# Biomedical Annotation for Gut-Brain Interplay Information Extraction Meets Terminology Work: Insights and Perspectives

Vanessa Bonato

Department of Linguistic and Literary Studies, University of Padua, Via E. Vendramini 13, Padova, 35137, Italy.

Contributing authors: vanessa.bonato@phd.unipd.it

**Abstract**

Within the context of the Conference and Labs of the Evaluation Forum (CLEF) 2025, the GutBrainIE @ CLEF 2025 challenge has been proposed. This natural language processing challenge represents one of the initiatives developed under the European-supported project HEREDITARY (HetERogeneous sEmantic Data integratIon for the guT-bRain interplaY). The focus of the challenge is on performing named entity recognition and relation extraction on a corpus of PubMed abstracts on the gut-brain interplay. The resulting annotations will be used to train information extraction systems, tailored to the biomedical domain under investigation. In this paper, we illustrate the concept systems that supported the development of the annotation schema, which served as the reference framework for both named entity recognition and relation extraction. Then, we examine the potential of reusing entity mentions and entity relations identified during the gold-standard dataset annotation process as terminological data in a medical terminology resource. Crucially, the task of named entity recognition differs from the process of term extraction that is adopted in terminology science to extract terminological data from textual documents. By examining annotated entity mentions and entity relations, we highlight how data obtained through named entity recognition and relation extraction can be partially reused for terminology work.

**Keywords:** Medical Terminology, Information Extraction, Biomedical Annotation, Gut-Brain Interplay

## 1 Introduction

Information extraction is described as "the process of automatically extracting structured pieces of information from unstructured or semi-structured text documents" (Chiticariu et al., 2018). Within this domain, named entity recognition and relation extraction represent two core tasks that are typically distinguished (Nasar, Jaffry, & Malik, 2021). Named entity recognition is "[a]n act of recognizing and categorizing named entities that are presented in a text document", whose objective is "to find references of rigid designators in the text which belong to established semantic kinds like a person, place, organization, etc." (Jehangir, Radhakrishnan, & Agarwal, 2023). Taking the medical domain as an example, the task involves the identification of text spans such as "familial hypercholesterolemia" or "autosomal dominant hypercholesterolemia", that are both named entities corresponding to medical conditions (Luo, Lai, Wei, Arighi, & Lu, 2022). Relation extraction, instead, deals with "identifying the relations between entities from underlying content" (Zhao et al., 2024). Specifically, the relations established between different entities convey

semantic content (Nasar et al., 2021). For instance, when applied to the medical domain, relation extraction allows to identify relations such as "treat_disease" or "cause_disease" (Qi et al., 2021), which express different connections established between named entities.

Named entity recognition and relation extraction can be carried out by information extraction systems, supporting the automatic extraction of knowledge from textual documents that can also be domain specific. Information extraction systems can be assessed through evaluation campaigns, such as CLEF (Conference and Labs of the Evaluation Forum).[1] At CLEF 2025, held on 9-12 September 2025, a specific task has been presented that is directly related to the European-supported project HEREDITARY (HetERogeneous sEmantic Data integratIon for the guT-bRain interplaY).[2] One of the goals of the HEREDITARY project is to foster the development of information extraction systems capable of automatically extracting knowledge about the gut-brain interplay from biomedical literature. These systems are designed to support healthcare professionals in acquiring specialized knowledge extracted from scientific texts that focus on the gut microbiota and its connections to various health conditions. These conditions include a wide range of mental health-related states (Dziedzic et al., 2024; Gulas, Wysiadecki, Strzelecki, Gawlik-Kotelnicka, & Polguj, 2018; Kim et al., 2023), Parkinson's disease (Bai, You, Lei, & Li, 2024; Feng et al., 2024; Liang et al., 2021) and neurological disorders (He et al., 2024; Oroojzadeh, Bostanabad, & Lotfi, 2022; You et al., 2024) (Nentidis et al., 2025).

The GutBrainIE @ CLEF 2025 challenge (Martinelli et al., 2025) was launched within the thirteenth edition of the large-scale biomedical semantic indexing and question answering challenge (BioASQ) (Nentidis et al., 2025) with the objective of creating a high-quality dataset aimed at improving the accuracy of targeted information extraction systems .[3] This natural language processing (NLP) challenge focuses on PubMed abstracts concerning the gut-brain interplay, and requires participants to perform both named entity recognition and relation extraction. In this context, named entity recognition involves identifying the text spans corresponding to entity mentions in documents dealing with the gut-brain interplay, and labeling them to predefined categories of interest within this specific domain of activity. For example, within the text, different types of microbiota are mentioned, such as "gut microbiota" and "fecal microbiota". These entity mentions are labeled to the predefined category "Microbiome". Relation extraction, instead, is concerned with determining the relations that link different entity mentions pertaining to the biomedical domain within the texts (Nentidis et al., 2025). For instance, in cases where an interaction is mentioned between two entity mentions labeled as "Chemical" in the texts, the predefined predicate "*interact*" is used to link them. The defined list of categories and entity relations are provided in an annotation schema, built on work carried out on two different concept systems. The corpus was annotated by annotation experts for both named entity recognition and relation extraction, resulting in the generation of a gold-standard training dataset that provides the high-quality annotations needed to train information extraction systems. The annotation has been carried out by computer scientists and terminology experts, with expertise in the modelling and representation of medical specialized knowledge in ontologies and terminology resources. Within these resources, knowledge on medical concepts extracted from scientific texts is represented, along with the different types of relations established between them. More detailed information about the annotators team will be described in Section 3.2.

When compared with the process of term extraction from text corpora, which is a key activity in terminology work (Di Nunzio, Marchesin, & Silvello, 2023; Drouin, 2003; Giovagnoli, Vezzani, & Di Nunzio, 2024; Rigouts Terryn, Hoste, & Lefever, 2018; Vezzani, Di Nunzio, & Silecchia, 2022), named entity recognition is a clearly differentiated task. In fact, term extraction can be defined as "terminology work that involves the identification and excerption of terminological data by searching through a text corpus" (ISO 1087, 2019). Specifically, according to ISO 1087 (2019), terminological data is "data related to concepts and their designations". A concept is a "unit of knowledge created by a unique combination of characteristics", and it is linguistically designated in texts by a term. The term is a "designation that represents a general concept by linguistic means", thus it represents a "concept that corresponds to a potentially unlimited number of objects which form a group by reason of shared properties". Taking these definitions into account, in this work we assume that entity mentions do not correspond to concepts and, in some cases, an entity mention identified through named entity recognition does not qualify as a term. For example, the entity

---

mention "oral and gut microbiota" is not a term, because two terms designating two distinct concepts can be identified within it: 'oral microbiota' and 'gut microbiota'. In contrast, many entity mentions align with terms in the medical domain, such as "major depressive disorder" and "Autism Spectrum Disorder".

In this paper, which extends the work of Bonato (2025), we present the two concept systems on which the annotation schema was developed. We then analyze the annotated entity mentions and entity relations, to assess the extent to which they can be reused as terminological data in a terminology resource concerning the gut-brain axis and related medical conditions. Taking this discussion as a starting point, we specifically focus on concept relations, developing further reflections on the subject.

The remainder of the paper is organized as follows: in Section 2, we outline the concept systems developed to identify entity labels and entity relations defined in the annotation schema. In Section 3, we present the dataset on which the annotations were performed, and we describe the annotation process. In Section 4, we elaborate on the possibility of reusing entity mentions and entity relations in a medical terminology resource, with a particular focus on entity relations. In Section 5, we provide our final remarks and directions for future work.

