Knowledge Graphs in Information Retrieval

Jakub Dutkiewicz and Czesław Jędrzejek
Poznań University of Technology, Poznań, Poland

jakub.dutkiewicz@put.poznan.pl
czeslaw.jedrzejek@put.poznan.pl

Abstract: This paper introduces an information retrieval model that leverages knowledge graphs, specifically tailored for Clinical Trials. In these scenarios, the document in question takes the form of a semi-structured clinical trial, containing details about enrolled patients, descriptions of experiments and procedures conducted during the trial, relevant diseases, and specific enrollment criteria. While the document retains a semi-structured format, the majority of the information is expressed in natural language. Queries in this context consist of specific patient characteristics, such as disease type, genetic information, and demographic data. The primary aim of this paper is to develop and utilize a knowledge graph capable of storing this information, including links to external resources like the Disease Ontology. We propose an Object-Relational model, which is then transformed into a knowledge graph. This graph is subsequently employed to identify semantic connections between concepts present in the clinical trials and those in the queries. These connections are then utilized to formulate a retrieval model for each aspect of the query. To achieve this, we design a relevance formula that incorporates weights to account for ontological relationships between concepts. We evaluate the effectiveness of our model by comparing the results with manual annotations.

Keywords: Knowledge Extraction, Knowledge Modeling, TREC Clinical Trials Track

1. Introduction

Information Retrieval(IR) deals with searching and evaluating information from document repositories. Given a set of documents and a list of queries, the task is to return a list of documents sorted by the relevance of each document to the given query. TREC-PM (Roberts, et. Al 2017, 2020) was (and is in a TREC-CT version) a biomedical track, which evaluates the systems performing the IR tasks on a specific set of documents (clinical trials) and queries called topics. The TREC PM in 2017-2018 in total and partially in 2019 was dedicated to oncology related topics. The set of clinical trials originates from clinicaltrials.gov. In this paper we use the snapshot which was used in the The relative narrowness of content facilitates a use of a knowledge based model. The organizers of the TREC-PM track provide a manual, along with a meticulously crafted set of annotations, and an outline to guide annotators. In this paper, our goal is to automate this process. Specifically, we aim to automatically identify relevant entities within both clinical trials and the topics. We propose that carefully chosen relevant links can serve as the foundation for an exceptionally accurate Clinical Trial retrieval system. This paper presents a detailed description of our Knowledge Graph model, along with the publication of the generated graph. Additionally, we provide a code repository containing key components of the implementation, including the object-relational model of the data, a set of functions for populating the database, and functions for generating the knowledge graph. We utilize the data stored within the database in order to find entity links to external Disease Ontology as well as to our hand-crafted set of data. This indirectly enhances the information retrieval quality. We also provide an ad-hoc retrieval model, which utilizes the entity links.

2. Related Work

The concept of employing knowledge graphs in Information Retrieval is succinctly outlined in a survey (Reinanda et al, 2020). Specifically, the notion of directly utilizing knowledge graphs for Document Retrieval is expounded upon in section 4.1. Our perspective resonates with the authors of the survey, who assert that effectively leveraging entity annotations and text in tandem to enhance ad-hoc document retrieval remains a challenge yet to be fully addressed. This challenge serves as a primary motivation for our research endeavors. Within the survey, the authors cite two papers that exemplify the utilization of knowledge graphs in this particular task. Both papers make use of the Knowledge Graphs for query expansion. Authors of the first paper mentioned in the survey (Dalton et el, 2014) introduce the idea of knowledge base links to the Information Retrieval model. We expand on the idea by employing only a specific subset of the knowledge base entities, which corresponds to the query domain. We also make use of dedicated external knowledge sources, such as the Disease Ontology (Schmirl et al, 2018). In our research queries are structured, hence specific parts of the query are linked to various external knowledge sources - mentions in a field dedicated for diseases are linked to the disease entities, mentions in a gene field are linked to the gene entities. Second paper mentioned in the survey (Xiong and Callan, 2015) exploits entity links in the process of query expansion. The terms in the query are linked to various

Freebase entities. Each linked entity is a basis for a query expansion. The expansion is done via a general Freebase description of an entity. This approach is different to ours as it doesn't exploit entity linking in the documents.

