UNIVERSIDAD
DE GRANADA

FACULTAD DE TRADUCCIÓN E INTERPRETACIÓN

DEPARTAMENTO DE TRADUCCIÓN E INTERPRETACIÓN

PROGRAMA DE DOCTORADO EN LENGUAS, TEXTOS Y CONTEXTOS

DOCTORAL THESIS

Description, Categorization, and Representation of Hyponymy in Environmental Terminology

Juan Carlos Gil Berrozpe

GRANADA, 2022

Editor: Universidad de Granada. Tesis Doctorales
Autor: Juan Carlos Gil Berrozpe
ISBN: 978-84-1117-726-9
URI: https://hdl.handle.net/10481/80672

UNIVERSIDAD
DE GRANADA

FACULTAD DE TRADUCCIÓN E INTERPRETACIÓN

DEPARTAMENTO DE TRADUCCIÓN E INTERPRETACIÓN

PROGRAMA DE DOCTORADO EN LENGUAS, TEXTOS Y CONTEXTOS

DOCTORAL THESIS

Description, Categorization, and Representation of Hyponymy in Environmental Terminology

Juan Carlos Gil Berrozpe

Supervisor: Pamela Faber Benítez

GRANADA, 2022

Courage need not be remembered, for it is never forgotten.

ACKNOWLEDGEMENTS

To my supervisor, Pamela. For being one of the most wonderful people I have had the good fortune to meet, for awakening in me the passion for research and for marking my path in life forever.

To my parents, Javier and Cecilia. For allowing me to be who I am today and for having transmitted to me, from the most complete humility and simplicity, the best values as a person, as a son, and as a professional.

To my brother, Gustavo. For accompanying me from the first memory of my childhood and for being one of the greatest sources of inspiration, perseverance, and self-improvement that I have had in life.

To my partner, Rocío, my soul mate. For being able to understand me at all times with and without words, for helping me to be a better person every day, and for being the star that lights my path.

To all my LexiCon family, especially Pilar, Arianne, Melania, Juan, Amal, Clara, Maribel, Ana, and Miguel. For all the adventures lived, for all the troubles shared. and for everything I have learned from each one of them.

To all the other friends and colleagues who have surrounded me during these years. For the influence they have had on me and for the small grain of sand that, directly or indirectly, they have contributed to this thesis.

To the past, for having allowed me to get this far.

To the future... for all that is yet to come.

Thank you from the bottom of my heart.

AGRADECIMIENTOS

A mi directora, Pamela. Por ser una de las personas más maravillosas que he tenido la suerte de conocer, por despertar en mí la pasión por la investigación y por marcar para siempre mi camino en la vida.

A mis padres, Javier y Cecilia. Por permitirme ser quien soy hoy y por haberme transmitido, desde la más completa humildad y sencillez, los mejores valores como persona, como hijo y como profesional.

A mi hermano, Gustavo. Por acompañarme desde el primer recuerdo de mi infancia y por ser una de las mayores fuentes de inspiración, perseverancia y superación que he tenido en la vida.

A mi pareja, Rocío, mi alma gemela. Por ser capaz de comprenderme en todo momento con y sin palabras, por ayudarme a ser mejor persona cada día y por ser la estrella que ilumina mi camino.

A toda mi familia de LexiCon, en particular a Pilar, Arianne, Melania, Juan, Amal, Clara, Maribel, Ana y Miguel. Por todas las aventuras vividas, por todas las fatigas compartidas y por todo lo que he aprendido de cada uno de ellos.

A todos los demás amigos y compañeros que me han rodeado durante estos años. Por la influencia que han tenido en mí y por el granito de arena que, directa o indirectamente, han aportado a esta tesis.

Al pasado, por haberme permitido llegar hasta aquí.

Al futuro... por todo lo que queda por llegar.

Gracias de todo corazón.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AI: artificial intelligence
BIO: Biology
CAT: computer-assisted translation
CHEM: Chemistry
CIV: Civil Engineering
CQL: corpus query language
CTT: Communicative Terminology Theory
DTD: Document Type Definition
EE: environmental event
EEC: EcoLexicon English Corpus
ENVO: Environmental Ontology
ESSG: EcoLexicon Semantic Sketch Grammar
EU: European Union
FBT: Frame-based Terminology
GEO: Geology
GTT: General Terminology Theory
KP: knowledge pattern
KRC: knowledge-rich context
KWIC: keywords-in-context
MWT: multi-word term
NLP: natural language processing
OWL: Web Ontology Language
PoS: part-of-speech
RegEx: regular expression
SCTT: Sociocognitive Terminology Theory
ST: Socioterminology
SWT: single-word term
TKB: terminological knowledge base
WIPO: World Intellectual Property Organization
WS: word sketch

ABSTRACT

Terminology has evolved from static and prescriptive theories to dynamic and cognitive approaches. Thanks to these approaches, there have been significant advances in the design and elaboration of terminological resources. This has resulted in the creation of tools such as terminological knowledge bases, which are able to show how concepts are interrelated through different semantic or conceptual relations. Of these relations, hyponymy is the most relevant to terminology work because it deals with concept categorization and term hierarchies.

This doctoral thesis presents an enhancement of the semantic structure of EcoLexicon, a terminological knowledge base on environmental science. The aim of this research was to improve the description, categorization, and representation of hyponymy in environmental terminology. Therefore, we created HypoLexicon, a new stand-alone module for EcoLexicon in the form of a hyponymy-based terminological resource. This resource contains twelve terminological entries from four specialized domains (Biology, Chemistry, Civil Engineering, and Geology), which consist of 309 concepts and 465 terms associated with those concepts.

This research was mainly based on the theoretical premises of Frame-based Terminology. This theory was combined with Cognitive Linguistics, for conceptual description and representation; Corpus Linguistics, for the extraction and processing of linguistic and terminological information; and Ontology, related to hyponymy and relevant for concept categorization.

HypoLexicon was constructed from the following materials: (i) the EcoLexicon English Corpus; (ii) other specialized terminological resources, including EcoLexicon; (iii) Sketch Engine; and (iv) Lexonomy. This thesis explains the methodologies applied for corpus extraction and compilation, corpus analysis, the creation of conceptual hierarchies, and the design of the terminological template. The results of the creation of HypoLexicon are discussed by highlighting the information in the hyponymy-based terminological entries: (i) parent concept (hypernym); (ii) child concepts (hyponyms, with various hyponymy levels); (iii) terminological definitions; (iv) conceptual categories; (v) hyponymy subtypes; and (vi) hyponymic contexts. Furthermore, the features and the navigation within HypoLexicon are described from the user interface and the admin interface.

In conclusion, this doctoral thesis lays the groundwork for developing a terminological resource that includes definitional, relational, ontological and contextual information about specialized hypernyms and hyponyms. All of this information on specialized knowledge is simple to follow thanks to the hierarchical structure of the terminological template used in HypoLexicon. Therefore, not only does it enhance knowledge representation, but it also facilitates its acquisition.

Keywords: *terminology, terminography, hyponymy, conceptual relations, specialized knowledge*

RESUMEN

La terminología ha evolucionado desde teorías prescriptivas y estáticas hasta enfoques cognitivos y dinámicos. Gracias a estos puntos de vista modernos, se han realizado grandes avances en el diseño y en la elaboración de recursos terminológicos, lo que ha llevado a la creación de herramientas como las bases de conocimiento terminológico. Por ejemplo, estos recursos pueden mostrar cómo se relacionan los conceptos entre sí mediante diferentes relaciones semánticas o conceptuales. De todas estas relaciones, la hiponimia es la más importante para la terminología porque es la encargada de la categorización conceptual y de las jerarquías de términos.

La presente tesis doctoral introduce una mejora en la estructura semántica de EcoLexicon, una base de conocimiento terminológico medioambiental. El objetivo de la investigación fue mejorar la descripción, categorización y representación de la hiponimia en la terminología medioambiental. De esta forma, se creó HypoLexicon, un nuevo módulo independiente para EcoLexicon bajo la forma de un recurso terminológico centrado en la hiponimia. En total, este recurso contiene doce entradas terminológicas pertenecientes a cuatro dominios especializados distintos (biología, química, ingeniería civil y geología), que se componen de 309 conceptos y de 465 términos asociados a dichos conceptos.

La investigación se basó sobre todo en los principios teóricos de la terminología basada en marcos. También estuvo influenciada por la lingüística cognitiva, pertinente en cuanto a la descripción y representación conceptual; la lingüística de corpus, empleada para la extracción y el procesamiento de la información lingüística y terminológica; y la ontología, relacionada con la hiponimia e importante para la categorización conceptual.

HypoLexicon se construyó a partir de los siguientes materiales: (i) el corpus en inglés de EcoLexicon; (ii) otros recursos terminológicos especializados, incluyendo EcoLexicon; (iii) Sketch Engine; y (iv) Lexonomy. Esta tesis proporciona una explicación de las metodologías aplicadas durante la compilación de corpus, el análisis de corpus, la creación de las jerarquías conceptuales y el diseño de la plantilla terminológica. Por otro lado, se ofrece un análisis de los resultados obtenidos con la creación de HypoLexicon, haciendo énfasis en la información contenida en las entradas terminológicas hiponímicas: (i) concepto padre (hiperónimo); (ii) conceptos hijo (hipónimos, con varios niveles de hiponimia); (iii) definiciones terminológicas; (iv) categorías conceptuales; (v) subtipos de hiponimia; y (vi) contextos hiponímicos.

Además, también se describen las funciones y la navegación dentro de HypoLexicon tanto desde la interfaz de usuario como desde la interfaz de administrador.

En conclusión, esta tesis doctoral sienta las bases para el desarrollo de un recurso terminológico que incluya información definicional, relacional, ontológica y contextual sobre hiperónimos e hipónimos especializados. Toda esta información sobre conocimiento especializado es fácil de seguir gracias a la estructura jerárquica de la plantilla terminológica empleada en HypoLexicon, por lo que no solamente mejora la representación de dicho conocimiento, sino que también facilita su adquisición.

Palabras clave: *terminología, terminografía, hiponimia, relaciones conceptuales, conocimiento especializado*

Description, Categorization, and Representation of Hyponymy in Environmental Terminology

1. INTRODUCTION

1.1. RATIONALE AND INTRODUCTION TO THE RESEARCH

Terminology is the study of specialized language, namely, the terms and phrases used in scientific and technical domains. Though interpreted in different ways (Sager 1994), Terminology is an interdisciplinary domain that includes not only linguistic but also extralinguistic aspects, such as elements of human perception and computational processes. Terminology arose from the need to unify concepts and terms in specialized subject fields in order to facilitate professional communication and knowledge transfer (Cabré 2000).

Most Terminology theories have practical applications, such as encyclopedias, specialized dictionaries, knowledge bases and other terminological or translation resources (Faber 2012), which are the flagship for their approach. These resources ideally display their information so that it can be easily retrieved and used by different user profiles (Sager 1990). This practice-based facet of Terminology, aimed at systematically describing and representing previously collected terminological data, is also often referred to as Terminography (Temmerman 2000).

Terminology has evolved from static and prescriptive theories (Wüster 1968, 1979) to dynamic and cognitive approaches (Cabré 1999; Faber 2009). Thanks to these modern approaches, there have been significant advances in the design and elaboration of terminological resources. Over the years, traditional paper-based glossaries and dictionaries have been gradually replaced by electronic or digital versions, which can also be easily updated and modified. In recent years, terminological knowledge bases (TKBs) have become an important linguistic resource, showing a wide range of linguistic and non-linguistic information through intuitive interfaces (Meyer *et al.* 1992).

An example of a modern TKB is EcoLexicon (Faber *et al.* 2016). It is a multidimensional and dynamic TKB on environmental science that provides conceptual, linguistic, phraseological, and multimodal data in each entry. EcoLexicon, apart from its ontological approach, is characterized by its visualization of conceptual networks, showing how concepts are interrelated through different semantic or conceptual relations – generic-specific, part-whole, and non-hierarchical relations. Of these relations, generic-specific or hyponymic relations are particularly relevant to terminology because they deal with concept categorization and term hierarchies

(Murphy 2006). For this reason, the description of concepts and terms can be greatly improved by highlighting their hyponymic information.

Hyponymy describes a hierarchical relation between a generic or parent concept (represented by a hypernym) and a specific or child concept (represented by a hyponym). This conceptual relation is important for terminology work and for the development of linguistic resources because it is the backbone of all conceptual hierarchies (Barrière 2004a). Furthermore, it also plays an important role in our conscious thinking about word meaning.

The most common template for creating entries in terminological resources is intensional and based on the classic Aristotelian definition, which uses hyponymy to describe a concept by listing the *genus* or a hypernym and the *differentiae* or qualities that distinguish the hyponym from a larger class (Murphy 2003, 2006).

For example, in the terminological definition of the concept EARTHQUAKE, "geologic phenomenon" is the *genus*, and "involving a sudden movement of the Earth's crust" and "caused by the release of stress accumulated along geologic faults or by volcanic activity" are the *differentiae,* which specify the process, cause, and location of the earthquake. This definitional template makes it possible for users to directly associate the hyponyms with the hypernym and to distinguish possible co-hyponyms in conceptual hierarchies. In addition, it also facilitates categorization by focusing on the shared characteristics of concepts, thus enhancing the representation of the conceptual hierarchies and providing an ontological approach (Gil-Berrozpe *et al.* 2019).

However, hyponymic relations are complex, and thus hypernym-hyponym pairs can be studied by taking different criteria into account. This research follows the premises of Frame-based Terminology (Faber *et al.* 2005, 2006; Faber 2012, 2022) to study the description, categorization, and representation of hyponymy in Terminology, specifically focusing on environmental terminology. The description of hyponymy is based on analyzing its relevance in terminology work and in ontologies. Its categorization is based on its classification according to multiple perspectives and to different approaches to hyponymy refinement. Its representation is based on reviewing how it is commonly portrayed in traditional and contemporary terminological resources.

Furthermore, as an innovation to the study of hyponymy, this study also contributes to the description, categorization, and representation of hyponymy by introducing a new hyponymy-related enhancement for EcoLexicon. It is known as

HypoLexicon and it is a stand-alone module in the shape of a terminological resource designed to facilitate the representation of hyponymic information and its acquisition. HypoLexicon contains twelve entries from four specialized domains (Biology, Chemistry, Civil Engineering, and Geology) and includes definitional, relational, ontological, and contextual information about specialized hypernyms and hyponyms.

1.2. OBJECTIVES

The main aim of this research was to study the phenomenology of hyponymy by focusing on its description, categorization, and representation in environmental terminology. In particular, this implied analyzing the relevance of hyponymy in terminology work and in ontologies, the multiple perspectives of hyponymy, the different approaches to hyponymy refinement, and how generic-specific relations are displayed in traditional and contemporary terminological resources. For the purpose of achieving this main goal, the following specific objectives were established:

- To review the state of the art on hyponymy and address the specific characteristics and particularities of hyponymy in Terminology, namely in environmental terminology.
- To extract, compile, and process a series of corpora belonging to different subdomains within the environmental sciences for the subsequent corpus analysis.
- To extract, identify, and select a set of hypernyms and hyponyms by using advanced corpus query methodologies.
- To create a series of conceptual hierarchies and populate them with hyponymic, definitional, relational, and contextual information coming from the corpus analysis and from a selection of specialized terminological resources.
- To design a terminological template with a hierarchical structure, suitable for showcasing all the hyponymic information about specialized concepts.
- To develop a new module for EcoLexicon by implementing and displaying all the data of this research.

This thesis introduces an innovative proposal towards the description, categorization, and representation of the generic-specific relation with a view to developing a new module for EcoLexicon. This proposal materialized in the form of a stand-alone module named HypoLexicon, which is actually a hyponymy-based terminological resource on its own. All the steps followed during this research were based on disciplines such as Terminology, Terminography, Applied Linguistics, Cognitive Linguistics, Computational Linguistics, and Corpus Linguistics.

1.3. OUTLINE OF THE THESIS

The rest of this doctoral thesis is organized as follows. Chapter 2 reviews the theoretical framework for this research, which is based on Cognitive Linguistics (**§2.1**) and Frame-based Terminology (**§2.2**).

The section on Cognitive Linguistics describes conceptualization (**§2.1.1**), which includes conceptual systems, cognition, and Frame Semantics. Categorization (**§2.1.2**) is explained by focusing on Prototype Theory and the establishment of conceptual categories. Conceptual relations (**§2.1.3**), which include both hierarchical and non-hierarchical relations, are also studied.

The section on Frame-based Terminology first provides an overview of Terminology (**§2.2.1**). It explains the distinction between Lexicology and Terminology, and it reviews the most prominent Terminology theories: (i) the General Terminology Theory; (ii) social and communicative Terminology theories, represented by Socioterminology and the Communicative Terminology Theory; and (iii) cognitive-based Terminology theories, such as the Sociocognitive Terminology Theory and Frame-based Terminology.

This is followed by an explanation of the theoretical and practical foundations of Frame-based Terminology (**§2.2.2**), which focuses on conceptual organization, multidimensionality, frame-based definitional templates, and knowledge extraction. Finally, after first tracing the evolution from dictionaries to terminological knowledge bases, the following section describes EcoLexicon (**§2.2.3**), the practical application of Frame-based Terminology, and highlights the main features of this terminological resource.

Chapter 3 describes hyponymy (**§3.1**) as well as its categorization (**§3.2**) and representation (**§3.3**). It underlines the importance of hyponymy in Terminology (**§3.1.1**) by explaining its characteristics, including taxonomy, incompatibility, troponymy, and autohyponymy. However, hyponymy goes beyond Terminology (**§3.1.2**) since it is also basic to ontologies, which are the focus of this section along with termontography and ontoterminology. Still another link between ontologies and Terminology is the ontological knowledge enhancement in EcoLexicon. Accordingly, we explain both the conceptual categorization process and the conceptual category module that resulted from it.

The categorization of hyponymy is based on how this conceptual relation can be classified within the context of taxonomic and functional hyponymy (**§3.2.1**) and that of direct and indirect hyponymy (**§3.2.2**). Another possibility described is related to its refinement in EcoLexicon (**§3.2.3**) by identifying hyponymy subtypes and hyponymic knowledge patterns. Lastly, the presence of hyponymy in traditional resources (**§3.3.1**), such as dictionaries and encyclopedias, is compared to its presence in contemporary resources (**§3.3.2**), namely, IATE, TERMIUM Plus, WIPO Pearl, and EcoLexicon. The chapter concludes with a discussion of criteria for establishing a new comprehensive representation of hyponymy (**§3.3.3**).

Chapter 4 describes the materials (**§4.1**) and methods (**§4.2**) followed in this study. The EcoLexicon English Corpus (**§4.1.1**) was used to obtain a series of specialized subcorpora: a Biology subcorpus, a Chemistry subcorpus, a Civil Engineering subcorpus, and a Geology subcorpus. Additionally, a selection of specialized terminological resources (**§4.1.2**), including EcoLexicon as well as various dictionaries and encyclopedias, were also part of the materials. The software (**§4.1.3**) in this research included the EcoLexicon internal application, Sketch Engine, and Lexonomy.

The subcorpora were extracted and compiled with the EcoLexicon internal application (**§4.2.1**). Sketch Engine was used to process and analyze the subcorpora (**§4.2.2**) in order to extract, identify, and select the hypernyms and hyponyms. The information from both corpus analysis and the specialized terminological resources was employed to create the conceptual hierarchies (**§4.2.3**) in the following steps: (i) construction of terminological definitions; (ii) classification of conceptual categories; (iii) structuration of conceptual hierarchies with hyponymy subtypes; and (iv) extraction and insertion of the hyponymic contexts. Lexonomy was used to design the terminological template (**§4.2.4**) for the representation and description of all hyponymic information.

Chapter 5 presents the results of this study. These results include the twelve hyponymy-based terminological entries (**§5.1**), which are based on the terminological template designed as well as on the hyponymic information extracted. This is followed by a description of HypoLexicon, a hyponymy-based terminological resource (**§5.2**), with special focus on the home view and entry view of the user interface (**§5.2.1**). Also highlighted is the admin interface (**§5.2.2**), with the entry settings and the entry editor. Finally, a summary with statistical data pertaining to the results (**§5.3**) is provided with a focus on the following: (i) the general information in all hyponymy-based terminological entries (**§5.3.1**); (ii) ring diagrams showcasing all conceptual categories and hyponymy subtypes (**§5.3.2**), and (iii) comparison with the original information contained in EcoLexicon (**§5.3.3**).

Finally, Chapter 6 presents the conclusions derived from this research and explores ideas for new lines of research and future work.

2. THEORETICAL FRAMEWORK

This chapter reviews the theoretical premises of Cognitive Linguistics and Frame-based Terminology applied in this research. Cognitive Linguistics is relevant because of its emphasis on conceptualization, categorization, and conceptual relations. However, the foundation of this study is Frame-based Terminology because it is the source of the materials used and the methodology applied.

2.1. COGNITIVE LINGUISTICS

Cognitive Linguistics (Lakoff & Johnson 1980; Lakoff 1987, 1990, 1993; Langacker 1987, 1991; Gibbs 1996; Talmy 2000; Croft & Cruse 2004) focuses on cognitive explanations for grammatical structure. It is a functional theory that incorporates ideas from knowledge domains such as Linguistics, Philosophy, Psychology, and Neuroscience. Theories within the broad framework of Cognitive Linguistics include Construction Grammar (Fillmore & Kay 1987; Goldberg 1995, 2005; Kay & Fillmore 1999), Conceptual Semantics (Jackendoff 1983, 1990, 1997), Cognitive Semantics (Talmy 2000), and Frame Semantics (Fillmore 1968, 1975, 1977, 1982, 1985).

Cognitive Linguistics is particularly relevant to this research because of the *cognitive shift* (Evans & Green 2006) in linguistic theory that also resulted in the development of cognitive-based Terminology theories. These theories, which focus on the conceptual network underlying language (Faber 2009) explain conceptualization (i.e., concept description and representation) and categorization (i.e., category structure and ontology-like features), among other elements.

Cognitive Linguistics is not homogeneous and thus cannot be understood as a single theory by itself. However, it has the following basic premises: (i) language is not an autonomous cognitive faculty; (ii) grammar is a conceptualization process; and (iii) knowledge of languages emerges from language use (Croft & Cruse 2004).

Even more important is the fact that language is regarded as a reflection of the human mind. Analyzing language from a cognitive perspective can thus lead to new insights into linguistic phenomena that include studies in phonology, syntax, semantics, pragmatics and psychological aspects of language use (Gibbs 1996). However, this approach signifies agreeing with the *generalization commitment* and the *cognitive commitment*, as proposed by Lakoff (1990).

On the one hand, the *generalization commitment* states that general principles underlying the theoretical descriptions of linguistic phenomena must be characterized. For instance, in syntax there are generalizations about how grammatical morphemes, categories, and constructions are distributed. Another example can be found in semantics, where there are generalizations about phenomena such as polysemy, synonymy, semantic fields, and inferences, *inter alia*. In addition, metaphor is a phenomenon that clearly follows this commitment and reveals that linguistic elements are also affected by conceptual or cultural aspects (Evans & Green 2006; Evans *et al*. 2007). Cognitive Linguistics thus addresses how these domains are related and influenced by each other.

On the other hand, the *cognitive commitment* highlights the importance of incorporating a wide range of information from brain-oriented or mind-oriented disciplines into the theoretical description of language. This commitment encourages the use of empirical data from related disciplines such as Cognitive and Developmental Psychology, Psycholinguistics, Anthropology and Neuroscience (Gibbs 1996). According to Croft & Cruse (2004: 2), the organization and retrieval of linguistic knowledge is not so very different from the organization and retrieval of other knowledge in the mind. Moreover, the cognitive abilities applied to speaking and understanding language are not significantly different from those applied to other cognitive tasks, such as visual perception, reasoning, and motor activity.

In summary, the *generalization commitment* requires the representation of the general principles underlying human language, and the *cognitive commitment* requires such general principles to be described in accordance with cognitive processes within the human mind. If there is a contradiction between the two commitments, it is the *cognitive commitment* that prevails, because generalizations based on the reality of cognition are required (Lakoff 1990).

Apart from these premises of Cognitive Linguistics, there are also other approaches within its framework that are relevant to our research, such as Cognitive Semantics (Talmy 2000) and Frame Semantics (Fillmore 1968, 1975, 1977, 1982, 1985). The three cognitive phenomena that are at the core of this research are the following: conceptualization (§**2.1.1**), categorization (§**2.1.2**), and conceptual relations (§**2.1.3**).

2.1.1. CONCEPTUALIZATION

Within Cognitive Linguistics, it is Cognitive Semantics that studies knowledge representation through conceptual systems and meaning construction. In other words, this theory explains how language codifies and reflects conceptual structures, and hence emphasizes conceptualization (Evans & Green 2006; Evans *et al*. 2007). In fact, conceptualization is understood as the process by which certain phenomena or elements are organized through the description of their salient characteristics (Studer *et al*. 1998). Faber (2011: 10) states that knowledge of conceptualization processes as well as the organization of semantic information in the brain should underlie any theoretical assumptions concerning the access, retrieval, and acquisition of specialized knowledge as well as the design of specialized knowledge resources.

One of the objectives of this research was to describe and enhance the representation of hyponymy in a terminological resource. Since a useful terminological knowledge base (TKB) should reflect the way concepts are organized in the mind (Meyer *et al*. 1992), it is imperative that hyponymic relations portray conceptual structure and how knowledge is transmitted from generic to specific concepts.

2.1.1.1. CONCEPTUAL SYSTEMS

Before examining conceptual systems, it is necessary to clarify what a concept is. According to the studies in Cognitive Linguistics, a concept is a construct of the mind that allows us to organize classes of objects in a prototypical way (Lakoff 1990). A concept is thus an element of thought, a mental classification that a term designates to facilitate communication and knowledge sharing between humans. Terms have different dimensions depending on the theory describing them. They can be viewed from a linguistic (Sager 1990), communicative (Cabré 1993), social (Gambier 1987; Gaudin 1993), or cognitive (Temmerman 2000; Faber 2012) perspective.

In Linguistics (Croft & Cruse 2004) and Cognitive Psychology (Medin 1998; Medin *et al*. 2000; Murphy 2002), concepts and their classification reflect the categorization process (**§2.1.2**). However, there is no consensus of opinion as to the definition, typology or even the internal structure of concepts. This depends on the objectives of the model. Concepts are thus difficult to describe. They are often regarded as a cluster of features or characteristics. Nonetheless, meaning is not a stable entity and can vary. Features, even belonging to the same concept, can be more or less central, depending on the context and knowledge domain (León-Araúz 2009).

In this line, a set of interrelated concepts becomes a conceptual system, which is a more complex mental representation consisting of concepts linked by horizontal and vertical relations. Accordingly, conceptual relations are the cognitive or semantic links that associate and differentiate concepts from each other.

Conceptual systems can also be seen as a more accurate representation developed from a global frame such as ontologies or events (León-Araúz 2009). A conceptual system depicts the types of conceptual relations that link concepts. Moreover, it displays the way in which concepts interact within processes, depending on their internal structure and their nature. This type of representation follows what Talmy (2000) calls *cognitive representation*, since human conceptual system is composed of two interlinked systems: (i) the conceptual system that shows the structural properties of a situation or scenario (i.e., the macrostructure); and (ii) the conceptual system that stores the content, with a more detailed and comprehensive representation of the situation or scenario.

Conceptual systems have been studied in other knowledge fields, such as Cognitive Psychology, Artificial Intelligence and Ontology Engineering, which agree that knowledge can be modelled using a set of elements linked by relations (Leake *et al*. 2004). For instance, Novak & Gowin (1984) developed *conceptual maps* to graphically represent knowledge structure. In Computational Engineering, Quillian (1968) proposed a series of *semantic networks* in which nodes and arcs show a hierarchical organization, whereas Sowa (1984) proposed *conceptual graphs* to facilitate human-machine communication through first-order logic. In Terminology, even though some authors prefer to speak of *semantic relations* and *semantic networks* instead of *conceptual relations* and *conceptual systems* (L'Homme *et al*. 2003), it is clear that a conceptual system is composed of propositions that connect concepts through relations and which possess a hierarchical structure (Schmitz 2006).

Furthermore, an important feature related to conceptual systems is multidimensionality. Multidimensionality is a phenomenon of conceptual classification that arises when concepts are classified in more than one way within a conceptual system according to different characteristics (Bowker 2022). As explained by Kageura (1997: 120):

> It is generally accepted in terminological studies that a concept, or a unit of thought, consists of a set of characteristics (Felber 1984; Sager 1990). […] Since the characteristics of a concept are frequently specified from different points of view or facets (function, material, shape, weight, etc.), a set of

> characteristics that constitutes a concept is normally multidimensional. From this point alone, we can expect a concept system to be multidimensional.

Conceptual multidimensionality can affect a wide range of properties and attributes of concepts. The representation of multidimensionality is thus a major challenge in the design of knowledge resources such as TKBs. There may be contexts where a set of concepts need only be classified in a single dimension to meet a specific need. However, focusing on just one dimension is not efficient because the users of terminology resources may have several different needs (Bowker 2022). For this reason, it is necessary to identify and represent the occurrences of multidimensionality. Otherwise, the result would be a monodimensional system with simple relations between concepts, which, besides being unrealistic, does not facilitate *in vitro* knowledge acquisition (Dubuc & Lauriston 1997; Cabré 1999). The notion of multidimensionality (**§2.2.2.2**) is a key element in Frame-based Terminology (**§2.2**).

Not surprisingly, the configuration of specialized concepts in conceptual systems with conceptual relations has proven to be one of the most important aspects of current terminology work (León-Araúz *et al*. 2012). Nevertheless, this task is far from simple because, in certain cases, the semantics of the relations is too vague, as can be observed in thesauri, conceptual maps, and semantic networks (Jouis 2006). That is the reason why a wide range of methods for structuring knowledge have been applied in different approaches to terminology work. The best methodology used to create high-quality terminological products should be based on logical properties that facilitate the accurate organization of conceptual relations.

2.1.1.2. COGNITION

In relation to conceptualization, cognition is another phenomenon that is relevant to Cognitive Linguistics and Terminology. Cognition involves storing property information in sensory modalities so that this information can later be reactivated in context (Damasio & Damasio 1994; Barsalou 2008). Apart from sensory information, cognition is also a relational process in which meaning and intentions are emergent products of social interaction (Temmerman 2022). Conceptual structure is based on our experience, which leads to *experimentalism* (Lakoff & Johnson 1980) or *embodied experience* (Gibbs 2003). Embodied experience means that our vision of the world depends on what is experienced by our body, either physically or subjectively, and that our brain processes concepts through simulations (Faber & León-Araúz 2014).

The most relevant theories of cognition are related to what is known as *grounded cognition* (Barsalou *et al*. 1993; Barsalou 2003, 2008). According to theories on grounded cognition, conceptual representations underlying knowledge in the human mind are based on the sensory and motor systems instead of being represented and abstractly processed in amodal structures of conceptual information (Pezzulo *et al*. 2013).

On the one hand, traditional theories of cognition state that knowledge is stored within a semantic memory system independent of the modal systems of perception (e.g., sight, hearing), action (e.g., movement, proprioception), and introspection (e.g., mental states, feelings) (Barsalou 2008). Therefore, those representations from modal systems become amodal when stored in semantic memory. However, grounded theories of cognition state that amodal symbols do not represent knowledge in semantic memory, and even question the existence of amodal representations in the brain (Barsalou 2008).

One of the most important aspects shared by embodied and grounded cognition theories is that knowledge about concepts is obtained through simulation, especially sensory and motor simulation (Barsalou 2009; Faber 2011; Tercedor Sánchez *et al*. 2011). This hypothesis claims that interactions between sensorimotor systems and the physical world underlie cognition. When a physical object is encountered, our senses represent it during perception and action. Processing the object then involves partially capturing property information on these modalities so that this information can later be activated (Damasio & Damasio 1994).

In fact, the results of brain-imaging experiments show how simulation affects cognition. For instance, when perceiving everyday objects (e.g., a cup), simulations of potential actions are triggered (e.g., action of grasping such cup by the handle or drinking its content) (Tucker & Ellis 1998, 2001). Neuroimaging research has also been done with specialized knowledge (Faber *et al*. 2014b, 2017), thus confirming that simulation is a key part of cognition and conceptual processing at all levels. When conceptual knowledge about objects is represented, this activates the brain areas that represent their properties during perception and action, particularly, the areas that represent the shape and color of objects, their motion, and the actions that agents perform on them.

These reenactments not only occur in the presence of the object itself, but also in response to words and other symbols (León-Araúz *et al*. 2012). Accordingly, as soon as we want to reflect and communicate about our conceptualizations and cognitive findings (Krishna & Schwarz 2014), our sensory experience needs to be captured in linguistic expressions. Furthermore, the words or terms used are also influenced by culture or shared knowledge. For instance, expectation patterns when tasting food and drinks are culturally and linguistically determined because of previous experiences (Shapiro 2004), since we are taught to use words to express what we taste in a cultural environment. Apart from this, another way that embodiment influences conceptual systems according to grounded cognition is also through metaphorical projections (Johnson 1987), because conceptual metaphor is a mechanism by which a concept is understood and experimented through a different one (Lakoff & Johnson 1980).

Therefore, theories on cognition are relevant for understanding linguistic and conceptual phenomena. Given the fact that conceptual relations, such as hyponymy, are based on the nature of concepts and concept interaction, grounded cognition can optimize the description of characteristics and the representation of hypernym-hyponym pairs.