## 2 Insights into Gut-Brain Interplay Concept Systems

For the purpose of the GutBrainIE @ CLEF 2025 challenge, an annotation schema was created to define the set of entity labels to be associated by annotators with entity mentions identified in the abstracts. The schema also specified the entity relations that link entity mentions, along with their corresponding predicates. However, identifying the set of relevant entity labels and entity relations to be included in the annotation schema required prior knowledge of the concepts concerning the gut-brain interplay and related health conditions. In terminology work, the systematic organization and representation of knowledge about concepts in specialized domains is achieved through the creation of concept systems (Dubuc, 1978; Nuopponen, 2022; Roche, 2007; Vezzani, 2022). Within a concept system, different types of concept relations can be established, reflecting the specialized knowledge of the domain shared by experts.

The annotation schema was developed based on work carried out on two concept systems. The first concept system was created by terminologists, following terminology principles. The second concept system represents a reworking of the first, and it was developed by terminology experts and computer science experts. The second concept system provided the foundation for the final design of the annotation schema.

### 2.1 Prototype Concept System

The development of the first concept system (see Figure 3 in Annex 1) is based on terminology work performed on abstracts related to the gut-brain interplay. Indeed, domain-specific knowledge is acquired from texts written by domain experts, where terms designating specialized concepts are used. To this end, a corpus of over 700 abstracts from biomedical publications was built. The abstracts were retrieved from PubMed,[4] that is the largest database of biomedical literature, using the keywords "mental health" AND "gut microbiota". Thus, the abstracts focus on specific health conditions related to the gut-brain interplay, namely mental health-related conditions. In addition to the abstracts, the corpus also comprised the titles of the full-text publications from which the abstracts were extracted. The analysis of abstracts and titles of scientific papers constitutes the starting point for this preliminary study. However, we consider that it is crucial to include the full texts of scientific articles within the analyzed dataset, to access "knowledge-rich contexts" (Meyer, 2001), where definitions of concepts and explanations are often provided.

The terminology work started with an automatic term extraction. The corpus was processed using the corpus analysis software Sketch Engine,[5] which enabled the extraction of more than 7,000 candidate terms. After the extraction phase, we manually assessed all the candidate terms to ensure that only lexical units representing terms in the medical domain were retained. The manual assessment led to the identification of over 1,300 medical terms.

Within the set of medical terms, several term variants that linguistically designate the same concept were identified. The term variants were clustered in accordance with this criterion, and a single concept was assigned to the term variants. The number of identified concepts, therefore, differs from the number of terms identified within the corpus, as it amounts to more than 1,200.

---

[4]https://pubmed.ncbi.nlm.nih.gov/
[5]https://www.sketchengine.eu

For each medical concept, different definitions were collected from specialized resources, such as MeSH Terms,[6] Ontobee,[7] and OLS Ontology Search.[8] This methodological approach ensured that concept knowledge was not tied to a single definition from one source, but was instead based on multiple specialized sources of information. Moreover, this approach was adopted to observe which characteristics that create the concept recur across different definitions, and can thus be regarded as fundamental for understanding and representing the concept.

Subsequently, terminological definitions for each concept were manually formulated, specifically in the form of intensional definitions. According to the ISO 1087 (2019), an intensional definition is a "definition that conveys the intension of a concept by stating the immediate generic concept and the delimiting characteristics". In particular, the delimiting characteristics of a concept are defined as "essential characteristics used for distinguishing a concept from related concepts" (ISO 1087, 2019). The identification of the generic concept, that is the "concept in a generic relation that has the narrower intension" (ISO 1087, 2019) is crucial to establish hierarchical relations between concepts, distinguishing superordinate concepts from their respective subordinate concepts. For instance, the concept <Anxiety disorder> can be defined as "psychiatric disorder constituting a category of disorders, characterized by anxious feeling or fear". Within the intensional definition, two elements can be distinguished: 1) the immediate generic concept <Psychiatric disorder>, which is the superordinate concept of <Anxiety disorder>, and 2) the delimiting characteristics of the concept.

Delimiting characteristics are another fundamental source of knowledge, as they are the set of essential characteristics that allow a concept to be uniquely identified, distinguishing it clearly from other concepts that share the same generic concept. This emerges from the comparison of the intensional definition of <Anxiety disorder> with that of <Mood disorder>, which is as follows: "psychiatric disorder constituting a category of disorders, characterized by disturbance in mood as their predominant feature". As can be noted, the same immediate generic concept is included within the two intensional definitions. Consequently, the immediate generic concept does not function as a distinguishing element between concepts. Conversely, the delimiting characteristics enable to distinguish the two concepts.

Formulating intensional definitions, therefore, proves to be a highly important activity for the detailed analysis of concepts in a specialized domain. Indeed, intensional definitions "enable terminology users to reconstruct the concept system" (ISO 704, 2022) and are considered as "preferable to other types of definitions because they clearly reveal the characteristics of a concept within a concept system" (ISO 1087, 2019).

The corpus, consisting of 460 intensional definitions, was automatically processed using Sketch Engine. This process represents a key component of our terminological analysis. By implementing this additional step, knowledge could be extracted not only from abstracts and titles of scientific papers, but also from intensional definitions, providing an essential source of domain-specific knowledge on concepts related to the gut-brain interplay. In particular, the corpus has been processed with the specific objective of extracting the terms that linguistically designate the immediate generic concepts and the delimiting characteristics of the defined concepts from intensional definitions. The Keywords extraction function was used to extract single-word terms and multi-word terms. Specifically, the extraction was performed relative to the English Web 2021 (enTenTen21) corpus, used as a reference corpus for comparison.

The identification process of both single-word and multi-word terms pertaining to the medical domain followed the same methodology adopted for the identification of medical terms in the corpus composed of abstracts and titles of scientific papers. Specifically, an automatic term extraction was first conducted, followed by the manual validation of the extracted terms. Once the list of medical single-word terms and multi-word terms was obtained, we associated the corresponding generic concept to each term. The generic concepts were linked to one another in the concept system by terminologists, using both hierarchical relations and associative relations.

Concerning hierarchical relations, both generic relations and partitive relations are used in the concept system to link different concepts. As the ISO 1087 (2019) states, a generic relation is a "concept relation between a generic concept and a specific concept where the intension of the specific concept includes the

---

[6] https://www.ncbi.nlm.nih.gov/mesh/
[7] https://ontobee.org
[8] https://www.ebi.ac.uk/ols4/

intension of the generic concept plus at least one additional delimiting characteristic". In the concept system, a hierarchical generic relation conveyed by the relation marker is_a can be identified between the three subordinate concepts <Bacterium>, <Fungus>, <Virus>, and the superordinate concept <Microorganism>. The same relation is established between the subordinate concept <Microorganism> and the superordinate concept <Organism>. As can be observed, the direction of the arrow that links the concepts is indicative of the reading direction of the represented concept relation (see Figure 3 in Annex 1).

The other type of hierarchical relation that is used within the concept system is the partitive relation, defined as a "concept relation between a comprehensive concept and a partitive concept" (ISO 1087, 2019). Within this relation, the comprehensive concept is a "concept in a partitive relation that is viewed as a whole consisting of various parts" (ISO 1087, 2019). Instead, the partitive concept represents a "concept in a partitive relation that is viewed as a part of a whole" (ISO 1087, 2019). In the concept system, this relation is established between the partitive concept <Genome> and the comprehensive concept <Microbiome>, and it is made explicit through the use of the relation marker is_part_of. A partitive relation is also established between the partitive concept <Microbiota> and the comprehensive concept <Microbiome>. By observing these concept relations, for instance, it is possible to infer that the microbiome is composed of both the microbiota and the genome. Also in this case, the arrow visually represents the reading direction of the relation (see Figure 3 in Annex 1).

Associative relations are also represented within the concept system. In particular, associative relations encompass all relations that are neither generic nor partitive, and whose relation markers express other types of links between concepts. For example, an associative relation occurs between <Microbiome> and <Medical condition>. To signal the existence of a link between the concepts, we used the relation marker is_linked_to. An associative relation is also established between the concepts <Medical condition> and <Intervention>, by using the relation marker treated_by. As can be observed, the relation marker is_linked_to does not exhibit the same level of precision compared to the relation marker treated_by (see Figure 3 in Annex 1). However, the relation marker is_linked_to represents the state-of-the-art level of specificity regarding the connection between the microbiome and related medical conditions.