Information retrieval techniques are frequently employed in knowledge management (Rezgui, 2006). This is due to the connection between natural language and knowledge models—concepts are articulated using specific phrases that are part of both the knowledge model vocabularies and natural language. Ontology serves as a fundamental formal mechanism to link concepts from knowledge models to their natural language expressions. (Munir and Anjum, 2018) discuss effective methods for retrieving information from a formal ontology. However, this approach has limitations, as converting extensive natural language resources into a formal framework requires highly complex and detailed algorithms. A successful attempt at creating a Clinical Trials Knowledge Base (CTKB) for clinical trials resources was undertaken by (Chen, 2020). This work uses data provided by publicly available Aggregate Content of Clinical Trials (AACT) and focuses on the full description of each clinical trial. (Jayaram et al., 2015) propose an approach for querying the knowledge model. By analyzing their work, one can clearly deduce that information retrieval engineering should focus on key, elementary concepts, rather than graph nodes with extensive textual information. Consequently, we propose a cost-effective approach that incorporates the key low-level concepts into the model. In our work, the documents are interpreted as knowledge model entities. We make use of the natural language descriptions within Clinical Trials and link the documents to existing concepts in external ontologies. The knowledge graph in this work focuses on all aspects of the document retrieval process: documents (clinical trials), document entity links, queries (topics) and query entity links. (Otegi et al., 2015) propose an alternative approach, using concept relatedness for information retrieval. Unlike our method, they employ ad-hoc query expansion, linking documents to external resources each time a search is performed. We believe it is more effective to index all relevant concepts before the information retrieval process. In our opinion, a knowledge model-based approach allows for the creation of an environment where AI methods can be utilized in an explainable manner (Pascal et al., 2020).

3. Object-Relational Representation of a Clinical Trial

The dataset utilized for this research comprises a snapshot of clinical trials sourced from clinicaltrials.gov. We use the particular snapshots of the data, used by the human annotators. In this research, the trials were initially converted into an Object-Relational model using a dedicated Django application. This model comprises 21 object-oriented classes put into 25 relations. In addition to clinical trials, the model encompasses TREC PM topics, Entity Links to external ontologies, and hand-crafted resources such as lists of drugs and genes. These resources provide sufficient terminology to construct a T-Box ontology tailored specifically for this subset of clinical trials and queries.

Let's explore the primary sections of the model. The ***clinicaltrial*** relation is intended for storing data related to Clinical Trial documents. Clinical Trials may include various additional fields, such as lists of ***keywords***, ***conditions***, or ***outcomes*** they are associated with. The ***entity*** relation includes links to external ontologies, particularly the Disease Ontology. The ***gene*** and ***drug*** relations are designed to house hand-crafted data concerning drugs and genes relevant to TREC topics. Entities, drugs, and genes are linked to clinical trials through three connecting relations: ***entitylink***, ***druglink***, and ***genelink***. These relations establish associations between specific entities, drugs, or genes and the respective clinical trials in which they are referenced or utilized. Information regarding Information Retrieval queries is stored within the ***topic*** relation. Each instance of a ***rating*** relation pertains to a specific topic and a Clinical Trial. Details regarding matching genes and diseases (and indirectly drugs) are stored within the ***conditionannotation*** and ***geneannotation*** relations. These relations contain data on the associations between conditions, genes, and diseases referenced within the clinical trials and mentioned in a given topic.