2.1.1.3. FRAME SEMANTICS

Frame Semantics (Fillmore 1968, 1975, 1977, 1982, 1985) is a cognitive theory largely influenced by conceptualization and thus by conceptual systems and cognition. It describes lexical meaning by extracting contextual information from a large corpus of texts, and then structuring this meaning in cognitive frames. Frames are described as unified frameworks of knowledge or coherent schematizations of experience (Fillmore 1985: 223), which refers to the simulation present in embodied and grounded cognition theories. Moreover, Frame Semantics assumes that the meanings of lexical units or terms are construed against the background of experience, beliefs, and practices (Fillmore & Atkins 1992). Additionally, frames also differ with regard to the level of specialization of the situation, as in the following example (Fillmore 1982: 127):

> In both everyday language and legal language there is a contradictory opposition between INNOCENT and GUILTY. In everyday language, the difference depends on whether the individual in question did or did not

commit the crime in question. In legal language, by contrast, the difference depends on whether the individual in question has or has not been declared guilty by the court as a result of a legal action within the criminal system.

One of the core ideas of Frame Semantics is that meanings have to be described within a schematic representation of such structures (i.e., the frames). A frame is a system of concepts which are interrelated in such way that the activation of a single concept evokes the whole conceptual system. Concepts are thus not only related through shared characteristics, but also through frames that include properties and actions as well.

Consequently, frames are cognitive, dynamic, and flexible structures that can be applied to different fields and to different categorization levels. In most cases, frames reflect concepts that are codified in language and which are retrieved by a person based on their background or previous knowledge. Some of these frames are intrinsic, but others are learnt through experience or training (Kittay & Lehrer 1992). With regard to this, Lee (2001: 9) highlights the conceptual and cultural dimension of frames, and states that frames embrace the traditional concept of connotation. For instance, the frame activated by the concept HOLIDAY carries a complex range of associations to activities, time periods, emotional states, interaction with other entities, etc.

One of the most typical examples of frames is COMMERCIAL TRANSACTION (Fillmore & Atkins 1992), which involves a series of entities (e.g., the buyer, the seller, the product), actions (e.g., buying, selling, asking), and situations (e.g., a market, a shop). In this linguistic interpretation of frame, each word allows the speaker and the hearer to focus their attention on one part of the entire frame. Words evoke a frame, but by filling slots in the frame, they foreground only those parts of the frame closely linked to the slot.

The practical application of Frame Semantics is FrameNet (Fillmore & Atkins 1998; Fillmore *et al*. 2003; Ruppenhofer *et al*. 2010). It is the result of a computational lexicographic project based on extracting information about syntactic-semantic properties of words from a large English electronic corpus. The English version of FrameNet now contains over 1200 frames and over 200,000 manually annotated sentences. According to Fillmore *et al*. (2003: 235), FrameNet identifies and describes semantic frames, and analyzes the meanings of words by directly appealing to the frames that underlie their meanings. It also studies the syntactic properties of words by asking how their semantic properties are given syntactic form. This is accomplished by (i) schematically characterizing the type of entity or situation represented by the frame; (ii) labelling the components of the frame; and (iii) listing the words that belong to such frame, so that propositions containing these lexical units will allow for comparable semantic analyses.

Specialized knowledge is also structured and organized in accordance with the elements present in frames. Because of their combination of linguistic and cognitive elements, they have become a key element in contemporary theories of Terminology, namely Frame-based Terminology (**§2.2.2.1**).

2.1.2. CATEGORIZATION

Categorization is also an important topic in Cognitive Linguistics that is vital to language understanding and meaning representation. Categorization is the mental process that enables humans to classify elements of the world by perceiving similarities and differences between them, and mentally storing concepts that represent such elements (Taylor 1995). In other words, categorization makes it possible for humans to structure and understand the world (Lakoff 1987: 6):

> Without the ability to categorize, we could not function at all, either in the physical world or in our social and intellectual lives. An understanding of how we categorize is central to any understanding of how we think and how we function, and therefore central to an understanding of what makes us human.

Taylor (1995) states that categorization is both a question of motivation (i.e., categories have a perceptual base) and convention (i.e., learning or training), which is related to frames. Categorization is thus a cognitive process that involves the following (Faber & López Rodríguez 2012): (i) forming a structural description of the entity; (ii) identifying categories similar to the structural description; (iii) selecting the most similar representation; (iv) making inferences regarding the entity; and (v) storing information about the categorization in long-term memory.

However, categorization cannot be truly understood without knowing what a category is. Cruse (2006) defines a category as a class of the world's entities. Entity in this context refers to physical entities (e.g., items, animals), abstract entities (e.g., feelings, ideas), properties (e.g., color, size), and actions (e.g., writing, jumping). Similarly, Sloutsky (2003) defines a category as the equivalence class of different entities. Categorization is thus the ability to form categories and treat entities as members of an equivalence class.

In reference to concepts, conceptual structure should be grounded on a set of theoretical assumptions regarding categorization, more specifically, whether and to what extent sensory information is part of semantic representation and processing (Meteyard *et al.* 2012). In this sense, Patterson *et al*. (2007), propose a supramodal format for semantic representations, which is modality-invariant though derived from mappings across sensory and motor input. In Terminology, the correlate of this supramodal representation is a category schema or template as posited by various authors (Faber *et al*. 2014b; Roche *et al*. 2009; Leonardi 2010). This top-level schema constrains perceptual input, although, at the same time, it is also derived from sensorimotor mappings, as argued by the embodied or grounded theories of cognition (**§2.1.1.2**).

There are various theories of categorization (Evans & Green 2006). One of the most important is the Classical Theory, which holds that conceptual category membership is based on a shared set of necessary and sufficient features. Concepts that do not have these features do not belong to the category (Hampton 1991; Keizer 2007). This was the generally accepted approach to categorization until the late 1970s, when Rosch (1978) proposed Prototype Theory, in which conceptual structure is based on judgments of graded similarity.

2.1.2.1. PROTOTYPE THEORY

Prototype Theory (Rosch 1975, 1978; Rosch & Mervis 1975; Rosch *et al*. 1976) is based on research on family resemblance (Wittgenstein 1953), focal colors (Berlin & Kay 1969), and container artifacts such as cups (Labov 1973). According to Prototype Theory, categories are graded according to their similarity to an ideal member or *prototype*, which best represents this category. For instance, the category *vehicle* is more likely to evoke the image of a car in our minds than one of a bus or a truck.

Rosch (1978) argues that prototypes have many features in common with other members of the category, whilst the less prototypical concepts coincide in few characteristics with the rest of the members. The most recurrent features among the concepts of a category are conceived as the most representative. Thus, it is the degree of similarity to the most prototypical member that determines category membership. Categories are organized in concentric circles around the prototype. Least similar concepts are closer to the category limits, whereas the more prototypical concepts are closer the center.

Moreover, Rosch (1978) states that the categorization system has two dimensions: a vertical and a horizontal one. The vertical dimension refers to the level of inclusion of a category: the higher the elevation on the axis, the greater the inclusion of concepts in the category. This dimension alludes to hierarchical relations, such as hyponymy, since more general categories (i.e., hypernyms) have a greater capacity for inclusion. Thus, three levels are distinguished on this vertical axis: superordinate (higher level), basic (intermediate level), and subordinate (lower level) categories. The superordinate level is the most inclusive, but the most useful level for cognitive reasons is the basic level, since its categories have more features in common (Rosch *et al*. 1976: 428):

> Categories at higher levels of abstraction have lower cue validity than the basic because they have fewer attributes in common; categories subordinate to the basic have lower cue validity than the basic because they share most attributes with contrasting subordinate categories.

In this line, since basic categories arise from the interaction between human experience and the environment, they are influenced by embodied or grounded cognition. Accordingly, even though the organization of categories at the superordinate, basic and subordinate levels may be universal, the level at which a given category or concept is included may vary according to the language, culture or level of knowledge of the person (Evans & Green 2006).

On the other hand, the horizontal axis also pointed out by Rosch (1978) refers to the existing differences between categories that present the same degree of inclusion. In general, there are two principles that guide the formation of categories in the human mind: (i) the principle of *cognitive economy*, according to which an effort is made to store the maximum amount of information with a view to saving cognitive effort; and (ii) the principle of *perceived world structure*, which takes into account the concepts that are most frequently related to others.

In summary, Prototype Theory accounts for conceptual classification in the human mind and highlights the existence of representative models and relations between categories, the influence of experience, and the importance of the surroundings in conceptual structuring. Although Prototype Theory dictates a series of cognitive criteria relevant to the conceptual organization of reality, it remains to be determined the nature of the categories encountered.

2.1.2.2. CONCEPTUAL CATEGORIES

Conceptual categories (also known as *concept categories* or *semantic categories*) are ontological classifications to which concepts with a series of shared characteristics belong. The organization of conceptual space through categorization is crucial in Terminology because this discipline focuses on the mapping of specialized knowledge domains. This involves establishing the major conceptual categories in the subject field, ascertaining which concepts qualify for membership in each category, and specifying their relations with each other (Faber 2022).

Category membership and disjunction are guided by a set of conceptual characteristics and subdivision criteria. There are multiple reasons for grouping concepts together: (i) a similar function or purpose; (ii) similar physical, spatial or temporal characteristics; (iii) a similar role in a structure of other entities or concepts; and (iv) some combination of these or other factors (Michalski 1991).

To specify and classify the concepts in a specialized knowledge field, it is usual to begin with a set of terms that are regarded as the most characteristic of that domain. These can be extracted from a specialized knowledge resource or from a domain-specific corpus of texts, based on their frequency or salience. Depending on their meaning, the concepts designated by the terms tend to be first grouped in ontological starter categories such as *entity*, *process*, *attribute*, etc. (Mahesh & Nirenburg 1995; Moreno-Ortiz & Pérez-Hernández 2002). It is worth mentioning that, because of phenomena such as dynamism and multidimensionality (**§2.2.2.2**), certain concepts may be members of more than one category (Kageura 2002).

For instance, physical concepts tend to be natural or artificial objects or processes. Abstract concepts include theories, units of measurement, equations, etc. (Faber *et al*. 2009). These criteria are crucial in terminology work, where the focus is on what differentiates concepts from each other (Madsen & Thomsen 2009).

Although there is a certain correspondence in the starter or higher conceptual categories, the lists of more specific conceptual categories are therefore variable depending on the knowledge domain and the concepts involved.

2.1.3. CONCEPTUAL RELATIONS

Conceptual relations (also known as *concept relations* or *semantic relations*) are links created by human thought processes to describe a type of interaction between concepts. In the end, these relations enable the creation of conceptual systems (**§2.1.1.1**), which are sets of concepts structured according to domain knowledge, and which can be consulted by different user groups. Since a conceptual system is subject to time-induced changes of reality and cognition, it may show variable degrees of formality as it evolves during the steps of terminology work. As explained by Sager (1990: 29):

> Inside subject fields concepts are also related either by their nature or by the real-life connections of the objects they represent. As in real life between objects, the kinds of relationships which exist between concepts are numerous and varied.

Conceptual relations are instrumental in the extraction of conceptual information, the analysis and organization of concepts, the definition of concepts, the specification of connections between concepts and terms, and term formation and evaluation (Nuopponen 2022). Additionally, they are also relevant for presenting terminological information in vocabularies, databases, taxonomies, knowledge representation systems such as ontologies, and other resources such as TKBs. They guarantee the quality of terminological products and facilitate knowledge transmission and acquisition.

Authors such as Meyer *et al*. (1992) and León-Araúz *et al*. (2013a) insist on the idea that cognitive-oriented terminological resources such as TKBs should reflect the way that concepts are structured and related to each other in the human mind. Displaying information about conceptual relations in terminological resources is also didactic (Picht & Draskau 1985), because it provides a better understanding of concepts, terms, and definitions (Madsen *et al*. 2001). Accordingly, this is a harder task than only displaying unrelated terminological entries. However, the benefits of including such relations compensate the effort (Marshman *et al*. 2012), since this facilitates knowledge transfer. As stated by Faber (2011: 10):

> [...] knowledge of conceptualization processes as well as the organization of semantic information in the brain should underlie any theoretical

assumptions concerning the access, retrieval, and acquisition of specialized knowledge as well as the design of specialized knowledge resources.

ISO Standard 704 (2009) states that there are two main types of conceptual relations: (i) hierarchical relations, which include both generic relations and partitive relations; and (ii) non-hierarchical relations, which include associative relations. However, according to Nuopponen (2005: 127), the distinction between these relation types might be enough for traditional terminology work and term banks, but more advanced resources for terminology management or concept modelling could benefit from a wider range of conceptual relations. In fact, how conceptual relations are distinguished from each other depends on the theoretical background or discipline and purpose or context of use or need (Nuopponen 2014).

To illustrate this, Nuopponen (2022) compared three typologies of conceptual relations (**Table 1**): that of the OntoQuery project (Madsen *et al*. 2001; Andreasen *et al*. 2004), that of Maroto & Alcina (2009), and that of EcoLexicon (Faber *et al*. 2014a, 2016).

OntoQuery (Andreasen *et al*. 2004): ***semantic relations***	▪ **Hyponymy** (IS-A) ▪ **Location relation:** dynamic location (source relation: source-target ∣ event-source relation; target relation: source-target ∣ event-target relation) ∣ static location (event-static location ∣ entity-static location relation) ▪ **Purpose relation** ▪ **Event relation:** event-source ∣ event-target ∣ event-static location ∣ entity-static location ∣ event-agent ∣ event-patient ∣ event-theme ∣ event-instrument ∣ event-result relation ▪ **Function relation** (entity-way of working) ▪ **Partitive relation:** subpart relation ∣ partition relation ∣ material relation ∣ set-element relation ▪ **Causal relation** ▪ **Role relation:** agent relation (event-agent ∣ agent-patient ∣ agent-theme ∣ agent-result relation ∣ etc.) ∣ patient relation ∣ theme relation ∣ instrument relation ∣ result relation (all roles combined with each other) ▪ **Measurement relation** ▪ **Characteristic relation** ▪ **Temporal relation:** time-entity ∣ entity-entity temporal relation (phase relation ∣ development relation)

Maroto & Alcina (2009): ***conceptual relationships***	▪ **Logical relationships:** hypernym-hyponym ǀ hyponym-hyponym ▪ **Meronymic relationships:** functional component-object ǀ member-collection ǀ portion-mass ǀ material-object ǀ stage-process ǀ characteristic-activity ǀ space-area ǀ part-part ▪ **Sequential relationships:** concept-simultaneous concept (in space) ǀ concept-place it goes to ǀ concept-simultaneous concept (in time) ǀ previous concept-subsequent concept ▪ **Argumental and circumstantial relationships:** process-agent ǀ process-patient ǀ process-product ǀ cause-effect ǀ process-instrument ǀ process-method ǀ object-use ▪ **Other relationships:** phenomenon-measure ǀ object-characteristic ǀ associative relationship
EcoLexicon (Faber *et al.* 2014a): ***conceptual relations***	▪ **Generic-specific relations:** *type of* ▪ **Part-whole relations:** *part of* ǀ *made of* ǀ *located at* ǀ *delimited by* ǀ *takes place in* ǀ *phase of* ▪ **Non-hierarchical relations:** *opposite of* ǀ *result of* ǀ *has function* ǀ *attribute of* ǀ *affects* ǀ *studies* ǀ *represents* ǀ *measures* ǀ *effected by* ǀ *causes* ǀ *destroys*

Table 1. Typologies of conceptual relations (adapted from Nuopponen 2022)

2.1.3.1. HIERARCHICAL RELATIONS

The classical approach to conceptual relations considers that hierarchical relations are divided into generic relations and partitive relations. On the one hand, generic relations refer to *generic-specific* or *hyponymic* relations – a type of relation between a *hypernym* or *superordinate* concept and its *hyponym* or *subordinate* concept. According to ISO Standard 704 (2009: 9):

> A generic relation exists between two concepts when the intension of the subordinate concept includes the intension of the superordinate concept plus at least one additional delimiting characteristic. […] The superordinate concept in a generic relation is called the generic concept [hypernym] and the subordinate concept is called the specific concept [hyponym].

In the more recent ISO Standard 1087 (2019), they are understood as relations between a *generic concept* and a *specific concept* (e.g., ANIMAL – DOG), where the intension of the latter includes the intension of the former and at least one additional delimiting characteristic. Additionally, the set of specific or subordinate concepts on the same level of abstraction are known as *co-ordinate concepts* (e.g., DOG – CAT).

Nuopponen (2018) proposes a distinction between *direct generic super/subordination* and *indirect generic super/subordination*. Following this approach, direct superordinate and subordinate concepts are those located at subsequent abstraction levels (e.g., SOFTWARE – APPLICATION SOFTWARE), whilst indirect superordinate and subordinate concepts are those located at a different abstraction level further away (e.g., SOFTWARE – TEXT PROCESSING SOFTWARE).

Similarly, a distinction is also made between *direct generic co-ordination*, with co-ordinate concepts on the same level and below the direct superordinate concept and same criteria of division (e.g., APPLICATION SOFTWARE – SYSTEM SOFTWARE); and *indirect generic co-ordination*, with co-ordinate concepts on the same level, but under different direct superordinate concepts or criteria of division (e.g., APPLICATION SOFTWARE – SERVER SOFTWARE). Finally, *generic diagonal relation* refers to other pairs of concepts on different abstraction levels in the same concept system (e.g., SYSTEM SOFTWARE – TEXT PROCESSING SOFTWARE).

Generic-specific or hyponymic relations are the heart of this doctoral thesis. Given the complexity of their description, categorization, and representation, they are described in greater detail in Chapter 3.

Partitive relations refer to *part-whole* or *meronymic* relations – a type of relation between a *holonym* or *whole concept* and a *meronym* or *partial concept*. According to ISO Standard 704 (2000: 9):

> A partitive relation is said to exist when the superordinate concept represents a whole, while the subordinate concepts represent parts of that whole. The parts come together to form the whole. The superordinate concept in a partitive relation is called the comprehensive concept [holonym] and the subordinate concept is called the partitive concept [meronym]. Subordinate concepts at the same level and sharing the same dimension are also called coordinate concepts.

Following the more recent ISO Standard 1087 (2019), they are understood as relations between a *comprehensive concept* and a *partitive concept* (e.g., CAR – WHEEL), where the former is viewed as a whole consisting of various parts and the latter is viewed as part of a whole. However, in addition to relations between whole and part concepts, partitive relations also include *part-part* relations that are based on parts of the same whole (e.g., WHEEL – SUSPENSION SYSTEM).

Cruse (1986) proposes a division between *canonical meronyms* and *facultative meronyms* by determining if a concept needs the part or can exist without the part. Based on this, Nuopponen (1994) also distinguishes canonical or facultative partitive superordination and subordination in partitive conceptual systems. In this line, *canonical superordination* means that the entity needs the part (e.g., CAR – TIRE); *facultative superordination* means that the entity does not need the part (e.g., HOTEL – RESTAURANT); *canonical subordination* means that the part cannot exist without the whole (e.g., FINGER – HAND); and *facultative subordination* means that the part does not need a whole (e.g., TREE – FOREST).

Nonetheless, not all meronymic relations refer to parts of a whole. For instance, other types of part-whole relations include material-component relations (e.g., ALCOHOL – WINE), property relations (e.g., HOT – FIRE), and locative relations (e.g., BOOK – LIBRARY). In EcoLexicon, the basic meronymic relation *part of* is divided into subtypes: *made of*, *located at*, *delimited by*, *takes place in*, and *phase of*. For example, even though CONDENSATION is *part of* the HYDROLOGIC CYCLE, it is more accurate to say that CONDENSATION is a *phase of* the HYDROLOGIC CYCLE. This distinction was made because of domain-specific needs, ontological reasoning, and transitivity-related consistency (León-Araúz & Faber 2010). Interestingly, the other hierarchical relation, hyponymy, has a similar nature and can therefore be divided into subtypes, as is explained in the section regarding hyponymy refinement in EcoLexicon (**§3.2.3**).

2.1.3.2. NON-HIERARCHICAL RELATIONS

Non-hierarchical relations are also known as *associative relations,* and they refer to any other kind of conceptual relation that does not establish a hierarchical dependency between a superordinate concept and a subordinate concept. Instead, they describe a pragmatic relation based on experience or knowledge (e.g., functionality, causality, location, time, etc.). According to ISO Standard 704 (2009: 17–18):

> Associative relations are non-hierarchical. An associative relation exists when a thematic connection can be established between concepts by virtue of experience. Some associative relations exist when dependence is established between concepts with respect to their proximity in space or time. […] Some

> relations involve events in time such as a process dependent on time or sequence; others relate cause and effect.

In the more recent ISO Standard 1087 (2019), three types of associative relations are registered: (i) *spatial relations*, based on the criterion of relative location in space (e.g., FLOOR – CEILING); (ii) *temporal relations*, based on the criterion of following or preceding in time (e.g., PRODUCTION – CONSUMPTION); and (iii) *causal relations*, based on the criterion of cause and its effect (e.g., ACTION – REACTION).

Nuopponen (2022) also has her own proposal of non-hierarchical relations, which include the following: (i) *activity relations*, involving agents, patients, objects, and tools (e.g., RESEARCHER – METHOD); (ii) *origination relations*, referring to originators, ingredients, products, and purposes (e.g., BREAD – OVEN); (iii) *developmental relations*, including stages of individuals, species, generations or materials (e.g., CHILD – ADULT); (iv) *interactional relations*, which are a more diffuse category based on the interplay between objects of reference (e.g., EMPLOYER – EMPLOYEE); and (v) *causal relations*, related to causes, symptoms, effects, and results (e.g., VACCINATION – IMMUNITY).

Another example of a set of associative relations is found in EcoLexicon, with a series of eleven non-hierarchical relations (i.e., *opposite of*, *result of*, *has function*, *attribute of*, *affects*, *studies*, *represents*, *measures*, *effected by*, *causes*, and *destroys*) that are quite self-explanatory and that show a similar nature and reasoning to those of the other proposals.

There is thus no consensus on a closed inventory of non-hierarchical or associative relations because they are domain-dependent and related to knowledge acquisition through the simulation of human interaction with entities of the world (León-Araúz *et al*. 2012). Since they are thus also related to embodied or grounded cognition, they are just as important as the hierarchical relations (i.e., hyponymy and meronymy) from a cognitive point of view.

2.2. FRAME-BASED TERMINOLOGY

Frame-based Terminology (Faber *et al.* 2005, 2006; Faber 2012, 2022) is the theoretical approach on which this doctoral thesis is based. It is a cognitive and text-based approach to Terminology that was developed within the context of the LexiCon Research Group of the University of Granada. It focuses on specialized knowledge representation by paying special attention to its semantic and cognitive aspects. In the words of Faber (2022: 353):

> One of the most crucial issues that must be addressed in any theory of Terminology is how specialized concepts should be represented so as to help users understand their meaning and their relations with other concepts. This type of knowledge representation or conceptual modeling should capture and account for both the micro- and macrocontexts of concepts. It signifies not only representing individual concepts and their terminological designations, but also integrating them into larger knowledge structures or frames in order to map the relations that they hold with others. This is the basis for Frame-based Terminology (FBT).

In what follows, this theory is situated within the frame of Terminology studies (**§2.2.1**). After explaining its theoretical and practical foundations (**§2.2.2**), we then describe EcoLexicon (**§2.2.3**), which is its practical application.

2.2.1. BACKGROUND

In order to understand the origins and evolution of Terminology, it is necessary to clarify the distinction made between Lexicology and Terminology. Thereafter, a review of the main theories of Terminology is carried out, highlighting the following: the General Terminology Theory, social and communicative Terminology theories, and cognitive-based Terminology theories.

2.2.1.1. LEXICOLOGY AND TERMINOLOGY

Lexicology is the discipline that studies and describes the lexicon of a language, whilst Lexicography is the compilation and design of general language resources (Pérez 2002). Similarly, Terminology studies and describes specialized language, whereas Terminography focuses on the elaboration of specialized language resources.

As is well-known, *terminology* is polysemic since it can refer to three different concepts (Nkwenti-Azeh 1998): (i) a theory that explains the relation between concepts and terms; (ii) the activity of compiling, describing, and presenting terms; and (iii) the vocabulary of a specialized domain. In addition, sometimes the words *terminology* and *terminography* are indiscriminately used when referring to the practical application of Terminology (Montero-Martínez *et al*. 2011). In this thesis, *Terminology* (with a capital T) refers to the discipline, and *terminology* (with a lowercase T) or *terminology work* refers to the practical application of this type of activity.

The main difference between Lexicology and Terminology or Lexicography and Terminography evidently resides in the nature of the language. This distinction is analyzed by Cabré (1993), who states that the rules, units, and restrictions that form part of the knowledge of most speakers of a language constitutes the common or general language, which is *unmarked*. In contrast, specialized languages contain a series of subcodes dependent of specific criteria (e.g., subject field, interlocutor level, situation, context in which the communicative exchange occurs, etc.) that can be considered as *marked*.

Although the differences between these disciplines is not always clear, there is a certain agreement that lexicologists and lexicographers deal with general language, whilst terminologists and terminographers deal with specialized language (Bergenholtz & Tarp 2010).

2.2.1.2. General Terminology Theory

The General Terminology Theory (GTT) (Wüster 1968, 1979) was the first theoretical proposal in the field of Terminology. In fact, Terminology, as a discipline, began in the 1960's with Eugen Wüster, the author of *The Machine Tool, an Interlingual Dictionary of Basic Concepts* (Wüster 1968), a systematically organized French and English dictionary of standardized terms intended as a model for technical dictionaries. This work set out the initial set of principles for the compilation and description of terminological data with a view to the standardization of scientific language. The GTT itself was later developed in Vienna by Wüster's successors, who interpreted his ideas and carried on his work.

For Wüster (1968, 1979) and the Vienna School, terminology work is prescriptive in nature and oriented towards standardization. In accordance with the GTT, Terminology focuses on the delimitation of concepts and their association with a single term, thus avoiding synonymy and polysemy. Moreover, terms in specialized language are regarded as inherently different from general language words because of the monosemic reference between terms and concepts (Felber 1984).

The general claim is that a term or a specialized language unit can be distinguished from a general language word because of its single-meaning relation with the specialized concept that it designates and by the stability of the relation between form and content in texts dealing with this concept (Pavel & Nolet 2001). This ensures the univocity of professional communication, especially in the context of international communication.

In fact, the main objectives of the GTT were the following (Cabré 2003): (i) the elimination of ambiguity in technical language through the standardization of its terminology to facilitate technical communication; (ii) the dissemination of the benefits of standardized terminology among specialized language users; and (iii) the transformation of Terminology into a science. Interestingly, these premises are still reflected in the work of the International Organization for Standardization (ISO), and in the design and elaboration of modern terminological resources such as IATE (Zorrilla-Agut & Fontenelle 2019) or TriMED (Vezzani & Di Nunzio 2020).

However, the approach of the GTT towards the semantics of terminological units projected a rather limited representation of specialized knowledge concepts without allowing for their multidimensional nature (Faber & López Rodríguez 2012). The GTT did not account for the syntax and pragmatics of specialized language, which were not relevant. While it is true that the GTT was a major breakthrough in the consolidation of Terminology as a discipline, it did not address the social, pragmatic and cognitive elements of language, aspects that were the focus of subsequent theoretical frameworks.

Despite the limitations of the GTT, its importance is still widely acknowledged because, at the time, it offered the only set of principles and premises for compiling terminological data (Faber 2009).

2.2.1.3. SOCIAL AND COMMUNICATIVE TERMINOLOGY THEORIES

It was not until the early 1990's when more descriptive approaches arose as a reaction to the GTT. These new approaches are known as *social and communicative Terminology theories*. Unlike the GTT, these theories describe terminological units in discourse by analyzing the sociological and discourse conditions that give rise to different types of texts. Since they emphasize how terms are used in real communicative contexts, they provide a more realistic view of Terminology. These theories are Socioterminology and the Communicative Terminology Theory.

2.2.1.3.1. SOCIOTERMINOLOGY

Socioterminology (ST) (Gambier 1987, 1991; Boulanger 1991, 1995; Gaudin 1993, 2003, 2005) is a sociolinguistic approach to Terminology that was born in the 1980's in Quebec and France. ST came into existence because terminologists were dissatisfied with the limitations imposed by the GTT, whose primary aim was technical standardization. These restrictions were particularly evident in the terminology methods used for language planning and were a subject of intense debate at various international conferences.

ST (Gaudin 1993, 2003, 2005) recognizes the existence of denominative and conceptual variation in specialized language and studies these dynamic phenomena in their contexts of use. Sociological aspects are thus taken into account, such as usage, users, social status, professional status, ethnic aspects, power relations, geographical location, and even temporal location. Diachrony is also studied, and neology thus acquires an important role in Terminology. Accordingly, one of the most important innovations introduced by ST is the conception of terminology work as descriptive. This gave a more realistic picture of specialized communication in contrast to the closed universe of traditional approaches, where concepts were artificially delimited and univocally associated with a single term.

Furthermore, Pihkala (2001) stresses that ST focuses on the social and situational aspects of specialized language communication, which may affect expert communication and give rise to term variation. According to this approach and in contrast to the GTT, standardization is an extremely difficult goal to achieve because language is in constant change and evolution. Polysemy and synonymy are inevitably present in terminology work and specialized texts, and the use of one term instead of another can reflect sociological aspects of a group of users or participants in the communicative interaction. Terminological variation inevitably highlights the fact that concept systems and definitions are not static. This is a reality that any theory aspiring to explanatory adequacy has to deal with.

ST defends that specialized knowledge is not ordered in delimited compartments, as argued by the GTT, but is organized in the form of connection nodes (Gambier 1991). Nodal theory arose from the descriptive study of specialized communication, which shows that knowledge is in continuous evolution and that a domain is nothing more than an interconnection of nodes.

Although ST did not aspire to becoming an independent theory, its importance resides in the fact that it opened the door for other descriptive theories of Terminology, which also take social and communicative factors into account, and whose theoretical principles are based on the way terms are actually used in specialized discourse (Faber 2009).

2.2.1.3.2. COMMUNICATIVE TERMINOLOGY THEORY

The Communicative Terminology Theory (CTT) (Cabré 1993, 1999, 2000, 2003; Cabré *et al*. 1996; Cabré & Feliu 2001) is a linguistic-communicative-oriented theoretical proposal that emerged in the mid-1990's at the University Institute of Applied Linguistics (*Institut Universitari de Lingüística Aplicada*, IULA) of the Pompeu Fabra University in Barcelona. The CTT expands the scope of ST to reflect the complexity of terms or specialized language units in real situational contexts and from a social, linguistic and cognitive perspective (Cabré 2003).

In a similar way to ST, the CTT states that the boundary between specialized language and general language (i.e., terms and words) is not clear. Therefore, it is the context that determines whether a particular unit is a lexical unit or a terminological unit (Cabré 2003). The CTT points to the nature of Terminology as both a transdisciplinary discipline, because all specialized subjects make use of it, and an interdisciplinary discipline, because it uses different subjects to describe its object of study. More specifically, Terminology integrates theories of knowledge, communication and language (Cabré 1993).

Cabré (2003) proposes the *Theory of the Doors*, a metaphor representing the possible ways of accessing, analyzing, and understanding terminological units. This theory introduces the notion of multidimensionality in terminology and identifies terms as polyhedrons (i.e., three-dimensional solid figures) with three dimensions: (i) the linguistic dimension, which describes the features of a specialized language unit as it is used in language; (ii) the communicative dimension, which describes how a specialized language unit is used in different contexts or communicative situations; and (iii) the cognitive dimension, which focuses on a description of concepts and conceptual relations.

In its linguistic dimension, the CTT studies terminological units themselves, which are part of the natural language and grammar of each language. From the point of view of the CTT, the general and specialized competence of the subject are integrated and include lexical units which, out of context, are neither words nor terms. Terms are modules of features associated with lexical units, which are described as denominative-conceptual units, endowed with reference capacity, and which can exercise different functions. Thus, terminological units are only potentially terms or non-terms and can belong to different domains. The character of term is activated by their use in an appropriate context and situation (Cabré 1999).

In its communicative dimension, the CTT seeks to explicitly describe the types of situations that can be produced, and to represent the correlation between type of situation and type of communication. This involves explaining the characteristics, possibilities, and limits of the different expression systems of a concept and of its units (Cabré 2003). Moreover, this also implies that in order to better examine the communicative situation in which a specialized language unit is activated, terms are analyzed *in vivo*, based on their real use in context (Edo Marzá 2012).

In its cognitive dimension, the aim of the CTT is to account for the way in which reality is conceptualized. To this end, conceptual domains are understood as sets of terminological units linked by different types of relations, where terminological units constitute nodes. The CTT also stresses that concepts can belong to different disciplines, in which they can conserve or modify their characteristics (Cabré 2003). Therefore, the importance of describing hierarchical and non-hierarchical conceptual relations as well as multidimensionality are emphasized. This signifies representing the same concept from different perspectives and with different characteristics.

Furthermore, the CTT divides the discipline of Terminology into a theoretical approach and a practical one. On the one hand, theoretical Terminology focuses on the formal, semantic, and functional description of the units that can acquire terminological value by explaining their activation and relations with other types of signs with a view to advancing knowledge about specialized communication and the units used in it. On the other hand, practical Terminology deals with the collection and analysis of units with terminological value. In addition, although the CTT defends the descriptive nature of Terminology, it also recognizes the existence of situations in which prescriptive work is necessary, as in the case of minority or minoritized languages (Cabré 2001).

Unlike ST, the CTT is the first viable theory capable of replacing the GTT. Many of the theoretical and practical premises of the CTT are still valid and, above all, it laid the foundations for Terminology to integrate the knowledge and advances of other disciplines in addition to Linguistics. In this line, the CTT recognizes the importance of cognition and communication in Terminology; it contemplates different contexts of use; and it accounts for dynamic phenomena such as variation and multidimensionality. Nonetheless, despite its valuable contributions, it has a series of drawbacks, such as the fact that it does not opt for any specific linguistic model; it does not explain conceptual relations and their potential constraints (beyond saying that there are conceptual maps made up of nodes and relations); and it does not offer a clear explanation of specialized meaning and of semantic analysis (Faber 2009).