## 2.2 Gut-BrainIE @ CLEF 2025 Challenge Concept System

The concept system developed by terminologists provided the foundation for a second concept system, which includes the concepts and concept relations to be annotated for the Gut-BrainIE @ CLEF 2025 Challenge (see Figure 4 in Annex 2) in the form of entity labels and entity relations. This concept system is the result of an interdisciplinary collaboration between terminologists and computer science experts, who together form the team of expert annotators. Then, this concept system was reviewed by an external group of biomedical experts, who validated the concepts and concept relations.

As can be observed, the concept system includes the concepts and concept relations to be extracted for the challenge, as well as concepts and concept relations that participants are instructed to ignore. As a matter of fact, during the early annotation phases, the team of expert annotators noted that some entity mentions were underrepresented relative to others in the document collection. Consequently, these entities were excluded from the annotation task, ensuring that the focus remained on entity mentions and relations that are both frequent and highly relevant to the study of the gut-brain interplay.

A comparative analysis of the two concept systems shows that some concepts were omitted from the second concept system, while others were added into it. To give some examples, both the concepts <Mouse> and <Dog> are not present in the second concept system, which includes only the superordinate concept <Animal>. This choice is once again driven by the consideration that a limited number of instances of these concepts can be found in abstracts. The concepts <Anatomical location>, <Statistical technique> and <Biomedical technique> were added, instead, among others, to the second concept system. This decision, conversely, reflects the significance that the team of expert annotators attributed to these concepts within the analyzed abstracts, leading them to deem it essential to incorporate the concepts into the concept system and to use them as entity labels in the annotation task.

In particular, it can be observed that several modifications were made in the organization of concepts within the second concept system. In the prototype concept system, <Human> appears as a subordinate concept of the superordinate concept <Animal>, which in turn is a subordinate concept of the superordinate concept <Organism>. In the second concept system, however, <Human> is not a subordinate concept of the superordinate concept <Animal>. Indeed, both <Human> and <Animal> appear as subordinate

concepts of the superordinate concept <Organism>. Another example concerns the concept <Microorganism>, that was removed from the second concept system. As a consequence, the concepts <Bacterium>, <Fungus> and <Virus> are no longer subordinate concepts of the superordinate concept <Microorganism>. Instead, the concepts <Bacteria>, <Virus> and <Fungus> are linked by way of a partitive concept relation to the comprehensive concept <Microbiota>.

Some concept relations have been added within the second concept system. One example concerns the relation between <Bacteria> and <Gene>, which is not present in the prototype concept system. In this case, a unidirectional concept relation has been established between the concepts, whose relation marker is "change expression". Another concept relation that has been added links the concepts <Chemical> and <Anatomical location>, and it is verbalized through the relation marker "located in".

Other existing relations linking two concepts in the first concept system have been modified. It is the case of the concept relation established between the concepts <Microbiome> and <Neurotransmitter>. In the prototype concept system, the concept <Neurotransmitter> is indicated as a subordinate concept of the superordinate concept <Metabolic intermediate>. The concept <Metabolic intermediate>, in turn, is linked to the concept <Microbiome> by way of the associative relation "produced_by". In the second concept system, instead, the concept <Metabolic intermediate> is no longer present, and the associative relation "related to" is used to link the concept <Neurotransmitter> to the concept <Microbiome>.

Another significant modification concerns the concept <Medical condition>, which was included in the prototype concept system. In the second concept system, this concept is represented as "DDF", which stands for "Disease, Disorder, or Finding". This detail indicates that a terminological approach for the construction of the concept system was not fully adopted. In fact, from a terminological perspective, <Disease>, <Disorder> and <Finding> are three distinct concepts, each linguistically designated by the corresponding term. However, this is a choice made by the team of expert annotators specifically aimed at maintaining a level of granularity consistent with the goals of the first edition of the challenge.

Within the concept system designed for the challenge, the addition of some links within the rectangles used to represent the concepts can be noted. These links are unique URIs, that link the concepts represented within the concept system to the corresponding concepts contained in a reference specialized vocabulary. Moreover, additional links are provided for concepts such as <Chemical> and <Drug>. These links are meant to assist participants during the annotation process by directing them to specialized resources containing lists of chemicals or drugs considered valid for annotation purposes.

The concepts included in this concept system – indicated in blue – serve as entity labels in the annotation schema. Following the same reasoning, the concept relations established within the concept system between the concepts considered relevant for the challenge are also part of the annotation schema. Within the annotation schema, the relation marker is represented as the predicate of the relation between a head entity and a tail entity.

# 3 Dataset and Dataset Annotation Description

This section provides a detailed overview of the dataset used for the annotation process and further information on the expertise of the team of expert annotators. Subsequently, we describe the manual annotation process that resulted in the creation of the gold-standard training dataset.

## 3.1 Dataset

As previously mentioned, the dataset was compiled by extracting abstracts from PubMed. To collect the documents, two separate queries were performed, using the following keyword combinations: 1) "mental health" AND "gut microbiota", and 2) "Parkinson" AND "gut microbiota". After removing duplicated records, the corpus consisted of 1647 documents. During the annotation process, 7 expert annotators annotated 403 PubMed abstracts.

## 3.2 Annotators

The annotation team brought together experts from two distinct domains: terminology and computer science. Specifically, the group included three annotators with a background in terminology and four with expertise in computer science, creating a heterogeneous team. For the task of biomedical annotation, the

combination of expertise from terminologists and computer scientists was essential. Indeed, computer scientists possess expertise in the development of ontologies, in which concepts and concept relations are formally represented. In addition to these skills, the computer scientists involved in the annotation process possess expertise in the field of information retrieval, which deals with textual corpora and the terms contained within documents.

As touched upon, the initial terminology work carried out by terminology experts laid the foundation for creating the prototype concept system. This concept system subsequently guided the development of a second concept system, which ultimately shaped the annotation schema. Indeed, terminology experts possess the expertise to extract terms in textual documents, infer the corresponding general concepts, draft terminological definitions of concepts and identify concept relations. The manual terminology work, as mentioned earlier, aimed to create a highly refined gold-standard annotated dataset. By following this methodology, only terms relevant to the medical domain were extracted from the abstracts, while filtering out non-relevant candidate terms.

The work of the annotators team was highly collaborative and synergistic, with the exchange of knowledge between experts from two different domains proving essential for developing the challenge. The terminologists trained the computer science experts by clarifying what is meant by "term" and "concept" in terminology science, and providing notions about terminological definitions and concept relations. In turn, computer science experts trained the terminologists to gain knowledge in both named entity recognition and relation extraction.

The gold-standard training dataset, however, was not the only document collection developed for the challenge. As a matter of fact, a silver collection was also created, which served as an additional training dataset. In particular, the silver collection was annotated by a group of students enrolled in the Master's Degree course in Modern Languages for International Communication and Cooperation of the University of Padua. Specifically, 26 students who attended the Translation-oriented Terminography course were trained under expert supervision to annotate a subset of documents concerning the gut-brain interplay. Their dataset also comprised abstracts, as well as the titles of the full-text publications from which the abstracts were extracted. In this case as well, both the tasks of named entity recognition and relation extraction were performed, based on the annotation schema provided to the students.

## 3.3 Annotation Schema

Designed to guide the annotation process, the annotation schema specifies: 1) the entity labels to be assigned to each entity mention found in the abstracts, and 2) the set of entity relations that can be used to link these mentions, along with their respective relation labels.

