/COVID-19)
Coronavirus disease 2019 (COVID-19) is a contagious disease caused by a virus, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China, in December 2019. The disease quickly spread worldwide, resulting in the COVID-19 pandemic.

The symptoms of COVID-19 are variable but often include fever, cough, headache, fatigue, breathing difficulties, loss of smell, and loss of taste. Symptoms may begin one to fourteen days after exposure to the virus. At least a third of people who are infected do not develop noticeable symptoms. Of those who develop symptoms noticeable enough to be classified as patients, most (81%) develop mild to moderate symptoms (up to mild pneumonia), while 14% develop severe symptoms (dyspnea, hypoxia, or more than 50% lung involvement on imaging), and 5% develop critical symptoms (respiratory failure, shock, or multiorgan dysfunction). Older people are at a higher risk of developing severe symptoms. Some people continue to experience a range of effects (long COVID) for months after recovery, and damage to organs has been observed. Multi-year studies are underway to further investigate the long-term effects of the disease.

Keywords selected by program:

- Acquired (qualifier value)
- Affecting
- Air
- Blood (substance)
- Breathing process (qualifier value)
- Chest pain
- Classified (qualifier value)
- Combined (qualifier value)
- Condition
- Condition severity (attribute)
- Confirmatory technique
- Cough
- Diagnosis
- Difficult
- Difficulty
- Difficulty breathing (finding)
- Disease
- Disorder due to infection
- Domain Bacteria (organism)
- Dry
- Dry cough
- Examination - action (qualifier value)
- Fever (finding)
- Frequent (qualifier value)
- Identified
- Including (qualifier value)
- Inflammation
- Inflammatory disorder (disorder)
- Known (qualifier value)
- Lung structure
- Microbial culture (procedure)
- Microbiology
- Pain
- Pathogenic organism
- Physical
- Plain chest X-ray (procedure)
- Pneumonia
- Responsible to (attribute)
- Severity (attribute)
- Small
- Sputum (substance)
- Tests (qualifier value)
- Thoracic structure
- Variable

- Virus

1. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking:

1. Severe acute respiratory syndrome is a disease caused by the SARS coronavirus that affects the respiratory system structure and is characterized by an infectious process.
2. Pneumonia caused by the Severe Acute Respiratory Syndrome coronavirus is equivalent to a disease that is associated with inflammation and consolidation, caused by the SARS coronavirus, located in the structure of the lung, and resulting from an infectious process.
3. An acute disease is equivalent to a disease that has a sudden onset and/or short duration.
4. A disorder of the respiratory system is equivalent to a disease that is associated with the respiratory system structure.
5. "Acute respiratory disease" as a subclass of disease that has a sudden onset or short duration and affects the structures of the respiratory system.

2. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

3. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

4. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

5. With regards to what **you** know about **covid-19**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

88

General information of Covid-19 (B)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/COVID-19)

Coronavirus disease 2019 (COVID-19) is a contagious disease caused by a virus, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China, in December 2019. The disease quickly spread worldwide, resulting in the COVID-19 pandemic.

The symptoms of COVID-19 are variable but often include fever, cough, headache, fatigue, breathing difficulties, loss of smell, and loss of taste. Symptoms may begin one to fourteen days after exposure to the virus. At least a third of people who are infected do not develop noticeable symptoms. Of those who develop symptoms noticeable enough to be classified as patients, most (81%) develop mild to moderate symptoms (up to mild pneumonia), while 14% develop severe symptoms (dyspnea, hypoxia, or more than 50% lung involvement on imaging), and 5% develop critical symptoms (respiratory failure, shock, or multiorgan dysfunction). Older people are at a higher risk of developing severe symptoms. Some people continue to experience a range of effects (long COVID) for months after recovery, and damage to organs has been observed. Multi-year studies are underway to further investigate the long-term effects of the disease.

Keywords selected by program:

- Acquired (qualifier value)
- Affecting
- Air
- Blood (substance)
- Breathing process (qualifier value)
- Chest pain
- Classified (qualifier value)
- Combined (qualifier value)
- Condition
- Condition severity (attribute)
- Confirmatory technique
- Cough
- Diagnosis
- Difficult
- Difficulty
- Difficulty breathing (finding)
- Disease
- Disorder due to infection
- Domain Bacteria (organism)
- Dry
- Dry cough
- Examination - action (qualifier value)
- Fever (finding)
- Frequent (qualifier value)
- Identified
- Including (qualifier value)
- Inflammation
- Inflammatory disorder (disorder)
- Known (qualifier value)
- Lung structure
- Microbial culture (procedure)
- Microbiology
- Pain
- Pathogenic organism
- Physical
- Plain chest X-ray (procedure)
- Pneumonia
- Responsible to (attribute)
- Severity (attribute)
- Small
- Sputum (substance)
- Tests (qualifier value)
- Thoracic structure
- Variable

89

- Virus

Sentence Ranking:

1. Severe acute respiratory syndrome is a disease caused by the SARS coronavirus that affects the respiratory system structure and is characterized by an infectious process.
2. Pneumonia caused by the Severe Acute Respiratory Syndrome coronavirus is equivalent to a disease that is associated with inflammation and consolidation, caused by the SARS coronavirus, located in the structure of the lung, and resulting from an infectious process.
3. An acute disease is equivalent to a disease that has a sudden onset and/or short duration.
4. A disorder of the respiratory system is equivalent to a disease that is associated with the respiratory system structure.
5. "Acute respiratory disease" as a subclass of disease that has a sudden onset or short duration and affects the structures of the respiratory system.

6. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

7. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

8. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

9. With regards to what **you** know about **covid-19**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

90

General information of Covid-19 (C)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/COVID-19)

Coronavirus disease 2019 (COVID-19) is a contagious disease caused by a virus, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China, in December 2019. The disease quickly spread worldwide, resulting in the COVID-19 pandemic.

The symptoms of COVID-19 are variable but often include fever, cough, headache, fatigue, breathing difficulties, loss of smell, and loss of taste. Symptoms may begin one to fourteen days after exposure to the virus. At least a third of people who are infected do not develop noticeable symptoms. Of those who develop symptoms noticeable enough to be classified as patients, most (81%) develop mild to moderate symptoms (up to mild pneumonia), while 14% develop severe symptoms (dyspnea, hypoxia, or more than 50% lung involvement on imaging), and 5% develop critical symptoms (respiratory failure, shock, or multiorgan dysfunction). Older people are at a higher risk of developing severe symptoms. Some people continue to experience a range of effects (long COVID) for months after recovery, and damage to organs has been observed. Multi-year studies are underway to further investigate the long-term effects of the disease.

Keywords selected by program:

- Affecting
- Associated morphology (attribute)
- Chest pain
- Cough
- Degree findings
- Difficulty breathing (finding)
- Disease
- Dry cough
- Finding of ease of respiration (finding)
- Finding site
- General adjectival modifier
- Inflammation
- Inflammation and consolidation (morphologic abnormality)
- Inflammatory disorder (disorder)
- Inflammatory morphology
- Lower respiratory tract structure
- Lung structure
- Magnitudes
- Method (attribute)
- Pain
- Pathological process
- Plain X-ray imaging - action
- Plain chest X-ray (procedure)
- Pneumonia
- Procedure site - Direct
- Pulmonary structure including vessels and lymphoid tissue (body structure)
- Radiographic imaging procedure (procedure)
- Role group
- Small
- Structure of organ in respiratory system
- Structure of pulmopleural compartment
- Structure of thoracic viscus
- Thoracic structure

91

10. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking:

1. Severe acute respiratory syndrome is a disease caused by the SARS coronavirus that affects the respiratory system structure and is characterized by an infectious process.

2. "Coronavirus infection (disorder)" is equivalent to a disease caused by a member of the subfamily Orthocoronavirinae, and characterized by an infectious process

3. A healthcare-associated severe acute respiratory syndrome (SARS) is equivalent to a disease that is caused by the SARS coronavirus, affects the respiratory system, has an infectious process as a pathological process, and occurs during hospital care.

4. Severe acute respiratory syndrome of upper respiratory tract is a subclass of Disease caused by SARS coronavirus characterized by an infectious process and the location of the pathological process in the body structure of the upper respiratory tract.

5. "Acute respiratory disease" as a subclass of disease that has a sudden onset or short duration and affects the structures of the respiratory system.

11. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

12. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

13. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

14. With regards to what **you** know about **covid-19**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

92

General information of Covid-19 (D)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/COVID-19)

Coronavirus disease 2019 (COVID-19) is a contagious disease caused by a virus, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China, in December 2019. The disease quickly spread worldwide, resulting in the COVID-19 pandemic.

The symptoms of COVID-19 are variable but often include fever, cough, headache, fatigue, breathing difficulties, loss of smell, and loss of taste. Symptoms may begin one to fourteen days after exposure to the virus. At least a third of people who are infected do not develop noticeable symptoms. Of those who develop symptoms noticeable enough to be classified as patients, most (81%) develop mild to moderate symptoms (up to mild pneumonia), while 14% develop severe symptoms (dyspnea, hypoxia, or more than 50% lung involvement on imaging), and 5% develop critical symptoms (respiratory failure, shock, or multiorgan dysfunction). Older people are at a higher risk of developing severe symptoms. Some people continue to experience a range of effects (long COVID) for months after recovery, and damage to organs has been observed. Multi-year studies are underway to further investigate the long-term effects of the disease.

Keywords selected by program:

- Affecting
- Associated morphology (attribute)
- Chest pain
- Cough
- Degree findings
- Difficulty breathing (finding)
- Disease
- Dry cough
- Finding of ease of respiration (finding)
- Finding site
- General adjectival modifier
- Inflammation
- Inflammation and consolidation (morphologic abnormality)
- Inflammatory disorder (disorder)
- Inflammatory morphology
- Lower respiratory tract structure
- Lung structure
- Magnitudes
- Method (attribute)
- Pain
- Pathological process
- Plain X-ray imaging - action
- Plain chest X-ray (procedure)
- Pneumonia
- Procedure site - Direct
- Pulmonary structure including vessels and lymphoid tissue (body structure)
- Radiographic imaging procedure (procedure)
- Role group
- Small
- Structure of organ in respiratory system
- Structure of pulmopleural compartment
- Structure of thoracic viscus
- Thoracic structure

93

Sentence Ranking:

1. Severe acute respiratory syndrome is a disease caused by the SARS coronavirus that affects the respiratory system structure and is characterized by an infectious process.
2. "Coronavirus infection (disorder)" is equivalent to a disease caused by a member of the subfamily Orthocoronavirinae, and characterized by an infectious process
3. "Acute respiratory disease" as a subclass of disease that has a sudden onset or short duration and affects the structures of the respiratory system.
4. A disorder of the respiratory system is equivalent to a disease that is associated with the respiratory system structure.
5. An acute disease is equivalent to a disease that has a sudden onset and/or short duration.

15. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

16. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

17. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

18. With regards to what **you** know about **covid-19**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

94

Ranking the tables: Covid-19

In this section you will rank the 4 different tables on how which one describes the text best.

Text:
Coronavirus disease 2019 (COVID-19) is a contagious disease caused by a virus, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China, in December 2019. The disease quickly spread worldwide, resulting in the COVID-19 pandemic.

The symptoms of COVID-19 are variable but often include fever, cough, headache, fatigue, breathing difficulties, loss of smell, and loss of taste. Symptoms may begin one to fourteen days after exposure to the virus. At least a third of people who are infected do not develop noticeable symptoms. Of those who develop symptoms noticeable enough to be classified as patients, most (81%) develop mild to moderate symptoms (up to mild pneumonia), while 14% develop severe symptoms (dyspnea, hypoxia, or more than 50% lung involvement on imaging), and 5% develop critical symptoms (respiratory failure, shock, or multiorgan dysfunction). Older people are at a higher risk of developing severe symptoms. Some people continue to experience a range of effects (long COVID) for months after recovery, and damage to organs has been observed. Multi-year studies are underway to further investigate the long-term effects of the disease.

Covid-19 (A)

1. Severe acute respiratory syndrome is a disease caused by the SARS coronavirus that affects the respiratory system structure and is characterized by an infectious process.
2. Pneumonia caused by the Severe Acute Respiratory Syndrome coronavirus is equivalent to a disease that is associated with inflammation and consolidation, caused by the SARS coronavirus, located in the structure of the lung, and resulting from an infectious process.
3. An acute disease is equivalent to a disease that has a sudden onset and/or short duration.
4. A disorder of the respiratory system is equivalent to a disease that is associated with the respiratory system structure.
5. "Acute respiratory disease" as a subclass of disease that has a sudden onset or short duration and affects the structures of the respiratory system.

Covid-19 (B)

1. Severe acute respiratory syndrome is a disease caused by the SARS coronavirus that affects the respiratory system structure and is characterized by an infectious process.
2. Pneumonia caused by the Severe Acute Respiratory Syndrome coronavirus is equivalent to a disease that is associated with inflammation and consolidation, caused by the SARS coronavirus, located in the structure of the lung, and resulting from an infectious process.
3. An acute disease is equivalent to a disease that has a sudden onset and/or short duration.
4. A disorder of the respiratory system is equivalent to a disease that is associated with the respiratory system structure.
5. "Acute respiratory disease" as a subclass of disease that has a sudden onset or short duration and affects the structures of the respiratory system.

Covid-19 (C)

1. Severe acute respiratory syndrome is a disease caused by the SARS coronavirus that affects the respiratory system structure and is characterized by an infectious process.
2. "Coronavirus infection (disorder)" is equivalent to a disease caused by a member of the subfamily Orthocoronavirinae, and characterized by an infectious process
3. A healthcare-associated severe acute respiratory syndrome (SARS) is equivalent to a disease that is caused by the SARS coronavirus, affects the respiratory system, has an infectious process as a pathological process, and occurs during hospital care.
4. Severe acute respiratory syndrome of upper respiratory tract is a subclass of Disease caused by SARS coronavirus characterized by an infectious process and the location of the pathological process in the body structure of the upper respiratory tract.
5. "Acute respiratory disease" as a subclass of disease that has a sudden onset or short duration and affects the structures of the respiratory system.

Covid-19 (D)

1. Severe acute respiratory syndrome is a disease caused by the SARS coronavirus that affects the respiratory system structure and is characterized by an infectious process.
2. "Coronavirus infection (disorder)" is equivalent to a disease caused by a member of the subfamily Orthocoronavirinae, and characterized by an infectious process
3. "Acute respiratory disease" as a subclass of disease that has a sudden onset or short duration and affects the structures of the respiratory system.
4. A disorder of the respiratory system is equivalent to a disease that is associated with the respiratory system structure.
5. An acute disease is equivalent to a disease that has a sudden onset and/or short duration.

95

19. Rank the collections of sentences from best to worst *

Mark only one oval per row.

	A	B	C	D
Best	◯	◯	◯	◯
2nd Best	◯	◯	◯	◯
3rd Best	◯	◯	◯	◯
4th Best	◯	◯	◯	◯

Closing comments

Thank you for taking the survey!

20. If you have any comments to the survey please leave them here!

8.3.5 Survey: Pneumonia

Survey for Ontology Summarization: Pneumonia https://docs.google.com/forms/u/0/d/1QlgqcqYU74u2K0uRmdDJsM...

Survey for Ontology Summarization: Pneumonia

This is a collection of surveys made for the purpose of evaluating the results of Martin Lam's Master's project. The project is about summarizing **knowledge bases** (Ontologies).
To summarize knowledge bases a program has been created that ranks and extracts *axioms*. Axioms can be viewed as **sentences**. Do keep in mind that the selected sentences were selected from a pool of 360'000 sentences.
This is 1 of 5 available surveys.
This survey will cover the topic: **Pneumonia**

The survey is divided into **5** sections (A, B, C, D and ranking).

The **text** that the program evaluates is the same for **all** sections.

The first **4** sections each have their own enumerated list of ranked sentences. There may be shared sentences between them. The ranking in the lists are ordered in descending order, where rank 1 is the most relevant sentence and rank 5 is the least relevant sentence among those selected.

You will be presented with **2** different lists of keywords. The keyword list in section (A) and (B) is the same and the list in (C) and (D) is the same.