2.2.1.4. COGNITIVE-BASED TERMINOLOGY THEORIES

In the early 2000s, *cognitive-based Terminology theories* appeared on the horizon as a result of the *cognitive shift* (Evans & Green 2006) which affected linguistic theory as well as Terminology. Accordingly, the Sociocognitive Terminology Theory and Frame-based Terminology focus on the conceptual network underlying language, and implement premises from Cognitive Linguistics and Psychology with regard to concept description and category structure.

2.2.1.4.1. SOCIOCOGNITIVE TERMINOLOGY THEORY

The Sociocognitive Terminology Theory (SCTT) (Temmerman 2000, 2001, 2007; Temmerman *et al*. 2005) arose at the beginning of the 21st century, when insights from Cognitive Semantics (Talmy 2000) were applied to Terminology theory. The SCTT concentrates on the cognitive potential of Terminology in domain-specific language and on terminological variation as related to verbal, situational, and cognitive contexts in discourse and in a wide range of communicative environments (Temmerman *et al*. 2005).

From the viewpoint of the SCTT, concepts are *units of understanding* (Temmerman 2000) and their conception involves organization into categories, which are framed in idealized cognitive models and usually present a prototypical structure (Rosch 1978). Therefore, particular focus is placed on cognitive notions such as categorization and prototypes. In addition, a cognitive approach to metaphor is carried out, because it is present in categorization and gives rise to metaphorical lexicalizations (Temmerman 2000).

The SCTT integrates the notion of dynamism by considering that a category is a piece of knowledge with a core and a structure, which is in a process of continuous reformulation and in continuous transition. Temmerman (2000) states that dynamism arises from several active factors that simultaneously influence the conceptual system: (i) the need for more and better understanding; (ii) the interaction between users of different languages; and (iii) the prototypical structure that can be seen as both the result and one of the causes of meaning evolution.

This theory is also in consonance with ST and the CTT since it is descriptive rather than prescriptive, thus regarding terms as the starting point for terminological analysis. Therefore, many aspects of the GTT are criticized, namely the onomasiological and prescriptive approach, the existence of well-defined boundaries in concepts and categories, the univocity of terms, and synchronic studies (Temmerman 2000, 2007).

In contrast, the SCTT adopts a semasiological and descriptive approach, arguing that the boundaries between concepts and categories are not well delimited, as they usually present a prototypical structure. With regard to definitions, they depend on the type of unit of understanding, the specialized domain in which it takes place, and the user profiles. Moreover, it acknowledges the existence of multidimensionality and the importance of variation as well as of diachronic studies of terms and concepts. Based on these theoretical principles, Temmerman (2000) proposes new methods of terminological analysis borrowed from Cognitive Semantics (Talmy 2000): prototypical structure analysis, idealized cognitive model analysis, and diachronic analysis.

The SCTT also includes the notion of frame, but in the sense of Lakoff's (1987) idealized cognitive models. On this basis, Temmerman (2000) explains that units of understanding have both intracategorial and intercategorial structures. Regarding the intracategorial structure, the SCTT distinguishes different information modules that will vary in relevance according to the type of category. As for the intercategorial structure, it considers perspective, domain and intention within the idealized cognitive model.

Furthermore, the SCTT also introduces the idea of ontologies as a more viable way of implementing conceptual representations. This combination of terminology work and ontologies is called *termontography* (Temmerman & Kerremans 2003; Temmerman *et al*. 2005), a hybrid term derived from *terminology*, *ontology*, and *terminography*. Its objective is to link ontologies with multilingual terminological information, and to incorporate them into terminological resources. It is a multidisciplinary approach in which theories and methods for multilingual terminological analysis (Temmerman 2000) are combined with methods and guidelines for ontological analysis (Fernández *et al*. 1997; Sure & Studer 2003), since ontologies are conceived as a more practical way of making conceptual representations.

Another interesting approach that integrates ontologies into terminology work is *ontoterminology* (Roche 2007, 2012a, 2012b; Roche *et al*. 2009). In this model, ontologies are a means of conceptual representation in terminology work and a clear distinction is made between the definition of the term (i.e., in natural language) and the definition of the concept (i.e., in formal language). This highlights the importance that the underlying conceptual system has acquired in current terminological studies.

It is thus clear that the cognitive contributions made to Terminology thanks to the SCTT bring this discipline closer to a more reliable description and representation of conceptual systems and reality. However, there are still certain points that can be criticized. One of them, for example, is related to how the SCTT deliberately avoids syntactic issues. Faber & López Rodríguez (2012) claim that this is perhaps because any kind of syntactic analysis, whether of general or specialized language, has to be based on a syntactic theory, and the syntactic projections of terms have not as yet been studied in any depth.

2.2.2. Theoretical and Practical Foundations of Frame-based Terminology

Frame-based Terminology (FBT) (Faber *et al.* 2005, 2006; Faber 2012, 2022) is a recent theory that combines a descriptive approach with elements of Corpus Linguistics (Sinclair 1991, 1995, 1996a, 1996b), Cognitive Semantics (Talmy 2000), and Frame Semantics (Fillmore 1968, 1975, 1977, 1982, 1985) in order to create structured specialized domains and non-language-specific representations. This cognitive-based Terminology theory shares many of the aspects of the CTT and the SCTT.

For example, it is also a descriptive and text-driven approach which admits term variation and polysemy. However, FBT specifically focuses on specialized knowledge representation, category organization and description, as well as the semantic and syntactic behavior of terminological units or terms in one or various languages. More specifically, FBT assumes that in scientific and technical communication, specialized knowledge units activate domain-specific semantic frames that are in consonance with the users' background knowledge (Faber *et al*. 2016). These frames are cognitive structuring devices based on experience that provide the background knowledge for the words in a language, as well as the way that those words are used in discourse. FBT also operates on the premise that knowledge of conceptualization processes and the organization of semantic information in the brain should underlie all theoretical assumptions concerning the access, retrieval, and acquisition of specialized knowledge as well as the design of specialized knowledge resources (Faber 2011).

FBT is composed of a set of microtheories (Faber 2022): (i) a semantic microtheory focusing on term meaning, definitions, and conceptual organization; (ii) a syntactic microtheory that analyzes specialized phraseology and the structure of multi-word terms; and (iii) a pragmatic microtheory that explains the cultural and contextual parameters of specialized communication. Each microtheory is related to the information in term entries, the relations between specialized knowledge units, and the concepts that they designate. In the words of Faber (2022: 374):

> The semantic microtheory explains how specialized knowledge concepts are classified by determining degrees of specificity and conceptual similarity. Concepts are described by definitions, which should follow the template that characterizes their respective semantic categories. Such templates consist of a set of vertical and horizontal relations typical of a given category, some of which may be domain-specific. The syntactic microtheory highlights the importance of the interface between syntax and semantics in Terminology and specialized language. It focuses on the combinatorial value and distinctive syntactic projections of terms, stemming from their semantic category as well as their combinations with other categories. This is reflected in phraseological patterns in specialized texts as well as in multi-word terms (MWTs). [...] Finally, the pragmatic microtheory accounts for context-dependent aspects of specialized meaning. In FBT, such contexts can be linguistic, cultural, and graphical. Linguistic contexts are found in concordance lines, especially those that reflect Knowledge-Rich Contexts (KRCs). Their analysis is a key factor in the retrieval of semantic relations codified in knowledge patterns (KPs).

The theoretical and practical foundations of FBT that apply to this thesis are the following: (i) conceptual organization, reflected through frames or events; (ii) multidimensionality, expressed through both hierarchical and non-hierarchical relations; (iii) knowledge extraction, carried out through the use of multilingual corpora; and (iv) frame-based definitional templates, producing systematic definitions portraying the contextual frame.

2.2.2.1. CONCEPTUAL ORGANIZATION: FRAMES AND EVENTS

FBT uses an adapted version of basic principles of Frame Semantics (Fillmore 1968, 1975, 1977, 1982, 1985) to structure specialized domains and create non-language-specific representations. The idea that meaning is context-dependent is the basis of the notion of frame, which is in consonance with the encyclopedic approach to meaning in Cognitive Linguistics. Lexical items provide access to a structured body of non-linguistic or encyclopedic knowledge (Faber & López Rodríguez 2012). Conceptual organization and knowledge representation are central to Terminology because terminology work involves modeling specialized knowledge concepts within a given specialized domain so that users can understand them better. Conceptual modeling in knowledge resource design involves capturing both the micro- and macro-contexts of concepts.

As is well-known, frames are a type of cognitive structuring device based on experience that provides the background knowledge and motivation for the existence of words in a language as well as the way those words are used in discourse. However, frames have the advantage of making explicit both the potential semantic and syntactic behavior of specialized language units. This necessarily includes a description of conceptual relations and of a term's combinatorial potential.

The most characteristic frame in FBT is the event, understood as the template used for describing the processes occurring within a specialized domain (e.g., the commercial event, the medical event). It is grounded on the premise that the description of specialized domains is based on the processes that occur in them (Grinev & Klepalchenko 1999). FBT argues that the way concepts are represented affects the configuration of information in individual terminological entries and the contents of each data field, especially with regard to the definition of each concept (Faber *et al*. 2006). Therefore, in order to dynamically represent a specialized knowledge, the most generic categories of a domain should be organized in a prototypical event. Accordingly, FBT argues that every knowledge area will have an event providing a structure for the organization of concepts within it (Faber & López Rodríguez 2012).

The practical application of FBT is EcoLexicon (**§2.2.3**), an environmental knowledge base. This theory thus places special emphasis on describing what is known as the *environmental event* (EE) (Faber *et al*. 2005). This event underlies the texts related to the environment and it is where the analysis of concepts and their relations allows for acquiring specialized knowledge (Faber 2012). Thus, the EE is composed of the prototypical actions and processes that take place in the environment, as well as the entities that participate in them (**Figure 1**).

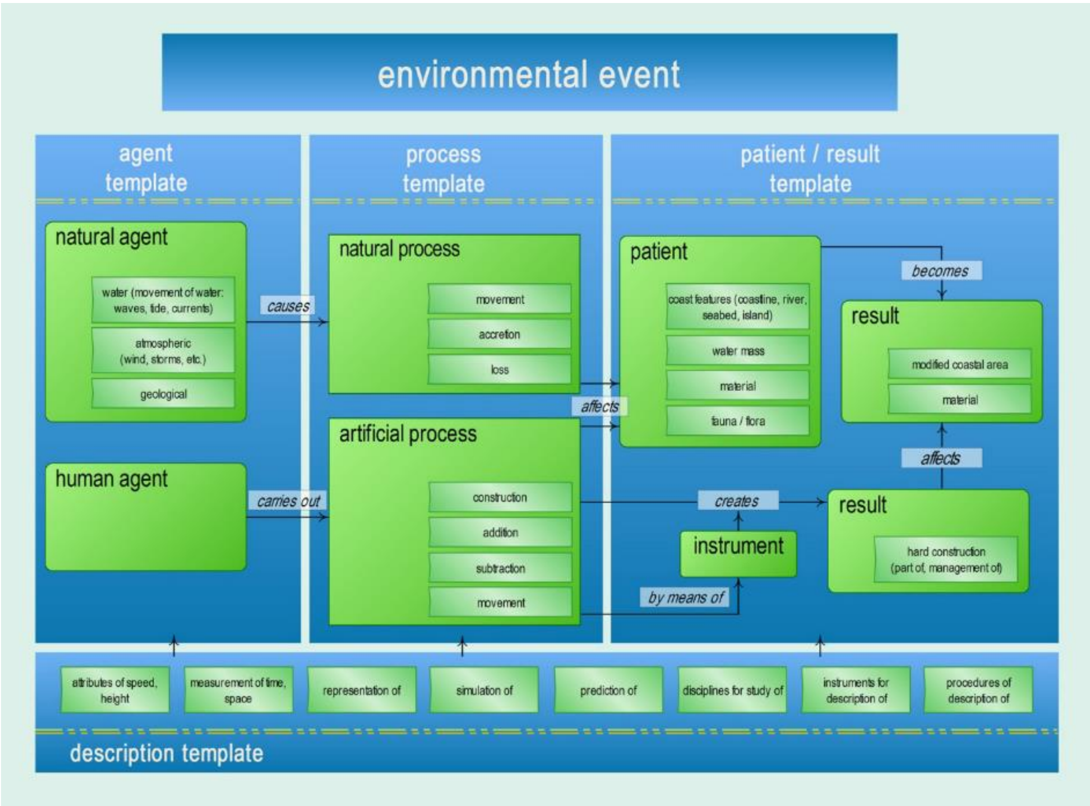


Figure 1. The environmental event

In the frame of the EE, general categories of environmental entities are linked by predicates codifying the states, processes, and events in which the entities can participate (Faber 2015). The EE contains basic meanings that relate concepts, roles, and categories pertaining to general environmental knowledge (León-Araúz *et al*. 2012). Moreover, the EE also links generic categories at the superordinate level and provides the basis for subframes that can be used to restrict contextual information to what is most relevant.

As shown in **Figure 1**, the EE has two types of *agents* that can initiate processes, i.e., *natural agents* (i.e., inanimate) and *human agents* (i.e., animate). On the one hand, *natural agents* (e.g., WATER MOVEMENT) cause *natural processes* (e.g., RIVER EROSION) in specific locations, commonly regarded as *patients* (e.g., river bed) which, as a *result*, may suffer alterations (e.g., DETERIORATION, MODIFICATION OF SIZE or SHAPE). On the other hand, *human agents* can also carry out *artificial processes* (e.g., CONSTRUCTION) to alter the *effects* normally caused by natural processes (e.g., PROTECTION), or to create new effects through the use of certain *instruments* (e.g., DEFENSE STRUCTURES).

Nevertheless, the conceptual representation of environmental knowledge cannot be achieved simply by assigning these generic semantic or conceptual roles to concepts as if all of them would belong to a universal type of event (León-Araúz *et al*. 2012). For this reason, the EE was originally used as a macrostructure for the further design of context-dependent microstructures (e.g., Coastal Engineering, Meteorology, Oceanography). Depending on the conceptual relation involved, emphasis is placed on one dimension or another of the same concept in different contexts (Evans & Green 2006). This leads to multidimensionality.

2.2.2.2. MULTIDIMENSIONALITY

Multidimensionality (Kageura 1997, 2002; Bowker 1997, 2022; Rogers 2004), is the conceptual classification that arises when concepts are classified in more than one way within a conceptual system based on different characteristics. It is thus extremely important in terminology work for its influence in conceptual systems (**§2.1.1.1**). Interestingly, multidimensionality is related to Cruse's idea (2000) that words have context-dependent meanings known as *facets*. Each of these individual meanings is activated by a specific utterance. Bowker (2022: 135) writes:

> [...] the concept RED WINE has a value of "red" for the characteristic "colour", and this will be inherited automatically by the subordinate concepts in the hierarchy, such as CHIANTI and SHIRAZ, which are types of RED WINE. What is of particular interest for multidimensionality, however, is the fact that these knowledge engineering tools include a mechanism for *multiple* inheritance. In other words, a subordinate concept can inherit characteristics from superordinate concepts that originate in different dimensions. For example, in addition to inheriting the characteristic "red" from the superordinate concept RED WINE (in the dimension based on the characteristic "colour"), CHIANTI could also inherit a characteristic from a superordinate concept in the dimension based on "country of origin", and it could inherit yet another characteristic from a superordinate concept in the dimension based on "sugar

> content". Therefore, CHIANTI could have three different superordinate concepts – RED WINE, ITALIAN WINE and DRY WINE – and it would inherit characteristics from each.

Multidimensionality is an important aspect of cognitive-oriented terminological resources such as TKBs, because it allows users to gain a better understanding of the concepts and to envisage a wider range of user needs. According to Bowker (2022), the ability to manage multiple inheritance was greatly facilitated by the mechanisms incorporated into knowledge engineering tools, with the ability to graphically display the results of multiple inheritance.

Not surprisingly, a major issue in the design of any knowledge resource is how to link the concepts since the inventory of conceptual relations affects the informativity of the terminological resource. Termbases that only display hierarchical relations are static, whereas knowledge bases are more dynamic with a wider set of relations when they include non-hierarchical or associative relations. In fact, there is a vast array of conceptual relations potentially available to enrich conceptual systems (Rogers 2004), and hence represent multidimensionality.

One of the key aspects of FBT is the holistic depiction of multidimensionality with a diverse inventory of conceptual relations. For this reason, in EcoLexicon, a total of 18 conceptual relations are available: one generic-specific relation (i.e., *type of*), six part-whole relations (i.e., *part of*, *made of*, *located at*, *delimited by*, *takes place in*, *phase of*), and eleven non-hierarchical relations (i.e., *opposite of*, *result of*, *has function*, *attribute of*, *affects*, *studies*, *represents*, *measures*, *effected by*, *causes*, *destroys*).

The existence of different dimensions within a concept is determined by various factors (e.g., conceptual category, semantic role, contextual domain, etc.), and thus the conceptual relations involved can vary from one dimension to another. In TKBs, multidimensionality enriches static representations through the inclusion of different perspectives in a conceptual system. However, this can also generate an information overload that prevents the user from acquiring knowledge. A possible solution to this problem is to enhance the representation of multidimensionality by refining hyponymy. This can be accomplished by correcting property inheritance, implementing umbrella concepts or establishing hyponymy subtypes (Gil-Berrozpe 2016).

Conceptual relations are also fundamental in the elaboration of definitions. According to FBT, each conceptual category has a definitional template composed of the conceptual relations that activate that category (Faber *et al*. 2001). These templates permit homogeneous definitions and the conceptual organization of terms.

2.2.2.3. FRAME-BASED DEFINITIONAL TEMPLATES

Specialized concepts are designated by terms and described by a definition, which is a natural language explanation of its location in the conceptual structure of the specialized domain (Faber 2022). A definition not only specifies the properties of a concept, but also links it to other realities (Antia 2000).

According to FBT (Faber 2012), a definition should thus be coherent with the concept itself, its conceptual category, and the set of concepts related to it. It should not only provide a comprehensive description of the concept itself, but also its entire context. In fact, coherent definitions should follow the pattern or template typical of the conceptual category.

ISO Standard 1087 (2019) understands a definition as the representation of a concept by an expression that describes it and differentiates it from related concepts. Furthermore, it makes a distinction between four different types of definitions: (i) *intensional definition*, conveying the intension of a concept by stating the immediate generic concept and the delimiting characteristics; (ii) *extensional definition*, enumerating all the subordinate concepts of a superordinate concept under one criterion of subdivision; (iii) *generic extensional definition*, enumerating all the specific concepts of a generic concept under one criterion of subdivision on the same hierarchical level; and (iv) *partitive extensional definition*, enumerating all the partitive concepts of a comprehensive concept on the same hierarchical level.

Frame-based definitions are intensional definitions (Durán-Muñoz 2016), which are based on the notion of Aristotelian definitions traditionally used in Terminology to define terms. Accordingly, most definitions of specialized knowledge units are composed of a generic or superordinate term and differentiating features (Eck & Meyer 1995; Sager 1990), reflecting an external semantic representation, which relates other concepts to the concept being defined. In other words, terminological definitions describe concepts designating by a term with a view to identifying necessary and sufficient features of such concepts within the limits of a specific domain (Sager & Ndi-Kimbi 1995).

In this line, frame-based definitional templates are composed of two elements: the *genus* or superordinate concept (i.e., its direct hypernym, also related to the conceptual category), and the *differentiae* or characteristics that make a concept unique and different to the concepts at the same hierarchical level (i.e., its co-hyponyms). For example, following this definitional template for the concept EARTHQUAKE, "geologic phenomenon" would be the *genus*, and "involving a sudden movement of the Earth's crust" plus "caused by the release of stress accumulated along geologic faults or by volcanic activity" would be the *differentiae*. Furthermore, various conceptual relations can be identified in the *differentiae* (e.g., "involving a sudden movement of the Earth's crust" – *causes* and *takes place in*; "caused by the release of stress accumulated along geologic faults or by volcanic activity" – *result of*).

Furthermore, FBT stresses that each conceptual category has a template that can be used to model the definitions of category members (Faber 2022). The set of conceptual relations in this template depends on the nature of the category. For example, basic characteristics underlying the representation of a human-made object or artifact such as a scientific instrument include its subtype, parts, function, and context of manipulation. As exemplified by Faber (2022: 362):

> For example, a RECORDING INSTRUMENT (MARIGRAPH, PLUVIOGRAPH, ANEMOGRAPH, etc.) is a subtype of INSTRUMENT. As an artifact, a RECORDING INSTRUMENT has a function (i.e., RECORDING) as well as a recorded phenomenon (TIDES, RAIN, WIND). As a tool, it is operated by humans and thus activates a simulation frame in which much of the perceiver's knowledge of the artifact involves his/her ability to manipulate it, place it at a certain location, and in some way, extract information from it.

However, all conceptual and contextual information described following frame-based definitional templates has a grounded origin, which is identified through knowledge extraction.

2.2.2.4. KNOWLEDGE EXTRACTION

The methodology followed by FBT for knowledge extraction is based on a study of conceptual systems and specialized domains through both a top-down and a bottom-up approach. On the one hand, the top-down approach consists of analyzing, filtering, combining, and restructuring information contained in specialized dictionaries and other reference material, which is then validated by experts. On the other hand, the bottom-up approach consists of extracting linguistic information from a corpus of texts in one or several languages related to the specialized domain in question.

Therefore, FBT bases its methodology for knowledge extraction on the premises of Corpus Linguistics (Sinclair 1991, 1995, 1996a, 1996b), particularly on corpus analysis strategies and techniques. Corpus analysis is thus commonly used by terminologists in first instance to find terms in large corpora and extract their syntactic and semantic information. Traditionally, corpora have been analyzed and processed by manually reading concordance lines related to a particular term. León-Araúz *et al*. (2016: 73) state that this time-consuming task led to the development of new corpus-based methods and applications to analyze and extract linguistic information.

Accordingly, one of the most validated approaches for the efficient extraction of information from a corpus is to search for knowledge-rich contexts (KRCs). They are described by Meyer (2001: 281) as "a context indicating at least one item of domain knowledge that could be useful for conceptual analysis". Such contexts are highly informative since they provide conceptual information and domain knowledge, and they usually codify conceptual relations in the form of knowledge patterns (KPs), which can be used to find KRCs in corpora (Meyer 2001). KPs are the lexico-syntactic patterns between terms encoded in a proposition in real texts (Meyer 2001).

Bielinskiene *et al*. (2012: 18) state that important conceptual characteristics are expressed in KRCs in the form of conceptual relations (e.g., hyponymy or meronymy) and that they can be identified through KPs. In this line, these elements make it possible to extract the relevant terminography-oriented knowledge about the concept from a corpus and then use that information to provide a starting point for any terminological purpose (Bielinskiene *et al*. 2012). However, this task is complicated by the fact there are no user-friendly publicly available applications that allow terminologists to find KRCs in their own corpora with ready-made KPs.

In this sense, KPs are regarded as one of the most reliable and effective methodologies for the extraction of semantic relations (Condamines 2002; Bowker 2003; Barrière 2004b; L'Homme & Marshman 2006; Mortchev-Bouveret 2006; Auger & Barrière 2008; Gödert *et al*. 2014; Lafourcade & Ramadier 2016; Gil-Berrozpe 2017; Lefeuvre *et al*. 2017; Rojas-García & Cabezas-García 2019). However, to a certain degree, terminologists still tend to rely on manual work to hand-select all the semantic information that they need for the description of specialized concepts (León-Araúz *et al*. 2016).

Because of their interest for terminology work, corpus-based analysis and KPs have become a major research topic over the years as a method of automatically or semi-automatically extracting linguistic information concerning different conceptual or semantic relations. **Table 2** shows a typology of KPs (herein referred to as *lexical patterns*) proposed by Bowker & Pearson (2002), along with the type of conceptual knowledge expressed and some examples of possible KRCs.

LEXICAL PATTERNS	CONCEPTUAL KNOWLEDGE	EXAMPLE
is a ǀ *kind of* ǀ *type of* ǀ *includes*	generic-specific relations	The tabor *is a* type of drum.
has a ǀ *contains* ǀ *consists of* ǀ *includes*	part-whole relations	A snare drum *has a* batter head and a snare head.
used for ǀ *used to* ǀ *employed to*	function relations	A wooden stick *is used to* strike the drum head.
causes ǀ *produces* ǀ *produced by* ǀ *results from*	cause-effect relations	Striking the drum head *causes* the snares to vibrate.
also called ǀ *also known as* ǀ *sometimes referred to as*	possible synonymy	The tambourine, *also known as* the tambourin provençale, is the largest of all the tabors.

Table 2. Lexical patterns and possible KRCs (Bowker & Pearson 2002)

Nonetheless, these conceptual relations have not received an equal amount of attention. For example, meronymic or part-whole relations have been widely researched (Berland & Charniak 1999, Girju *et al*. 2003). They can be codified by prepositional phrases, possessives, and partitive verbs. Moreover, non-hierarchical relations have also been studied and implemented as KPs.

For instance, causality (Marshman 2002, Marshman *et al*. 2002; Soler & Alcina 2008) can be expressed by passive, active, subject-object, nominal and verbal propositions, involving all kinds of causative nouns and verbs. However, the most commonly studied patterns are hyponymic KPs (Hearst 1992, 1998; Pearson 1998; Liu *et al*. 2006; Pantel & Pennacchiotti 2006; Bielinskiene *et al*. 2012; Nazar *et al*. 2012; Lefever *et al*. 2014; Li *et al*. 2014; Baisa & Suchomel 2015; Gil-Berrozpe *et al*. 2017; Faralli *et al*. 2018; Lewis 2019), because of their importance in relation to categorization and property inheritance.

FBT thus uses this corpus analysis methodology and these KP-based techniques for identifying relevant KRCs in order to obtain information about specialized knowledge (e.g., terms, patterns, contexts, etc.). In FBT, KP-based queries in the form of micro-grammars (León-Araúz *et al*. 2016; León-Araúz & San Martín 2018) are used to extract concordance lines for a term and analyze its combinatorial potential. **Table 3** shows the typical KPs that codify some of the conceptual relations studied by FBT (León-Araúz & Reimerink 2010).

CONCEPTUAL RELATION	KNOWLEDGE PATTERNS
is a	such as ǀ rang* from ǀ includ*
part of	includ* ǀ consist* of ǀ formed by/of
made of	consist* of ǀ built of/from ǀ constructed of ǀ formed by/of/from
located at	form* in/at/on ǀ found in/at/on ǀ tak* place in/at ǀ located in/at
result of	caused by ǀ leading to ǀ derived from ǀ formed when/by/from
has function	designed for/to ǀ built to/for ǀ purpose is to ǀ used to/for
effected by	carried out with ǀ by using

Table 3. Examples of conceptual relations and KPs (León-Araúz & Reimerink 2010)

2.2.3. ECOLEXICON

EcoLexicon[1] (Reimerink & Faber 2009; Reimerink *et al*. 2010; Faber & León-Araúz 2010, Faber *et al*. 2014a, 2016; San Martín *et al*. 2020) is a multidimensional, multimodal, and dynamic TKB on the environment developed by the Lexicon Research Group of the University of Granada. It is the practical application of FBT. EcoLexicon is organized according to conceptual and linguistic premises at the macrostructural and microstructural levels (Faber & León-Araúz 2021). To date, it has over 4,500 concepts and over 24,500 terms in seven different languages: English, Spanish, German, French, Modern Greek, Dutch, and Russian. This knowledge resource is conceived for language and domain experts as well as for the general public. It targets users such as translators, technical writers, and environmental experts who need to understand specialized environmental concepts with a view to writing and/or translating specialized and/or semi-specialized texts.

This section first overviews the evolution from traditional dictionaries to contemporary TKBs, focusing on the most valid criteria for the development of modern terminology resources. The main features and modules of EcoLexicon are then described.

2.2.3.1. FROM DICTIONARIES TO TERMINOLOGICAL KNOWLEDGE BASES

Advances in terminography have led to the development of terminological resources adapted to the standards of the digital era. While it is true that paper-based

[1] Available at: http://ecolexicon.ugr.es/

lexicological and terminological resources, such as dictionaries and encyclopedias, are still often consulted, new electronic resources as well as electronic versions of paper dictionaries are now more frequently used. Digital resources are not only easier to consult but they are also easier to update.

One example of this adaptation is the *Oxford English Dictionary* or OED (Oxford University Press 2010). Its second edition, originally published in 1989, was discontinued in favor of online publication in 2000 in the form of *OED Online* (Oxford University Press 2022). A third edition of the physical version of the OED was released in 2010, but only the online version can be easily updated.

In recent years, the creation of electronic terminological resources has led to theoretical, methodological, and technical advances, which have provided solutions for different issues related to content, search techniques, and resource maintenance (Roche *et al*. 2019a). At present, a large number of terminological resources are globally available at a single click in different formats and in different languages. According to Roche *et al*. (2019a: 139):

> In a globalised society, terminological dictionaries – including resources such as knowledge and terminological databases, ontologies, wordnets, "traditional" dictionaries, etc. – should comply with both human and machine needs. Changes regarding information and language processing brought forward by the evolution of society have led to a series of consequences in: (i) the design of terminological resources; (ii) the way data and knowledge are represented; (iii) the way data are interrelated, both within and between resources; (iv) the way users access data; and (v) users' expectations.

Additionally, the emphasis on the Semantic Web and on the Linked Data initiative have contributed to new perspectives and opportunities in Terminology, especially in relation to the operationalization of terminological products in order to conceive and build different types of electronic dictionaries (Roche *et al*. 2019b). As a result, today, not only are dictionaries and encyclopedias available in electronic format, but a whole new repertory of informational and terminological resources has emerged (e.g., thesauri, termbases, databanks, knowledge bases, ontologies, taxonomies, etc.), including TKBs.

TKBs stem from COGNITERM (Meyer *et al*. 1992) and are conceived as a hybrid between term banks and knowledge bases. Accordingly, they represent the specialized knowledge of a certain field through related concepts and the terms that designate them in one or various languages, reflecting both linguistic and cognitive processes. Optimally, TKBs should reflect how conceptual networks are established and structured in our minds. They must also be designed to meet the needs of a specific group of users, whether they are experts or lay public.

According to León-Araúz *et al*. (2013b), TKBs should account for the representation of natural and contextual knowledge dynamism. Various issues must thus be considered when designing and creating a TKB. On the one hand, the organization of the knowledge field should accurately represent the concepts and the semantic relations linking them. On the other hand, access to information and its retrieval should facilitate knowledge acquisition.

However, one of the main problems in concept representation is that concepts are multidimensional and their characteristics may vary depending on the perspective taken. However, the representation of multidimensionality must also follow rules. In this sense, conceptual relations cannot be created on demand, but should be systematically derived from a set inventory (León-Araúz *et al*. 2012). For this reason, a logical methodology should be followed when expanding the existing conceptual relations in a TKB.

In line with this, structuring specialized concepts in networks with both hierarchical and non-hierarchical relations is one of the key elements in modern terminological resources. Nonetheless, this process can produce an extremely basic resource if designers do not envisage methods for structuring knowledge (e.g., establishing subtypes of conceptual relations, extending non-hierarchical relations, specifying the properties of the relations, etc.).

A conceptually-structured TKB, in which terms are linked to concepts based on non-language-specific criteria, is thus a useful resource for end users. In this way, not only is there coherent cross-referencing, but also linguistic data can be added and manipulated without altering the quality and consistency of the conceptual design (Giacomini 2014). Moreover, their representation of metalinguistic and encyclopedic data contributes to the enhancement of knowledge acquisition by allowing to search for corpus concordances and parallel texts. Not surprisingly, this is an extremely valuable feature for many users, such as translators, since it allows them to avoid extra-terminographic searches and queries, which can be time-consuming tasks.

Furthermore, according to Giacomini (2015), any type of electronic terminological resource, including TKBs, should respond to three requirements: (i) conceptual structure availability and properties, with a multi-level depth of conceptual structures and multi-vocal relations; (ii) ease of access to conceptual data, with direct access via the conceptual structure and the microstructure, and with specified relations; and (iii) consistency of concept-term correspondences, in the search by concept or by term. Therefore, these criteria reveal the importance of having a conceptual structure in the form of an ontology.

Moreover, any ontology-based terminological database geared to the fulfillment of these requirements should reflect dynamic phenomena such as multidimensionality and natural and contextual knowledge dynamism (León-Araúz *et al*. 2013b), which are basic to specialized knowledge representation and acquisition. Methods for implementing multidimensionality in a TKB include the addition or the deletion of certain concepts or relations in specific nodes or in the system, the modification of certain characteristics or relations in specific nodes or in the system, and the implementation of new ways to represent knowledge (linguistic, conceptual, visual, interactive, etc.). On the other hand, contextual dynamism can be achieved by showing how concepts — and therefore, terms — modify their features and use depending on their context and depending on the level of user expertise or knowledge. In line with this, Tercedor Sánchez *et al*. (2012) also highlight the intimate link between dynamicity and multidimensionality. Since it is now possible to represent concepts from different perspectives or dimensions, lexicological and terminological practice should thus envisage the elaboration of more dynamic representations.

Furthermore, another salient characteristic of TKBs is multimodality. Evidently, linking multimedia information (e.g., images, videos) to the linguistic, conceptual, and contextual information of a TKB helps to satisfy user needs with regard to the reception, production, and translation of specialized texts (Prieto Velasco 2009). In this way, TKBs can be represented as visual thesauri, merging multimodal information and highlighting the multidimensional character of knowledge representations. These are all thus the essential characteristics of a TKB such as EcoLexicon.

2.2.3.2. DESCRIPTION AND KEY FEATURES OF ECOLEXICON

EcoLexicon has an intuitive visual interface with a series of modules that provide conceptual, linguistic, and multimodal information. Instead of viewing all this information simultaneously, users can browse and select the data that are most relevant to their needs. Once users select a concept, it is represented in the center of an interactive map representing its conceptual system. Also displayed are the multilingual terms for that concept, as well as different conceptual relations between all the concepts belonging to the same contextual domain (**Figure 2**).