For entity labeling, the schema provides 13 categories under which the entity mentions can be classified, namely "Anatomical Location", "Animal", "Statistical Technique", "Biomedical Technique", "Bacteria", "Chemical", "Dietary Supplement", "Disease, Disorder, or Finding" also called "DDF", "Drug", "Food", "Gene", "Human" and "Microbiome". Taking the label "Microbiome" as an example, from a terminological standpoint, a hierarchical relation can be established with respect to the entity mentions "gut microbiome" and "gut microbiota". At this stage of the work on the challenge, indeed, the use of entity labels is meant to reach a more general level of abstraction.

This study centers on 3 out of the 12 labels listed in the GutBrainIE@CLEF25 Annotation Guidelines: "Disease, Disorder, or Finding", "Microbiome" and "Chemical".[9] This work focuses on these entity labels for two distinct reasons. First, based on the statistics regarding the count of entity labels identified in the training collections, these labels are the most representative.[10] Second, with reference to the entity labels "Disease, Disorder, or Finding" and "Microbiome", these are key elements of study within the gut-brain interplay field. Indeed, one of the main goals of the HEREDITARY project is to investigate the nature of the connection between health conditions and the microbiota or microbiome.

Regarding entity relations, the annotation schema allows for 57 distinct types. Each relation is verbalized through a specific predicate, which serves as a label defining the connection between two annotated entity mentions. For instance, when an entity labeled "Microbiome" is linked to an entity labeled "Disease,

---

[9]For a detailed description of entity labels, relation labels and the annotation rules for entity mentions and entity relations, see https://hereditary.dei .unipd.it/challenges/gutbrainie/2025/#

[10]https://hereditary.dei.unipd.it/challenges/gutbrainie/2025/
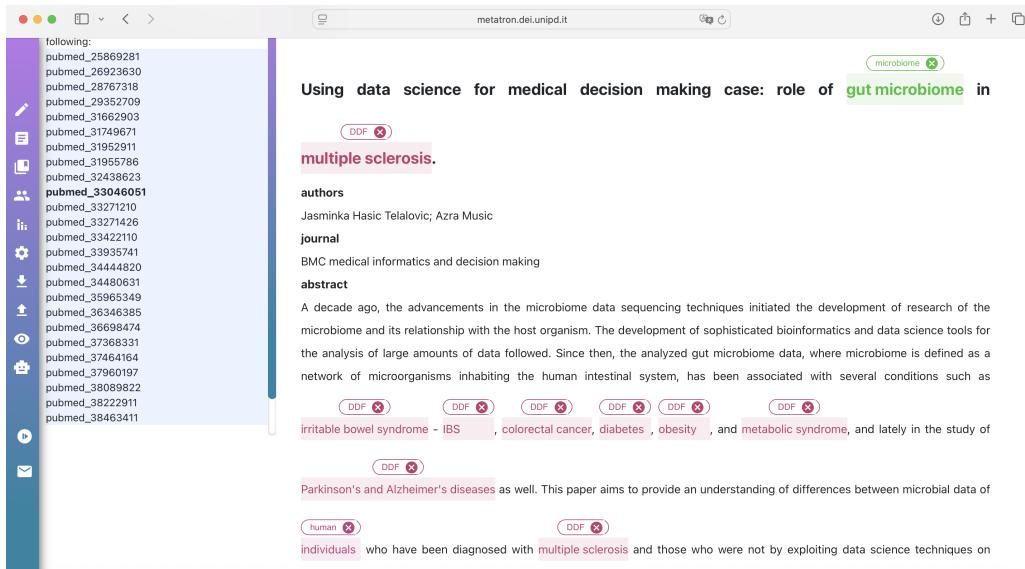
**Fig. 1** Screenshot of the MetaTron interface for abstract annotation

Disorder, or Finding", the relation is expressed with the predicate "*is linked to*". In this relation, "Microbiome" functions as the head entity and "Disease, Disorder, or Finding" as the tail entity. Another example involves the relation between the head entity "Chemical" and the tail entity "Disease, Disorder, or Finding", with the predicate "*change effect*" verbalizing their relation. In some cases, more than one predicate is associated with an entity relation. This is due to the fact that some entity mentions are linked through different types of relations in texts. In particular, within the Guidelines, the relation between a head entity and a tail entity can, in some instances, be either a hierarchical relation or an associative relation. It is the case of the entity relations that link the head entity "Chemical" with the tail entity "Chemical", they are respectively expressed by using the predicates "*interact*" and "*part of*". Indeed, two chemicals may either interact with each other or one may constitute part of the other. In the case of the relations between the head entity "Chemical" and the tail entity "Microbiome", they are expressed through two predicates both signaling an associative type of relation, namely "*impact*" and "*produced by*".

## 3.4 MetaTron

Annotation was performed using MetaTron (Irrera, Marchesin, & Silvello, 2024), which is a tool tailored to support the annotation process of biomedical corpora (Figure 1). Using this tool, annotators were able to proceed systematically, allowing to first perform named entity recognition and then relation extraction for each abstract.

## 3.5 Annotation Process

For each abstract, the annotation process was carried out by sequentially performing: 1) named entity recognition, and 2) relation extraction (Figure 2). Named entity recognition required annotators to identify text spans corresponding to entity mentions and assign to each mention a label chosen from a predefined list. The relation extraction task, in contrast, focused on determining the relations connecting pairs of labeled entities, whether explicitly stated in the text or inferred within the abstract. To support this work, annotators followed detailed guidelines that outline the complete list of entity labels and entity relation types defined for the challenge.

Before starting the individual annotation, the annotators jointly annotated a set of abstracts, to develop a single agreed-upon annotation and to address potential ambiguous cases or issues in identifying entity mentions and entity relations. Subsequently, two distinct annotation phases were conducted individually by the expert annotators. The first annotation phase was carried out on a dataset of 25 abstracts, 5 of which constituted the honeypot documents. The honeypot documents are the abstracts that were shared among the annotators, and they were used to compute the inter-annotator agreement, thereby assessing
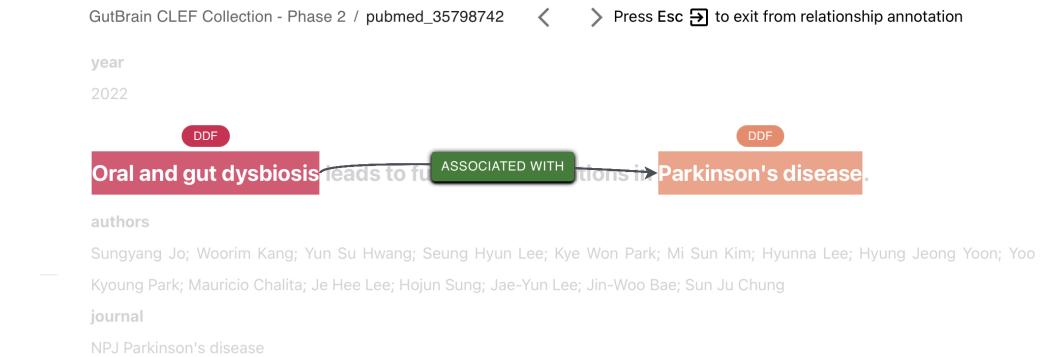
**Fig. 2** Screenshot of entity mentions and entity relation annotation

annotation consistency. In this phase, 148 abstracts were annotated manually, since no pre-annotations for entity mentions and entity labels were provided. This fully manual procedure resulted in the identification of 4,860 entity mentions and 2,360 relations, providing a carefully curated set of abstracts. In the second annotation phase, 255 abstracts were annotated. Unlike the first round, however, annotators worked with the support of pre-annotations automatically generated by unsupervised algorithms. Specifically, the named entity recognition model GLiNER (Zaratiana, Tomeh, Holat, & Charnois, 2024) was used to provide the pre-annotations, and it was fine-tuned on the basis of the annotations performed during the first annotation phase. In this second phase, the task of the annotators consisted of refining, correcting or validating the automatically generated annotations. In this phase, 6317 entity mentions and 3045 entity relations were identified.