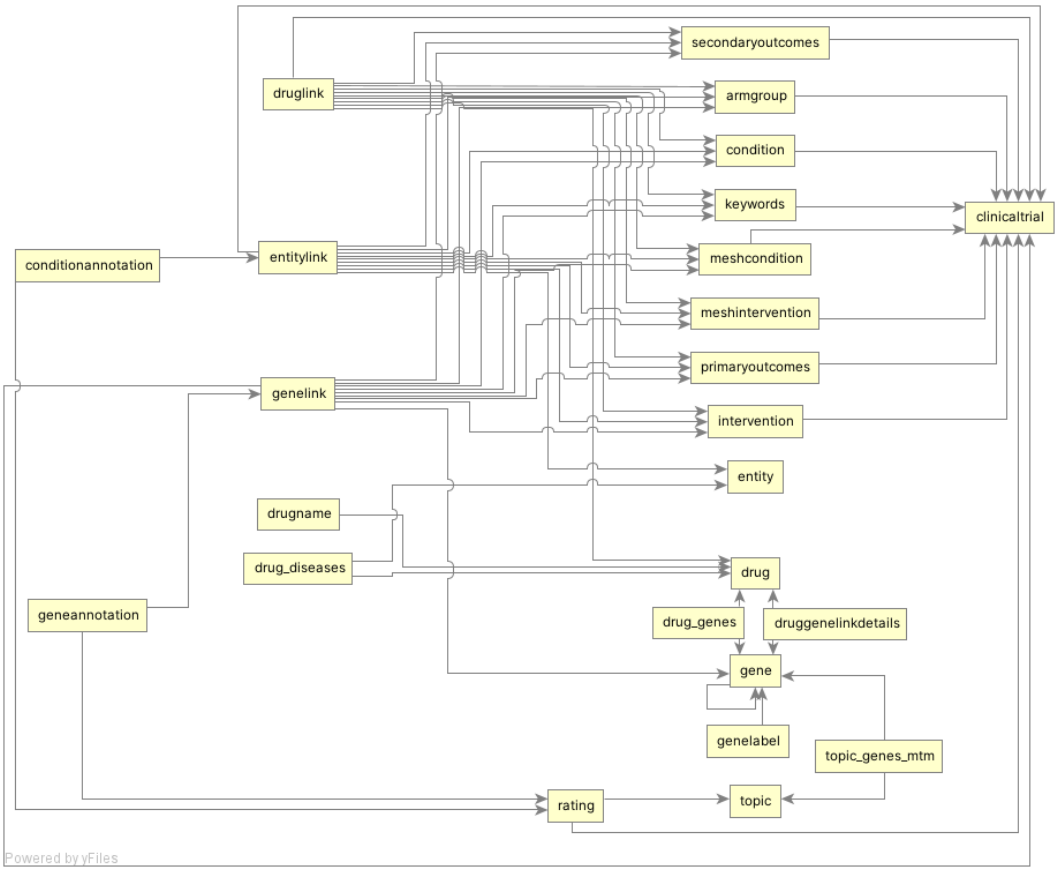

Figure 1: Relational model of the data.

4. Knowledge Model for a Clinical Trial

In this section, we examine different components of the Knowledge Model. We specifically address various **entities** and the *relationships* that exist between them. We propose a simple model for describing a clinical trial, where each document within the database is instantiated as an example of this class. Similarly, we outline a model for topic description. The connection between Clinical Trials and Topics is established through links to external resources or a manually curated vocabulary.

4.1 Clinical Trial Model

The focal points of the graph are the **Clinical Trial Documents**. Each document consists of a set of fields of a **Clinical Trial Document Part** type. As this is a knowledge graph, the vast majority of the data is stored as the individuals. Each individual consists of a set Data Properties, for which every **Clinical Trial** has exactly one value. The **Clinical Trial** model is depicted in Fig. 2 and Fig. 3. The first figure illustrates all of the Clinical Trial document properties. A model of **Clinical Trial Document Parts** in the second figure is depicted on the latter figure. Such parts are modeled as Object Properties, this is due to the fact the documents can be related to more than one such part. The second figure illustrates only two document part types, however in fact the graph incorporates Mesh Terms, descriptions of the Arm Groups, Keywords, Interventions and Outcomes in a similar manner. The data we used to create a graph consists of 18883 unique clinical trials.

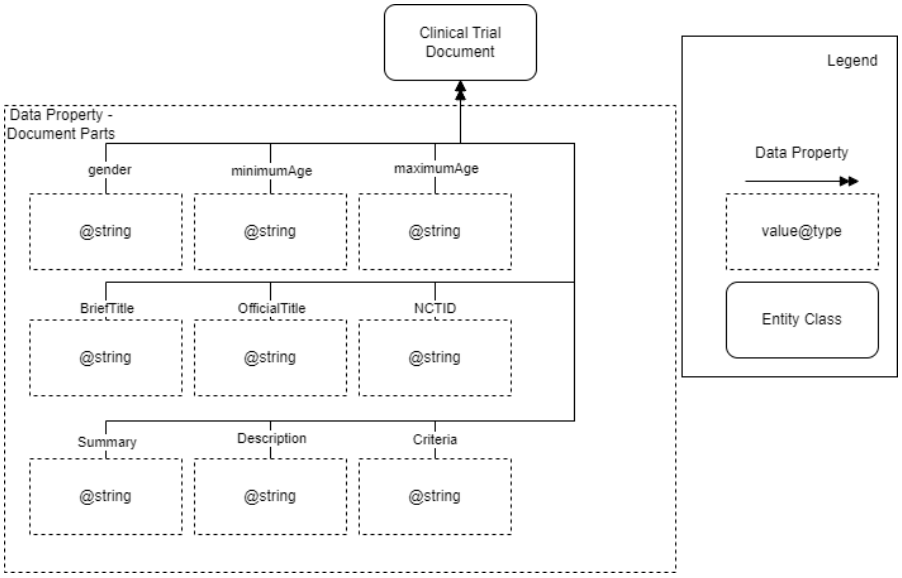

Figure 2: The clinical trial model with a list of document parts, which appear only once in the trial.