Lastly you will be asked to rank the 4 enumerated lists.

For the survey to have the best results, I kindly ask that you answer as honestly as you can.

97

* Indicates required question

General information of Pneumonia (A)

Text the program evaluates
(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Pneumonia)
Pneumonia is an inflammatory condition of the lung primarily affecting the small air sacs known as alveoli. Symptoms typically include some combination of productive or dry cough, chest pain, fever, and difficulty breathing. The severity of the condition is variable.

Pneumonia is usually caused by infection with viruses or bacteria, and less commonly by other microorganisms. Identifying the responsible pathogen can be difficult. Diagnosis is often based on symptoms and physical examination. Chest X-rays, blood tests, and culture of the sputum may help confirm the diagnosis. The disease may be classified by where it was acquired, such as community- or hospital-acquired or healthcare-associated pneumonia

2 of 18 08/05/2023, 12:10

Keywords selected by program:

- 1 (qualifier value)
- 14
- Accident due to exposure to weather conditions (event)
- Acute disease (disorder)
- Acute respiratory disease (disorder)
- Acuteness (qualifier value)
- Body organ structure (body structure)
- Breath (substance)
- Case - unit of product usage
- China (geographic location)
- Classified (qualifier value)
- Concept model range
- Contagious disease (navigational concept)
- Coronavirus infection (disorder)
- Cough
- Damage
- Difficulty
- Disease
- Disease caused by severe acute respiratory syndrome coronavirus 2 (disorder)
- Disorder due to infection
- Disorder of respiratory system (disorder)
- Dyspnea (finding)
- Effect (qualifier value)
- Fatigue (finding)
- Fever (finding)
- First
- Frequent (qualifier value)
- Headache (finding)
- Hypoxia
- Identified
- Imaging
- Including (qualifier value)
- Investigates (attribute)
- Involvement
- Just noticeable
- Known (qualifier value)
- Long
- Lung structure
- Mild (qualifier value)
- Moderate (severity modifier) (qualifier value)
- Old (qualifier value)
- Patient (person)
- Pneumonia
- Rapidly
- Recovering from
- Respiratory function (observable entity)
- Resulting in
- Risk of (contextual qualifier) (qualifier value)
- SARS coronavirus
- Sense of smell
- Severe
- Severe acute respiratory syndrome
- Shock (disorder)
- Spread (attribute)
- Study (environment)
- Taste
- Third (qualifier value)
- Variable
- Virus
- day (qualifier value)
- function (observable entity)
- function (observable entity)
- month

1. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Selected Sentences

1. Post-inflammatory pulmonary fibrosis is equivalent to a disease that is associated with Fibrosis in the lungs, which occurs after an inflammatory disorder.
2. Pneumonia is equivalent to a Disease associated with Inflammation and consolidation in the lungs.
3. Chest pain is equivalent to Pain in the thoracic structure.
4. Atypical pneumonia is a subclass of Disease associated with inflammation caused by Domain bacteria in the lungs characterized by an infectious process.
5. Bacterial pneumonia is equivalent to a Disease associated with Inflammation and consolidation caused by Domain bacteria in the lungs characterized by an infectious process.

2. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

3. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

4. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

5. With regards to what **you** know about **pneumonia**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

General information of Pneumonia (B)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Pneumonia)

Pneumonia is an inflammatory condition of the lung primarily affecting the small air sacs known as alveoli. Symptoms typically include some combination of productive or dry cough, chest pain, fever, and difficulty breathing. The severity of the condition is variable.

Pneumonia is usually caused by infection with viruses or bacteria, and less commonly by other microorganisms. Identifying the responsible pathogen can be difficult. Diagnosis is often based on symptoms and physical examination. Chest X-rays, blood tests, and culture of the sputum may help confirm the diagnosis. The disease may be classified by where it was acquired, such as community- or hospital-acquired or healthcare-associated pneumonia

Keywords selected by program:

- 1 (qualifier value)
- 14
- Accident due to exposure to weather conditions (event)
- Acute disease (disorder)
- Acute respiratory disease (disorder)
- Acuteness (qualifier value)
- Body organ structure (body structure)
- Breath (substance)
- Case - unit of product usage
- China (geographic location)
- Classified (qualifier value)
- Concept model range
- Contagious disease (navigational concept)
- Coronavirus infection (disorder)
- Cough
- Damage
- Difficulty
- Disease
- Disease caused by severe acute respiratory syndrome coronavirus 2 (disorder)
- Disorder due to infection
- Disorder of respiratory system (disorder)
- Dyspnea (finding)
- Effect (qualifier value)
- Fatigue (finding)
- Fever (finding)
- First
- Frequent (qualifier value)
- Headache (finding)
- Hypoxia
- Identified
- Imaging
- Including (qualifier value)
- Investigates (attribute)
- Involvement
- Just noticeable
- Known (qualifier value)
- Long
- Lung structure
- Mild (qualifier value)
- Moderate (severity modifier) (qualifier value)
- Old (qualifier value)
- Patient (person)
- Pneumonia
- Rapidly
- Recovering from
- Respiratory function (observable entity)
- Resulting in
- Risk of (contextual qualifier) (qualifier value)
- SARS coronavirus
- Sense of smell
- Severe
- Severe acute respiratory syndrome
- Shock (disorder)
- Spread (attribute)
- Study (environment)
- Taste
- Third (qualifier value)
- Variable
- Virus
- day (qualifier value)
- function (observable entity)
- function (observable entity)
- month

Sentence Ranking:

1. Chest pain is equivalent to Pain in the thoracic structure.
2. Dry cough is a subclass of Cough
3. Pneumonia is equivalent to a Disease associated with Inflammation and consolidation in the lungs.
4. Inflammatory disorder is equivalent to Disease associated with Inflammatory morphology
5. Disorder of lung is equivalent to Disease found in the lungs.