### 3.5.1 Entity Mentions

To ensure consistency in how entity mentions were selected, a set of annotation rules was introduced. Among them, annotators were given the following instruction: "[a]nnotate composite entities as a single entity if they belong to the same category. However, if entities belong to the same category but appear as a sequence, annotate them separately."

According to this annotation rule, "Parkinson's and Alzheimer's diseases" is annotated as a single entity mention, because it comprises two composite entities that belong to the same label. This logic also applies to entity mentions such as "Oral and gut dysbiosis", "mineralocorticoid and N-methyl-D-aspartate receptors" and "oral and gut microbiome". Indeed, these pairs of composite entities are respectively treated as a single entity mention, due to the fact that both entities fall under the same category. Conversely, as exemplified within the guidelines, within the sentence "Cytochrome P-450 genes (CYP1A1, CYP2A6, CYP2D6, and CYP2E1)" five entity mentions need to be identified: "Cytochrome P-450 genes", "CYP1A1", "CYP2A6", "CYP2D6", "CYP2E1", each of them labeled as "Gene".

### 3.5.2 Relations

The annotation guidelines also addressed how to handle relations between labelled entities. In some instances, the predefined predicate may not precisely capture the connection between entity mentions that emerges from the text. To accommodate such cases, annotators were asked to use the predicate "*associated with*" to indicate a kind of relation that does not fit the predefined relation types. This relation was additionally applied to link entity mentions for which no specific relation had been defined in the guidelines, ensuring that additional relevant connections were captured.

For instance, whenever an explicit or implicit connection between two entity mentions labeled "Disease, Disorder, or Finding" is identified within an abstract, the relation can be annotated using the predicate "*associated with*". Figure 2 provides an example of this application. The title of the analyzed abstract is "Oral and gut dysbiosis leads to functional alterations in Parkinson's disease". The predicate "*associated with*" is adopted to signal the existence of a link between the entity mentions "Oral and gut dysbiosis" and "Parkinson's disease".

# 4 Can Named Entity Recognition and Relation Extraction Annotations be reused as Terminological Data?

The annotation carried out for information extraction made it possible to identify entity mentions and relations linking them within abstracts concerning the gut-brain interplay and related health conditions. However, as previously noted, named entity recognition operates differently from term extraction.

In the context of the annotation process, ensuring consistency across manually performed named entity recognition annotations was crucial to achieve a high inter-annotator agreement. To address this, the team introduced a set of internal annotation rules that all annotators were required to follow. These rules ensured the establishment of a ground truth for training information extraction systems. Indeed, by defining annotation rules, a gold-standard training dataset consisting of high-quality annotations was developed, to be used as a reliable reference for training information extraction systems specifically tailored to the biomedical domain under investigation. However, it is relevant to note that these rules were specifically meant to guarantee consistency in annotation, and do not always align with the terminological approach used to extract terms from documents. In the following sections, we will assess the extent to which both named entity recognition annotations (Section 4.1) and relation extraction annotations (Section 4.2) can be reused in a medical terminology resource as terminological data.

## 4.1 Analysis of Named Entity Recognition Annotations

The analysis of the annotations highlights a key distinction between entity mentions and terms. In the GutBrainIE dataset, a text span is considered a single entity mention whenever it groups multiple entities under the same label. A term, on the other hand is the linguistic designation of a concept, that can either be a single-word term or a multi-word term.

For example, the entity mentions "Oral and gut dysbiosis", "mineralocorticoid and N-methyl-D-aspartate receptors" and "oral and gut microbiome", labeled as "Disease, Disorder, or Finding", "Chemical" and "Microbiome" respectively, are not terms. In fact, each of these text spans encompasses two distinct terms designating different concepts. Specifically, the terms that can be identified are: 1) 'oral dysbiosis', 2) 'gut dysbiosis', 3) 'mineralocorticoid receptor', 4) 'N-methyl-D-aspartate receptor', 5) 'oral microbiome', and 6) 'gut microbiome'.

As illustrated in Figure 1, "Parkinson's and Alzheimer's diseases" is a single entity mention, since both composite entities share the same label "Disease, Disorder, or Finding". From a terminological standpoint, however, this text span is not a term. As a matter of fact, the text span covers two separate concepts, designated at the linguistic level by two terms: 'Parkinson's disease' and 'Alzheimer's disease'.

However, Figure 1 also highlights several additional entity mentions whose label is "Disease, Disorder, or Finding", namely "multiple sclerosis", "irritable bowel syndrome", "IBS", "colorectal cancer", "diabetes", "obesity" and "metabolic syndrome". Unlike the previous example, each of these mentions correspond to a medical term, as each designates a medical concept. More specifically, in the case of the entity mentions "irritable bowel syndrome" and "IBS", these are terminological variants, that is, two different terms that designate the same concept. These terms, therefore, could all form part of lexical networks, where relations between terms are outlined.

Moreover, in order to further analyze the difference between entity mentions and terms, it is relevant to mention another annotation rule provided to the participants of the challenge. As stated in the Guidelines, participants were asked to "[i]nlcude relevant contextual modifiers needed to capture the full and precise meaning of the entity mentions. Adjectives should be included with the entity mention, while nouns used as modifiers should be annotated separately." Following this rule, "EC-12 supplemented mice" is indicated as an entity mention within the Guidelines. From the annotated abstracts, it can be observed that other text spans are annotated as entity mentions by virtue of this annotation rule. For example, with regard to the entity label "Disease, Disorder, or Finding", the following entity mentions were annotated: "corticosterone-induced stress", "ampicillin-induced inflammation" and "antibiotic-induced gut dysbiosis". For what concerns the entity label "Chemical", the entity mention "plant-origin and microbially-formed neuroactive compounds" was annotated. These entity mentions, however, would not constitute terms in a medical terminology resource. Indeed, in a medical terminology resource, the considered terms would be 'stress', 'inflammation', 'gut dysbiosis' and 'neuroactive compound'.

The results of this analysis highlight that entity mentions do not always correspond to terms. However, through a selection process, entity mentions can be reused as terminological data to be included in a terminology resource. Conducting a terminological analysis that considers both the linguistic and conceptual dimensions of terminology (Costa, 2013), therefore, allows to differentiate entity mentions that can be reused as terms in a terminology resource from entity mentions that cannot be reused for this specific purpose.

The pivotal role of terminological analysis in this context also becomes evident when analyzing the entity label "Disease, Disorder, or Finding", defined in the guidelines by adopting the following definition provided in the National Cancer Institute Thesaurus (NCIT): "A condition that is relevant to human neoplasms and non-neoplastic disorders. This includes observations, test results, history and other concepts relevant to the characterization of human pathologic conditions".[11] A terminological analysis, indeed, enables the differentiation of the three concepts grouped together under the single entity label, which are respectively designated by the terms 'disease', 'disorder' and 'finding'. This distinction is crucial for representing specialized medical knowledge from a conceptual and a linguistic perspective at the terminological level.

## 4.2 Analysis of Relation Extraction Annotations

The entity relations annotated within the GutBrainIE gold-standard dataset can also be partially reused in a medical terminology resource as terminological data. Specifically, in the context of terminology work, some of these relations can serve as concept relations to be systematically represented within concept systems. For instance, the guidelines define a relation between the head entity "Bacteria" and the tail entity "Microbiome", using the predicate "*part of*". This corresponds to a part-whole relation, also called partitive relation, in which the concept <Microbiome> is the comprehensive concept and <Bacteria> is the corresponding partitive concept. The specialized knowledge acquired through the terminology work preceding the annotation process corresponds therefore to the knowledge that can be captured through information extraction, illustrating how entity relations can be directly used as concept relations in a terminological framework.

Another hierarchical relation that is present within the guidelines is the is-a relation, also called generic relation. In particular, this is the kind of relation established between the head entity "Disease, Disorder, or Finding" and the tail entity "Disease, Disorder, or Finding", whose predicate is "*is a*". This relation, in fact, is used to indicate that, within an abstract, a particular disease, disorder, or finding is considered a subordinate concept relative to its respective superordinate concept, mentioned in the same text. This type of annotation is therefore essential for terminology work, as these relations between entity mentions can be mapped onto concept relations between two concepts within a concept system.