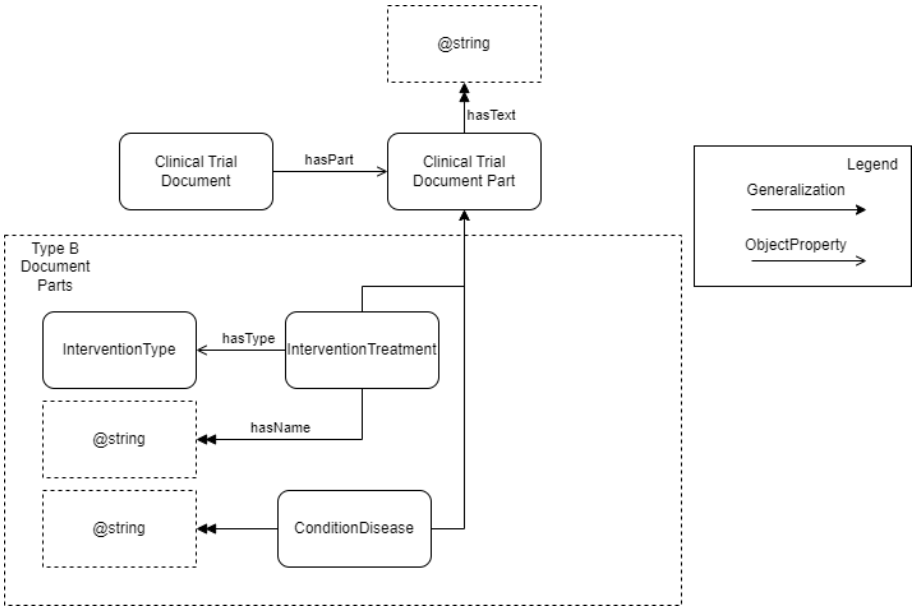

Figure 3: The clinical trial model with sample document parts, which can appear more than once in the trial.

4.2 External Terminology

This research encompasses three types of external terminology: a collection of phrases utilized to represent Disease, Drug, and Gene concepts. It's important to note that frequently, existing ontological resources prioritize proper topology and relational properties of the described classes, while overlooking the aspect of a rich and extensive vocabulary. This oversight is particularly crucial in Entity Linking, and indirectly affects Information Retrieval when utilizing such resources.

For disease-related vocabulary, we rely on Disease Ontology. This standardized ontology offers consistent, reusable, and sustainable descriptions of human diseases. Considered a core ontology, it features external links to many popular resources such as MeSH, ICD, NCI's thesaurus, SNOMED, and OMIM. We extensively utilize the *label* and *hasExactSynonym* annotation properties to model the vocabulary knowledge of concepts. Efforts are made to manually expand this knowledge to accommodate the language used in clinical trials for referring to disease concepts. The concepts described in the Disease Ontology are integrated into our model using the *entityLink* object property. This property has a domain of Clinical Trial Documents and Clinical Trial Document Parts, and a range of the flat Entity class, essentially serving as a reference to the Disease Ontology. While this relation may seem somewhat redundant as the link could directly refer to an external class, it proves convenient

in practical applications where we utilize only a portion of the external resource. The Entity class functions as a hyper reference, facilitating streamlined access to relevant disease-related concepts within our model depicted on Fig. 3.

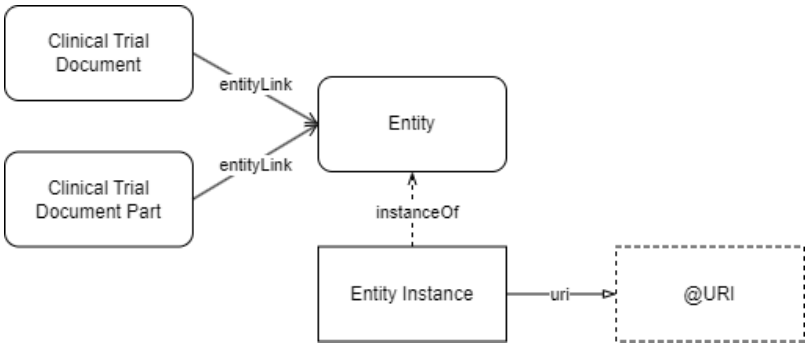

Figure 3: Entity Linking Knowledge Model

The vocabulary for **Gene** and **Drug** entities is hand-crafted. Both genes and drugs are stored in the ontological T-Box. The information stored in the Knowledge Model does not refer to specific instances of neither Drugs nor Genes. **Clinical Trials Documents** relate to the concepts via **GeneLinks** and **DrugLinks**. The model for hand-crafted vocabulary is presented on Fig. 4.