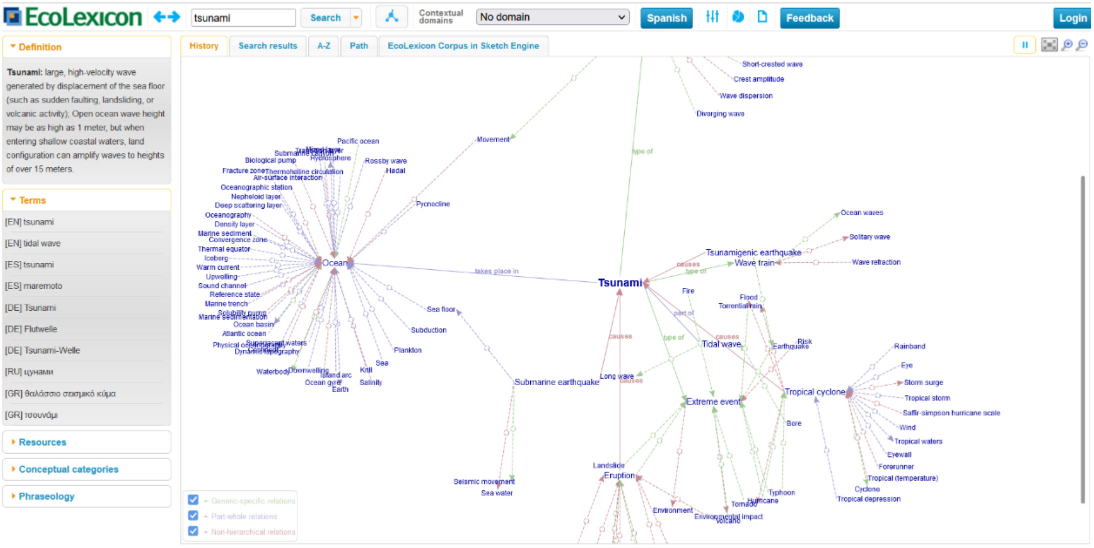


Figure 2. Visual interface of EcoLexicon (conceptual system: TSUNAMI)

As for the macrostructure, when EcoLexicon is accessed, three sections appear: (i) the top horizontal bar, which gives access to the term/concept search engine and to the contextual domain filter, among other settings; (ii) the vertical bar on the left, which provides the information about the five main modules (i.e., definition, terms, resources, conceptual categories, and phraseology); and (iii) the central area, which is divided into five tabs (i.e., history, search results, alphabetical list, shortest path between two concepts, and corpus concordances) and the main area of the screen, showing a conceptual system.

As previously mentioned, conceptual relations are classified in three main groups in EcoLexicon: generic-specific relations, part-whole relations, and non-hierarchical relations (**Figure 3**). As can be observed, hierarchical relations are divided into two groups to distinguish between hyponymic relations and meronymic relations. The set of generic-specific relations only comprises *type of*. In contrast, the set of part-whole relations contains *part of*, *made of*, *delimited by*, *located at*, *takes place in*, and *phase of*. In the last place, the set of non-hierarchical relations includes *affects*, *causes*, *attribute of*, *opposite of*, *studies*, *measures*, *represents*, *result of*, *effected by*, *has* *function*, and *destroys*. The set of all conceptual relations in EcoLexicon comes to a total of 18.

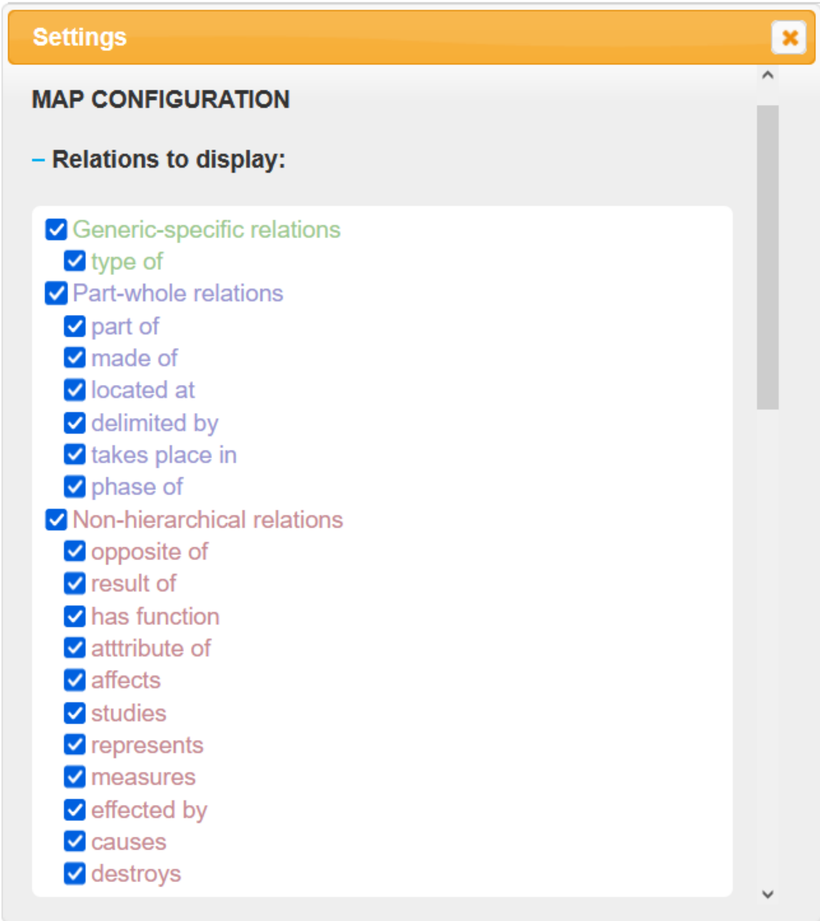


Figure 3. Conceptual relations in EcoLexicon

With regard to the microstructure of this EcoLexicon, when a concept is selected, five modules are displayed on the left side of the interface:

- **Definition:** this module provides an intensional or terminological definition following a frame-based definitional template. This type of definition is composed of the *genus* (i.e., hypernym or superordinate) and one or many *differentiae* (i.e., characteristics that vary in each co-hyponym) (**Figure 4**). In the case of TSUNAMI, "wave" is the *genus*, and the *differentiae* are "large" [*attribute of*], "high-velocity" [*attribute of*], "generated by displacement of the sea floor" [*result of*], etc.

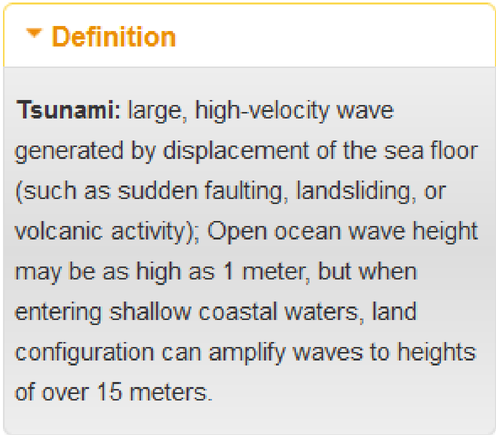


Figure 4. Definition module in EcoLexicon (concept: TSUNAMI)

- **Terms:** this module displays the lexical denominations for a concept in the different languages in EcoLexicon, information regarding the term type and the part of speech, and the option to display a list of corpus concordances (**Figure 5**). For example, for TSUNAMI, it shows term variations in English (i.e., *tsunami* and *tidal wave*), in Spanish (i.e., *tsunami* and *maremoto*), in German (i.e., *Tsunami*, *Flutwelle*, and *Tsunami-Welle*), in Russian (i.e., *цунами*), and in Modern Greek (i.e., *θαλάσσιο σεισμικό κύμα* and *τσουνάμι*). Interestingly, these terms are listed according to language codes (i.e., EN, ES, DE, RU, GR) instead of flags so as to avoid cultural issues.

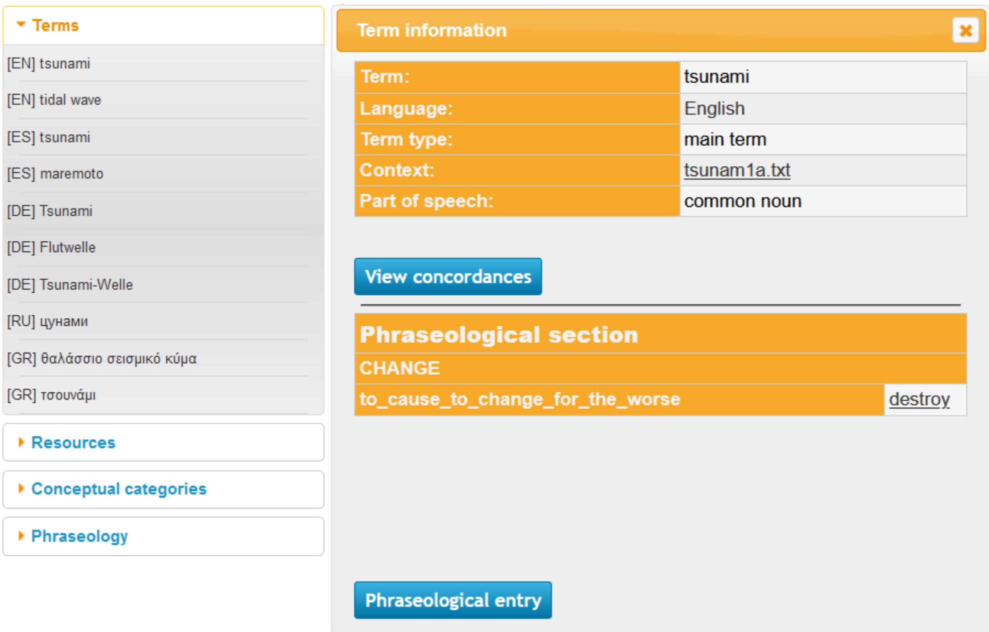


Figure 5. Term module in EcoLexicon (concept: TSUNAMI)

- **Resources:** this module offers a list of multimodal resources (e.g., images, videos, hyperlinks to external websites) for the chosen concept (**Figure 6**). In this case, TSUNAMI presents a wide variety of resources, including pictures, photos, diagrams, academic websites with explanations on the topic, and even satellite images of tsunamis.

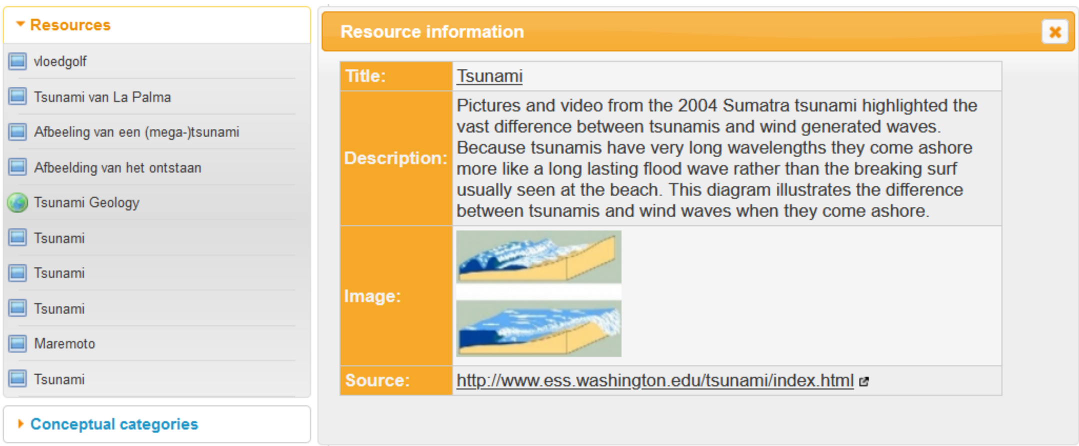


Figure 6. Resource module in EcoLexicon (concept: TSUNAMI)

- **Conceptual categories:** this module provides a list of the conceptual categories in EcoLexicon and classifies the concept as a member of one or various of them (**Figure 7**). For example, TSUNAMI is classified as *water movement* (P-11.3.1). Furthermore, when the conceptual category is selected in the hierarchy, a list of all concepts in EcoLexicon belonging to that same category is displayed (e.g., URBAN RUNOFF, TORRENT, ATMOSPHERIC TIDE, INTRUSION, etc.).

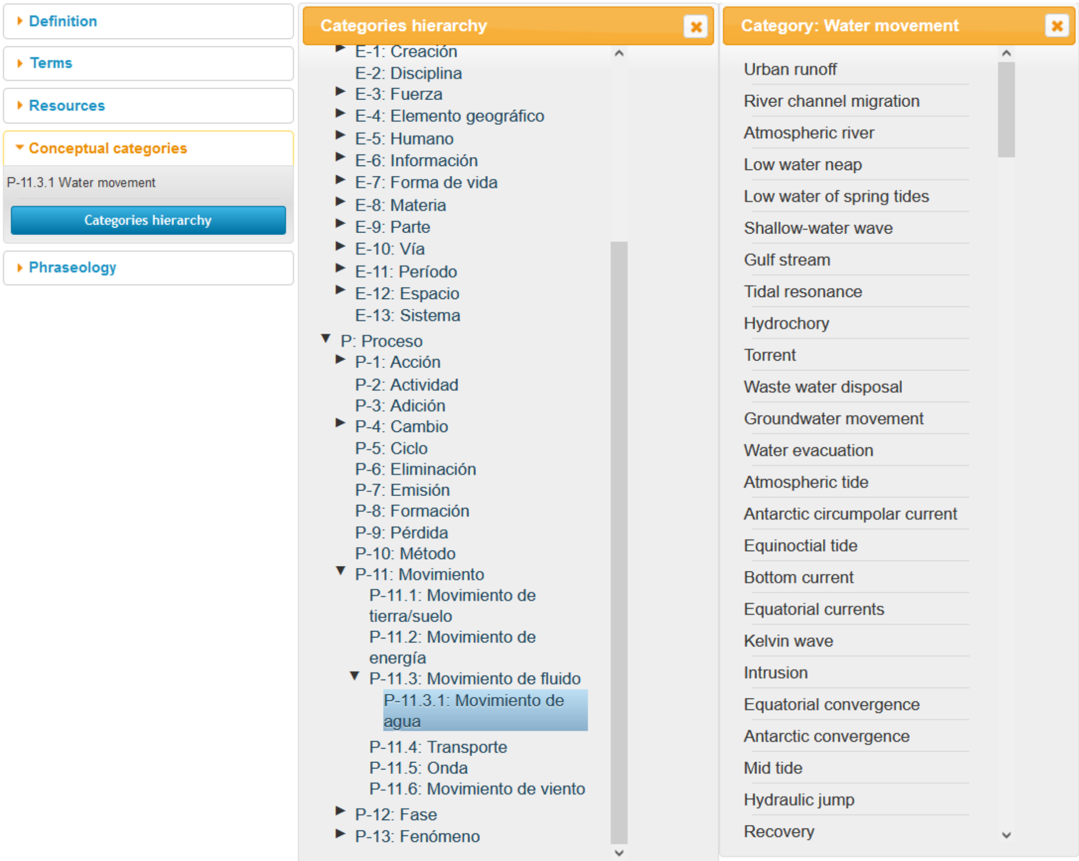


Figure 7. Conceptual category module in EcoLexicon (concept: TSUNAMI)

- **Phraseology:** this module displays syntactic and phraseological information, showing the nuclear meaning, the meaning dimension, the phraseological pattern, and the verbs related to a certain concept (**Figure 8**). TSUNAMI, for example, has a negative semantic prosody, since it is described as a "NATURAL DISASTER that causes a PATIENT to change for the worse". Moreover, it is related to the verb *destroy*, which further increases the negative connotation of the concept.

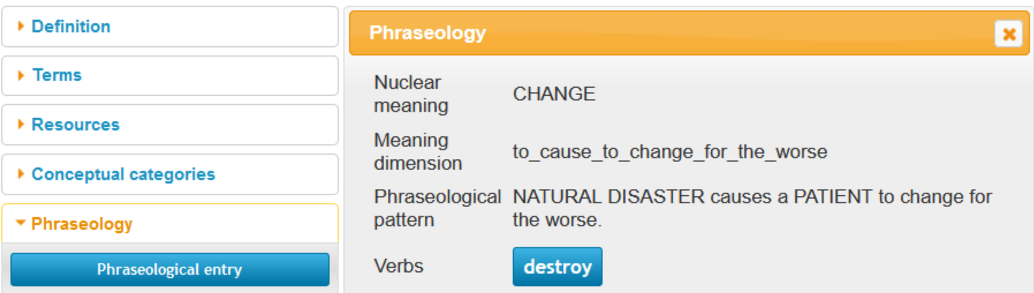


Figure 8. Phraseology module in EcoLexicon (concept: TSUNAMI)

These are the five main features that can be found within the microstructure of EcoLexicon, which highlight its linguistic, conceptual, and multimodal nature. Thanks to all of these modules, EcoLexicon is a resource that enhances knowledge acquisition because of its many-faceted knowledge representation: (i) through conceptual relations that codify conceptual propositions (i.e., concept-relation-concept) based on hierarchical and non-hierarchical criteria; (ii) through terminological definitions that reflect the salience of those conceptual relations, drawing from a central *genus*; and (iii) through multimodal resources that complement the conceptual and linguistic information. Furthermore, studies by García Aragón *et al*. (2014) and Giacomini (2014) have corroborated and validated the effectiveness of this terminological resource.

Moreover, EcoLexicon is related to another resource developed by the Lexicon Research Group: EcoLexiCAT[2] (León-Araúz *et al*. 2017; León-Araúz & Reimerink 2018). EcoLexiCAT is a web-based computer-assisted translation (CAT) tool developed by the Lexicon Research Group that facilitates the translation of environmental texts in English and Spanish by incorporating knowledge resources and corpus analysis tools. It is based on the free CAT software MateCat to enrich the source text with information from EcoLexicon, BabelNet (Navigli & Ponzetto 2012), and Sketch Engine (Kilgarriff *et al*. 2004, 2014).

As is evident from its different features and the by-products, EcoLexicon is a TKB that accounts for dynamic phenomena in specialized knowledge, such as variation, multidimensionality, and phraseology. As EcoLexicon has been a work in progress over the last twenty years, there is no doubt that in the future it will continue to offer cutting-edge terminological solutions to users.

[2] Available at: https://ecolexicon.ugr.es/EcoLexiCAT/index.htm

3. HYPONYMY

This chapter provides an overview of the state of the art on the study of hyponymy. As such, it focuses on the description, categorization, and representation of hyponymy. Firstly, it reviews the conception and role of hyponymy in Terminology and Ontology. Secondly, hyponymy is categorized by analyzing different ways of classifying it. Thirdly, the representation of hyponymy is compared in different terminological resources.

3.1. DESCRIPTION OF HYPONYMY

Hyponymy is the conceptual or semantic relation between a hypernym (i.e., a term referring to a generic, superordinate or parent concept) and a hyponym (i.e., a term referring to a specific, subordinate or child concept). Accordingly, the hyponyms of a same hypernym that are located at the same hierarchical level are regarded as co-hyponyms (i.e., terms referring to sibling concepts). The inverse relation of hyponymy is hyperonymy. Murphy & Koskela (2010: 80) write:

> Hyponymy is the lexical relation that expresses a relationship of inclusion between two lexemes, such as *bird* and *swan* or *cup* and *teacup*. The lexeme with the more general or inclusive meaning is called a hypernym (or, in some texts, hypernym), while the lexeme with the more specific or less inclusive meaning is a hyponym. Thus *swan* is a hyponym of *bird*, and conversely, *bird* is the hypernym of *swan*. Lexemes that are hyponyms of the same hypernym, at the same level of categorization (and that are therefore 'semantic sisters' and in a relationship of contrast) are called co-hyponyms – thus, for example, *swan*, *robin* and *pigeon* are all co-hyponyms.

Since all conceptual hierarchies are based on hyponymy, it naturally has an important role in our conscious thinking about word meaning. For this reason, it has been widely studied not only in Linguistics and Terminology, but also in Computer Science and Ontology Engineering.

3.1.1. HYPONYMY IN TERMINOLOGY

Hyponymy has been a major topic of study and research for both linguists and terminologists (Cruse 1995, 2000, 2002; Taylor 1995; Murphy 2003, 2006, 2010; Goddard & Schalley 2010; Murphy & Koskela 2010; L'Homme 2020). It is central to many models of the lexicon for the following reasons (Murphy 2003): (i) its inference-invoking nature; (ii) its importance in definition; and (iii) its relevance to selectional restrictions in grammar.

Hyponymy is defined in terms of inclusion, but the content that is inherited is dependent on whether hyponymy is viewed in terms of extensions (i.e., the categories that the words refer to), or in terms of intensions or senses (i.e., the semantic content associated with the words). Following Murphy & Koskela's (2010) example of birds, from the extensional perspective the category BIRD includes all the members of the category SWAN. However, from an intensional perspective, the inclusion relation is reversed and thus the hyponymic sense includes the sense of the hypernym. This means that, if BIRD is defined as "a winged animal that lays eggs", then SWAN would include all of these characteristics plus a few others (e.g., having a long neck, being usually white). Since this property inheritance does not happen in reverse, hyponymy gives rise to transitivity or unilateral entailment, by which the hypernym entails the hyponym, but not vice versa (Murphy & Koskela 2010).

Hyponymic relations tend to be represented in hierarchical or tree structures, which reveals their relevance towards conceptual organization. For instance, Murphy (2006) illustrates this kind of visual representation with a summarized version of the hyponymic relations in the lexical field of FRUIT (**Figure 9**).

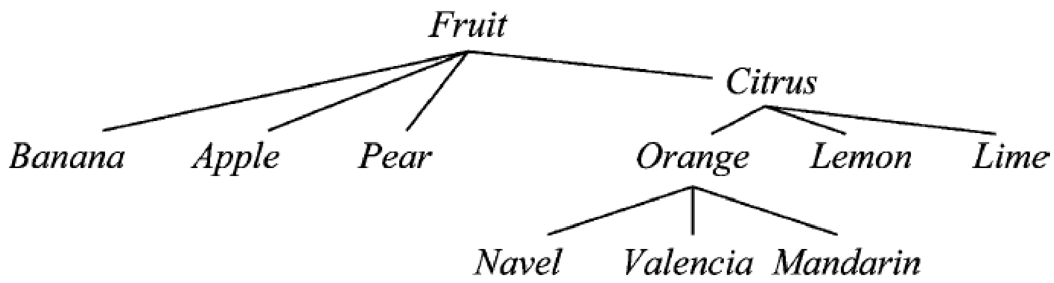


Figure 9. Hierarchy and hyponymic relations in the lexical field of FRUIT (Murphy 2006)

The conceptual hierarchy and its hyponymic relations are asymmetrical. This means that any word may have many hyponyms, but in most cases, only one immediate hypernym. When a hyponym has more than one hypernym, it is multidimensional. However, this is more frequent when a concept belongs to different contextual domains.

For instance, a MORAINE is a type of SEDIMENT because of its composition, but it is also a type of LANDFORM because of its impact on the terrain. The example shown in **Figure 9**, however, represents a monodimensional conceptual system. In it, ORANGE has various hyponyms (i.e., NAVEL, VALENCIA, MANDARIN), but each of them has only one hypernym (ORANGE). ORANGE has only one immediate hypernym (CITRUS), which is then associated with the most general hypernym, FRUIT.

In relation to multidimensionality, phenomena that affect hyponymy are *facets* and *microsenses* (Cruse 1995, 2002). Facets are dimensions or aspects of a concept that show a high degree of autonomy, and which make it possible to describe that concept from any of those perspectives. For instance, Cruse (2002) highlights two facets or dimensions in the hyponyms of BOOK, and divides them into two sets: *physical object* (e.g., HARDBACK, PAPERBACK) and *abstract text* (e.g., NOVEL, BIOGRAPHY). In these cases, the co-hyponyms of the same hypernym display within-set incompatibility, but between-set compatibility (a certain BOOK can be simultaneously a NOVEL and a HARDBACK, but a HARDBACK cannot be a PAPERBACK at the same time).

In contrast, a microsense is a specific meaning of a concept (i.e., regarding its properties, attributes or functions) which is only activated in a certain context. For example, although KNIFE generally has a single sense, it can be classified in different domains under a variety of hypernyms (WEAPON, TOOL, SURGICAL INSTRUMENT, etc.).

Hyponymy is a paradigmatic relation (Murphy 2003, 2006, 2010), and paradigmatic relations hold between members of the same syntactic category (e.g., nouns, adjectives, verbs). Nevertheless, there is a phenomenon known as *quasi-hyponymy* (Lyons 1977), which refers to cross-categorical relations by which hypernyms are sometimes of a different syntactic category than the hyponyms.

For example, adjectives may have nominal hypernyms (e.g., EMOTION – HAPPY, SAD, ANGRY). However, even in nominal hierarchies, there can be differences in syntactic categories (Murphy 2006). In the FRUIT hierarchy, whilst CITRUS can be a noun, it is more often an adjective (i.e., CITRUS FRUIT), as are the types of ORANGE (i.e., NAVEL ORANGE, MANDARIN ORANGE).

While quasi-hyponymy may seem like a problem for hyponymy because it describes transitive inclusion relations, it is also true that the line between hyponymy and quasi-hyponymy is far from clear (Murphy 2003). In fact, this problem disappears completely when hyponymic relations are analyzed from a cognitive perspective, since the focus is on the concept and not the term. Accordingly, knowledge-based approaches refer to hypernyms as generic concepts and to hyponyms as specific concepts (L'Homme 2020).

According to Cruse (2000, 2002), hyponyms are prototype-based categories. In this line, hypernyms or even the hyponyms at the top of a hierarchy in a conceptual system could allude to conceptual categories, depending on the level of abstraction. The hyponyms and hypernyms in *type of* statements are among the most prototypical cases of hyponymy in the sense that they participate in entailment relations, and the extra specificity of the hyponym in relation to the hypernym is central to the meaning of the hyponym (Cruse 2002). In relation to this, Cruse (2000) states that hyponymy should not only be understood as a definition of necessary and sufficient criteria (i.e., *a horse is an animal*), but also as a relational concept with a prototype structure and no clear definition or boundaries (i.e., *a horse is a type/kind/sort of animal*).

Another concept related to hyponymy is that of hyponymic enrichment (Cruse 2000). Contexts add semantic content to words or terms, thus enriching the meaning and making it more specific. Enrichments arise as a result of inferences which are no different from those operating more generally in language understanding (Cruse 2000). In this way, it is possible to hyponymically enrich a term (i.e., by narrowing it down to a subclass) and meronymically (i.e., by narrowing it down to a subpart).

Hyponymic enrichment occurs when the context adds features of meaning to a word or term which are not made explicit by the lexical item itself. Cruse (2000) exemplifies this with cases of hyponymic enrichment such as gender (e.g., *the teacher is on maternity leave*), height (e.g., *my brother bumps his head when he goes through the door*), and temperature (e.g., *the coffee burnt my tongue*). On the other hand, contextual determination may be to a specific kind of the class normally denoted by the lexical item (e.g., *I wish that animal would stop meowing*).

3.1.1.1. TAXONOMY

Taxonomy is a variant of hyponymy which describes a classification relation by which the hypernym and hyponym can be expressed in the frame *X is a type/kind/sort of Y* (Murphy & Koskela 2010). This phenomenon in Linguistics and, thus, Terminology, was first described by Cruse (1986), when he stated that taxonomy is an inclusion relation upon which well-formed taxonomies (i.e., conceptual hierarchies) are based. Accordingly, most hyponymic relations are taxonomic because they describe a relation between a generic concept and a specific concept. There are also other types of hyponyms, but Murphy (2010) considers that taxonyms are prototypical hyponyms.

A hypernym-hyponym pair is taxonomic if it fits the *X is a type of Y* proposition instead of the *X is a Y* proposition. For instance, APPALOOSA *is a type of* HORSE because it is a specific breed with particular characteristics (e.g., place of origin, color, appearance, physical traits). However, it is impossible to say that a MARE *is a type of* HORSE, whereas a MARE *is a* HORSE is correct, because this is a distinction based on gender and not on taxonomic membership. APPALOOSA is a taxonym of HORSE because it subdivides the species of horse into a distinctive and internally coherent class; however, MARE is not as distinctive or internally coherent, since it can include any horse as long as it is female. This is related to incompatibility.

3.1.1.2. INCOMPATIBILITY

Incompatibility is a relation of exclusion (Cruse 2000), in opposition to hyponymy, which is a relation of inclusion. Incompatibility manifests itself when a superordinate has more than one immediate hyponym (i.e., without any intermediate concepts or terms) and, among them, there is a set of terms that are not related to the others. In other words, incompatibles are terms which denote classes that have no shared members.

Cruse (2000) exemplifies this phenomenon through hypernym-hyponym pairs related to animals (e.g., ANIMAL – DOG, CAT, MOUSE, LION, SHEEP). If a concept is a DOG, it is an ANIMAL, but it cannot be a CAT, MOUSE, LION or SHEEP. Accordingly, there is no real entity that can belong simultaneously to the category of DOG and the category of CAT. This is even more evident in the case of hyponyms which are not taxonyms. As in the previous example, a MARE is a HORSE, but so is a STALLION. However, a MARE cannot be a STALLION at the same time, since it already has a female gender attribute.

Nonetheless, there are certain co-hyponyms that can be compatible, and this is directly related to the notion of facets (Cruse 2002). As previously explained, the hypernym BOOK has two sets of hyponyms. HARDBACK and PAPERBACK are co-hyponyms dependent on physical format, whereas NOVEL and BIOGRAPHY are co-hyponyms dependent on the textual content. This is an example of between-set compatibility.

Cruse (2000) also mentions the notion of co-taxonomy as a relation that is influenced by both taxonomy and incompatibility. It appears as *X is a different kind of Y from Z* (e.g., *a CAT is a different kind of ANIMAL from a DOG*). Therefore, taxonomy, co-taxonomy and incompatibility all correspond to a fundamental mode of categorization of experience: successive subdivision into prototypically mutually exclusive subcategories (Cruse 2000).

3.1.1.3. TROPONYMY

Troponymy is a temporally inclusive relation (i.e., occurring in the same time interval) that describes hyponymy between verbs (Fellbaum 2002, Murphy 2010). It refers to relations in which one verb expresses a specific manner in which the other verb's action is carried out. For instance, WALK, RUN, CRUISE, and SWIM are troponyms of MOVE because they mean *to move in a certain way* (L'Homme 2020). In other words, troponymy subclassifies verbs that represent the effects of processes or actions.

Moreover, like hyponymy, troponymy is also an asymmetrical relation since the relation can only happen in one direction (e.g., RUN is a troponym of MOVE, but MOVE is not a troponym of RUN). In this case, *hyperonymy* is still used to refer to the relation between a generic process or action (MOVE) and a specific process or action (RUN). Troponymy behaves like hyponymy in the sense that the subordinate actions or processes still contain the meaning of the superordinate actions or processes, but add characteristics (e.g., speed, manner, length, volume, angle) to it.

Additionally, Murphy (2010) considers that verbs in the troponymy relation are co-extensive. In other words, any time a manner-specifying verb like FRY is used, its hypernym COOK also describes the situation. Thus, every moment within a *frying* event can also be described as *cooking* (i.e., *frying* inhabits the same time interval as *cooking*). Likewise, every time someone is *running*, that person is also *moving* (i.e., *running* is temporally included within *moving*).

3.1.1.4. AUTOHYPONYMY

Autohyponymy, also known as vertical polysemy, occurs when a word or term has both a default general sense and a contextually restricted sense which denotes a subvariety of the general sense (Cruse 1995, 2000). For example, DOG refers simultaneously to the superordinate of DOG, understood as an *animal*, and to the subordinate of DOG, understood as a *male dog* (in opposition to BITCH, which is a *female* *dog*). Curiously, the more specific sense demonstrates autonomy, since the second clause can contradict the general sense of DOG: if the animal is a BITCH, then it *is a* DOG (Cruse 1995).

Furthermore, in some cases, autohyponyms duplicate both existing hyponyms and different ones (Horn 1984). For instance, the hypernym SMELL (i.e., *to emit any smell*) has a hyponym which is STINK (i.e., *to emit a bad smell*). However, it is autohyponymous because SMELL can also mean *to emit a bad smell*, even though there is no hyponym that means *to emit a smell that is not bad*.

3.1.2. BEYOND TERMINOLOGY: ONTOLOGIES AND HYPONYMY

Hyponymy has also been analyzed by experts in Computer Science, namely, Ontology Engineering, because of its relevance to the configuration of conceptual systems (Veres 2005; Stock 2010; Babaie 2011; Szostak 2012; Zhan & Wang 2015; Kleineberg 2017).

Models of the lexicon often involve the organization of conceptual or semantic relations, not only in semantic field theory, but also in network models such as ontologies (Murphy 2006). In such models, hyponymy is used to organize lexical items, or rather concepts. Accordingly, the representation of hyponymy is useful in ontology-based computational lexicons, such as WordNet (Miller 1986; Fellbaum 1998), whose aim is to represent the knowledge and conceptual relations that make the lexicon meaningful.

Veres (2005) also states that hyponymy (i.e., *is-a* or *is-a-kind-of*) is of primary importance to organize concepts in knowledge-based systems such as ontologies, and that well-structured taxonomies bring order to conceptual models. Since conceptual hierarchy is the backbone of ontologies, conceptual hierarchy acquisition has become a hot topic in ontology learning (Zhan & Wang 2015). There is thus the need for theories that guide the construction of class hierarchies, which are the ontological version of taxonomies, especially if this process is to be automated in computational resources. In Computer Science, names for hyponymy include AKO (*A Kind Of*), IS, SUPERC, and SUBSET. Nevertheless, in this domain, hyponymy is usually referred to as the *is-a* link (Veres 2005). Moreover, it implements inheritance as a way of optimizing the storage of information.