The other types of relations considered in the annotation schema are associative relations. As previously mentioned, however, some predicates are more expressive than others, which convey a less precise and connection between two entity mentions. For example, the predicates "*change expression*", "*administered*" and "*strike*" express a more significant relation compared to "*influence*", "*interact*" and "*is linked to*". The predicate "*influence*", indeed, used to verbalize the relation between the head entity "Bacteria" and the tail entity "Disease, Disorder, or Finding", does not specify the nature of the influence exerted by bacteria on a medical condition or a symptom, nor whether this influence is beneficial. The same reasoning applies to the relation between the head entity "Chemical / Dietary Supp. / Drug / Food" and "Bacteria / Microbiome", whose predicate is "*impact*". In this case as well, it is not clear what type of impact – nor whether it is positive or negative – is exerted by a chemical, dietary supplement, drug or food on a bacterium or on the microbiome. Within this study, relations such as "*is linked to*" and "*influence*" are considered too general, and therefore cannot be used as such. The rationale behind this consideration is that, as mentioned earlier, one of the main goals of the HEREDITARY project is to specify the nature of the possible correlations between medical conditions and the microbiota or microbiome. Considering this, we assume that the mere presence of a connection between two concepts such as the one expressed by the predicate "*is linked to*" serves as a prerequisite for the existence of a concept relation. However, such a relation must be specified precisely in order to accurately represent the knowledge of medical concepts shared by domain experts.

---

[11]http://purl.obolibrary.org/obo/NCIT_C7057

An additional consideration involves the predicate "*associated with*", which is used whenever the predefined relation labels fail to fully capture the link between two entity mentions in a specific abstract. Rather than specifying a particular type of relation, this predicate exclusively signals that a link exists. For instance, in Figure 2, the relation between "Oral and gut dysbiosis" and "Parkinson's disease" is conveyed by the predicate "*associated with*", indicating a connection between these entity mentions without defining its exact nature. Considering all these reflections, it is possible to state that the entity relations included in the annotation schema can be partially reused in a concept system as terminological data. However, it is necessary to adopt a finer level of analysis in labeling entity mentions, with a view to systematizing the representation of conceptual knowledge. In this sense, terminology work can support the identification of additional relations that capture the links between head entities and tail entities for which the generic predicates "*influence*", "*interact*", "*is linked to*" and "*associated with*" are used in the challenge. An important consideration, however, is that one of the goals of the HEREDITARY project is to identify the precise relation between the gut-brain interplay and various neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease and Multiple Sclerosis, as well as mental health-related conditions. Hence, advancing specialized knowledge on the gut-brain axis within the project will enable the precise characterization of relations that are unspecified at this stage.

## 5  Conclusions and Future Research

In this paper, we presented the work underlying the development of the Gut-BrainIE @ CLEF 2025 Challenge, with a particular focus on the related terminology work. In the first place, we illustrated the two different concept systems that supported the design of the annotation schema. A subsequent step of the study involved assessing the potential of reusing data obtained through the manual annotation of the GutBrainIE gold-standard training dataset aimed at information extraction as terminological data. The analysis revealed that knowledge obtained from the performance of information extraction tasks can be leveraged in terminology work, to support the acquisition and representation of specialized biomedical knowledge. Indeed, reusing entity mentions and entity relations as terminological data in a terminology resource enables the representation of additional specialized knowledge, both from a linguistic and a conceptual viewpoint.

Specifically, the study revealed that both entity mentions and entity relations can be partially reused in the form of terminological data concerning the gut-brain interplay, whether integrated in a terminology resource, domain-specific lexical networks or concept systems. However, not all entity mentions are suitable for this purpose, making it necessary to filter and select only those that can be reused as terms. In addition, to enable the reusability of entity relations into a concept system, greater specificity must characterize some of the predicates used to verbalize associative relations. Nevertheless, annotations on biomedical abstracts represent a valuable additional source of domain-specific knowledge. As previously discussed, in the concept system developed by terminologists, only generic concepts and the concept relations established among them are represented. The annotation process performed with MetaTron enables the identification and annotation of relations not only between generic concepts, but also among the specific concepts occurring in the abstracts in the form of terms, in the cases in which entity mentions correspond to terms. Considering this, although not all identified associative relations are considered specific to the analysis of the gut-brain interplay in this study, the annotation process carried out with MetaTron enables rapid access to a substantial amount of highly-curated data, annotated and validated by a team of expert annotators. Comprehensively, this makes the process of acquisition of specialized knowledge more efficient.

As previously noted, this preliminary work is based on a dataset consisting of the abstracts and titles of scientific articles. To enable a more comprehensive terminological analysis, we intend to expand the dataset to include the full texts of the articles. Indeed, the full texts can provide additional knowledge useful for terminology work, and lead to the identification of further associative relations.

As future work, we also plan to specify in more detail the relations established between head entities and tail entities. In particular, it should be noted that the guidelines list head entities such as "Chemical / Dietary Supp. / Drug / Food" as referring to four distinct head entities, linked to the same tail entity through the same predicate. Considering this, we aim to investigate the associative relations that hold between each individual head entity and the same tail entity indicated in the Guidelines, to test whether

this leads to the identification of additional and more specific types of relations. For the same reason, we also intend to evaluate the possibility of splitting the individual entity "Disease, Disorder, or Finding" into three distinct entities.

We also plan to further investigate concept relations established between concepts related to the gut-brain interplay analyzing them in light of the classification of concept relations developed by Nuopponen (2022). Taking this classification as a reference, the investigation will be aimed at developing a typology of concept relations specific to the analysis of the gut-brain interplay, with a particular focus on the concept relations established between the microbiota and disorders and diseases that impact mental and neurological health.

As a last objective, we will compare the gold-standard annotated dataset with data produced by automatic term extractors, focusing on the analysis of precision and recall. Specifically, the goal will be to assess the performance of different automatic term extractors in extracting terms pertinent to the gut-brain interplay, by benchmarking them against the terms identified in the highly curated gold-standard annotated dataset in the form of entity mentions.

## Declaration on Generative AI

During the preparation of this work, the author(s) used ChatGPT-4 in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

## Acknowledgements

## References

Bai, F., You, L., Lei, H., Li, X. (2024). Association between increased and decreased gut microbiota abundance and parkinson's disease: A systematic review and subgroup meta-analysis. *Experimental Gerontology*, *191*, 112444, https://doi.org/10.1016/j.exger.2024.112444

Bonato, V. (2025). Reusability of biomedical annotations for gut-brain interplay information extraction as terminological data. F. Vezzani, G.M. Di Nunzio, E. Loupaki, G. Meditskos, & M. Papoutsoglou (Eds.), *Proceedings of the 4rd international conference on multilingual digital terminology today (mdtt 2025), thessaloniki, greece, june 19-20, 2025.* Thessaloniki, Greece: CEUR-WS. Retrieved from https://ceur-ws.org/Vol-3990/short13.pdf

Chiticariu, L., Danilevsky, M., Ho, H., Krishnamurthy, R., Li, Y., Raghavan, S., … Zhu, H. (2018). Web information extraction. In L. Liu & M.T. Özsu (Eds.), *Encyclopedia of database systems* (pp. 4620–4629). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-8265-9_459

Costa, R. (2013). Terminology and specialised lexicography: two complementary domains. *Lexicographica*, *29*, 29-42, https://doi.org/10.1515/lexi-2013-0004

Di Nunzio, G.M., Marchesin, S., Silvello, G. (2023). A systematic review of automatic term extraction: What happened in 2022? *Digital Scholarship in the Humanities*, *38*, i41-i47, https://doi.org/10.1093/llc/fqad030

Drouin, P. (2003). Term extraction using non-technical corpora as a point of leverage. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, *9*(1), 99-115, https://doi.org/10.1075/term.9.1.06dro

Dubuc, R. (1978). *Manuel pratique de terminologie.* Montréal: Linguatech.