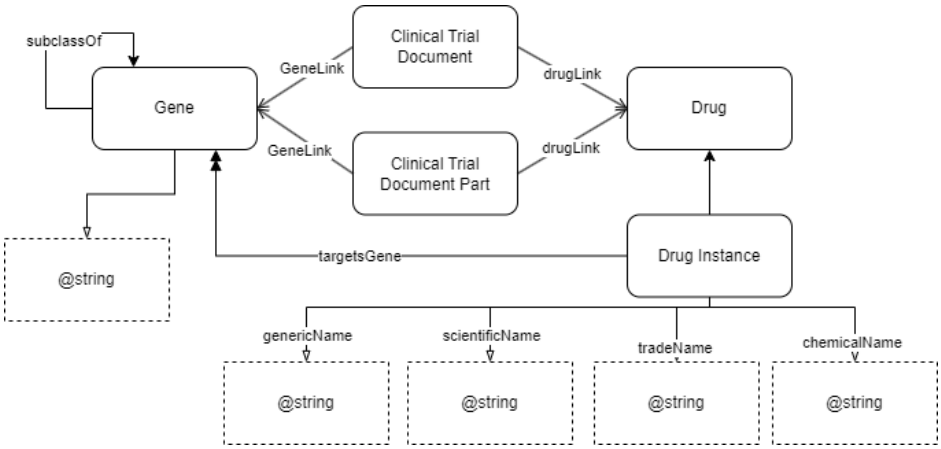

Figure 4: Gene and Drug Knowledge Model

4.3 Topic and Rating Model

In the knowledge graph, each TREC **Topic** is associated with a single disease, a list of genes, patients' demographics, and a single textual field for more information. Each **Topic** is linked via EntityLink with a **CrossReferenceEntity** and via a **GeneLink** with a set of Gene entities. The demographic description and the additional information are stored as a Data Properties. This structure facilitates efficient organization and retrieval of information related to TREC topics within the knowledge graph.

Rating encompasses information about the relevance between a given **Topic** and a **Clinical Trial Document**. There are two versions of the relevance assessments, the one delivered by TREC organizers and one calculated with use of this knowledge model. The fifth section of this work covers the details on obtaining the relevance score.

5. Using the Graph in Information Retrieval

In this section we describe the processes of using the Knowledge Graph for direct information retrieval. We provide description of annotation process, which we want to recreate.

5.1 TREC Annotation Guidelines

We implement our retrieval system according to the guidelines for annotators provided by the TREC organizers. There are several steps in the annotation process. The first step is ought to determine whether the clinical trial

is related to precision medicine. Here the organizers provide a list of steps, which after detailed evaluation seems vague. Document is related to Precision Medicine if: it is related to humans, involves any form of cancer, focuses on treatment or prognosis of cancer and contains any mention of a genes from the relevant topic. This seems not obvious as the PM property, in our eyes should not be decided in correlation with the formulated query. On top of that we find some annotations, which deem the documents as related to the Precision Medicine and annotate all of the genes as missing. We assume that this intermediate step should not be interpreted very directly, as it uses specific vocabulary – mention of a gene could not be interpreted as a relation to any given gene. In fact, we refrain from its complete evaluation. We intend to work on this step of the evaluation, as it is particularly important for the total evaluation. In this stead we focus on the analysis of clinical trial treatment type.

If a document passes the first phase – is annotated as relevant to precision medicine – it enters the second phase. In the second phase annotator is supposed to determine whether the document is relevant to a given query. It comes in three parts: determining appropriate category for a Disease part of a query (Exact, More General, More Specific and Not Disease), appropriate category for a Gene part of a query (Exact, Missing Variant, Different Variant, Missing Gene) and assessing the score for demographic criteria. We assume that the patient described in the query should be eligible for the trial, as it is expressed in eligibility criteria.

5.2 Treatment Evaluation

The first item in the TREC guidelines underscores the importance of documents in the field of precision medicine. This requirement prompts an analysis of the treatment type, which constitutes the focal point of the clinical trial.

Definition of “treatment” in TREC PM/clinical trials is the most difficult among the clinical trials’ concepts. In the wide sense “treatment” encompasses: diagnosis, treatment, and prevention. In general, Screening tests are not “treatment”, because they are primarily used for early detection of disease or risk factors whereas diagnostic tests are used to establish the presence or absence of disease. Medical “treatment” should be differentiated from clinical research.