Hyponymy or the *is-a* link is the most important relation for nouns in WordNet (Miller 1986; Fellbaum 1998), a database in which lexical items are structured in semantic or conceptual networks. The lexical items in WordNet are nouns, adjectives, verbs and adverbs. They are organized in *synsets* (i.e., sets of synonyms), which in turn designate their underlying structure. WordNet is onomasiologically structured, which means that its information is organized based on meaning rather than word morphology. According to Miller (1986), in the design of WordNet, several possibly different senses of *is-a* were combined into the single hyponymy relation between nouns.

In this line, certain ontology-based applications in computational linguistics distinguish between hyponymy (i.e., *is-a*) and taxonomy (i.e., *is-a-type-of* or *is-a-kind-of*), as happens in Linguistics and Terminology (**§3.1.1.1**). However, WordNet was never intended as an ontology, simply a lexical database in which general language words are taxonomically classified. Nevertheless, it came to be regarded as a linguistic ontology, whose function is limited to describing semantic constructs rather than modeling specific domains (Gómez-Pérez *et al*. 2003).

Furthermore, in the same way as other knowledge representation schemas, ontologies emphasize taxonomies (Veres 2005). The concept of class and sub-class are primary in the Web Ontology Language (OWL). Their meanings are defined via a model theoretic semantics in which taxonomy is modeled by *rdfs:subClassOf* or class inclusion. This discipline also considers that a good taxonym must have as its essence a specification of the essence of the hypernym. Therefore, hyponymy or generic-specific relations play an important role in the knowledge structure of ontologies.

3.1.2.1. ONTOLOGIES

Ontologies are important in Computer Science and Computer Engineering (Weigand 1997; Gómez-Pérez *et al*. 2003; Nickles *et al*. 2007; Eckard *et al*. 2012; Seppälä 2015; Stratogiannis *et al*. 2021; Montiel-Ponsoda 2022). An ontology is usually regarded as a database that describes the concepts of a knowledge field, their properties or characteristics, and how concepts are related to each other (Weigand 1997). Ontologies originated in Philosophy, which through the centuries has asked questions about the origin and essence of things, as well as their classification in the world (León-Araúz 2009). In this sense, the initial questions remain the same, though formulated in different ways depending on the model, and always for a practical purpose.

Ontologies can be defined from a wide range of perspectives, depending on their application or construction process. Different authors have thus offered their own classifications of ontologies. For example, Mizoguchi *et al*. (1995) distinguish between *content ontologies*, *communication* (*'Tell & Ask'*) *ontologies*, *indexing ontologies*, and *meta-ontologies*. Another example is Van Heijst *et al*. (1997), who speak of *terminological ontologies*, *information ontologies*, and *knowledge-modelling ontologies*. On the other hand, Guarino (1998) distinguishes between *top-level ontologies*, *domain ontologies*, *communication ontologies*, and *terminological ontologies*. These examples demonstrate that there are many kinds of ontologies and the lack of consensus as to their possible typification.

In general, ontologies are networks of cross-connected conceptualizations (Goddard & Schalley 2010), linked with *ontological relations* which correspond to the conceptual relations in Terminology. In a similar way to conceptual relations, there are ontological relations that are not reflected in the lexicon because they allude to abstractions (i.e., to concepts). Unfortunately, the distinction between lexical and ontological relations is often not clear in linguistics, natural language processing (NLP), and artificial intelligence (AI) (Goddard & Schalley 2010).

On the one hand, NLP studies the interactions between computers and human language, analyzing how to program computers to process and analyze natural language data. On the other hand, AI studies the simulation of human intelligence processes by machines or computer systems. The most representative case is that of WordNet (Miller 1986; Fellbaum 1998), which was designed as a lexical resource and built on the basis of sense relations, but is often treated as an ontology.

Ontologies have also been developed in Computer Science and Computer Engineering, namely in AI and knowledge representation (Goddard & Schalley 2010). This has led to models such as Ontological Semantics (Nirenburg & Raskin 2004), which lists definitions of concepts not only to describe the meanings of lexical items of natural languages, but also to specify the meanings of the text-meaning representations that are part of an interlingua for machine translation (Nickles *et al*. 2007). According to Nirenburg & Raskin (2004: 10):

> Ontological semantics is a theory of meaning in natural language and an approach to natural language processing (NLP) which uses a constructed world model, or ontology, as the central resource for extracting and representing meaning of natural language texts, reasoning about knowledge

> derived from texts as well as generating natural language texts based on representations of their meaning.

In Linguistics and Terminology, ontology-based approaches are becoming increasingly popular, and TKBs are an evident proof of that. As a matter of fact, ontologies and terminologies have many aspects in common (Montiel-Ponsoda 2022). Ontologies represent and structure knowledge in a way that can be processed by machines, so that reasoning can be automated and information can be inferred. In contrast, terminologies represent and structure knowledge in a way that can be understood and communicated by humans, so that real-world entities can be explained and knowledge can be stored and transferred. Therefore, it is not surprising that models combining both disciplines have emerged.

3.1.2.2. TERMONTOGRAPHY

Termontography (Temmerman & Kerremans 2003; Kerremans *et al*. 2004, 2005; Temmerman *et al*. 2005; Temmerman 2007) is a multidisciplinary approach in which theories and methods for multilingual terminological analysis typical of the SCTT are combined with methods and guidelines typical of Ontology Engineering (Temmerman 2007). A clear distinction is thus made between conceptual modeling at a language-independent level, and a language-specific analysis of units of understanding. The name of this model combines *terminology*, *ontology*, and *terminography*.

Termontography uses both top-down and bottom-up approaches to capture and represent domain expertise in collaboration with field specialists, and the information in a domain-specific corpus (Temmerman & Kerremans 2003). The result of the top-down approach is an initial set of categories, whilst the result of the bottom-up approach is knowledge retrieved from texts reflecting culture-specific categorizations. The combination of the two results in a terminological database that can be directly exported as an ontology. In this approach, domain-experts play an essential role in the initial stages of the terminography work, namely, when compiling the domain-specific corpus and when agreeing on the main classes to be represented in the resource (Montiel-Ponsoda 2022).

Therefore, the methodology involves knowledge management and representation in specific domains of experience (**Figure 10**). Based on application-specific requirements, an initial framework of categories, concepts, sets and relations is developed. This framework, which is a template for the manual and semi-automatic extraction of knowledge from corpora, gradually evolves to become an enriched and more fine-grained network of semantic relations (Temmerman 2007).

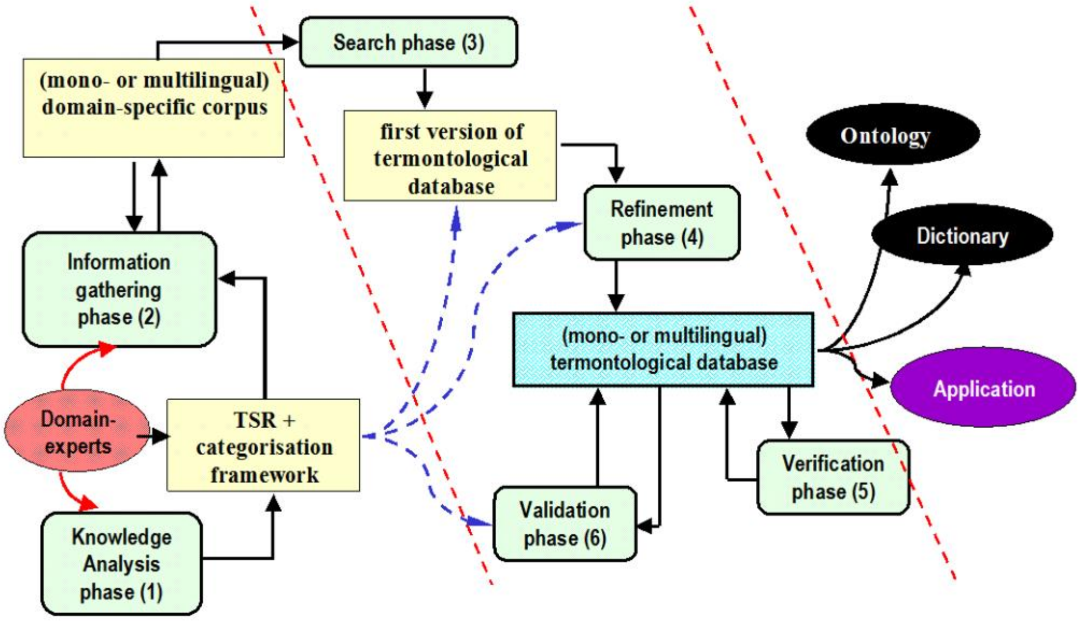


Figure 10. Methodology applied in termontography (Temmerman 2007)

The practical application of this model is the Termontography Workbench (Temmerman 2007). It is a software suite that distinguishes between a language-independent analysis resulting in a categorization framework plus ontology on the one hand, and a language-and-culture-dependent terminological description on the other. The results of this analysis are stored in a termontological database which can support different applications.

Furthermore, in accordance with the SCTT and thus termontography, before building a domain-specific conceptual model or ontology, it is necessary to have an excellent grasp of the categories (i.e., units of understanding) and their interrelationships, independent of any culture or language in the domain of interest (Kerremans *et al*. 2004). Even though termontography was initially a by-product of the SCTT, over the years it evolved to become an independent entity. It thus underwent a transformation and now bears little resemblance to the initial premises of the SCTT (Faber & López Rodríguez 2012).

For example, the conceptual representations proposed are in the form of computer-implemented ontologies, with no mention of prototypes and idealized cognitive models. However, the relation to Terminology is still evident. Accordingly, the types of termontographic conceptual relations mentioned by Kerremans *et al.* (2005) (e.g., *has_subtype*, *is_kind_of*) are directly equivalent to traditional hyponymic or generic-specific relations.

3.1.2.3. ONTOTERMINOLOGY

Ontoterminology (Roche 2007, 2012a, 2012b, 2015; Roche *et al.* 2009) describes a terminology whose conceptual system is a formal ontology. As such, it relies on epistemological principles, which focuses on the difference between the linguistic and conceptual dimensions of terminology and unifies them (Roche 2012b).

In this paradigm, a double semiotic triangle links linguistic notions to ontological ones. **Figure 11** shows that both dimensions are present in ontoterminological projects, though they rely on two different semiotic systems. Thus, signified (i.e., meaning) and signifier (i.e., term), related to Linguistics and natural language, are separated from the concept and its name (i.e., identifier), referring to a formal or conceptual system.

In this way, ontoterminology distinguishes between the definition of the term, written in natural language, and the definition of the concept, written in a formal language (Roche 2015). Accordingly, this distinction is particularly important in subject fields where concepts can be both represented and defined in a non-verbal way, as in Medicine.

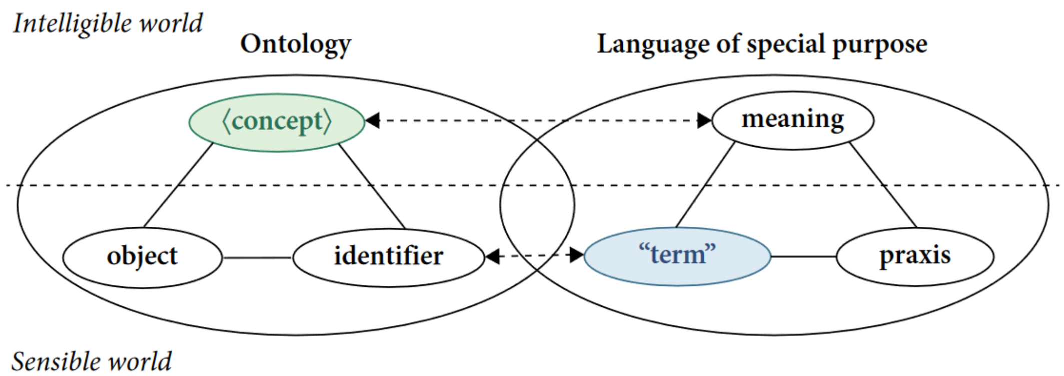


Figure 11. Double semiotic triangle according to ontoterminology (Roche 2015)

Therefore, ontoterminology combines ontologies with terminologies, focusing on the premise that a concept does not need a term to exist (Roche 2015). In this approach, formal ontologies are conceptualized by experts and represented in a formal language before natural language is used to designate and describe concepts (**Figure 12**). According to Roche (2015), if ontologies are derived from scientific discourse, then they cannot be considered formal ontologies, but rather textual ontologies. This evidently reduces the notion of terminology to a specialized lexicography.

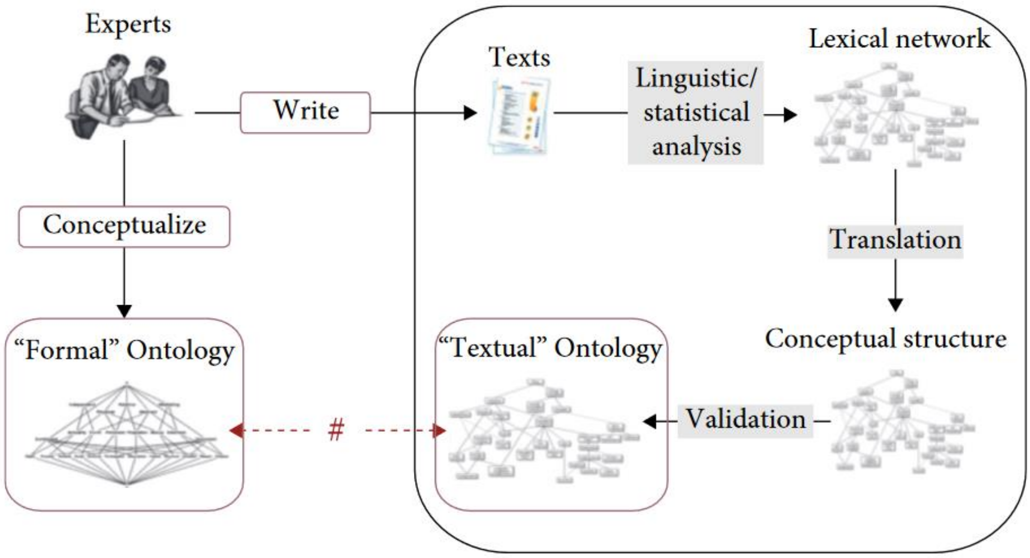


Figure 12. Formal ontology vs. textual ontology (Roche 2015)

In this approach, conceptualizations model a reality and are standardized (in a formal, logical language). It is also necessary to reach a consensus regarding the definitions and concept identifiers. Only then can terms and term definitions be accounted for and mapped onto the ontology concepts (Montiel-Ponsoda 2022). In fact, Roche (2012b) emphasizes the complementary nature of Terminology and AI through definitions written in natural language in order to clarify or standardize concepts. Terminology is relevant to AI because of its effort to represent concepts in natural or semi-formal language, whilst AI uses formal language in order to manipulate or compute concepts.

Ontoterminology and termontography are complementary. Whereas Temmerman's (2007) proposal focuses on specialized vocabulary and gives priority to the semasiological approach, Roche's (2015) theory is clearly onomasiological and in a certain way closer to the GTT (Wüster 1968, 1979).

3.1.2.4. ONTOLOGICAL KNOWLEDGE ENHANCEMENT IN ECOLEXICON

In recent years, EcoLexicon has been enriched by conceptual and semantic information that interrelates and nourishes its content. Because of this expansion in conceptual meaning, there is now the need for an enhanced ontology of environmental categories because the original environmental event (Faber 2015; León-Araúz *et al*. 2012) has no specific category types that can be used to annotate environmental concepts ontologically. For this reason, Gil-Berrozpe *et al*. (2019) carried out an ontological knowledge enhancement by categorizing all concepts in the database. This involved an in-depth revision of the ontology underlying EcoLexicon, and the implementation of new features for the new conceptual category module.

3.1.2.4.1. CONCEPTUAL CATEGORIZATION PROCESS

Ontologies are often organized as classification hierarchies and tend to be as universal as possible so that they can be used and reused for different applications. Such hierarchies tend to position the three most basic ontological categories at the top level: *entities* or *objects*, *processes* or *events*, and *attributes* or *properties* (Mahesh & Nirenburg 1995; Moreno-Ortiz & Pérez-Hernández 2000).

In this line, various ontology-based projects for categorizing environmental knowledge have been carried out, such as the Environmental Ontology (ENVO) (Buttigieg *et al*. 2013, 2016). More specifically, ENVO defines itself as "a community-led, open project which seeks to provide an ontology for specifying a wide range of environments relevant to multiple life science disciplines and, through an open participation model, to accommodate the terminological requirements of all those needing to annotate data using ontology classes" (Buttigieg *et al*. 2013). Although this project initially focused on the representation of biomes, environmental features, and environmental materials, it was expanded to include ontological information related to a multitude of interrelated fields (Buttigieg *et al*. 2016).

Similarly, the conceptual categorization process in EcoLexicon followed the premises behind ENVO's ontological reasoning by adapting the conceptual categories and hierarchies to the specific needs of the environmental knowledge in EcoLexicon. Because of the dynamism of environmental sciences (León-Araúz *et al*. 2012), it was essential to account for the multifaceted nature of concepts, since concepts can belong to more than one category depending on their salient features

(Kageura 1997). For this reason, the conceptual categorization process was carried out from a multidimensional perspective.

A series of conceptual categories or semantic classes belonging to different top-down categorization levels were established to determine degrees of specificity (Murphy & Lassaline 1997) and conceptual similarity (Hahn & Chater 1997). This meant that every concept was tagged with a category showing its interrelation with ontologically-similar elements. These conceptual categories were mainly based on definitions and on the contextual information in the EcoLexicon corpus, but they were also contrasted with the ontological classes in ENVO (Buttigieg *et al*. 2013, 2016). Consequently, an enhanced category system for EcoLexicon was specified and hierarchically organized (**Figure 13**).

Process	Loss		
Process	Method		
Process	Movement		
Process	Movement	Earth / Soil movement	
Process	Movement	Energy movement	
Process	Movement	Fluid movement	
Process	Movement	Fluid movement	Water movement
Process	Movement	Transport	
Process	Movement	Wave	
Process	Movement	Wind movement	
Process	Phase		
Process	Phase	Phase of cycle	
Process	Phase	Phase of treatment	
Process	Phenomenon		
Process	Phenomenon	Atmospheric phenomenon	
Process	Phenomenon	Atmospheric phenomenon	Precipitation
Process	Phenomenon	Optical phenomenon	

Figure 13. Segment of the hierarchy established in the conceptual categorization process

In this way, the 4,500 concepts then in EcoLexicon were classified in 152 categories, distributed in five categorization levels. To begin with, at the most general level are the three starter ontological categories (Mahesh & Nirenburg 1995; Moreno-Ortiz & Pérez-Hernández 2000):

- A: *attribute* – properties of entities and processes
- E: *entity* – physical and mental objects
- P: *process* – events extending over time and involving different participants

However, depending on the ontological nature of concepts, they can be subclassified in up to five levels of specificity, as shown in the category hierarchy of *creation* concepts:

E: *entity*
- E-1: *creation*
  - E-1.1: *artifact* (e.g., DC BUS)
    - E-1.1.1: *conduit* (e.g., DUCT)
    - E-1.1.2: *container* (e.g., SEDIMENTATION TANK)
    - E-1.1.3: *instrument* (e.g., CENTRIFUGAL PUMP)
      - E-1.1.3.1: *measuring instrument* (e.g., ACCELEROMETER)
      - E-1.1.3.2: *recording instrument* (e.g., ALBEDOGRAPH)
      - E-1.1.3.3: *sampling instrument* (e.g., AUTOMATIC SAMPLER)
      - E-1.1.3.4: *transforming instrument* (e.g., SOLAR CELL)
    - E-1.1.4: *vehicle* (e.g., DREDGER)
  - E-1.2: *software* (e.g., COMPUTER APPLICATION)
  - E-1.3: *structure* (e.g., PIER)
    - E-1.3.1: *building* (e.g., OIL REFINERY)
    - E-1.3.2: *defense structure* (e.g., REEF BREAKWATER)

Additionally, those multidimensional concepts (Kageura 1997) were classified in as many categorization hierarchies as necessary, depending on the salient features in their definitions and the corpus. For instance, one of the most multifaceted concepts is PORT, which is a member of four categories:

- **Concept:** PORT
- **Definition (from EcoLexicon):** place along a river or sea coast that gives ships and boats protection from storms and rough water, and where ships can load and unload cargo. It can be natural or artificial.
- **Conceptual category:**
  - E-1.3: *structure*
  - E-4.1: *artificial geographic feature*
  - E-4.2: *natural geographic feature*
  - E-12.1.2: *facility*

Figure 14 shows a segment of the categorization table that summarizes the classification process. The first column contains the concept analyzed; the second column indicates whether the concept is multidimensional; the third column describes the number of categories applied to a single concept; and the remaining columns contain the top-down categories applied to each concept.

survival	***NO***	***FIRST CATEGORY:***	Process	Activity			
rill	***NO***	***FIRST CATEGORY:***	Entity	Part	Part of landform		
longshore trough	***NO***	***FIRST CATEGORY:***	Entity	Geographic feature	Natural geographic feature	Landform	
exsurgence	***NO***	***FIRST CATEGORY:***	Entity	Geographic feature	Natural geographic feature	Landform	Natural water body
suspension	***NO***	***FIRST CATEGORY:***	Entity	Matter	Solid matter		
suspension feeder	***NO***	***FIRST CATEGORY:***	Entity	Lifeform	Animal		
substance	***NO***	***FIRST CATEGORY:***	Entity	Matter			
ergastic substance	***NO***	***FIRST CATEGORY:***	Entity	Matter	Chemical substance		
ozone-depleting substance	***NO***	***FIRST CATEGORY:***	Entity	Matter	Chemical substance		
tide table	***YES***	***FIRST CATEGORY:***	Entity	Information	Representation		
		SECOND CATEGORY:	Entity	Information	Document		
life table	***YES***	***FIRST CATEGORY:***	Entity	Information	Representation		
		SECOND CATEGORY:	Entity	Information	Document		
sheet pile	***NO***	***FIRST CATEGORY:***	Entity	Part	Part of structure		
taiga	***NO***	***FIRST CATEGORY:***	Entity	Geographic feature	Natural geographic feature	Landscape	
stem	***NO***	***FIRST CATEGORY:***	Entity	Part	Part of lifeform	Part of plant	
thallus	***YES***	***FIRST CATEGORY:***	Entity	Part	Part of lifeform	Part of plant	
		SECOND CATEGORY:	Entity	Part	Part of lifeform	Part of fungus	
slope	***NO***	***FIRST CATEGORY:***	Entity	Part	Part of landform		
continental slope	***YES***	***FIRST CATEGORY:***	Entity	Space	Layer		
		SECOND CATEGORY:	Entity	Part	Part of landform		
grain size	***NO***	***FIRST CATEGORY:***	Attribute	Physical attribute	Size		
drum	***YES***	***FIRST CATEGORY:***	Entity	Creation	Structure		
		SECOND CATEGORY:	Entity	Creation	Artifact	Container	

Figure 14. Segment of the table used in the conceptual categorization process

From an ontological point of view, 16 categories were associated with attributes, 93 with entities, and 43 with processes. See **Appendix I** for a full list of the conceptual category hierarchy in EcoLexicon and examples of each category.

In this research, conceptual categories played an important role in the categorization of the knowledge conveyed in the conceptual hierarchies. For this reason, they formed an essential part of the terminological entries that were designed, because they provide a better understanding of the ontological nuance of each concept.

3.1.2.4.2. Conceptual Category Module

The ontological enhancement of EcoLexicon was mainly based on the multidimensional categorization of its concepts in these categories. As a result, not only was it possible to improve the structure and organization of the environmental knowledge in the resource, but also to offer new practical applications and functionalities so that the end user could make the most of the ontological information. Essentially, the ontologically-enhanced functions implemented in EcoLexicon involved the design of a new conceptual categories module, which included (i) the revised category hierarchy, and (ii) a conceptual combination function, which will soon be available.

The original conceptual category module in EcoLexicon only classified concepts according to the semantic roles in the environmental event (Faber 2015;

León-Araúz *et al*. 2012). For this reason, after categorizing the concepts, it was necessary to redesign this module. This involved the following: (i) modification and update of the category hierarchy function; and (ii) implementation of the conceptual combinations function. **Figure 15** displays the conceptual category module for the concept PORT. Four conceptual categories (E-1.3: *structure*, E-4.1: *artificial geographic feature*, E-4.2: *natural geographic feature*, and E-12.1.2: *facility*) are showcased, as well as the buttons for category hierarchy and conceptual combinations.

Figure 15. Conceptual category module in EcoLexicon (concept: PORT)

The enhanced conceptual category hierarchy function of this new module includes a hierarchically-organized list of all 152 conceptual categories (**Figure 16**). The members of each category can be accessed by clicking on the triangle to the left, enlarging the list to view the more specific subcategories. When a category is selected, a new window pops up with all the concepts belonging to it. This provides easy access to each entry, its information, and its ontologically-interrelated concepts in EcoLexicon. For example, in **Figure 16** the concepts belonging to the *defense structure* category are listed alphabetically, and clicking on any of them (e.g., COFFERDAM, DIKE) leads to its full entry with all the information.

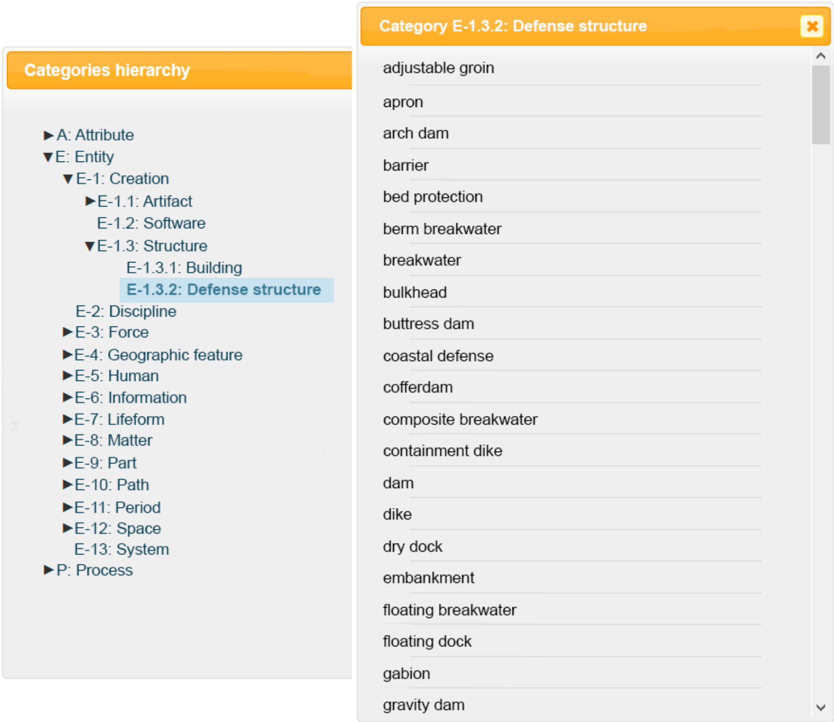


Figure 16. Category hierarchy function in EcoLexicon (category: *defense structure*)

On the other hand, in the conceptual combination function of the new conceptual category module, users can perform a simple or advanced query. **Figure 17** and **Figure 18** show the query screen and the results screen of the simple query "hard structure". The simple query box (**Figure 17**) can be used to perform a proximity search, since it then autocompletes with the available concepts.

As shown in the results screen (**Figure 18**), the system automatically converts the search into a query expression ("hard structure [CONCEPT]") and displays a list of results in EcoLexicon that show the combinatorial potential of the search concept with other concepts through specific conceptual relations. By default, these results are collected under conceptual propositions composed of conceptual categories (in black) linked by conceptual relations (in orange).

For instance, the fourth result in **Figure 18** is listed as "[Defense structure] *made of* [Material]". However, in order to see the specific concepts codified under those categories, it is necessary to click on the "+ Show specific results" option (in blue) next to this conceptual proposition, and thus the actual results of the query appear: "HARD STRUCTURE *made of* CONCRETE", "HARD STRUCTURE *made of* STEEL", "HARD STRUCTURE *made of* QUARRY STONE", etc.

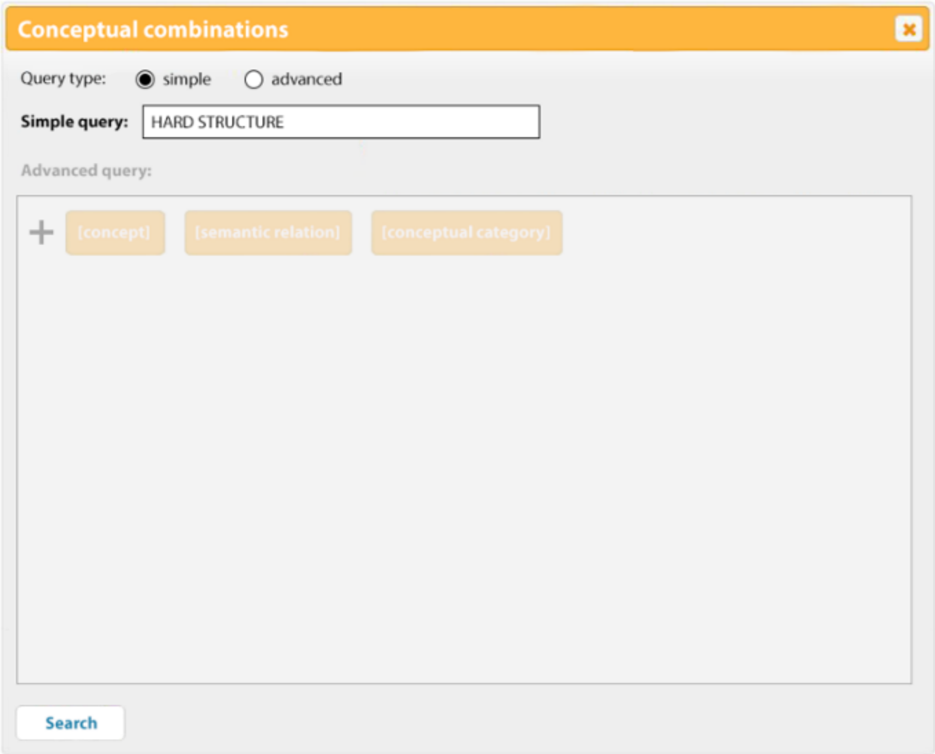


Figure 17. Example of a simple query in the conceptual combinations function

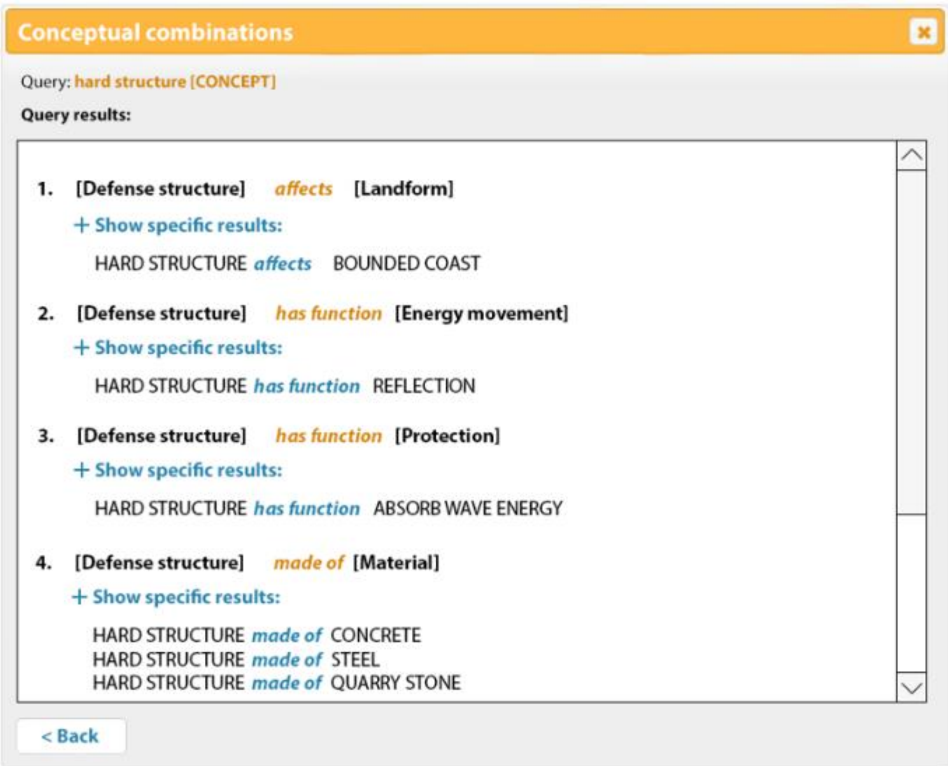


Figure 18. Results of a simple query in the conceptual combinations function

In contrast, the advanced query allows users to perform more complex searches. As shown in **Figure 19**, this type of query is based on three elements: (i) concepts; (ii) conceptual relations (herein referred to as *semantic relations*); and (iii) conceptual categories.

By clicking on the orange bubbles next to the "+" symbol, users can add as many elements to the query as they wish in any order, since this query allows for free element combination (e.g., "category + relation", "concept + relation + category", "category + relation + category", etc.). Similarly, any element can also be deleted.

The concept bubble also has a free text box where users can type anything, whilst the conceptual relation and the conceptual category bubbles display a picklist showing all the relations or categories in EcoLexicon. However, it is also possible to choose the option "ANY" in the conceptual relation and conceptual category bubbles. In fact, displaying all the possibilities with a picklist is the simplest way for users to find and select the most suitable option for their query. In addition, each bubble contains an "AND" and an "OR" button, which are useful if users want to look for more than one concept, relation, and/or category found in the same position.