Dziedzic, A., Maciak, K., Bliźniewska-Kowalska, K., Gałecka, M., Kobierecka, W., Saluk, J. (2024). The power of psychobiotics in depression: A modern approach through the microbiota–gut–brain axis: A literature review. *Nutrients*, *16*(7), 1054, https://doi.org/10.3390/nu16071054

Feng, M., Zou, Z., Shou, P., Peng, W., Liu, M., Li, X. (2024). Gut microbiota and parkinson's disease: potential links and the role of fecal microbiota transplantation. *Frontiers in Aging Neuroscience*, *16*, 1479343, https://doi.org/10.3389/fnagi.2024.1479343

Giovagnoli, A., Vezzani, F., Di Nunzio, G.M. (2024). Doctor-patient communication: terminological and simplification analysis in the domain of female oncology. *Umanistica Digitale*, *8*(17), 143–164, https://doi.org/10.6092/issn.2532-8816/19265

Gulas, E., Wysiadecki, G., Strzelecki, D., Gawlik-Kotelnicka, O., Polguj, M. (2018). Can microbiology affect psychiatry? a link between gut microbiota and psychiatric disorders. *Psychiatria Polska*, *52*(6), 1023–1039, https://doi.org/10.12740/PP/OnlineFirst/81103

He, Y., Wang, K., Su, N., Yuan, C., Zhang, N., Hu, X., … Zhao, F. (2024). Microbiota–gut–brain axis in health and neurological disease: Interactions between gut microbiota and the nervous system. *Journal of Cellular and Molecular Medicine*, *28*(18), e70099, https://doi.org/10.1111/jcmm.70099

Irrera, O., Marchesin, S., Silvello, G. (2024). Metatron: advancing biomedical annotation empowering relation annotation and collaboration. *BMC Bioinformatics*, *25*, 112, https://doi.org/10.1186/s12859-024-05730-9

ISO 1087 (2019). *Terminology work and terminology science — vocabulary.* International Organization for Standardization. Retrieved from https://www.iso.org/standard/62330.html

ISO 704 (2022). *Terminology work – principles and methods.* International Organization for Standardization. Retrieved from https://www.iso.org/standard/79077.html

Jehangir, B., Radhakrishnan, S., Agarwal, R. (2023). A survey on named entity recognition —datasets, tools, and methodologies. *Natural Language Processing Journal*, *3*, 100017, https://doi.org/10.1016/j.nlp.2023.100017

Kim, S.-Y., Woo, S.-Y., Raza, S., Ho, D., Jeon, S.W., Chang, Y., … Kim, H.-N. (2023). Association between gut microbiota and anxiety symptoms: A large population-based study examining sex differences. *Journal of Affective Disorders*, *333*, 21-29, https://doi.org/10.1016/j.jad.2023.04.003

Liang, Y., Cui, L., Gao, J., Zhu, M., Zhang, Y., Zhang, H.-L. (2021). Gut microbial metabolites in parkinson's disease: Implications of mitochondrial dysfunction in the pathogenesis and treatment. *Molecular Neurobiology*, *58*(8), 3745–3758, https://doi.org/10.1007/s12035-021-02375-0

Luo, L., Lai, P.-T., Wei, C.-H., Arighi, C.N., Lu, Z. (2022, 07). Biored: a rich biomedical relation extraction dataset. *Briefings in Bioinformatics*, *23*(5), bbac282, https://doi.org/10.1093/bib/bbac282

Martinelli, M., Silvello, G., Bonato, V., Di Nunzio, G.M., Ferro, N., Irrera, O., … Vezzani, F. (2025, September). Overview of GutBrainIE@CLEF 2025: Gut-Brain Interplay Information Extraction. G. Faggioli, N. Ferro, P. Rosso, & D. Spina (Eds.), *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025)* (Vol. 4038, pp. 65–98). Madrid, Spain: CEUR. Retrieved from https://ceur-ws.org/Vol-4038/paper_5.pdf

Meyer, I. (2001). Extracting knowledge-rich contexts for terminography. *Recent advances in computational terminology* (p. 279-302). John Benjamins Publishing Company. https://doi.org/10.1075/nlp.2.15mey

Nasar, Z., Jaffry, S.W., Malik, M.K. (2021). Named entity recognition and relation extraction: State-of-the-art. *ACM Comput. Surv.*, *54*(1), 1-39, https://doi.org/10.1145/3445965

Nentidis, A., Katsimpras, G., Krithara, A., Krallinger, M., Ortega, M.R., Loukachevitch, N., ... Paliouras, G. (2025). Bioasq at clef2025: The thirteenth edition of the large-scale biomedical semantic indexing and question answering challenge. C. Hauff et al. (Eds.), *Advances in information retrieval* (pp. 407–415). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-88720-8_61

Nuopponen, A. (2022). Conceptual relations: From the general theory of terminology to knowledge bases. P. Faber & M.-C. L'Homme (Eds.), *Theoretical perspectives on terminology: Explaining terms, concepts and specialized knowledge* (p. 63-86). John Benjamins Publishing Company. https://doi.org/10.1075/tlrp.23.03nuo

Oroojzadeh, P., Bostanabad, S.Y., Lotfi, H. (2022). Psychobiotics: the influence of gut microbiota on the gut-brain axis in neurological disorders. *Journal of Molecular Neuroscience*, *72*, 1952–1964, https://doi.org/10.1007/s12031-022-02053-3

Qi, T., Qiu, S., Shen, X., Chen, H., Yang, S., Wen, H., ... Huang, Y. (2021). Kemre: Knowledge-enhanced medical relation extraction for chinese medicine instructions. *Journal of Biomedical Informatics*, *120*, 103834, https://doi.org/10.1016/j.jbi.2021.103834

Rigouts Terryn, A., Hoste, V., Lefever, E. (2018, May). A gold standard for multilingual automatic term extraction from comparable corpora: Term structure and translation equivalents. N. Calzolari et al. (Eds.), *Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018).* Miyazaki, Japan: European Language Resources Association (ELRA). Retrieved from https://aclanthology.org/L18-1284.pdf

Roche, C. (2007). Le terme et le concept : Fondements d'une ontoterminologie. *TOTh 2007 : Terminologie et Ontologie : Théories et Applications.* (p. 1-22). Anncey, France.

Vezzani, F. (2022). *Terminologie numérique : conception, représentation et gestion.* Lausanne, Switzerland: Peter Lang Verlag. https://doi.org/10.3726/b19407

Vezzani, F., Di Nunzio, G.M., Silecchia, S. (2022). The phraseology of international disarmament treaties: the ditto terminology resource. *Umanistica Digitale*, *6*(14), 91–117, https://doi.org/10.6092/issn.2532-8816/14796

You, M., Chen, N., Yang, Y., Cheng, L., He, H., Cai, Y., ... Hong, G. (2024). The gut microbiota–brain axis in neurological disorders. *MedComm*, *5*(8), e656, https://doi.org/10.1002/mco2.656

Zaratiana, U., Tomeh, N., Holat, P., Charnois, T. (2024, June). GLiNER: Generalist model for named entity recognition using bidirectional transformer. K. Duh, H. Gomez, & S. Bethard (Eds.), *Proceedings of the 2024 conference of the north american chapter of the association for computational linguistics: Human language technologies (volume 1: Long papers)* (pp. 5364–5376). Mexico City, Mexico: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2024.naacl-long.300/