100

6. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

7. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

8. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

9. With regards to what **you** know about **pneumonia**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

101

General information of Pneumonia (C)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Pneumonia)

Pneumonia is an inflammatory condition of the lung primarily affecting the small air sacs known as alveoli. Symptoms typically include some combination of productive or dry cough, chest pain, fever, and difficulty breathing. The severity of the condition is variable.

Pneumonia is usually caused by infection with viruses or bacteria, and less commonly by other microorganisms. Identifying the responsible pathogen can be difficult. Diagnosis is often based on symptoms and physical examination. Chest X-rays, blood tests, and culture of the sputum may help confirm the diagnosis. The disease may be classified by where it was acquired, such as community- or hospital-acquired or healthcare-associated pneumonia

Keywords selected by program:

- Acute disease (disorder)
- Acute respiratory disease (disorder)
- Causative agent
- China (geographic location)
- Clinical course (attribute)
- Clinical finding (finding)
- Contagious disease (navigational concept)
- Coronavirus infection (disorder)
- Disease
- Disorder of respiratory system (disorder)
- Far east country
- Finding site
- Human coronavirus (organism)
- Infectious process (qualifier value)
- Navigational concept
- Pathological process
- Role group
- SARS coronavirus
- Severe
- Severe acute respiratory syndrome

10. The selected keywords describe the text well *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

Sentence Ranking:

1. Pneumonia is equivalent to a Disease associated with Inflammation and consolidation in the lungs.
2. Post-inflammatory pulmonary fibrosis is equivalent to a disease that is associated with Fibrosis in the lungs, which occurs after an inflammatory disorder.
3. Lung structure is a subclass of Structure of thoracic viscusand Pulmonary structure including vessels and lymphoid tissue and Structure of pulmopleural compartment and Structure of organ in respiratory system and Lower respiratory tract structure.
4. Non-infectios pneumonia is a subclass of Disease and is associated with inflammation and consolidation found in the lungs.
5. Community acquired pneumonia is a subclass of Disease and is associated with inflammation and consolidation found in the lungs.

11. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

12. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

13. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

14. With regards to what **you** know about **pneumonia**, do you agree or disagree with the statement: *
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

General information of Pneumonia (D)

Text the program evaluates

(First paragraph of the wikipedia article: https://en.wikipedia.org/wiki/Pneumonia)

Pneumonia is an inflammatory condition of the lung primarily affecting the small air sacs known as alveoli. Symptoms typically include some combination of productive or dry cough, chest pain, fever, and difficulty breathing. The severity of the condition is variable.

Pneumonia is usually caused by infection with viruses or bacteria, and less commonly by other microorganisms. Identifying the responsible pathogen can be difficult. Diagnosis is often based on symptoms and physical examination. Chest X-rays, blood tests, and culture of the sputum may help confirm the diagnosis. The disease may be classified by where it was acquired, such as community- or hospital-acquired or healthcare-associated pneumonia

Keywords selected by program:

- Acute disease (disorder)
- Acute respiratory disease (disorder)
- Causative agent
- China (geographic location)
- Clinical course (attribute)
- Clinical finding (finding)
- Contagious disease (navigational concept)
- Coronavirus infection (disorder)
- Disease
- Disorder of respiratory system (disorder)
- Far east country
- Finding site
- Human coronavirus (organism)
- Infectious process (qualifier value)
- Navigational concept
- Pathological process
- Role group
- SARS coronavirus
- Severe
- Severe acute respiratory syndrome

103

Sentence Ranking:

1. Pneumonia is equivalent to a Disease associated with Inflammation and consolidation in the lungs.
2. Lung structure is a subclass of Structure of thoracic viscusand Pulmonary structure including vessels and lymphoid tissue and Structure of pulmopleural compartment and Structure of organ in respiratory system and Lower respiratory tract structure.
3. Plain chest X-ray is equivalent to Radiographic imaging procedure and is a method of Plain X-ray imaging of the Thoracic
4. Inflammatory disorder is equivalent to Disease and is associated with Inflammatory morphology.
5. Pneumonitis is equivalent to Disease and is associated with Inflammatory morphology found in the lungs.

15. The selected sentences make sense according to the keywords *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

16. The selected sentences provide an overview of what the entire knowledge base has to offer. *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

17. The selected sentences are informative *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

18. With regards to what **you** know about **pneumonia**, do you agree or disagree * with the statement:
The ranking of the sentences are good.

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly Agree

104

Ranking the tables: Pneumonia

In this section you will rank the 4 different tables on how which one describes the text best.

Text:
Pneumonia is an inflammatory condition of the lung primarily affecting the small air sacs known as alveoli. Symptoms typically include some combination of productive or dry cough, chest pain, fever, and difficulty breathing. The severity of the condition is variable.