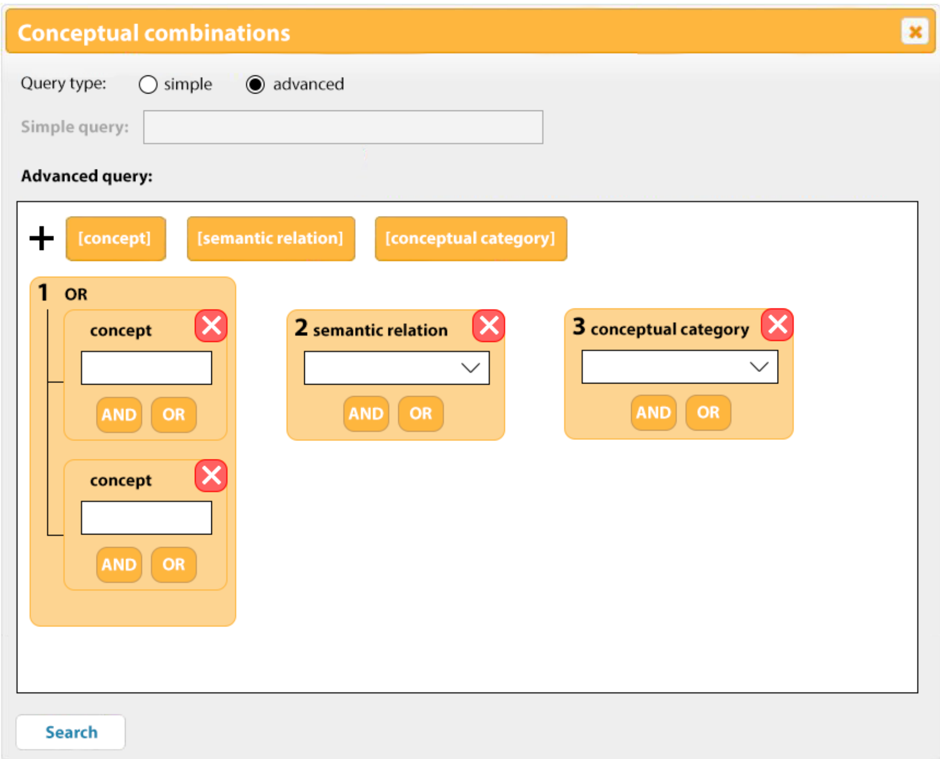


Figure 19. Options of the advanced query in the conceptual combinations function

Finally, **Figure 20** and **Figure 21** show the query screen and the results screen of the advanced query "Water movement [CATEGORY] + any [SEMANTIC RELATION] + Natural water body [CATEGORY]". In order to perform this search, users must select the option "advanced" next to "Query type". This activates the advanced query box, where the user then creates a conceptual category bubble in order to select "Water movement", a conceptual relation bubble in order to select "ANY", and a conceptual category bubble in order to select "Natural water body".

As a consequence, this expression displays a series of results that include conceptual propositions linking concepts belonging to the *water movement* category and the *natural water body* category through any conceptual relation. For instance, the first case is the conceptual proposition "[Water movement] affects [Natural water body]", including examples such as "FLOOD CURRENT affects BAY", "TIDE affects TIDAL RIVER", and "REGRESSION affects SEA".

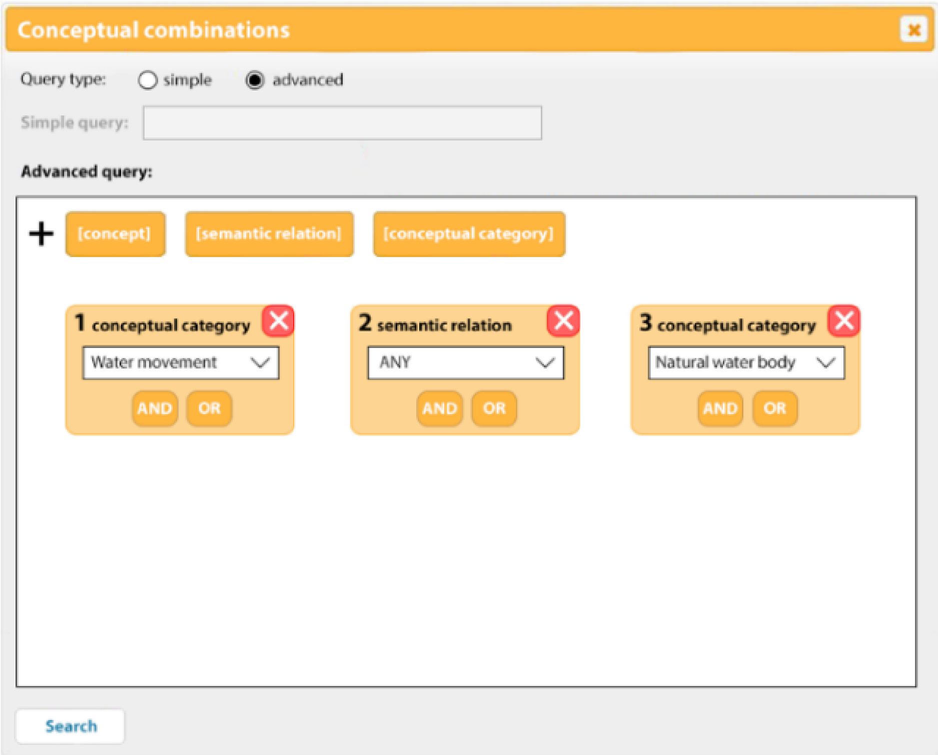


Figure 20. Example of an advanced query in the conceptual combinations function

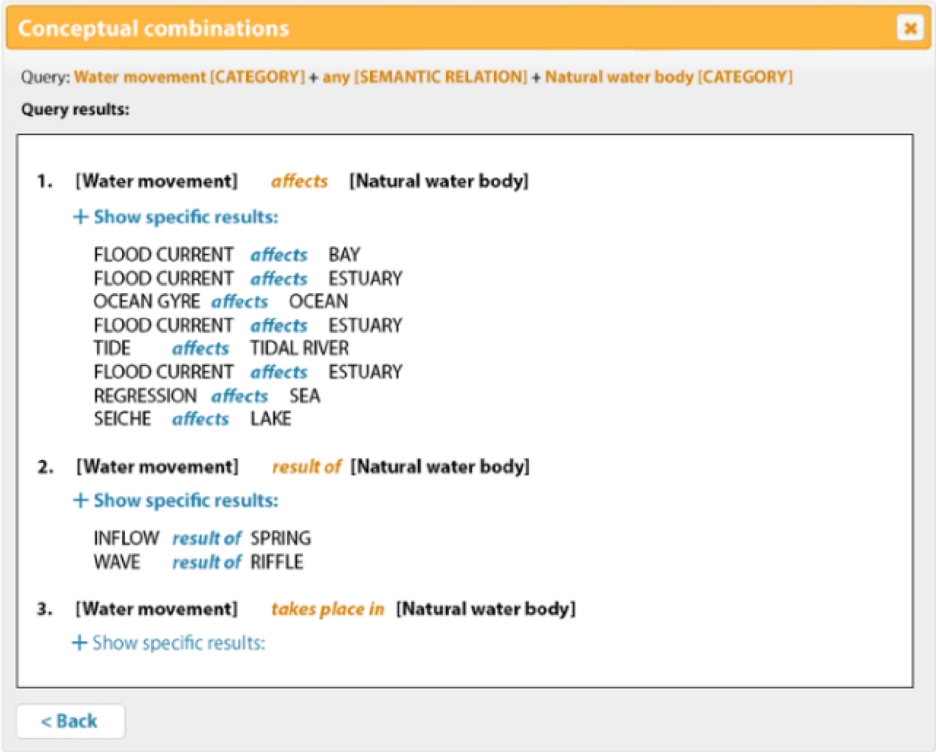


Figure 21. Results of an advanced query in the conceptual combinations function

3.2. CATEGORIZATION OF HYPONYMY

The categorization of hyponymy is based on its classification according to different typologies with specific nuances. The representation of specialized knowledge is complex because hyponymy is the conceptual relation on which the conceptualization of real-world entities and the hierarchization of all entities and processes are based. However, this complexity is not exclusive to hyponymy since other hierarchical and associative relations also have more specific subtypes.

Meronymic or part-whole relations have been widely researched in this regard (Winston *et al*. 1987; Berland & Charniak 1999; Cruse 2000; Girju *et al*. 2003.; Murphy 2003, 2010; L'Homme 2020). As a practical example of the categorization of meronymy, León-Araúz *et al*. (2012) specify that the relations *made of*, *phase of*, *delimited by*, *located at*, and *attribute of* can be regarded as subtypes of meronymy (i.e., *part of*).

In this sense, not all parts (i.e., meronyms) interact in the same way with their wholes (i.e., holonyms). This categorization, which includes one general meronymy relation and five meronymy subtypes, was incorporated into EcoLexicon because of domain-specific needs, but also to improve ontological reasoning and transitivity-related consistency.

For instance, if both processes and entities were connected through the same general *part of* relation, there would be no restrictions on conceptual category membership. In this line, even though CONDENSATION is a *part of* the HYDROLOGIC CYCLE, it is more accurate to say that CONDENSATION is a *phase of* the HYDROLOGIC CYCLE. Thus, the same argument could be applied to hyponymy as well.

For this reason, this section analyzes the different types of hyponymy categorization. The theoretical proposals discussed are based on the following distinctions: (i) taxonomic and functional hyponymy; and (ii) direct and indirect hyponymy. In addition, this section presents a case study of hyponymy refinement, which involves the specification of hyponymy subtypes typical of environmental terminology, and the identification of hyponymic knowledge patterns.

3.2.1. TAXONOMIC AND FUNCTIONAL HYPONYMY

The most generally accepted distinction in hyponymy subtypes is between taxonomic and functional hyponymy (Wierzbicka 1984; Cruse 1986; Miller 1998; Murphy 2003). Taxonomic hyponymy is thus considered the *is a type of* relation, whilst functional hyponymy is regarded as the *is used as a type of* relation (Miller 1998). In this distinction, taxonomic hyponymy essentially refers to the phenomenon of taxonomy.

With regard to functional hyponymy, many nouns incorporate functional features in their definitions and the function, or at least an aspect of it, is frequently captured by the meaning of a hyponym (Cruse 2002). In these cases, the hyponym simultaneously specifies the function more precisely and adds perceptual features which are largely absent in the hypernym.

For example, COW is a taxonomic hyponym of ANIMAL (i.e., a COW *is a type of* ANIMAL), but a functional hyponym of LIVESTOCK (i.e., a COW *is used as a type of* LIVESTOCK). Interestingly, this distinction is related to the notion of microsenses (Cruse 1995, 2002). According to a microsense in Biology or Physiology, there is a hyponymic relation between ANIMAL and COW. However, from the perspective of a microsense in the field of farm and livestock production, there is a hyponymic relation between LIVESTOCK and COW. Moreover, whereas taxonomic hyponymy is always analytical, functional hyponymy is vaguer since it does not refer to logically necessary relations (e.g., not every COW is LIVESTOCK) (Murphy 2003).

Another example mentioned by Murphy (2006) is GAME CONSOLE, which *is a type of* COMPUTER, which in turn *is a type of* OFFICE EQUIPMENT. Because of property inheritance and transitivity, this would mean that a GAME CONSOLE *is a type of* OFFICE EQUIPMENT. Though strange at first glance, it can be explained when different hyponymy subtypes, such as taxonomic hyponymy and functional hyponymy, are taken into account.

More specifically, the first relation states that a GAME CONSOLE is a COMPUTER (i.e., taxonomic hyponymy), whereas the second premise indicates what a COMPUTER is used for (i.e., functional hyponymy). Furthermore, there are situations where even a GAME CONSOLE can be directly considered as a type of OFFICE EQUIPMENT in a functional sense, as in the offices of game developers who might use them as tools for testing their games.

This reveals that hyponymic relations are richer than the simple *is a* conceptual relation and that, depending on the context, any term might have hypernyms of at least taxonomic and functional hyponymy. In fact, this was one of the motivations to enrich the categorization of hyponymy in EcoLexicon by means of subtypes.

Wierzbicka (1984) points out three categories of superordinate or hypernym that extend the notion of taxonomic and functional hyponymy: *collectiva-singularia tantum*, *collective-pluralia tantum*, and *pseudo-contables*. All of these categories are marked by their particular morpho-semantic status with regard to countability and number (Murphy 2003). The collectiva-singularia tantum category involves non-countable, singular superordinates that have countable subordinates defined in terms of "what for and where" (e.g., FURNITURE – CHAIR). The collectiva-pluralia tantum category includes non-countable, plural superordinates defined in terms of "where and why" (e.g., LEFTOVER – no specific hyponym). The pseudo-countable category refers to hypernyms such as VEGETABLE and NARCOTIC, defined in terms of "what for and where from".

Finally, and also in relation to the distinction between taxonomic and functional hyponymy, Cruse (2002) specifies three hyponymy subtypes: *natural kind hyponymy*, *nominal kind hyponymy*, and *functional hyponymy*. In fact, the notions of natural kind hyponymy and functional hyponymy as proposed by Cruse (2002) are equivalent to those of taxonomic hyponymy and functional hyponymy as proposed by Miller (1998), respectively.

That said, it remains to be seen what nuances the other remaining subtype adds to the categorization of hyponymy. More specifically, nominal kind hyponymy refers to a relation that can be captured in terms of a single differentiating feature (e.g.,

HORSE – MARE, CAT – KITTEN, WOMAN – BLONDE). According to Cruse (2002), this type of relation is like that of a morphologically derived term and its base (e.g., LION – LIONESS, DUCK – DUCKLING). Therefore, nominal kind hyponymy is parallel to non-taxonomic hyponymy in the same way that natural kind hyponymy is parallel to taxonomy.

3.2.2. DIRECT AND INDIRECT HYPONYMY

A more recent categorization of hyponymic or generic relations is proposed by Nuopponen (2018, 2022), who distinguishes between *direct generic super/subordination* and *indirect generic super/subordination*. In this thesis, this refers to *direct hyponymy* and *indirect hyponymy*.

According to Nuopponen (2018, 2022), superordinate concepts (i.e., hypernyms) have a higher level of abstraction, whereas subordinate concepts (i.e., hyponyms) possess a lower one. The differences and proximity or distance in abstraction levels is precisely what determines direct generic super/subordination (i.e., direct hyponymy) and indirect generic super/subordination (i.e., indirect hyponymy).

On the one hand, direct generic super/subordination is present when super- and subordinate concepts are located at subsequent abstraction levels (e.g., SOFTWARE – APPLICATION SOFTWARE). This means that they are in a relation of direct hyponymy because they are conceptually as close as possible to each other, without any intermediate concept that further specifies any characteristics inherited through the conceptual hierarchy.

On the other hand, indirect generic super/subordination occurs when super- and subordinate concepts are at a different abstraction level that is further away (e.g., SOFTWARE – TEXT PROCESSING SOFTWARE). In other words, they are in a relation of indirect hyponymy because they are conceptually distant and lack an intermediate concept that narrows the characteristics between the superordinate and the subordinate concept in the conceptual hierarchy.

Similarly, Nuopponen (2018, 2022) also alludes to *generic co-ordination* (i.e., co-hyponymy), and distinguishes between *direct generic co-ordination* (i.e., direct co-hyponymy) and *indirect generic co-ordination* (i.e., indirect co-hyponymy). In the case of direct generic co-ordination, co-ordinate concepts are located at the same level below the same direct superordinate concept and have the same criteria of division (e.g., APPLICATION SOFTWARE – SYSTEM SOFTWARE). However, in the case of indirect generic co-ordination, co-ordinate concepts are on the same level too, but under different direct superordinate concepts or with different criteria of subdivision (e.g., APPLICATION SOFTWARE – SERVER SOFTWARE).

In addition, there is the *generic diagonal relation,* which refers to other pairs of concepts at different abstraction levels in the same concept system (e.g., SYSTEM SOFTWARE – TEXT PROCESSING SOFTWARE). **Table 4** shows all of the generic concept relations indicated by Nuopponen (2022).

Generic concept relations	Relation participants	Examples
Generic super/ subordination	superordinate (higher level of abstraction) subordinate (lower level of abstraction)	
a. Direct generic super/subordination	super- and subordinate concepts on subsequent abstraction levels	SOFTWARE – APPLICATION SOFTWARE
b. Indirect generic super/subordination	super- and subordinate concepts on a different abstraction level further away	SOFTWARE – TEXT PROCESSING SOFTWARE
Generic co-ordination		
a. Direct generic co-ordination	co-ordinate concepts on the same level and below same direct superordinate concept and same criteria of division	APPLICATION SOFTWARE – SYSTEM SOFTWARE
b. Indirect generic co-ordination	co-ordinate concepts on the same level, but under different direct superordinate	APPLICATION SOFTWARE – SERVER SOFTWARE
Generic diagonal relation	other pairs of concepts on different abstraction levels in the same concept system	SYSTEM SOFTWARE – TEXT PROCESSING SOFTWARE

Table 4. Types of generic concept relations (Nuopponen 2022)

3.2.3. HYPONYMY REFINEMENT IN ECOLEXICON

Terminological knowledge bases can be more coherent and dynamic when the range of conceptual relations is wider than the traditional generic-specific and part-whole relations (León-Araúz *et al*. 2012). This means considering non-hierarchical relations and expanding the original sense of both hyponymy and meronymy. Nevertheless, in EcoLexicon, the *type of* relation had still not been divided into more specific subtypes. This was a source of problems in the representation of conceptual hierarchies (**Figure 22**). More specifically, the representation of different dimensions of co-hyponyms tended to produce noise, information overload, redundancy, and transitivity problems regarding property inheritance (Gil-Berrozpe & Faber 2016).

For instance, at the first hyponymy level it was possible to find both a general concept such as METAMORPHIC ROCK and a specific concept such as ALPUJARRA DOLOMITE, which is incorrect because they have different granularity levels. Another example is LIMESTONE, originally represented as a hyponym of both ROCK and SEDIMENTARY ROCK at the same level. This is impossible in terms of transitivity because, in reality, LIMESTONE is a direct hyponym of SEDIMENTARY ROCK, but an indirect hyponym of ROCK.

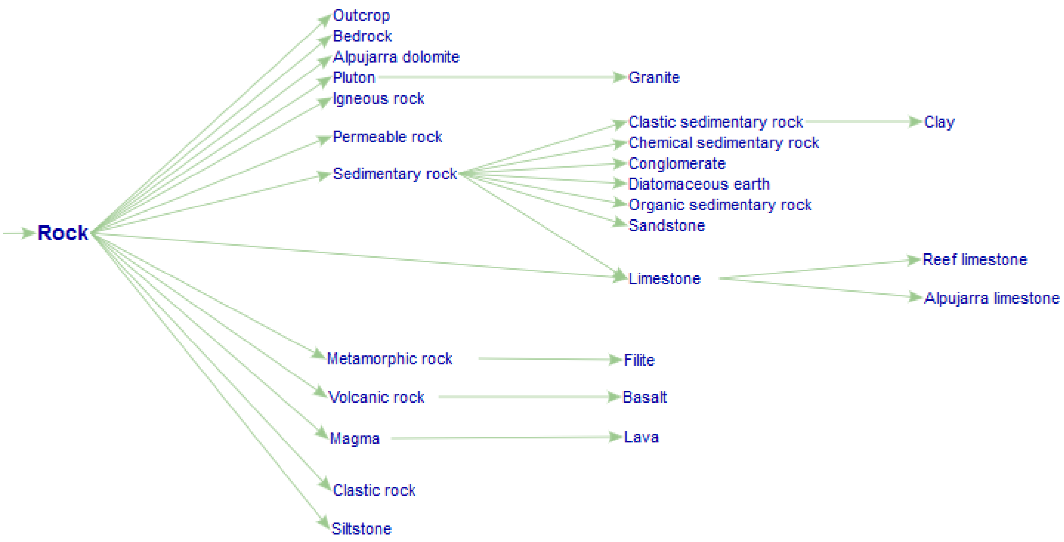


Figure 22. Problems in the conceptual hierarchies in EcoLexicon

To alleviate these problems, EcoLexicon allowed – and still allows – the filtering of conceptual systems according to a series of contextual domains (**Figure 23**). This feature is related to the microsenses typical of hypernym-hyponym pairs as pointed out by Cruse (1995, 2002). Therefore, it consists of representing concepts by means of conceptual propositions in the contextual discipline-based domains in which they are activated.

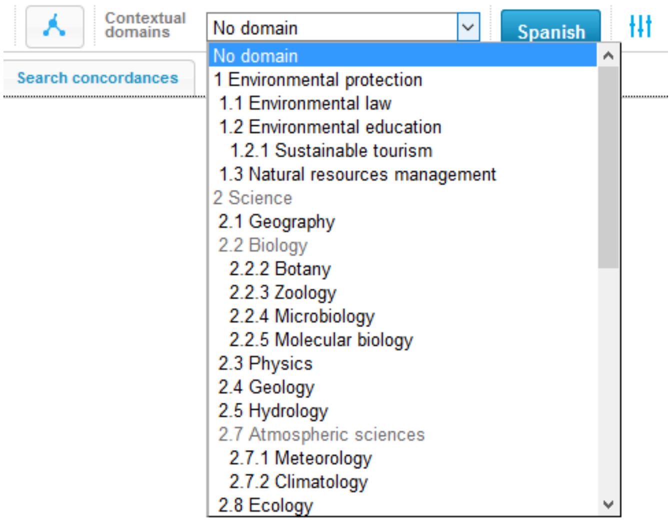


Figure 23. Contextual domains in EcoLexicon

For example, based on the information in EcoLexicon, CHLORINE is a concept with multiple perspectives (San Martín 2016). Therefore, it has two microsenses: a taxonomical microsense and a functional microsense. In all contextual domains, CHLORINE is only taxonomically classified as a *type of* HALOGEN (**Figure 24**). However, in the domains of Water Treatment and Supply, and Chemical Engineering, CHLORINE is also functionally categorized as a *type of* WATER DISINFECTANT, apart from being taxonomically considered a HALOGEN as well (**Figure 25**).

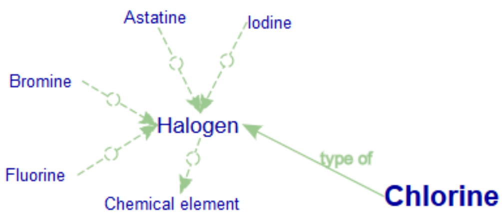


Figure 24. Taxonomical microsense of CHLORINE in EcoLexicon

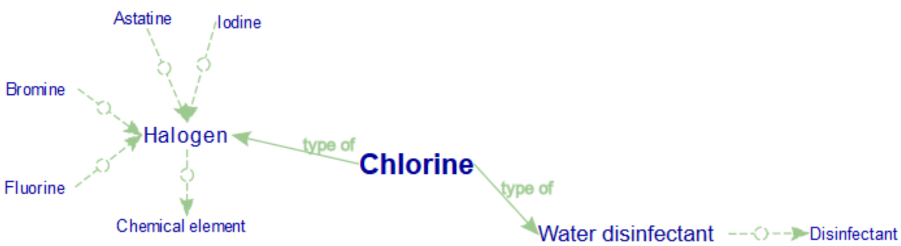


Figure 25. Taxonomical and functional microsenses of CHLORINE in EcoLexicon

However, with regard to hyponymy refinement, the classification of concepts in contextual domains only makes it possible to filter the query and show context-dependent hypernyms and hyponyms. In other words, the original sense of hyponymy remains the same and still needs to be decomposed in a certain way so as to guarantee a more accurate representation of generic-specific relations. For this reason, a pilot study on hyponymy refinement (Gil-Berrozpe 2016; Gil-Berrozpe *et al*. 2018) was carried out, based on the following criteria: (i) the correction of property inheritance according to concept definitions; (ii) the creation of umbrella concepts; and (iii) the decomposition of hyponymy into subtypes.

In the first place, accurate property inheritance was regarded as the initial step towards dividing the *type of* relation into subtypes because, before obtaining the final results, it was necessary for concept definitions to be both correct and coherent. In this way, it was possible to show how hyponyms inherited the features or traits of their respective hypernyms.

Table 5 shows an example of erroneous property inheritance in the original conceptual hierarchy from ROCK to BASALT. BASALT was defined as a "rock of igneous origin", but its hypernym (VOLCANIC ROCK) was also defined as an "igneous rock". Furthermore, the hypernym of VOLCANIC ROCK was assumed to be ROCK, regardless of the fact that the only types of rock mentioned in its definition were "igneous, sedimentary and metamorphic".

ROCK: consolidated or unconsolidated aggregate or mass of minerals or organic materials. The three types of rock are igneous, sedimentary, and metamorphic.		
	VOLCANIC ROCK: extrusive <u>igneous rock</u> solidified near or on the surface of the Earth, resulting from volcanic activity.	
		BASALT: very hard <u>rock of igneous origin</u>, consisting of augite and triclinic feldspar, with grains of magnetic or titanic iron, and also bottle-green particles of olivine. It is formed by decompression melting of the Earth's mantle.

Table 5. Original conceptual hierarchy of ROCK – BASALT (former definitions)

Table 6 then shows how property inheritance was improved in the new conceptual hierarchy. In this case, all senses were respected. As a result, BASALT is a *type of* VOLCANIC ROCK, which is a *type of* IGNEOUS ROCK, which is a *type of* SOLID ROCK, which is a *type of* ROCK. In other words, BASALT in the end reflects the inheritance of the characteristics possessed by all of its hypernyms.

ROCK: consolidated or unconsolidated aggregate or mass of minerals or organic materials.
SOLID ROCK: rock in solid state, formed by the compression of sediments or the solidification of molten material.
IGNEOUS ROCK: solid rock formed by solidification of molten magma either beneath or at the Earth's surface.
VOLCANIC ROCK: extrusive igneous rock solidified near or on the surface of the Earth, resulting from volcanic activity.
BASALT: very hard volcanic rock, consisting of augite and triclinic feldspar, with grains of magnetic or titanic iron, and also bottle-green particles of olivine. It is formed by decompression melting of the Earth's mantle.

Table 6. Enhanced conceptual hierarchy of ROCK – BASALT (new definitions)

This correction of property inheritance led to the creation and implementation of umbrella concepts in the conceptual hierarchies, which was the second step in the hyponymy refinement process. Umbrella concepts are artificial concepts which can be introduced at intermediate levels of a hierarchy to further specify the sense of the expressed hyponymic relation (Gil-Berrozpe & Faber 2017). Therefore, their main objective is to narrow the link that connects parent concepts to child concepts by implementing an intermediate abstract concept, often characterized by an essential adjective (e.g., MEASURING INSTRUMENT).

For instance, the original conceptual hierarchy of INSTRUMENT (**Table 7**) had a significant amount of information overload because of the large number of subordinates. More specifically, 68 hyponyms (e.g., ANEMOGRAPH, BAROMETER, SOLAR PANEL) were linked to the same superordinate (i.e., INSTRUMENT).

<table>
<tr><td>INSTRUMENT
(68 hyponyms)</td><td>ACCELEROMETER | AIR SAMPLER | ALBEDOGRAPH | ALBEDOMETER | ALTI-ELECTROGRAPH | ALTIMETER | ANEMOCLINOMETER | ANEMOGRAPH | ANEMOMETER | ANEROID CAPSULE | ATMORADIOGRAPH | AUTOMATIC SAMPLER | BAR SCREEN | BAROMETER | BATHYMETER | CLINOMETER | CLOUD CHAMBER | COMPASS | CREST GAGE | CTD | CURRENT METER | DEGREASER | DEPTH | FINDER | DREDGE | ECHO SOUNDER | EKMAN WATER BOOTLE | ELECTROSONDE | EMANOMETER | EVAPORIMETER | EVAPOTRANSPIROMETER | FLOWMETER | HYGROMETER | IMPEDOMETER | INCLINED GAUGE | INFILTROMETER | MARIGRAPH | METEOROGRAPH | PERMEAMETER | PHOTOMETER | PIEZOMETER | PLUVIOGRAPH | PLUVIOMETER | PSYCHROMETER | RADAR | SALINOMETER | SAND FILTER | SECCHI DISK | SEDIMENT SAMPLER | SEDIMENT TRAP | SEISMOGRAPH | SEISMOMETER | SEXTANT | SNOW GAUGE | SOLAR CELL | SOLAR PANEL | SOUNDING BALLOON | SOUNDING LEAD | STADIMETER | STAFF GAUGE | TENSIOMETER | THERMOMETER | THICKENER | TIDE STAFF | VENTILATED THERMOMETER | WATER SAMPLER | WATER-LEVEL RECORDER | WATER-TREATMENT PLANT</td></tr>
</table>

Table 7. Original conceptual hierarchy of INSTRUMENT (without umbrella concepts)

For this reason, and because of the semantics of the concept INSTRUMENT, a set of five umbrella concepts specifying a functional hyponymic relation (i.e., FILTERING INSTRUMENT, MEASURING INSTRUMENT, RECORDING INSTRUMENT, SAMPLING INSTRUMENT, TRANSFORMING INSTRUMENT) was introduced to provide a more accurate classification of the hyponyms (**Table 8**).

<table>
<tr><td rowspan="5">INSTRUMENT
(5 hyponyms)</td><td>FILTERING INSTRUMENT</td><td>BAR SCREEN | DEGREASER | SAND FILTER | SIEVE</td></tr>
<tr><td>MEASURING INSTRUMENT</td><td>ALBEDOGRAPH | ALTI-ELECTROGRAPH | ANEMOGRAPH | ATMORADIOGRAPH | COMPASS | MARIGRAPH | METEOROGRAPH | PLUVIOGRAPH | RADAR | SEISMOGRAPH | WATER-LEVEL RECORDER</td></tr>
<tr><td>RECORDING INSTRUMENT</td><td>ACCELEROMETER | ALBEDOMETER | ALTIMETER | ANEMOCLINOMETER | ANEMOMETER | ANEROID CAPSULE | BAROMETER | BATHYMETER | CLINOMETER | CREST GAGE | CTD | CURRENT METER | DEPTH FINDER | ECHO SOUNDER | ELECTROSONDE | EMANOMETER | EVAPORIMETER | EVAPOTRANSPIROMETER | FLOWMETER | HYGROMETER | IMPEDOMETER | INCLINED GAUGE | INFILTROMETER | PERMEAMETER | PHOTOMETER | PIEZOMETER | PLUVIOMETER | PSYCHROMETER | SALINOMETER | SECCHI DISK | SEDIMENT TRAP | SEISMOMETER | SEXTANT | SNOW GAUGE | SOUNDING BALLOON | SOUNDING LEAD | STADIMETER | STAFF GAUGE | TENSIOMETER | THERMOMETER | TIDE STAFF | VENTILATED THERMOMETER</td></tr>
<tr><td>SAMPLING INSTRUMENT</td><td>AIR SAMPLER | AUTOMATIC SAMPLER | DREDGE | EKMAN WATER BOTTLE | SEDIMENT SAMPLER | WATER SAMPLER</td></tr>
<tr><td>TRANSFORMING INSTRUMENT</td><td>CLOUD CHAMBER | SOLAR CELL | SOLAR PANEL | THICKENER | WATER-TREATMENT PLANT</td></tr>
</table>

Table 8. Enhanced conceptual hierarchy of INSTRUMENT (with umbrella concepts)

After correcting property inheritance and enriching the hierarchies with new concepts, a fine-grained set of hyponymy subtypes was specified as the final step in the hyponymy refinement process.

What can be inferred from this process is that terminological definitions that accurately reflect property inheritance and conceptual hierarchies that show the links between hypernyms and hyponyms are key to the identification of hyponymy subtypes. These steps are thus a fundamental part of our methodology.

3.2.3.1. HYPONYMY SUBTYPES

The results of the pilot study by Gil-Berrozpe (2016) indicated that hyponymy subtypes were constrained by the nature of the concept, namely whether it was an entity (i.e., ROCK) or a process (i.e., EROSION). As a result, the following five entity-related hyponymy subtypes were specified according to the dimensions triggered by each entity: (i) *state-based hyponymy*, depending on the state of matter of the hyponyms (e.g., SOLID ROCK); (ii) *formation-based hyponymy*, depending on the formation process of the hyponyms (e.g., SEDIMENTARY ROCK); (iii) *composition-based hyponymy*, depending on the components or the constituents of the hyponyms (e.g., SILTSTONE); (iv) *location-based hyponymy*, depending on the physical situation or location of the hyponyms (e.g., PLUTONIC ROCK); and (v) *attribute-based hyponymy*, depending on the traits or features of the hyponyms (e.g., PERMEABLE ROCK).

In addition, four process-related hyponymy subtypes were established in relation to the characteristics of each process: (i) *agent-based hyponymy*, depending on the agent or the promoter that causes the hyponyms (e.g., SEA EROSION); (ii) *patient-based hyponymy*, depending on the entity or location affected by the hyponyms (e.g., CHANNEL SCOUR); (iii) *result-based hyponymy*, depending on the results and effects of the hyponyms (e.g., GULLY EROSION); and (iv) *attribute-based hyponymy*, depending on the traits or features of the hyponyms (e.g., POTENTIAL EROSION).

Therefore, Gil-Berrozpe (2016) provided insights into hyponymy refinement. Nonetheless, corpus analysis was only used to expand the hierarchies of a single entity and a single process. For this reason, it was necessary to focus on the linguistic markers for hyponymy and for a wider series of conceptual relations.

For this reason, a more comprehensive corpus-based study (Gil-Berrozpe 2017; Gil-Berrozpe *et al*. 2017) was performed so as to identify all hyponymy subtypes and hyponymic KPs in EcoLexicon by analyzing a randomized portion of the EcoLexicon

English Corpus. This corpus analysis explored the correlation of concepts in a variety of different categories with KPs as well as with hyponymy subtypes. These constraints led to a more comprehensive inventory of generic-specific relations in the environmental domain, as well as to a more accurate way of extracting them.