Clinical research is characterized by (*Clinical Research Versus Medical Treatment*, access 2024):

1. Generally designed and intended to benefit future patients.
2. Involves periodic and systematic assessment of patient data.
3. Tests products and procedures of unproven benefit to the patient.

Medical “treatment” is characterized by:

1. Generally designed and intended to benefit future patients.
2. Address the needs of individual patients.
3. Intended to benefit the individual patient.
4. Requires real-time decisions.
5. Based on as-needed patient assessment.

This analysis encompasses the following fields within a clinical trial: **Intervention/Treatment** and **Study Type**. A single **Clinical Trial** document has one **Study Type**. This is represented with *hasStudyTypeAssignment* object property. **Clinical Trial** is capable of encompassing several **Intervention/Treatment** fields. There are four distinct **Study Types:** Interventional, Observational, Expanded Access and Observational (Patient Registry). The knowledge model for the study types is represented on Fig. 5. The knowledge model is implemented with the owlready2 library, which does not support class punning, which would be very useful in this case. We use a similar model, in which we create a single instance for each of the study type classes, which represents the respective class.

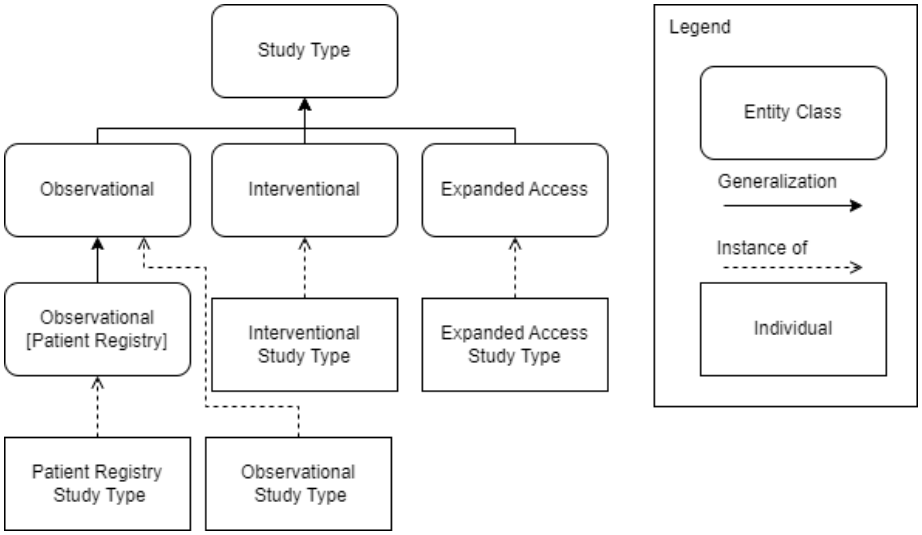

Figure 5: Study Type knowledge model

The **Intervention/Treatment** field comprises two sections: a type and a description. Here we analyze a set of types. There are a total of 8 distinct Intervention/Treatment types: Drug, Procedure, Other, Biological, Behavioral, Genetic, Radiation, Device.

The scoring depends on what weight we ascribe to each of the treatment types> If a Drug name appears in a document (alone or with other types) the score has value 1. For observational study type the score is 0.

For 2017-2019 years (2019 only cancer related topics) the total number of evaluated document is 41444 (15481 labeled by TREC as Human PM, and 25963 Not PM). We can directly compare TREC Human PM cases with our score.

When each of the 8 distinct Intervention/Treatment types: Drug, Procedure, Other, Biological, Behavioral, Genetic, Radiation, Device is equal 0 (corresponding to non-intervention types, mostly observational we get 1255 cases (1144 cases for only observational type). There are 2218 cases where a Drug is absent but there might be other non-zero types. Using our method we get 1615 non-treatment cases for TREC Human PM. This amounts to a 10.4% discrepancy with TREC in the treatment Human PM category. For the Not PM category the direct comparison with TREC is not possible (because a possible cause: disease, gene, treatment and demography are not separated). This will be the goal of future analysis.

5.3 Demography Evaluation

Demographic evaluation follows a straightforward process. Both TREC topics and clinical trials adhere to standardized demographic information. In TREC topics, demographic information is represented using a specific format: NN-TIMEUNIT-years-old gender. Here, NN denotes the numerical quantity of the age property, while TIMEUNIT indicates the unit of time used. To manage this format, we utilize a regular expression interpreter. If the TIMEUNIT is anything other than "year," we interpret the specified age as zero. The TREC topics cover only two genders: male and female.