Pneumonia is usually caused by infection with viruses or bacteria, and less commonly by other microorganisms. Identifying the responsible pathogen can be difficult. Diagnosis is often based on symptoms and physical examination. Chest X-rays, blood tests, and culture of the sputum may help confirm the diagnosis. The disease may be classified by where it was acquired, such as community- or hospital-acquired or healthcare-associated pneumonia

Pneumonia (A)

1. Post-inflammatory pulmonary fibrosis is equivalent to a disease that is associated with Fibrosis in the lungs,
   which occurs after an inflammatory disorder.
2. Pneumonia is equivalent to a Disease associated with Inflammation and consolidation in the lungs.
3. Chest pain is equivalent to Pain in the thoracic structure.
4. Atypical pneumonia is a subclass of Disease associated with inflammation caused by Domain bacteria in the lungs characterized by an infectious process.
5. Bacterial pneumonia is equivalent to a Disease associated with Inflammation and consolidation caused by Domain bacteria in the lungs characterized by an infectious process.

Pneumonia (B)

1. Chest pain is equivalent to Pain in the thoracic structure.
2. Dry cough is a subclass of Cough
3. Pneumonia is equivalent to a Disease associated with Inflammation and consolidation in the lungs.
4. Inflammatory disorder is equivalent to Disease associated with Inflammatory morphology
5. Disorder of lung is equivalent to Disease found in the lungs.

Pneumonia (C)

1. Pneumonia is equivalent to a Disease associated with Inflammation and consolidation in the lungs.
2. Post-inflammatory pulmonary fibrosis is equivalent to a disease that is associated with Fibrosis in the lungs, which occurs after an inflammatory disorder.
3. Lung structure is a subclass of Structure of thoracic viscusand Pulmonary structure including vessels and lymphoid tissue and Structure of pulmopleural compartment and Structure of organ in respiratory system and Lower respiratory tract structure.
4. Non-infectios pneumonia is a subclass of Disease and is associated with inflammation and consolidation found in the lungs.
5. Community acquired pneumonia is a subclass of Disease and is associated with inflammation and consolidation found in the lungs.

Pneumonia (D)

1. Pneumonia is equivalent to a Disease associated with Inflammation and consolidation in the lungs.
2. Lung structure is a subclass of Structure of thoracic viscusand Pulmonary structure including vessels and lymphoid tissue and Structure of pulmopleural compartment and Structure of organ in respiratory system and Lower respiratory tract structure.
3. Plain chest X-ray is equivalent to Radiographic imaging procedure and is a method of Plain X-ray imaging of the Thoracic
4. Inflammatory disorder is equivalent to Disease and is associated with Inflammatory morphology.
5. Pneumonitis is equivalent to Disease and is associated with Inflammatory morphology found in the lungs.

105

19. Rank the collections of sentences from best to worst *

Mark only one oval per row.

	A	B	C	D
Best				
2nd Best				
3rd Best				
4th Best				

Closing comments

Thank you for taking the survey!

20. If you have any comments to the survey please leave them here!

Bibliography

[1] Sergey Astanin. *python-tabulate*. Accessed: Nov 17, 2022. 2021. URL: https://github.com/astanin/python-tabulate.

[2] Michael K Bergman. 'A common sense view of knowledge graphs'. In: (2019). URL: https://www.mkbergman.com/2244/a-common-sense-view-of-knowledge-graphs/.

[3] Stephen Bonner et al. 'Understanding the performance of knowledge graph embeddings in drug discovery'. In: *Artificial Intelligence in the Life Sciences* 2 (Dec. 2022), p. 100036. DOI: 10.1016/j.ailsci.2022.100036. URL: https://doi.org/10.1016%2Fj.ailsci.2022.100036.

[4] Antoine Bordes et al. 'Translating Embeddings for Modeling Multi-relational Data'. In: *Neural Information Processing Systems (NIPS)*. South Lake Tahoe, United States, Dec. 2013, pp. 1–9. (Visited on 05/07/2022).

[5] Ricardo Campos et al. *YAKE! Keyword extraction from single documents using multiple local features*. en. DOI: 10.1016/j.ins.2019.09.013. URL: https://www.sciencedirect.com/science/article/pii/S0020025519308588 (visited on 14/02/2023).

[6] Jiaoyan Chen et al. 'OWL2Vec*: Embedding of OWL Ontologies'. en. In: *arXiv:2009.14654 [cs]* (Jan. 2021). arXiv: 2009.14654. (Visited on 14/01/2022).

[7] Jieying Chen, Michel Ludwig and Dirk Walther. 'Computing Minimal Subsumption Modules of Ontologies'. In: *Global Conference on Artificial Intelligence*. 2018.

[8] Jieying Chen, Yue Ma and Dirk Walther. 'Computing Best Ontology Excerpts via Weighted Partial Max-SAT Solving'. en. In: (), p. 12.

[9] Jieying Chen et al. 'Towards Extracting Ontology Excerpts'. en. In: *Knowledge Science, Engineering and Management*. Ed. by Songmao Zhang, Martin Wirsing and Zili Zhang. Vol. 9403. Series Title: Lecture Notes in Computer Science. Cham: Springer International

Publishing, 2015, pp. 78–89. ISBN: 978-3-319-25159-2. DOI: 10.1007/978-3-319-25159-2_7. (Visited on 14/01/2022).

[10] G. Cheng, T. Tran and Y. Qu. 'RELIN: Relatedness and Informativeness-based Centrality for Entity Summarization'. In: *Proc. 10th International Semantic Web Conference (ISWC), Part I*. 2011, pp. 114–129. DOI: 10.1007/978-3-642-25073-6_8.

[11] KENNETH WARD CHURCH. 'Word2Vec'. In: *Natural Language Engineering* 23.1 (2017), pp. 155–162. DOI: 10.1017/S1351324916000334.