Zhao, X., Deng, Y., Yang, M., Wang, L., Zhang, R., Cheng, H., ... Xu, R. (2024). A comprehensive survey on relation extraction: Recent advances and new frontiers. *ACM Comput. Surv.*, *56*(11), 293, https://doi.org/10.1145/3674501
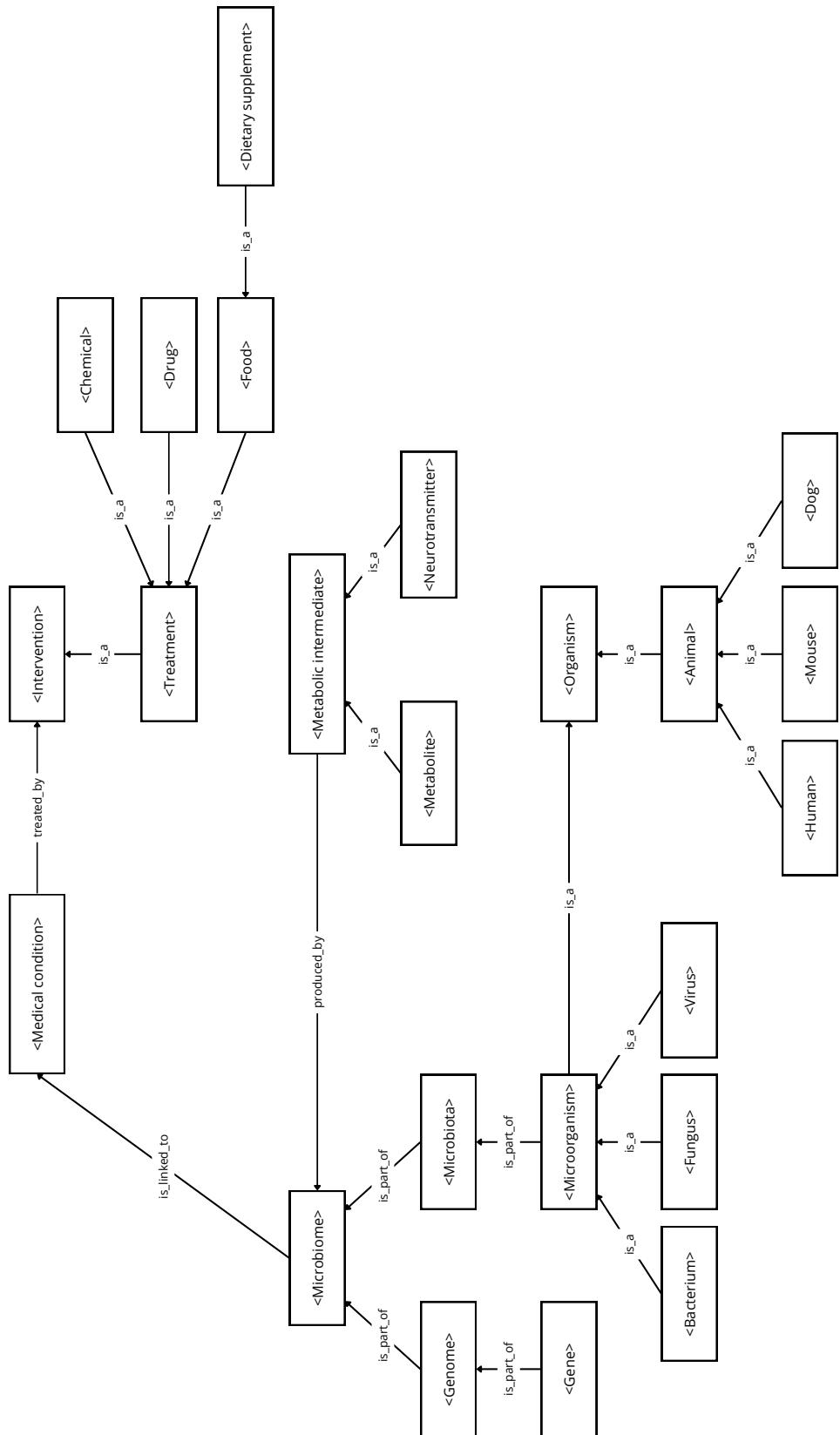
# Annex 1



**Fig. 3** Prototype concept system on gut-brain interplay developed by terminologists

# Annex 2
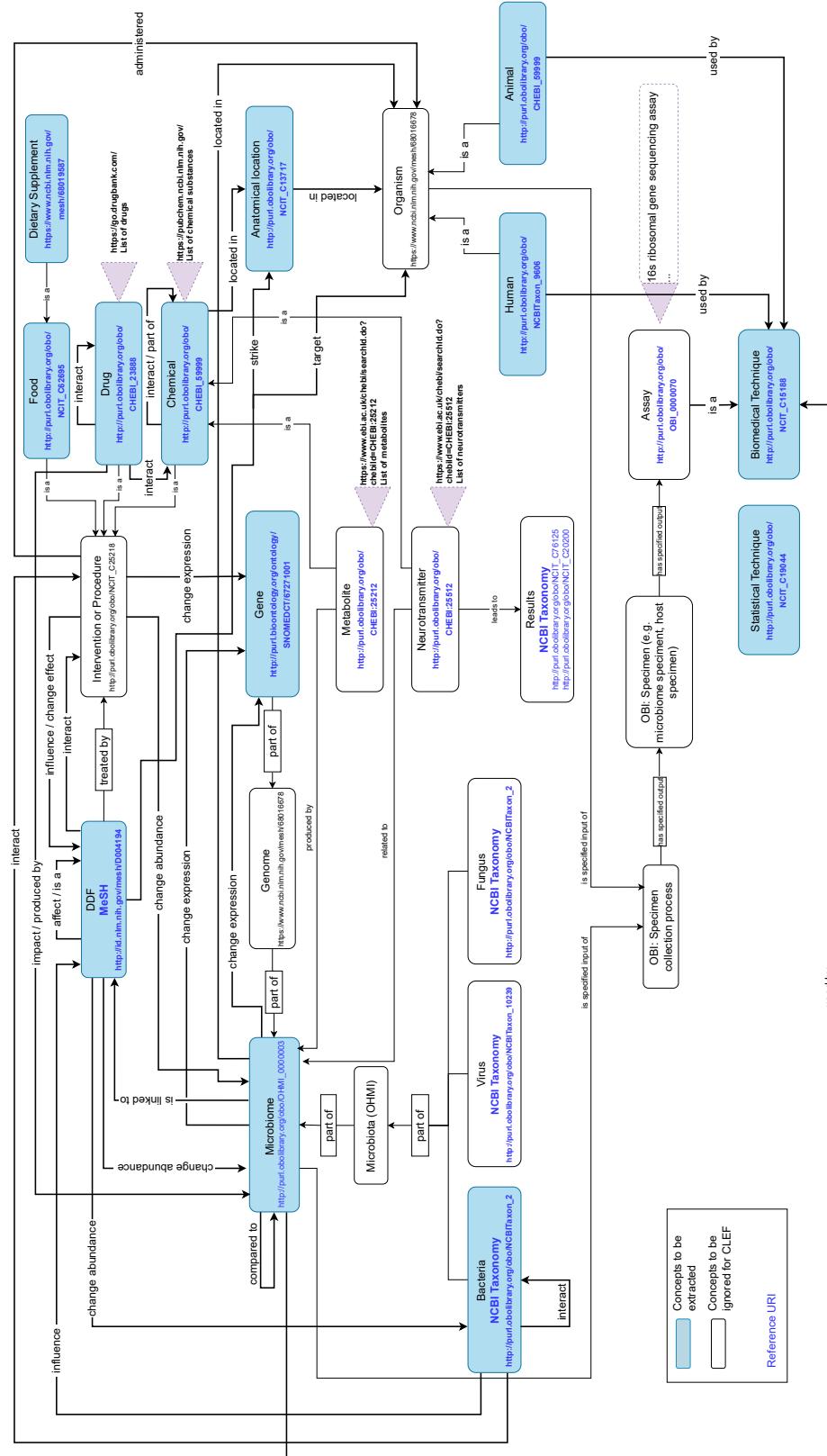


**Fig. 4** Concept system developed for the Gut-BrainIE @ CLEF 2025 Challenge