Each Clinical Trial includes properties such as minimum age, maximum age, and gender, represented as text strings following a standardized format. If any of these properties are unspecified, we assume that every patient meets the relevant demographic criteria. The outcome of the evaluation is binary. If a particular topic describes a patient who does not fall within the specified criteria range outlined in the trial, the trial is considered irrelevant to that topic. Conversely, if the patient described in the topic fits within the specified criteria range, the demography criterion is deemed fulfilled. Table 1 presents the agreement between our system and human annotators on the Demography assessment.

Table 1: Annotation agreement between human TREC annotators and our automated system on the demography assesment.

Annotation\Demography score	0	1
Excludes	371	370
Not Discussed	21	651
Matches	67	8 319

5.4 Disease and Gene Evaluation

Another aspect of the relevance between a given query and a clinical trial is assessed by considering the number of relevant entity links. Each topic is associated with a single assigned entity from the Disease Ontology and can have up to three Gene links.

In the Clinical Trial, various entity links are present: disease links, gene links, and drug links. Each link is annotated with the part of the document it originates from. We posit that a higher number of links in a document, which either match the topic's disease link, match one of the topic's gene links, or are drug links for drugs targeting one of the genes within the topic, indicate greater relevance between the topic and the document. Conversely, we hypothesize that links within the document that do not match the topic links may lower the relevance score.

We also consider the source of the links within the document. Links originating in the title (or in the condition/disease field for diseases) of the clinical trial are deemed more important than those from other parts of the document. Additionally, we utilize the taxonomic properties of the ontology. Diseases found in the trial that are more specific than the one mentioned in the title are considered fully relevant. Conversely, diseases that are more general are treated as partially relevant and contribute to the total score with partial impact. If there are any disease links in the title of the document, the disease part of the relevance score is calculated as follows:

$$rel_{disease} = \frac{\#disease_{rel,title} + \#disease_{rel,condition} + 0.5\#disease_{part-rel,title} + 0.5\#disease_{part-rel,condition}}{\#disease_{rel,title} + \#disease_{rel,condition} + \#disease_{part-rel,title} + \#disease_{part-rel,condition} + \#disease_{non-rel,title} + \#disease_{non-rel,condition}}$$

If there are no diseases mentioned in the title, we rely on other parts of the document. In this scenario, we operate under the assumption that if the ratio between relevant and non-relevant diseases appearing in the same document is sufficiently high, the document is relevant. Therefore, the relevance score should be adjusted using a non-linear, sigmoid-shaped function:

$$rel_{disease} = \sigma(w_0 \cdot \frac{w_1 \cdot \#disease_{rel,other} + w_2 \cdot \#disease_{part-rel,other}}{1 + w_3 \cdot \#disease_{non-rel,other}} - b)$$

The parameters are selected manually after investigating the shape of the relevance function with synthetic data. Empirically assigned values are as follows: w0=3, w1=1, w2=0.5, w3=1, b=3. Moreover, the gene relevance score incorporates information on the number of relevant drugs mentioned in the document. However, it's important to note that genes are not searched in the condition field. The coefficient d represents the number of mentioned relevant drugs in the entire clinical trial:

$$rel_{gene} = \sqrt[d+1]{\frac{\#gene_{rel,title} + 0.5\#gene_{part-rel,title}}{\#gene_{rel,title} + \#gene_{part-rel,title} + \#gene_{non-rel,title}}}$$

Similarly, if there are no genes in the title, the gene score is given by:

$$rel_{gene} = \sqrt[d+1]{\sigma(w_0 \cdot \frac{w_1 \cdot \#gene_{rel,other} + w_2 \cdot \#gene_{part-rel,other}}{1 + w_3 \cdot \#gene_{non-rel,other}} - b)}$$

We use the same parameter values as for the disease score.To complete the relevance assessment between a document and a topic, we need to implement a total score, which should be calculated as the multiplication of the component relevances.

Table 2: Annotation agreement between human TREC annotators and our automated system on the disease assesment.

Annotation\Disease Score	0.25 > Score	0.25 <= Score < 0.5	0.5 <= Score < 0.75	Score >= 0.75
Not Disease	5 025	57	60	44
More General	2 271	250	273	156
More Specific	236	129	260	1 969
Exact	450	272	646	3 364

Table 3: Annotation agreement between human TREC annotators and our automated system on the gene assesment.