[12] Cipher.AI. *Gold Standards Webinar Slides*. June 2020. URL: https://cipher.ai/wp-content/uploads/2020/07/Gold-Standards-webinar-slides-2020-06-29-F_compressed.pdf.

[13] Andrew Collette and contributors. *H5Py*. Accessed: March 8, 2023. 2021. URL: https://docs.h5py.org/en/stable/.

[14] Bernardo Cuenca Grau et al. 'A Logical Framework for Modularity of Ontologies'. In: *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI)*. 2007. URL: https://www.ijcai.org/Proceedings/07/Papers/046.pdf.

[15] 'Chapter Three - RDF and the Semantic Web Stack'. In: *RDF Database Systems*. Ed. by Olivier Curé and Guillaume Blin. Boston: Morgan Kaufmann, 2015, pp. 41–80. ISBN: 978-0-12-799957-9. DOI: https://doi.org/10.1016/B978-0-12-799957-9.00003-1. URL: https://www.sciencedirect.com/science/article/pii/B9780127999579000031.

[16] Sarah L. Harris David Harris. *Digital design and computer architecture*. 2nd ed. Morgan Kaufmann, 2012, pp. 127–129. ISBN: 978-0123944245.

[17] Gianluca Demartini, Djellel Difallah and Philippe Cudre-Mauroux. 'ZenCrowd: Leveraging Probabilistic Reasoning and Crowdsourcing Techniques for Large-Scale Entity Linking'. In: Apr. 2012, pp. 469–478. DOI: 10.1145/2187836.2187900.

[18] Jacob Devlin et al. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv:1810.04805 [cs]. May 2019. (Visited on 09/07/2022).

[19] FoodOn. *About FoodOn*. Website. 2021. URL: https://foodon.org/about/.

[20] Dinesh Garg et al. 'Quantum Embedding of Knowledge for Reasoning'. In: *Advances in Neural Information Processing Systems*. Vol. 32. Curran Associates, Inc., 2019. (Visited on 14/01/2022).

[21] Genet Asefa Gesese, Russa Biswas and Harald Sack. 'A Comprehensive Survey of Knowledge Graph Embeddings with Literals: Techniques and Applications'. en. In: (), p. 10.

[22] Genet Asefa Gesese et al. 'A Survey on Knowledge Graph Embeddings with Literals: Which model links better Literal-ly?' en. In: *arXiv:1910.12507 [cs]* (May 2020). arXiv: 1910.12507. (Visited on 14/01/2022).

[23] Bernardo Grau et al. 'Modular Reuse of Ontologies: Theory and Practice'. In: *J. Artif. Intell. Res. (JAIR)* 31 (Jan. 2008), pp. 273–318. DOI: 10.1613/jair.2375.

[24] Bernardo Cuenca Grau et al. 'OWL 2: The next step for OWL'. In: *Journal of Web Semantics* 6.4 (2008). Semantic Web Challenge 2006/2007, pp. 309–322. ISSN: 1570-8268. DOI: https://doi.org/10.1016/j.websem.2008.05.001. URL: https://www.sciencedirect.com/science/article/pii/S1570826808000413.

[25] Thomas R. Gruber. 'A translation approach to portable ontology specifications'. In: *Knowledge Acquisition* 5.2 (1993), pp. 199–220. ISSN: 1042-8143. DOI: https://doi.org/10.1006/knac.1993.1008. URL: https://www.sciencedirect.com/science/article/pii/S1042814383710083.

[26] Kalpa Gunaratna et al. 'Relatedness-based Multi-Entity Summarization'. In: *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*. 2017. URL: https://www.ijcai.org/proceedings/2017/0147.pdf.

[27] Trevor Hastie, Robert Tibshirani and Jerome Friedman. *The Elements of Statistical Learning*. en. Springer Series in Statistics. New York, NY: Springer New York, 2009, pp. 14–18. ISBN: 978-0-387-84857-0. DOI: 10.1007/978-0-387-84858-7. URL: http://link.springer.com/10.1007/978-0-387-84858-7 (visited on 04/04/2023).

[28] Guoliang Ji et al. 'Knowledge Graph Embedding via Dynamic Mapping Matrix'. In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Beijing, China: Association for Computational Linguistics, July 2015, pp. 687–696. DOI: 10.3115/v1/P15-1067. (Visited on 10/08/2022).

[29] Dan Jurafsky. *Speech and Language Processing, 3rd Edition*. 2021. URL: https://web.stanford.edu/~jurafsky/slp3/.

[30] Maxat Kulmanov et al. 'EL Embeddings: Geometric Construction of Models for the Description Logic EL++'. en. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*. Macao, China: International Joint Conferences on Artificial Intelligence Organization, Aug. 2019, pp. 6103–6109. ISBN: 978-0-9992411-4-1. DOI: 10.24963/ijcai.2019/845. (Visited on 14/01/2022).

[31] Martin Tri Vien Lam. *OwlTest*. GitHub repository. 2021. URL: https://github.uio.no/mtlam/OwlTest.

[32] Jean-Baptiste Lamy. 'Owlready: Ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies'. In: *Artificial Intelligence in Medicine* 80 (2017), pp. 11–28. ISSN: 0933-3657. DOI: https://doi.org/10.1016/j.artmed.2017.07.002. URL: https://www.sciencedirect.com/science/article/pii/S0933365717300271.