The decomposition of the generic-specific relation was based on common features in the cases analyzed. This led to the identification of the following 32 hyponymy subtypes for environmental concepts: *ability-based* hyponymy, *activity-based* hyponymy, *agent-based* hyponymy, *amount-based* hyponymy, *color-based* hyponymy, *composition-based* hyponymy, *degree-based* hyponymy, *denomination-based* hyponymy, *density-based* hyponymy, *domain-based* hyponymy, *effect-based* hyponymy, *function-based* hyponymy, *hardness-based* hyponymy, *height-based* hyponymy, *location-based* hyponymy, *method-based* hyponymy, *moisture-based* hyponymy, *movement-based* hyponymy, *origin-based* hyponymy, *patient-based* hyponymy, *relation-based* hyponymy, *result-based* hyponymy, *shape-based* hyponymy, *size-based* hyponymy, *speed-based* hyponymy, *state-based* hyponymy, *status-based* hyponymy, *technology-based* hyponymy, *temperature-based* hyponymy, *texture-based* hyponymy, *time-based* hyponymy, and *weight-based* hyponymy. See **Appendix II** for a full description and examples of the inventory of hyponymy subtypes in the environmental domain.

The analysis of hyponymic relations carried out by Gil-Berrozpe (2017) showed that certain subtypes (e.g., *agent-based*, *patient-based*, *result-based*, *method-based*, and *degree-based* hyponymy) closely correlated with process-related conceptual categories (e.g., *activity*, *phenomenon*, *process*, and *change of state*). In contrast, other hyponymy subtypes (e.g., *composition-based*, *technology-based*, and *function-based* hyponymy) were directly linked to entity-related conceptual categories (e.g., *substance*, *landform*, *construction*, and *instrument*). Furthermore, the results also demonstrated that a distinction can be made between relational hyponymy subtypes (i.e., those involving another concept, such as *agent-based*, *result-based*, and *location-based* hyponymy) and attributional hyponymy subtypes (i.e., those involving the intrinsic characteristics of the concept, such as *shape-based*, *texture-based*, and *moisture-based* hyponymy).

Relational hyponymy subtypes were generally associated with processes, whilst attributional hyponymy subtypes were mainly related to entities. In addition, further distinctions were made not only depending on whether concepts were entities or processes, but also between natural and artificial concepts. In this line, some hyponymy subtypes were shown to be mostly exclusive to natural concepts (e.g., *origin-based*, *state-based*, and *time-based* hyponymy), whereas other hyponymy subtypes were generally attributed to artificial concepts (e.g., *function-based*, *technology-based*, and *weight-based* hyponymy).

Not only were these hyponymy subtypes used to enhance the representation of hyponymy representation in EcoLexicon, but they were also the typology of subtypes applied to categorize the generic-specific relations analyzed in the corpus study of this research.

3.2.3.2. HYPONYMIC KNOWLEDGE PATTERNS

As previously explained, one of the applications of Corpus Linguistics to Terminology is the use of KP-based customized corpus query language (CQL) search expressions so as to obtain KRCs with valid terminological content for specific purposes (Condamines 2002; Bowker 2003; Barrière 2004b; L'Homme & Marshman 2006; Mortchev-Bouveret 2006; Auger & Barrière 2008; Gödert *et al*. 2014; Lafourcade & Ramadier 2016; Gil-Berrozpe 2017; Lefeuvre *et al*. 2017; Rojas-García & Cabezas-García 2019).

In particular, corpus-based analysis and KPs have become a major research topic as a method of automatically or semi-automatically extracting linguistic information concerning different conceptual relations. The most commonly studied patterns are hyponymic KPs (Hearst 1992, 1998; Pearson 1998; Liu *et al*. 2006; Pantel & Pennacchiotti 2006; Bielinskiene *et al*. 2012; Nazar *et al*. 2012; Lefever *et al*. 2014; Li *et al*. 2014; Baisa & Suchomel 2015; Gil-Berrozpe *et al*. 2017; Faralli *et al*. 2018; Lewis 2019), because of their importance for categorization and property inheritance.

Examples of hyponymic KPs are *comprise(s)*, *consist(s)*, *define(s)*, *denote(s)*, *designate(s)*, *is/are*, *is/are called*, *is/are defined as*, and *is/are known as* (Pearson 1998). Moreover, these hyponymic KPs identify what are known as *definitional KRCs* (Meyer 2001), *expository definitional contexts* (Pearson 1998), and *definitional contexts* (Malaisé *et al*. 2005; Sierra *et al*. 2008). Interestingly, definitional contexts tend to have the same components as a terminological definition (i.e., *genus* and *differentiae*).

For instance, Hearst (1998) proposed the automatic extraction of hyponymic KPs from texts by looking at sentences that contain hypernym-hyponym pairs in WordNet. In this way, six patterns in English (including simplified examples like *X such as Y* or *X and other Y*) were identified by observation and by analyzing the context between a hyponym and its superordinate in the corpus. However, these hyponymic

KPs generated a large number of mistakes, either because the extracted relation was too far away, because there were subjective opinions of no interest, or because of parsing errors.

Another example is Bielinskiene *et al*. (2012), who proposed a series of 18 definitional patterns as a means of obtaining hypernym-hyponym pairs in Lithuanian. These patterns included prototypical examples such as *X is Y* or *X is considered as Y*. However, their proposal also included polysemic KPs such as *X constitutes Y* or *X includes Y*, which can also be used to find meronymic KRCs. Unfortunately, their results showed that 55% of all cases (127 contexts from 227) were irrelevant because of grammar, the absence of hyponymy, and other inaccuracies related to the expressions.

Furthermore, León-Araúz *et al*. (2016) went one step further by creating a total of 56 sketch grammars for the Sketch Engine corpus analysis software to automatically extract semantic information from large corpora: 18 generic-specific grammars, 17 part-whole grammars, ten cause grammars, seven function grammars, and four location grammars. These sketch grammars are dynamic and have different permutations or variations so as to encompass all the different aspects that can involve every relation independently, and avoid possible problems such as noise or loops. As explained by León-Araúz *et al*. (2016: 76):

> In the development of our sketch grammars (a total of 56), we […] considered different issues that are specific to each relation. For instance, there are certain patterns that always take the same form and order (e.g. such as), whereas others show such a diverse syntactic structure that the directionality of the pattern must also be accounted for. We also had to take into account the fact that a single sentence could produce more than one term pair because of the enumerations that are often found on each side of the pattern (e.g. x, y, z and other types of w). This entails performing non-greedy queries in order to allow any of the enumerated elements fill the target term. However, this may also cause endless noisy loops. Sometimes it is necessary to limit the number of possible words on each side of the pattern. In this sense, we observed that enumerations are more often found on the side of hyponyms, parts, and effects than on the side of hypernyms, wholes, and causes. Consequently, the loops were constrained accordingly in the latter case.

In this line, hyponymic KPs have been extensively analyzed and used to obtain information about specialized knowledge. KP-based queries in the form of micro-grammars (León-Araúz *et al*. 2016; León-Araúz & San Martín 2018) are able to extract concordance lines for a term and analyze its combinatorial potential. León Araúz *et al*.'s (2016) sketch grammar was eventually made available as the EcoLexicon Semantic Sketch Grammar (León-Araúz & San Martín 2018). **Table 9** depicts a summarized and simplified version of the hyponymic KPs in the customized sketch grammar (León-Araúz *et al*. 2016).

1. HYPONYM, ǀ (ǀ : ǀ is ǀ belongs (to) (a ǀ the ǀ …) type ǀ category ǀ … of HYPERNYM // **2.** types ǀ kinds ǀ … of HYPERNYM include ǀ are HYPONYM // **3.** types ǀ kinds ǀ … of HYPERNYM range from (…) (to) HYPONYM // **4.** HYPERNYM (type ǀ category ǀ …) (, ǀ () ranging (…) (to) HYPONYM // **5.** HYPERNYM types ǀ categories ǀ … include HYPONYM // **6.** HYPERNYM such as HYPONYM // **7.** HYPERNYM including HYPONYM // **8.** HYPERNYM, ǀ (especially ǀ primarily ǀ … HYPONYM // **9.** HYPONYM and ǀ or other (types ǀ kinds ǀ …) of HYPERNYM // **10.** HYPONYM is defined ǀ classified ǀ … as (a ǀ the ǀ …) (type ǀ kind ǀ …) (of) HYPERNYM // **11.** classify ǀ categorize ǀ … (this type ǀ kind ǀ … of) HYPONYM as HYPERNYM // **12.** HYPERNYM is classified ǀ categorized in ǀ into (a ǀ the ǀ …) (type ǀ kind ǀ …) (of) HYPONYM // **13.** HYPERNYM (, ǀ () (is) divided in ǀ into (…) types ǀ kinds ǀ … : ǀ of HYPONYM // **14.** type ǀ kind ǀ … of HYPERNYM (is ǀ , ǀ () known ǀ referred ǀ … (to) (as) HYPONYM // **15.** HYPONYM is a HYPERNYM that ǀ which ǀ … // **16.** define HYPONYM as (a ǀ the ǀ …) (type ǀ category ǀ …) (of) HYPERNYM // **17.** HYPONYM refers to (a ǀ the ǀ …) (type ǀ category ǀ …) (of) HYPERNYM // **18.** (a ǀ the ǀ one ǀ two…) (type ǀ category ǀ …) (of) HYPERNYM: HYPONYM

Table 9. Hyponymic KPs in the customized sketch grammar (León-Araúz *et al.* 2016)

Table 10 shows the actual CQL representation of a hyponymic KP rule, followed by an explanation and three natural language examples of concordances matched with the grammar (León-Araúz & San Martín 2018).

<table>
<tr><td colspan="2">1:"N.*"[word=",|\("]?[tag="IN/that|WDT"]?"MD"*[lemma="be|,|\("]"RB.*"* [word="classified|categori.ed"]([word="by"][tag!="V.*"]+)?[word="in|into"][tag!="V.*"]* [lemma="type|kind|example|group|class|sort|category|family|species|subtype|subfamily| subgroup|subclass|subcategory|subspecies"]?[tag!="V.*"]*2:[tag="N.*"&lemma!="type|kind| example|group|class|sort|category|family|species|subtype|subfamily|subgroup|subclass| subcategory|subspecies"]</td></tr>
<tr><td>1:"N.*"</td><td>The hypernym is a noun.</td></tr>
<tr><td>[word=",|\("]?</td><td>An optional comma or bracket.</td></tr>
<tr><td>[tag="IN/that|WDT"]?</td><td>Optionally “that” or “which”.</td></tr>
<tr><td>"MD"*</td><td>Any modal verb from zero to infinite times.</td></tr>
<tr><td>[lemma="be|,|\("]</td><td>Lemma “be” or a comma or a bracket.</td></tr>
<tr><td>"RB.*"*</td><td>Any adverb from zero to infinite times.</td></tr>
<tr><td>[word="classified|categori.ed"]</td><td>Classified, categorized, or categorized.</td></tr>
<tr><td>([word="by"] [tag!="V.*"]+)?</td><td>Optionally, “by” followed by anything from one to infinite times that does not contain a verb.</td></tr>
<tr><td>[word="in|into"]</td><td>In or into.</td></tr>
<tr><td>[tag!="V.*"]*</td><td>Anything from zero to infinite times that does not contain a verb.</td></tr>
<tr><td>[lemma="type|kind|example|group|class| sort|category|family|species|subtype| subfamily|subgroup|subclass|subcategory| subspecies"]?</td><td>Optionally any of the lemmas “type”, “kind”, “example”, “group”, “class”, “sort”, “family”, etc.</td></tr>
<tr><td>[tag!="V.*"]*</td><td>Anything from zero to infinite times that does not contain a verb.</td></tr>
<tr><td>2:[tag="N.*"lemma!="type|kind|example| group|class|sort|category|family|species| subtype|subfamily|subgroup|subclass| subcategory|subspecies"]</td><td>The hyponym is any noun other than “type”, “kind”, “example”, “group”, “class”, “sort”, “family”, etc.</td></tr>
<tr><td colspan="2">Stony-iron meteorites are classified into pallasites and mesosiderites.
Modern reefs are classified into several geomorphic types: atoll, barrier, fringing, and patch.
Littoral materials are classified by grain size in clay, silt, sand, gravel, cobble, and boulder.</td></tr>
</table>

Table 10. CQL representation of a hyponymic KP rule (León-Araúz & San Martín 2018)

The corpus-based study to identify hyponymy subtypes and hyponymic KPs in EcoLexicon (Gil-Berrozpe 2017; Gil-Berrozpe *et al*. 2017) was based on the application of León-Araúz *et al*.'s (2016) sketch grammars for the automatic extraction of hyponymic information. A randomized portion of the EcoLexicon English Corpus was examined, from which 3,133 positive hyponymic concordances were selected for KP analysis with a view to identifying common patterns in order to categorize them.

A total of 125 hyponymic KPs were retrieved, and those expressing hyponymy in a similar way were classified in the same category. The result was the following ten categories of hyponymic KPs: *classification*, *definition*, *denomination*, *enumeration*, *exemplification*, *identification*, *inclusion*, *itemization*, *range*, and *selection* hyponymic KPs. See **Appendix III** for a full description and listing of the inventory of hyponymic KPs in the EcoLexicon English Corpus.

The analysis of hyponymic KPs in Gil-Berrozpe (2017) showed that certain KPs (e.g., *exemplification*, *selection*, *itemization*, and *inclusion* hyponymic KPs) were linked to conceptual categories that are the basis of scientific classifications (e.g., *lifeform* and *chemical element*). Furthermore, other KPs (e.g., *identification*, *denomination*, and *definition* hyponymic KPs) showed a more explanatory structure, and were thus most frequently linked to more complex conceptual categories involving various participants (e.g., *phenomenon*, *process*, and *technology*). Therefore, they required a more detailed description and/or explanation. *Range* hyponymic KPs were mostly associated with *time period* and *measure* since these categories are generally composed of values that are characterized by the distance between them in terms of time, space, and intensity.

In our research, hyponymic KPs were also taken into account, though with a lower level of relevance than conceptual categories or hyponymy subtypes. For this reason, their usefulness was reduced to the expression of hyponymic contexts for certain terms depending on their availability in the corpus analysis.

3.3. REPRESENTATION OF HYPONYMY

This section reviews the representation of hyponymy in terms of its visualization and expression in various terminological resources. As previously mentioned, hyponymy is an extremely complex conceptual relation not only because of its importance in terms of transitivity or property inheritance between different concepts, but also because it classifies real-world entitles in hierarchies or taxonomies. Hyponymy is thus best reflected in digital resources.

Apart from the term itself, another essential element in any terminological resource is the definition. As the natural language explanation of the location of a concept in the conceptual structure of the specialized domain (Faber 2022), definitions not only specify the properties of concepts, but also link them to other realities (Antia 2000). Since the most basic way of associating concepts is by alluding to their hypernym, hyponymy is always present in all terminological resources with intensional definitions. However, this indirect way of representing hyponymy does not fully exploit all its possibilities since it does not reflect all its complexity.

3.3.1. HYPONYMY IN TRADITIONAL RESOURCES

This section explains how hyponymy is represented in more traditional terminological resources, regardless of whether the resources are physical or electronic.

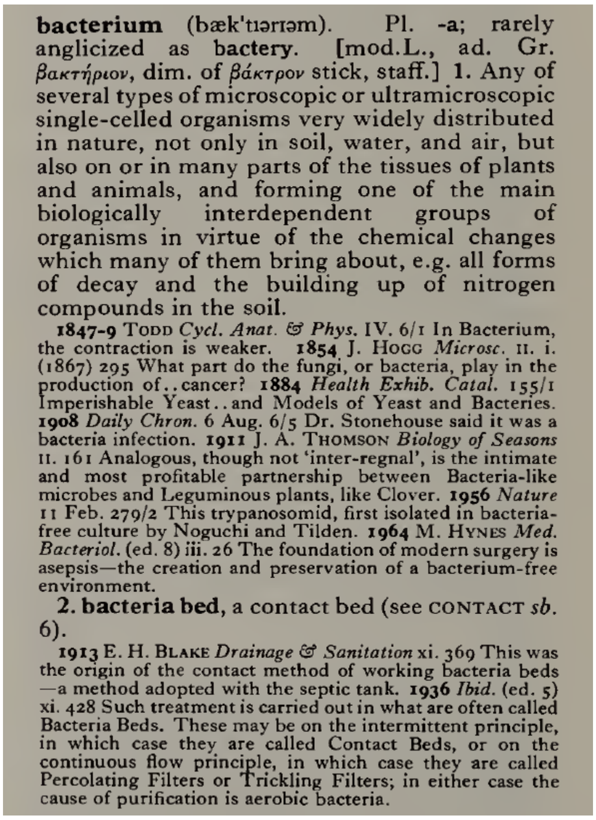
3.3.1.1. HYPONYMY IN DICTIONARIES

Dictionaries are lexicological or terminological resources (i.e., depending on their specialization) that list lexemes from the vocabulary or terminology of a language, or more languages in the case of bilingual and multilingual dictionaries. They often arrange this information alphabetically and include data regarding definitions, usage contexts, etymologies, pronunciations, and other elements. One of the most common classifications is the distinction between general language dictionaries and specialized language dictionaries, but here both types are reviewed from the perspective of terminology work. The emphasis is thus on terms instead of words.

General language dictionaries also include terms, but with very concise definitions. A good example of a general language dictionary is the *Oxford Dictionary of English* (Oxford University Press 2010). Since the entries in this and similar dictionaries focus more on meanings, various definitions are displayed. The linguistic data in these resources include usage contexts, etymological (i.e., origin or historical development of the term), phonetic (i.e., pronunciation of the term), and collocational information (e.g., compound nouns with the term acting as subject). Furthermore, electronic versions of general dictionaries, such as *OED Online* (Oxford University Press 2022), may even include multimedia information (e.g., audio files to check the pronunciation of the term according to different diatopic variants).

Figure 26 shows the terminological entry for BACTERIUM in the *Oxford Dictionary of English* (Oxford University Press 2010). This entry includes the following linguistic information: (i) singular form; (ii) pronunciation (using phonemes); (iii) plural ending; (iv) an alternative anglicized term; and (v) etymology. In addition, there are two intensional definitions of the term. The first refers to BACTERIUM itself (i.e., a microscopic single-celled organism) and the second to BACTERIA BED (i.e., a contact bed). Diachronic information is also provided to identify the time period when the term was used as well as in which type of text.

However, there is no mention of hyponymy or any other conceptual relation. Neither is there any information regarding complex nominals or compound nouns whose head is BACTERIA. This is another way to codify hyponymy by specifying characteristic attributes with hyponymic nuances (Gil-Berrozpe 2020). Therefore, the only way to identify hyponymy in this type of resource is through the intensional definition displayed, which alludes to the hypernym of the concept (e.g., BACTERIUM – ORGANISM).



bacterium (bæk'tɪərɪəm). Pl. **-a**; rarely anglicized as **bactery**. [mod.L., ad. Gr. *βακτήριον*, dim. of *βάκτρον* stick, staff.] **1.** Any of several types of microscopic or ultramicroscopic single-celled organisms very widely distributed in nature, not only in soil, water, and air, but also on or in many parts of the tissues of plants and animals, and forming one of the main biologically interdependent groups of organisms in virtue of the chemical changes which many of them bring about, e.g. all forms of decay and the building up of nitrogen compounds in the soil.

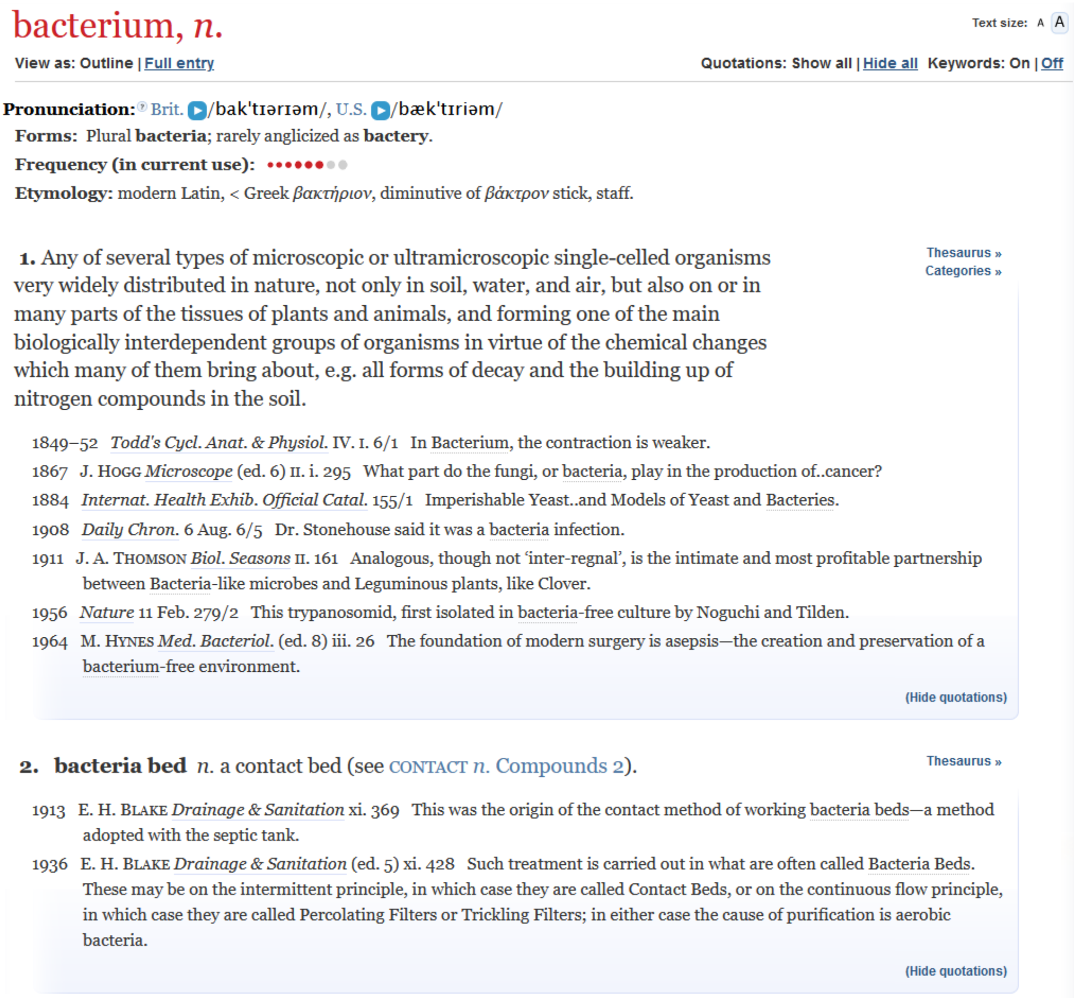
1847-9 TODD *Cycl. Anat. & Phys.* IV. 6/1 In Bacterium, the contraction is weaker. **1854** J. HOGG *Microsc.* II. i. (1867) 295 What part do the fungi, or bacteria, play in the production of..cancer? **1884** *Health Exhib. Catal.* 155/1 Imperishable Yeast..and Models of Yeast and Bacteries. **1908** *Daily Chron.* 6 Aug. 6/5 Dr. Stonehouse said it was a bacteria infection. **1911** J. A. THOMSON *Biology of Seasons* II. 161 Analogous, though not 'inter-regnal', is the intimate and most profitable partnership between Bacteria-like microbes and Leguminous plants, like Clover. **1956** *Nature* 11 Feb. 279/2 This trypanosomid, first isolated in bacteria-free culture by Noguchi and Tilden. **1964** M. HYNES *Med. Bacteriol.* (ed. 8) iii. 26 The foundation of modern surgery is asepsis—the creation and preservation of a bacterium-free environment.

2. bacteria bed, a contact bed (see CONTACT *sb.* 6).

1913 E. H. BLAKE *Drainage & Sanitation* xi. 369 This was the origin of the contact method of working bacteria beds —a method adopted with the septic tank. **1936** *Ibid.* (ed. 5) xi. 428 Such treatment is carried out in what are often called Bacteria Beds. These may be on the intermittent principle, in which case they are called Contact Beds, or on the continuous flow principle, in which case they are called Percolating Filters or Trickling Filters; in either case the cause of purification is aerobic bacteria.

Figure 26. Terminological entry in the *Oxford Dictionary of English* (Oxford University Press 2010)

Figure 27 shows the terminological entry for BACTERIUM in *OED Online* (Oxford University Press 2022), the digital version of the *Oxford Dictionary of English.* This entry includes the same information as the paper version of the dictionary, but its more intuitive structure facilitates knowledge acquisition by structuring the data to allow user interactivity. The entry is also enriched with the following information: (i) the British English and American English pronunciation in audio files, (ii) the frequency of the term; (iii) hyperlinks to access other entries; and (iv) thesaurus access. Apart from that, the terminological entry is exactly the same and, hence, there is no hyponymic representation or any specification of the hyponyms of BACTERIUM.


bacterium, *n.*

Text size: A A

View as: Outline | Full entry

Quotations: Show all | Hide all Keywords: On | Off

Pronunciation: Brit. /bak'tɪərɪəm/, U.S. /bæk'tɪriəm/

Forms: Plural **bacteria**; rarely anglicized as **bactery**.

Frequency (in current use):

Etymology: modern Latin, < Greek *βακτήριον*, diminutive of *βάκτρον* stick, staff.

1. Any of several types of microscopic or ultramicroscopic single-celled organisms very widely distributed in nature, not only in soil, water, and air, but also on or in many parts of the tissues of plants and animals, and forming one of the main biologically interdependent groups of organisms in virtue of the chemical changes which many of them bring about, e.g. all forms of decay and the building up of nitrogen compounds in the soil.

Thesaurus »
Categories »

1849–52 *Todd's Cycl. Anat. & Physiol.* IV. I. 6/1 In Bacterium, the contraction is weaker.

1867 J. Hogg *Microscope* (ed. 6) II. i. 295 What part do the fungi, or bacteria, play in the production of..cancer?

1884 *Internat. Health Exhib. Official Catal.* 155/1 Imperishable Yeast..and Models of Yeast and Bacteries.

1908 *Daily Chron.* 6 Aug. 6/5 Dr. Stonehouse said it was a bacteria infection.

1911 J. A. Thomson *Biol. Seasons* II. 161 Analogous, though not 'inter-regnal', is the intimate and most profitable partnership between Bacteria-like microbes and Leguminous plants, like Clover.

1956 *Nature* 11 Feb. 279/2 This trypanosomid, first isolated in bacteria-free culture by Noguchi and Tilden.

1964 M. Hynes *Med. Bacteriol.* (ed. 8) iii. 26 The foundation of modern surgery is asepsis—the creation and preservation of a bacterium-free environment.

(Hide quotations)

2. **bacteria bed** *n.* a contact bed (see CONTACT *n.* Compounds 2).

Thesaurus »

1913 E. H. Blake *Drainage & Sanitation* xi. 369 This was the origin of the contact method of working bacteria beds—a method adopted with the septic tank.

1936 E. H. Blake *Drainage & Sanitation* (ed. 5) xi. 428 Such treatment is carried out in what are often called Bacteria Beds. These may be on the intermittent principle, in which case they are called Contact Beds, or on the continuous flow principle, in which case they are called Percolating Filters or Trickling Filters; in either case the cause of purification is aerobic bacteria.

(Hide quotations)

Figure 27. Terminological entry in *OED Online*
(Oxford University Press 2022)

However, specialized language dictionaries, such as *A Dictionary of Biology* (Hine 2019), *The Dictionary of Cell and Molecular Biology* (Lackie 2013), and the *Henderson's Dictionary of Biology* (Lawrence 2008), provide less linguistic information and focus on more detailed definitions of the terms. In contrast to general dictionaries, each entry has a single definition. Not surprisingly, the amount of specialized knowledge in these resources is greater than in general language dictionaries, and thus contain a wider range of more specific terms. Interestingly, there are certain specialized language dictionaries, such as the *Dictionary of Geology and Mineralogy* (McGraw-Hill 2003b), which also specify the subdomain (e.g., Geophysics, Paleontology, Mineralogy) to which the term belongs. This makes it possible to delimit the concept and to better differentiate its microsenses in relation to contextual domains.

Figure 28 shows a frequent type of definition in dictionaries. More specifically, this entry in *A Dictionary of Biology* (Hine 2019) has a partitive or collective definition that defines the concept in terms of its membership in a broader category. This means that the user is redirected to the more general entry that helps him/her to understand the meaning of the more specific one.

bacteriostatic Capable of inhibiting or slowing down the growth and reproduction of bacteria. Some *antibiotics are bacteriostatic. *Compare* BACTERICIDAL.

bacterium (*pl.* **bacteria**) A member of the domain *Bacteria.

bacteroid The form adopted by a nitrogen-fixing bacterium when active within the root nodule of a host plant. For example, *Rhizobium* bacteria change into enlarged irregularly shaped branching cells when they infect the root cells of their legume host. These bacteroids become surrounded by a **peribacteroid membrane**, derived from membranes of their host cells, and differentiate to produce the key enzymes and other components of nitrogen fixation, such as *nitrogenase. As the infection process progresses, the bacteroids become housed in a *root nodule, where they are totally dependent on their host for the energy required for nitrogen. In return they supply their host with an assimilable form of nitrogen, i.e. ammonia, which is incorporated into amino acids. *See* NITROGEN FIXATION.

Figure 28. Partitive definition in an entry in *A Dictionary of Biology* (Hine 2019)

Figure 29 shows a terminological entry in *A Dictionary of Biology* (Hine 2019), where BACTERIA is explained not only with a definition, but also with additional sentences and paragraphs that describe related characteristics. Also specified is its relation to other concepts (e.g., MOLECULAR SYSTEMATICS, AEROBIC RESPIRATION, COCCUS, BACILLUS, NITROGEN CYCLE).

When these extra concepts are in the dictionary, they appear in small capitals so that users know in which entry they will be able to obtain more information about each concept. Given the length and detail of these explanations, specialized language dictionaries seem to be midway between general language dictionaries and encyclopedias.

Although there is no explicit allusion to hyponymy, it is in the intensional definition, as well as in the extensional explanation that alludes to other interrelated concepts. Accordingly, some of them appear as hyponyms to the entry term (e.g., COCCUS, BACILLUS, SPIRILLUM, VIBRIO, SPIROCHAETE, CYANOBACTERIA). In any case, these resources do not aspire to provide any representation of conceptual relations.

Bacteria **(Eubacteria)** A domain of life containing a diverse group of ubiquitous microorganisms all of which consist of only a single *cell that lacks a distinct nuclear membrane and has a *cell wall of a unique composition (see illustration). Bacteria constitute the prokaryotic organisms of the living world. However, their classification is a controversial issue. It is now recognized, on the basis of differences in ribosomal RNA structure and nucleotide sequences (*see* MOLECULAR SYSTEMATICS), that prokaryotes form two evolutionarily distinct domains: *Archaea (the archaea) and Bacteria. Defining characteristics of bacteria include the possession of cell walls containing peptidoglycan, and membrane lipids containing fatty acids in ester linkage to glycerol, whereas archaea lack peptidoglycan and have ether-linked lipids. However, in general parlance, the term 'bacteria' can still, erroneously, encompass both archaea and bacteria.

Bacteria can be characterized in a number of ways, for example by their reaction with *Gram's stain, their *GC content, or on the basis of their metabolic requirements (e.g. whether or not they require oxygen: *see* AEROBIC RESPIRATION; ANAEROBIC RESPIRATION) and shape. A bacterial cell may be spherical (*see* COCCUS), rodlike (*see* BACILLUS), spiral (*see* SPIRILLUM), comma-shaped (*see* VIBRIO), corkscrew-shaped (*see* SPIROCHAETE), or filamentous, resembling a fungal cell. The majority of bacteria range in size from 0.5 to 5 μm. Many are motile, bearing *flagella, possess an outer slimy *capsule, and produce resistant spores (*see* ENDOSPORE). In general bacteria reproduce only asexually, by simple division of cells, but a few groups undergo a form of sexual reproduction (*see* CONJUGATION) and *lateral gene transfer is common. Bacteria are largely responsible for decay and decomposition of organic matter, producing a cycling of such chemicals as carbon (*see* CARBON CYCLE), oxygen, nitrogen (*see* NITROGEN CYCLE), and sulphur (*see* SULPHUR CYCLE). A few bacteria obtain their food by means of *photosynthesis, including the *Cyanobacteria; some are saprotrophs; and others are parasites, causing disease. The symptoms of bacterial infections are produced by *toxins.

Figure 29. Terminological entry in *A Dictionary of Biology* (Hine 2019)

Figure 30 shows a specialized dictionary entry in *The Dictionary of Cell and Molecular Biology* (Lackie 2013). Its definition of BACTERIA also explains the concept itself with no mention of its linguistic features. Since the definition is much narrower, it is evident that not all specialized language dictionaries have the same type of specificity. In fact, this entry refers to fewer concepts related to the main entry term in bold type.

bacteraemia *US bacteremia* The presence of living bacteria in the circulating blood: usually implies the presence of small numbers of bacteria that are transiently present without causing clinical effects, in contrast to **septicaemia**.

Bacteria *formerly* Eubacteria One of the two major subdivisions of the prokaryotes that includes most Gram-positive bacteria, cyanobacteria, proteobacteria and mycoplasmas. Unlike the **Archaea** (*formerly* Archaebacteria) they have ester-linked lipids in the cytoplasmic membrane, **peptidoglycan** in the cell wall, and do not have **introns**. Bacteria are small with linear dimensions around 1 μm, do not have internal compartments, have circular DNA, and ribosomes of 70 S. Protein synthesis differs from that of eukaryotes so that it is possible to use anti-bacterial antibiotics that interfere with protein synthesis without affecting the eukaryotic host.

bacterial artificial chromosome *BAC* A vector used to construct a **genomic library** that has the sites necessary for the DNA to be handled and replicated as a bacterial chromosome. Like **YACs**, this allows clones to contain very large pieces of DNA (around 200 kb), so aiding rapid, low resolution **physical mapping.**

Figure 30. Terminological entry in *The Dictionary of Cell and Molecular Biology* (Lackie 2013)

Figure 31 shows entries from *Henderson's Dictionary of Biology* (Lawrence 2008), in which the definitions are more concise and synthetic without information regarding related concepts.

bacteriostatic *a.* inhibiting the growth of, but not killing, bacteria.
bacterium *n.* individual member of the Bacteria. The term is also in general use for any prokaryotic microorganism, including members of the Archaea. *plu.* **bacteria**.
bacterivorous *a.* feeding on bacteria.
bacteroid *n.* irregularly shaped bacterial cell, the form in which rhizobia are found in root nodules of legumes.