Annotation\Gene Score	0.25 > Score	0.25 <= Score < 0.5	0.5 <= Score < 0.75	Score >= 0.75
Missing Gene	5838	289	123	132
Different Variant	329	143	126	537
Missing Variant	1 336	341	400	710
Exact	1 882	734	834	1 949

6. Conclusions

In this work, we combine a knowledge graph-based approach with information retrieval for clinical trial documents. We make this knowledge graph publicly available, along with the object-relational model and graph creation code (https://github.com/dudenzz/ClinicalTrialsKGBasedIR). For convenience, we also publish a database filled with data. We compare our results with TREC annotations for individual concepts within the Human PM category. In contrast to (Otegi et al. 2015), who use concept relatedness and ad-hoc query expansion for information retrieval, we advocate for indexing relevant concepts prior to retrieval, believing this to be a more efficient approach. Additionally, our strategy supports the creation of an environment where AI methods can be employed in an explainable manner, a perspective also shared by Jayaram et al. (2015), who emphasize focusing on key, elementary concepts over graph nodes with extensive textual information. Our incorporates these key concepts into the model. In this work, we compare the outputs of our automated retrieval system, which is based on the knowledge graph, with hand-crafted annotations. We have shown that such a system can implement simple retrieval models. The proposed models appear to partially align with annotators' assessments; however, we refrain from a comparative analysis of retrieval quality at this point. Our major concern, which we leave to future work, is the formal and strict definition of a non-PM clinical trial. We aim to address this issue so we can ultimately compute and evaluate final relevance scores.

References

Chen, Z., Peng, B., Ioannidis, V.N. et al. (2022) *A knowledge graph of clinical trials (CTKG).* Sci Rep 12, 4724. https://doi.org/10.1038/s41598-022-08454-z

Dobbins, N.J., Mullen, T., Uzuner, Ö. et al. *The Leaf Clinical Trials Corpus: a new resource for query generation from clinical trial eligibility criteria.* (2022) Sci Data 9, 490.

Dalton J., Dietz L., and Allan J. (2014) *Entity query feature expansion using knowledge base links.* In Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval (SIGIR '14). Association for Computing Machinery, New York, NY, USA, 365–374. https://doi.org/10.1145/2600428.2609628

Reinanda R.; Meij E.; Maarten de Rijke (2020) *Knowledge Graphs: An Information Retrieval Perspective* , Foundations and Trends® in Information Retrieval: Vol. 14: No. 4, pp 289-444. http://dx.doi.org/10.1561/1500000063

Pascal Hitzler, Krzysztof Janowicz, and Freddy Lecue. 2020. *On the role of knowledge graphs in explainable AI*. Semant. web 11, 1 (2020), 41–51. https://doi.org/10.3233/SW-190374

Rezgui, Yacine. (2006). *Ontology-Centered Knowledge Management Using Information Retrieval Techniques*. Journal of Computing in Civil Engineering. 20. 10.1061/(ASCE)0887-3801(2006)20:4(261).

Roberts K., Demner-Fushman D., Voorhees EM, Hersh WR, Bedrick S, Lazar AJ, Pant S. (2017) *Overview of the TREC 2017 Precision Medicine Track*. Text Retr Conf. 2017 Nov;26. PMID: 32776021; PMCID: PMC7410346.

Roberts, Kirk et al. (2020) *Overview of the TREC 2020 Precision Medicine Track.* The text REtrieval conference : TREC. Text REtrieval Conference vol. 1266 (2020).

Munir, Kamran & Anjum, M.. (2017). *The use of Ontologies for Effective Knowledge Modelling and Information Retrieval*. Applied Computing and Informatics. 14. 10.1016/j.aci.2017.07.003.

Otegi, Arantxa & Arregi, Xabier & Ansa, Olatz & Agirre, Eneko. (2014). Using knowledge-based relatedness for information retrieval. Inf. Syst. 44. 1-30.

Schriml, L. M., Mitraka, E., Munro, J., Tauber, B., Schor, M., Nickle, L., Greene, C. (2018). *Human Disease Ontology 2018 update: classification, content and workflow expansion*. Nucleic acids research, 47(D1), D955–D962. doi:10.1093/nar/gky1032; PMID:30407550; PDF.

Xiong C. and Callan J. (2015) *Query Expansion with Freebase.* In Proceedings of the 2015 International Conference on The Theory of Information Retrieval (ICTIR '15). Association for Computing Machinery, New York, NY, USA, 111–120. https://doi.org/10.1145/2808194.2809446

Website access: 12.04.2024 *Clinical Research Versus Medical Treatment* https://www.fda.gov/patients/clinical-trials-what-patients-need-know/clinical-research-versus-medical-treatment