[33] Yankai Lin et al. 'Learning Entity and Relation Embeddings for Knowledge Graph Completion'. en. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 29.1 (Feb. 2015). ISSN: 2374-3468, 2159-5399. DOI: 10.1609/aaai.v29i1.9491. (Visited on 10/08/2022).

[34] Edward Loper and Steven Bird. 'NLTK: The Natural Language Toolkit'. In: *CoRR* cs.CL/0205028 (2002). URL: https://arxiv.org/abs/cs/0205028.

[35] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze. *Introduction to Information Retrieval*. New York, NY, USA: Cambridge University Press, 2008. ISBN: 0521865719. URL: https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf.

[36] Tomas Mikolov et al. *Distributed Representations of Words and Phrases and their Compositionality*. arXiv:1310.4546 [cs, stat]. Oct. 2013. (Visited on 05/07/2022).

[37] Tomas Mikolov et al. *Efficient Estimation of Word Representations in Vector Space*. arXiv:1301.3781 [cs]. Sept. 2013. (Visited on 01/07/2022).

[38] NumPy developers. *NumPy*. Accessed: Nov 17, 2022. 2021. URL: https://numpy.org/.

[39] *Ontologies - W3C*. URL: https://www.w3.org/standards/semanticweb/ontology (visited on 21/03/2023).

[40] Nils Reimers and Iryna Gurevych. *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks*. en. arXiv:1908.10084 [cs]. Aug. 2019. URL: http://arxiv.org/abs/1908.10084 (visited on 01/03/2023).

[41] Nils Reimers and Iryna Gurevych. 'Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks'. In: *CoRR* abs/1908.10084 (2019). arXiv: 1908.10084. URL: http://arxiv.org/abs/1908.10084.

[42] scikit-learn contributors. *scikit-learn: CountVectorizer*. Accessed: Nov 17, 2022. 2021. URL: https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.CountVectorizer.html.

[43] Wei Shen, Jianyong Wang and Jiawei Han. 'Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions'. In: *IEEE Transactions on Knowledge and Data Engineering* 27.2 (2015), pp. 443–460. DOI: 10.1109/TKDE.2014.2327028.

[44] Wei Shen et al. 'LINDEN: Linking named entities with knowledge base via semantic knowledge'. In: *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web* (Apr. 2012). DOI: 10.1145/2187836.2187898.

[45] SNOMED International. *About SNOMED International*. Last accessed on: April 4, 2021. Accessed 2021. URL: https://www.snomed.org/about.

[46] Harold Solbrig. *FunOwl Documentation*. 2021. URL: https://funowl.readthedocs.io/en/latest/.

[47] Harold Solbrig. *Ontology Validation with Python using FunOwl*. 2019. URL: https://www.slideshare.net/HaroldSolbrig/ontology-validation-with-python-using-funowl.

[48] Qi Song et al. 'Mining Summaries for Knowledge Graph Search'. In: *IEEE Transactions on Knowledge and Data Engineering* 30.10 (2018), pp. 1887–1900. DOI: 10.1109/TKDE.2018.2807442.

[49] Stanford Center for Biomedical Informatics Research. *Protege*. https://protege.stanford.edu/about.php. Website. Accessed 2020.

[50] M.Q. Stearns et al. 'SNOMED Clinical Terms: Overview of the Development Process and Project Status'. In: *Proceedings of the AMIA Symposium* 2001 (2001).

[51] Fabian M. Suchanek, Gjergji Kasneci and Gerhard Weikum. 'YAGO: A Large Ontology from Wikipedia and WordNet'. In: *Journal of Web Semantics* 6.3 (2008). World Wide Web Conference 2007Semantic Web Track, pp. 203–217. ISSN: 1570-8268. DOI: https://doi.org/10.1016/j.websem.2008.06.001. URL: https://www.sciencedirect.com/science/article/pii/S1570826808000437.

[52] Zhiqing Sun et al. 'RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space'. en. In: (2019), p. 18.

[53] *The Description Logic Handbook: Theory, Implementation and Applications*. 2nd ed. Cambridge University Press, 2007. DOI: 10.1017/CBO9780511711787.

[54] tqdm contributors. *tqdm*. Accessed: Mar 30, 2023. 2021. URL: https://github.com/tqdm/tqdm.

[55] Wikipedia. *COVID-19*. [Online; accessed 01-April-2023]. 2023. URL: https://en.wikipedia.org/wiki/COVID-19.

[56] Wikipedia. *Heart disease*. [Online; accessed 01-April-2023]. 2023. URL: https://simple.wikipedia.org/wiki/Heart_disease.

[57] Wikipedia. *Mean reciprocal rank*. [Online; accessed 01-April-2023]. 2023. URL: https://en.wikipedia.org/wiki/Mean_reciprocal_rank.

[58] Wikipedia. *Paracetamol*. [Online; accessed 01-April-2023]. 2023. URL: https://en.wikipedia.org/wiki/Paracetamol.

[59] Wikipedia. *Pneumonia*. [Online; accessed 01-April-2023]. 2023. URL: https://en.wikipedia.org/wiki/Pneumonia.

[60] Wikipedia. *Tylenol (brand)*. [Online; accessed 01-April-2023]. 2023. URL: https://en.wikipedia.org/wiki/Tylenol_(brand).

[61] Zhanqiu Zhang et al. 'Learning Hierarchy-Aware Knowledge Graph Embeddings for Link Prediction'. en. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 34.03 (Apr. 2020), pp. 3065–3072. ISSN: 2374-3468, 2159-5399. DOI: 10.1609/aaai.v34i03.5701. (Visited on 10/08/2022).