Figure 31. Terminological entry in *Henderson's Dictionary of Biology* (Lawrence 2008)

These examples indicate that the representation of hyponymy in both general and specialized language dictionaries is very limited. It is only indirectly present in intensional definitions or in related concepts mentioned in the explanations. Not surprisingly, the representation of conceptual relations is neither a priority nor a concern of this type of terminological resource.

3.3.1.2. HYPONYMY IN ENCYCLOPEDIAS

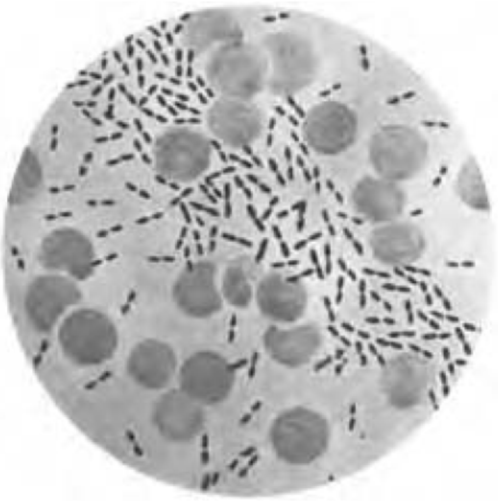
Encyclopedias are large compendia summarizing the general or specialized knowledge in a particular domain. They do not have either a linguistic or a definitional approach, because their objective is to provide as much information (e.g., chronological, cultural, social, technical) as possible about a certain topic or domain. Terminological information can be extracted from their entries, but it is more difficult to structure than in the case of dictionaries. Nevertheless, they usually contain graphical information, not only pictures but also diagrams or flow charts, which facilitate comprehension and knowledge acquisition.

Figure 32 shows an encyclopedic entry in the *Encyclopedia of Biology* (Rittner & McCabe 2004) for the concept BACTERIA. In this case, the entry is very similar to those in specialized language dictionaries, since it begins with an intensional definition (i.e., "Bacteria are microscopic, simple, single-cell organisms") followed by an explanation of the relation of the concept with other entities and processes (e.g., AEROBIC DECOMPOSITION, COLONIES, GRAM'S STAIN, etc.). Encyclopedias thus combine both an intensional and an extensional description of the concept.

In this case, hyponymy is again indirectly reflected in the definition, which begins with the hypernym (i.e., ORGANISM) as well as in certain parts of the text, which explain different classifications of BACTERIA (e.g., COCCAL, BACILLARY, SPIROCHETAL, VIBRO, AEROBIC, ANAEROBIC, GRAM-POSITIVE, GRAM-NEGATIVE). In addition, this encyclopedic entry is accompanied by an illustration that provides visual information about the concept. This is the distinguishing characteristic of encyclopedias in comparison to other traditional resources.

bacteria One of two prokaryotic (no nucleus) domains, the other being the ARCHAEA. Bacteria are microscopic, simple, single-cell organisms. Some bacteria are harmless and often beneficial, playing a major role in the cycling of nutrients in ecosystems via aerobic and anaerobic decomposition (saprophytic), while others are pathogenic, causing disease and even death. Some species form symbiotic relationships with other organisms, such as legumes, and help them survive in the environment by fixing atmospheric nitrogen. Many different species exist as single cells or colonies, and they fall into four shapes based on the shape of their rigid cell wall: coccal (spherical), bacillary (rod-shaped), spirochetal (spiral/helical or corkscrew), and vibro (comma-shaped). Bacteria are also classified on the basis of oxygen requirement (aerobic vs. anaerobic).

In the laboratory, bacteria are classified as gram-positive (blue) or gram-negative (pink) following a laboratory procedure called a Gram's stain. Gram-negative bacteria, such as those that cause the plague, cholera, typhoid fever, and salmonella, for example, have two outer membranes, which make them more resistant to conventional treatment. They can also easily mutate and transfer these genetic changes to other strains, making them more resistant to antibiotics. Gram-positive bacteria, such as those that cause anthrax and listeriosis, are more rare and are treatable with penicillin but can cause severe damage by either releasing toxic chemicals (e.g., clostridium botulinum) or by penetrating deep into tissue (e.g., streptococci). Bacteria are often called germs.



Photomicrograph of *Streptococcus (Diplococcus) pneumoniae* bacteria, using Gram's stain technique. *Streptococcus pneumoniae* is one of the most common organisms causing respiratory infections such as pneumonia and sinusitis, as well as bacteremia, otitis media, meningitis, peritonitis, and arthritis. *(Courtesy of Centers for Disease Control and Prevention, 1979)*

bacteriochlorin (7,8,17,18-tetrahydroporphyrin) A reduced PORPHYRIN with two pairs of nonfused saturated

Figure 32. Encyclopedic entry in *Encyclopedia of Biology* (Rittner & McCabe 2004)

However, not all encyclopedias are composed of alphabetical entries that describe individual concepts, similarly to dictionaries. **Figure 33** shows an excerpt from the *Encyclopedia of Cell Biology* (Bradshaw & Stall 2016), which consists of explanatory chapters on broad topics within the knowledge domain.

Cell Biology: An Overview

RA Bradshaw, University of California, Irvine, CA, USA
PD Stahl, Washington University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

All living organisms can be divided into three principal categories: archaebacteria, eubacteria, and eukaryotes; although they differ in structure and organization, they are all composed of cells as the fundamental life unit. At the molecular level, there is also a great deal of similarity in the basic materials that make up these entities because they use the same kinds of molecules to store and reproduce information, to run the cellular metabolism and machinery and to provide the structural framework. Thus nucleic acids, proteins, lipid membranes, and carbohydrates – alone and in various combinations – are universally present, albeit in distinguishable forms, along with innumerable metabolites and ions. There are components that are apparently essential for life and are found in one form or another in all species and there are many unique moieties and associated activities that are highly specialized and are found in relatively few organisms. Indeed, the similarities have underpinned the development of our understanding of cellular the rapid advances in microscopy both electron and light microscopy – the former allowing investigators to fractionate and characterize the components of the cell and the latter to literally see them – either *in situ* or in isolated form. This progress is illustrated by a series of Nobel Prizes starting in the early 1970s that chronicle the grand confluence of classical biochemistry and general physiology that created modern cell biology (Claude *et al.*, 1974) or the discovery and characterization of intracellular organelles – lysosomes and endoplasmic reticulum (ER) among others (Mitchell, 1978), for the chemisomotic hypothesis (Brown and Goldstein, 1985), for endocytosis (Ciechanover *et al.*, 2004), for ubiquitination and protein degradation, and just recently (Rothman *et al.*, 2013), for vesicle trafficking, and (Betzig *et al.*, 2014) for advances in light microscopy. The early cell biologists insisted on quantitative application of these new techniques and laid the groundwork for modern cell biology. As with bio-

Figure 33. Encyclopedic entry in *Encyclopedia of Cell Biology* (Bradshaw & Stall 2016)

However, since the chapters are thematically instead of alphabetically, concepts can only be searched in the index at the end of the encyclopedia, which lists the pages on which each concept is discussed (**Figure 34**).

Bacillus subtilis (*B subtilis*) **1**:57
back focal plane (BFP) **2**:67, **2**:122–123
backscattered electrons **2**:16
BACs *see* bacterial artificial chromosomes (BACs)
bacteria **1**:162, **2**:149, **2:588–596**, **2**:800
  bacterial filament
    involved in DNA segregation **2**:591–592*F*
    structure involved in cell division **2**:589*F*
  cell shape determination **2**:593, **2**:594*F*
    actin-like MreB in rod-shaped cells **2**:593
    filamentous proteins promoting curvature **2**:593
  cytokinesis **2**:588–590
  cytoskeletal elements functions **2**:588*F*

Figure 34. Alphabetical index in *Encyclopedia of Cell Biology* (Bradshaw & Stall 2016)

Consequently, the representation of hyponymy in encyclopedias is as limited as it is in dictionaries. It is only with the birth of new ways of representing knowledge that this becomes possible, though not all digital resources incorporate it.

3.3.2. Hyponymy in Contemporary Resources

This section explains how hyponymy is represented in digital resources such as term banks and TKBs. On the one hand, term banks (also known as *terminology databases* or *termbases*) provide direct access to terms as well as to their related linguistic information. Each entry may include data fields such as definition, correspondences in one or various languages, synonyms, abbreviations, status of each term (e.g., preferred, reliable, not recommended, etc.), usage contexts (and their corresponding references), and even the domain and subdomain to which the concept belongs.

Many TKBs go beyond term banks by implementing a wide range of features that enhance terminology. Such features include a dynamic knowledge representation, a graphic visualization of conceptual relations between concepts, and the integration of multimedia information, *inter alia*. The following sections present case studies of how hyponymy is represented in two term banks (i.e., IATE and TERMIUM Plus) and in two TKBs (i.e., WIPO Pearl and EcoLexicon).

3.3.2.1. Hyponymy in IATE

IATE[3] (Johnson & MacPhail 2000; Rummel & Ball 2001; Zorrilla-Agut & Fontenelle 2019) is the official terminology database of the European Union (EU), developed and supervised by the Translation Centre for the Bodies of the EU in collaboration with other European institutions. It is the largest multilingual term bank in the world, with around 900,000 concept entries and eight million terms in the 24 official languages of the EU.

The project was launched in 1999 with the aim of providing a web-based infrastructure for all EU terminology resources, enhancing the availability and standardization of the information. It has been used in the EU institutions and agencies since 2004 for the collection, dissemination, and management of EU-specific terminology. The new version of IATE, also publicly available, was released in 2018 following a full rebuild of the system with state-of-the-art technologies, the latest software development standards, best practices on usability and accessibility, and a new look and feel (Zorrilla-Agut & Fontenelle 2019). In fact, its basic use remains easy and intuitive.

[3] Available at: https://iate.europa.eu/

The IATE homepage (**Figure 35**) displays a search bar where the user can enter the concept and select both the source and target languages for the query.

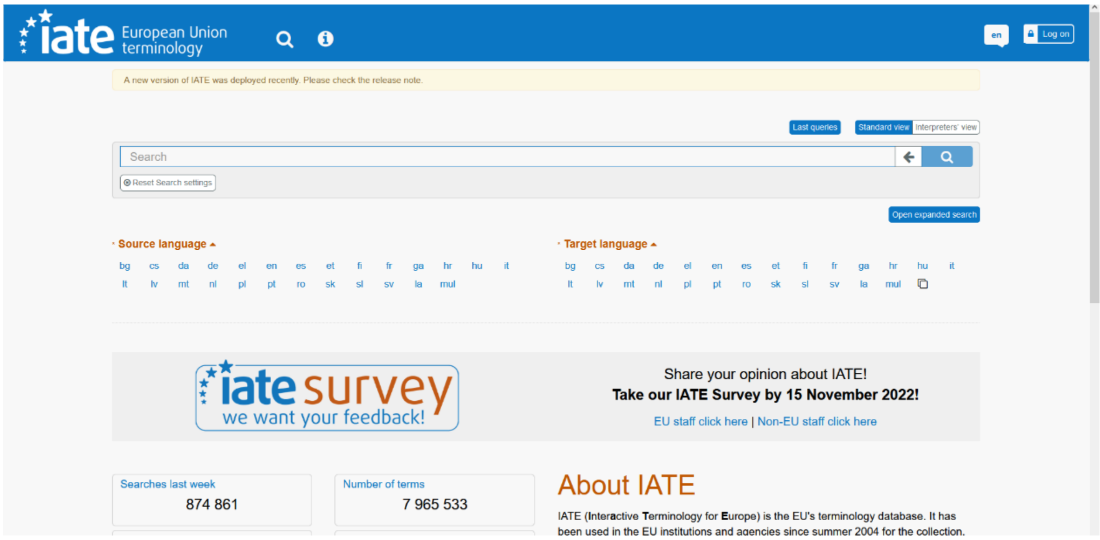


Figure 35. Homepage of IATE

After performing this search, the entries are grouped in knowledge domains. In this way, the same concept may have more than one entry depending on the field in which it is classified. **Figure 36** shows the terminological entry in IATE for BACTERIUM in the SCIENCE domain (i.e., Natural and Applied Sciences, Life Sciences, Biology).

The information in the entry is provided in English and Spanish, according to the search criteria selected. Within the entry, the following main elements are distinguished: (i) term; (ii) term reference; (iii) term reliability; (iv) definition; (v) definition reference; and (vi) creation and modification dates.

These terminological entries can contain usage contexts and observation notes. However, since no mention is made of any kind of relation between terms or concepts, no direct way of accessing hyponymic information, as is also the case of traditional resources.

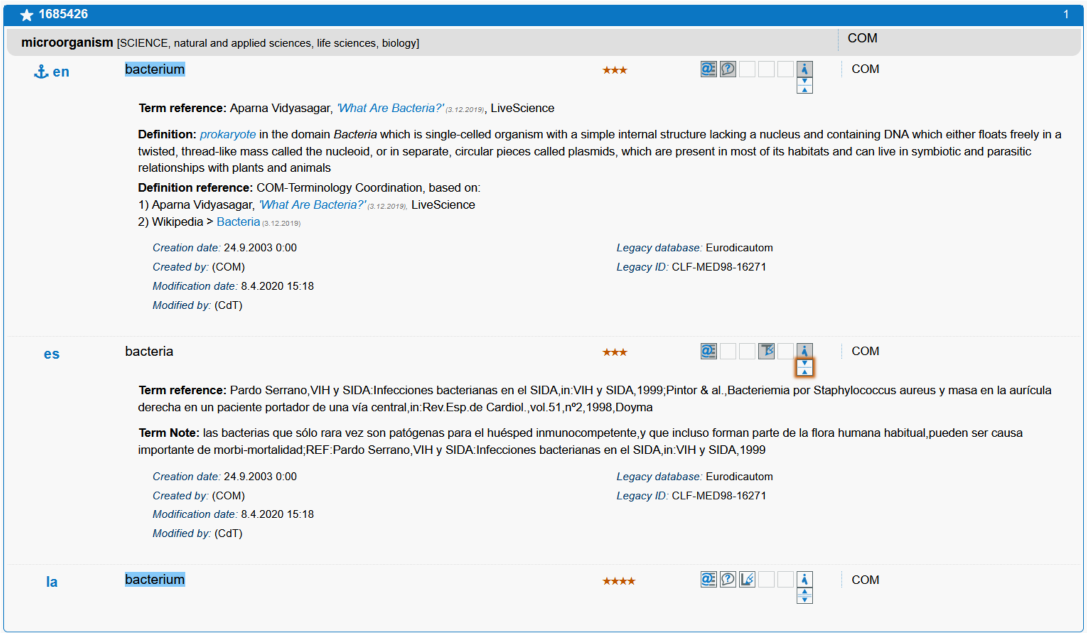


Figure 36. Terminological entry in IATE

An interesting feature of the new version of IATE is that it offers two display modes for the results: the *standard view* (**Figure 37**) and the *interpreters' view* (**Figure 38**). The standard view sequentially lists the entries of each concept per domain and offers parameters to display more or less terminological information about each entry. In contrast, the interpreters' view only provides direct access to multilingual terms, which are also displayed in parallel columns for easier viewing and comparison.

Interestingly, in a general search for a broader term, the interpreters' view makes it easier to visualize the complex nominals using the search term as a head. This favors the retrieval of hyponyms composed of two or more components.

However, IATE does not represent hyponymy. In fact, there are not even any hyperlinks within the entries that redirect users to other associated concepts. Therefore, once again, hyponymy representation is only indirectly reflected through intensional definitions and partially in the multi-word hyponyms in the interpreters' view.

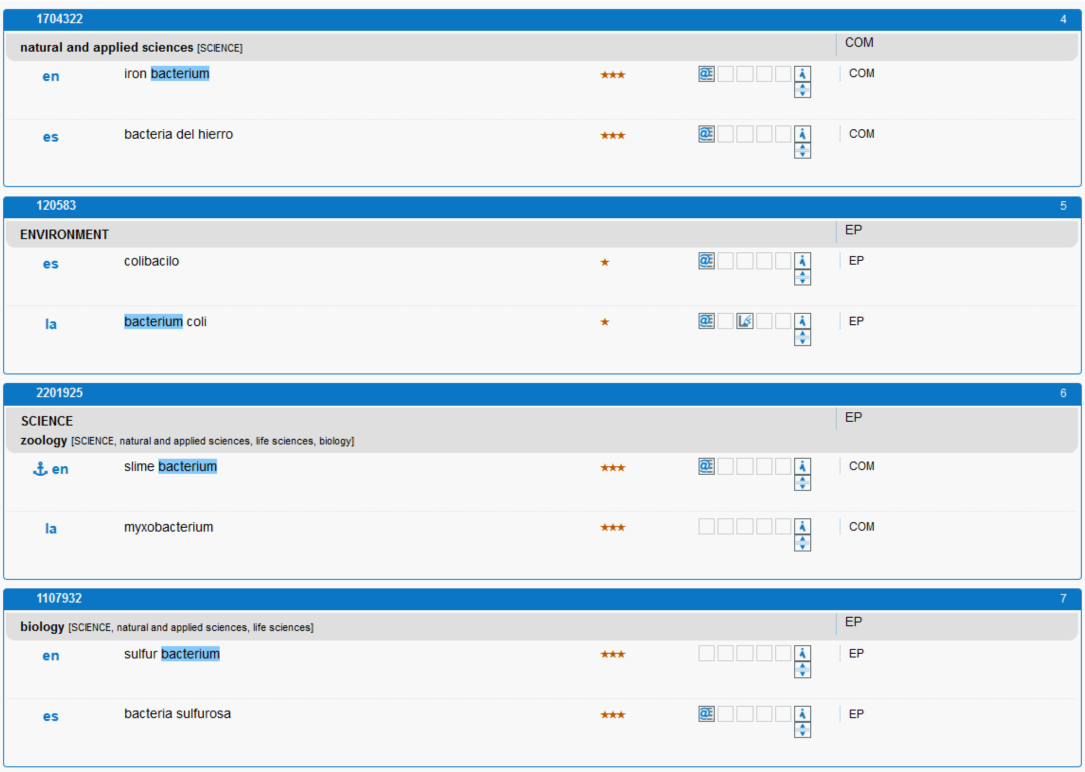


Figure 37. List of entries with the standard view of IATE

Figure 38. List of entries with the interpreters' view of IATE

3.3.2.2. Hyponymy in TERMIUM Plus

TERMIUM Plus[4] (Landry 1988; Bernier-Colborne *et al*. 2017) is a terminology and linguistic databank that is operated and maintained by the Translation Bureau of Public Services and Procurement of Canada. This term bank includes almost four million English and French terms, more than 200,000 Spanish terms, more than 18,000 Portuguese terms, and a variety of writing tools. It not only contains terms, but also abbreviations, definitions, and usage examples in a wide range of specialized fields.

This resource was conceived with different linguistic functionalities, such as understanding an acronym, checking an official title, and finding an equivalent in another language. It can be accessed through a simple interface where a query can be performed, with the possibility of filtering it by subject field (**Figure 39**).

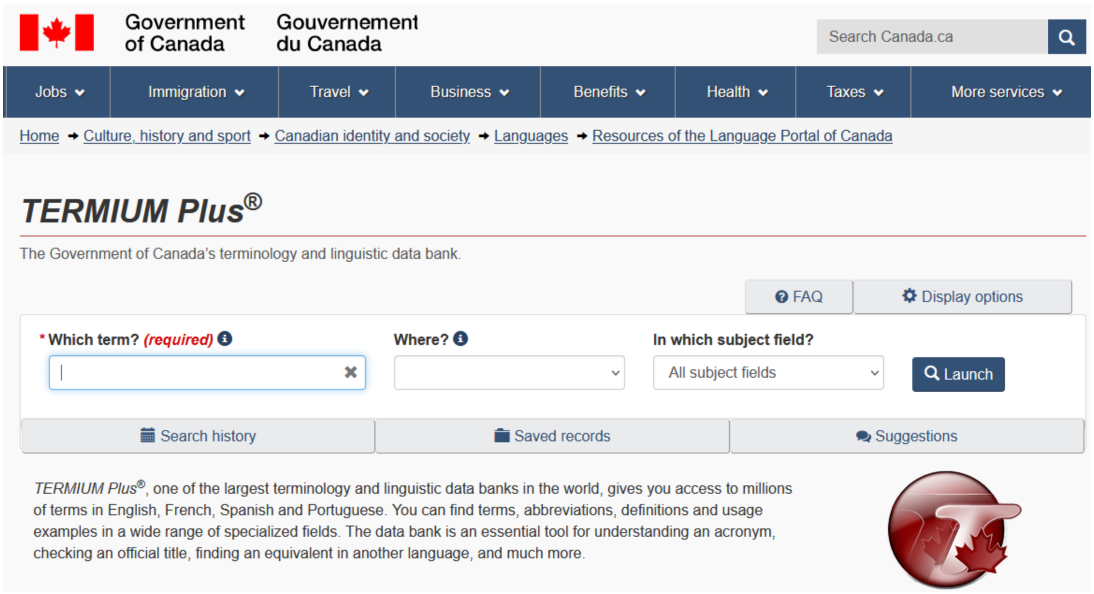


Figure 39. Homepage of TERMIUM Plus

The terminological entries in TERMIUM Plus are called *records*, and they correspond to the meanings of a term in different specialized knowledge fields. These entries indicate the subject fields, term(s), definition (DEF), context (CONT), observation (OBS), and phraseology (PHR). They thus focus on providing the definition of the term; an example of the term in a text fragment; terminological, linguistic or technical information; and/or common combination of a term with a noun, adjective or verb.

[4] Available at: https://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-eng.html

Figure 40 shows the entry in TERMIUM Plus for BACTERIUM in the Sciences domain (i.e., Microbiology and Parasitology, Bacterial Diseases, Water Pollution, and CBRNE Weapons). Definitions are intensional, which means that hyponymy is implicitly represented in the *genus* of the definition. However, there are no links between concepts nor is there any reference to generic-specific information in any of the elements available.

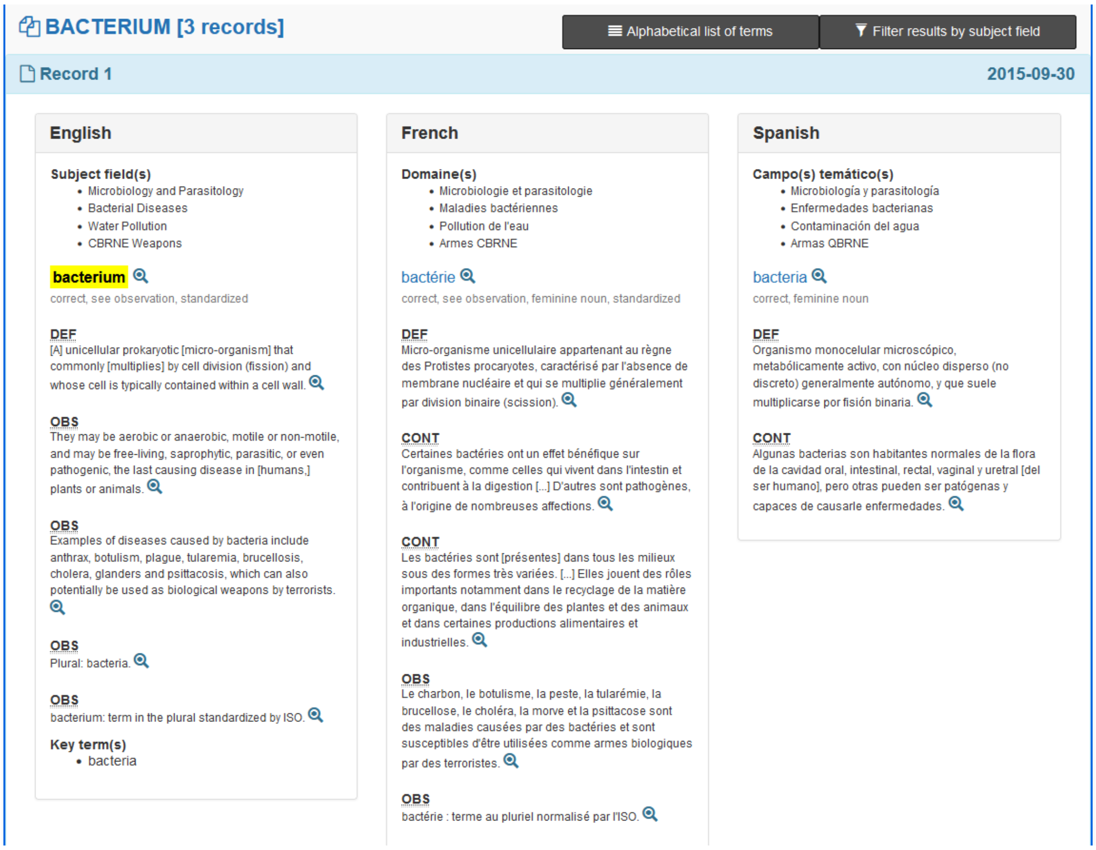


Figure 40. Terminological entry in TERMIUM Plus

3.3.2.3. HYPONYMY IN WIPO PEARL

WIPO Pearl[5] (Valentini 2013; Valentini *et al*. 2016; Caffrey & Valentini 2019) is a multilingual TKB developed by the PCT Translation Division of the World Intellectual Property Organization (WIPO). The WIPO Pearl database contains around 25,500 concepts and 230,000 terms in the ten languages used in this institution: Arabic, Chinese, English, French, German, Japanese, Korean, Portuguese, Russian, and Spanish.

[5] Available at: https://www.wipo.int/reference/en/wipopearl/

This terminological resource allows users to access the scientific and technical terms in patent documents, thus promoting the accurate and consistent use of terms across different languages, and making it easier to search and share scientific and technical knowledge. WIPO Pearl uses the software PATENTSCOPE to search patent documents and international patent applications. This means that it is possible to access the entire PATENTSCOPE corpus for terms and their equivalents in other languages. When accessing the WIPO portal, users can choose whether to perform a *linguistic search* or a *concept map search* (**Figure 41**).

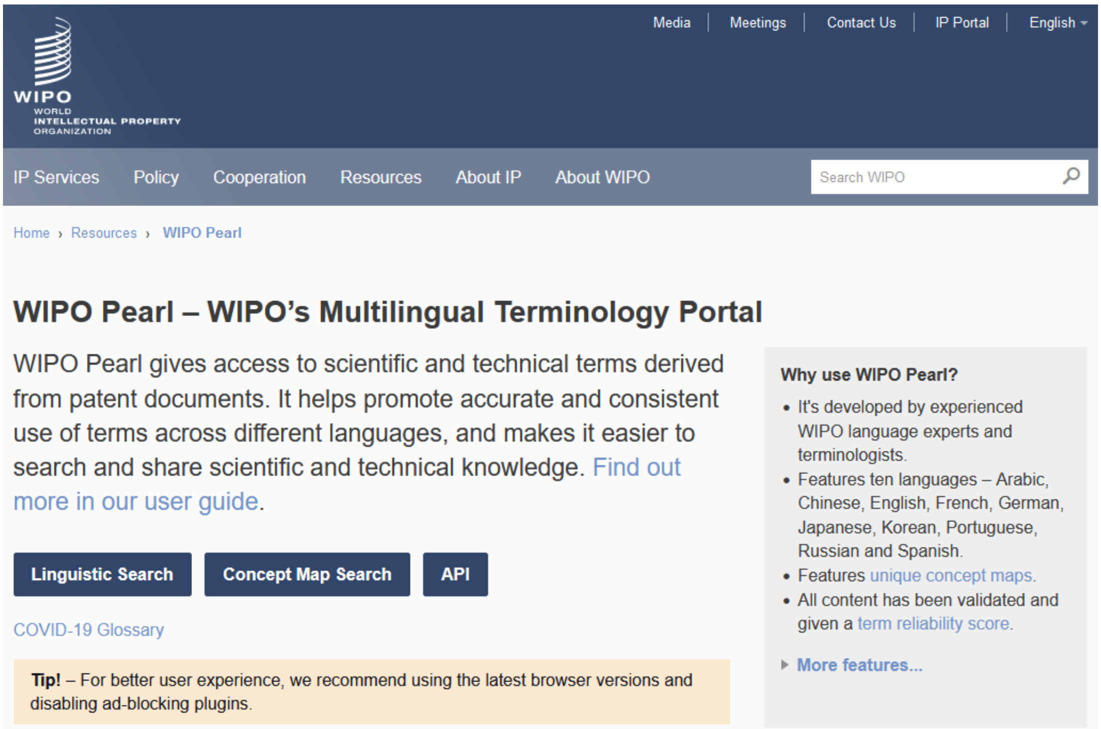


Figure 41. Homepage of WIPO Pearl

The linguistic search option allows users to perform conventional searches for terms by selecting both the source and target languages (**Figure 42**). As usual in this kind of terminological resources, the search can be filtered by subject field. Once the search is performed, the terminological entries (also called *records*) are grouped by subject fields and display the terms in all languages available on the platform, along with their reliability (**Figure 43**).

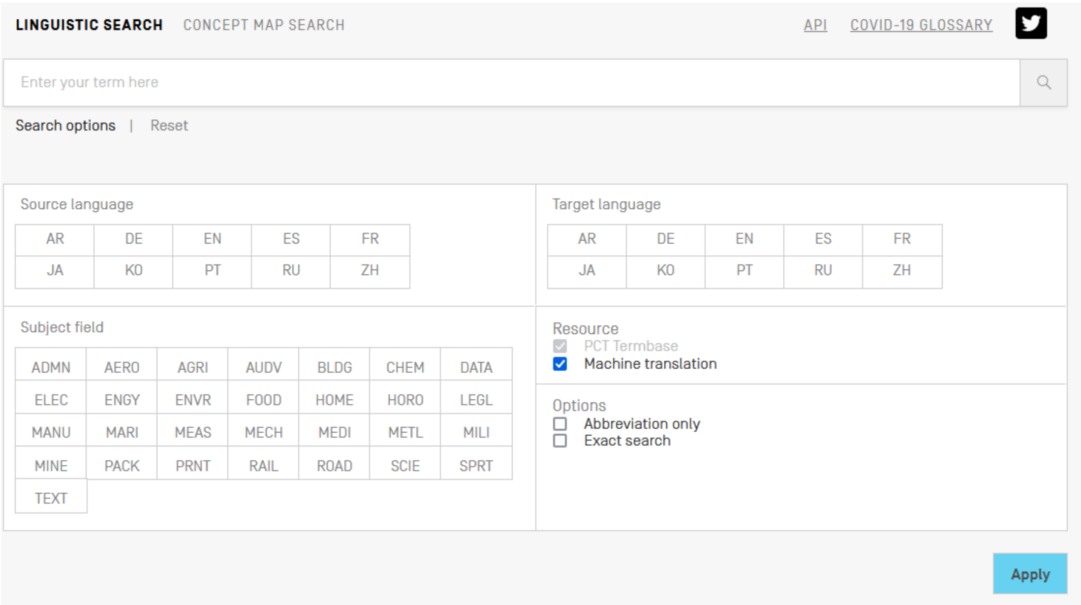


Figure 42. Linguistic search in WIPO Pearl

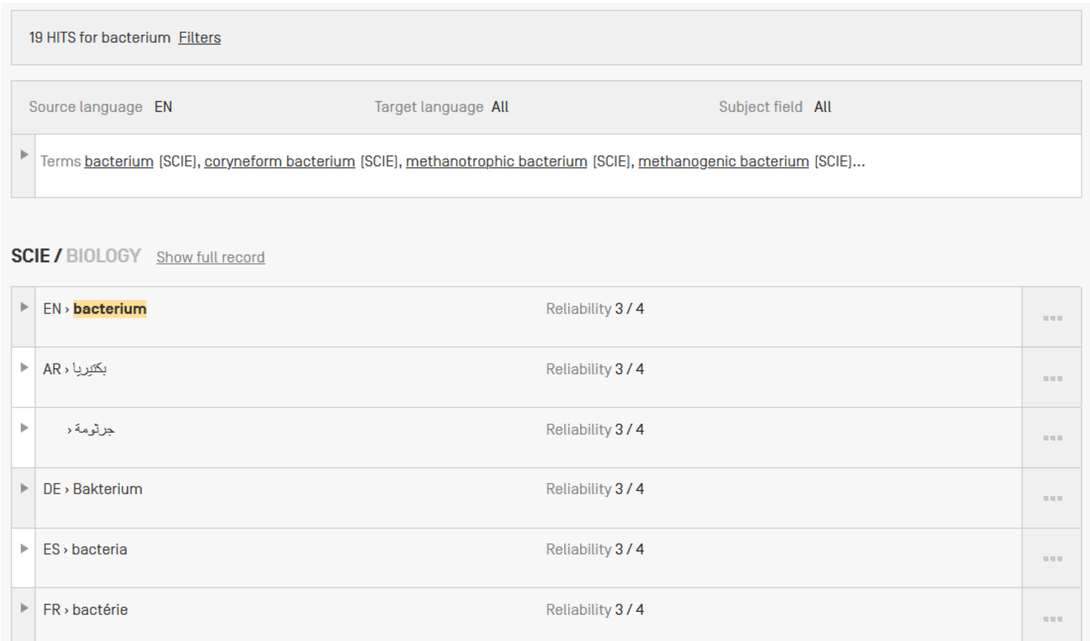


Figure 43. List of entries with the linguistic search in WIPO Pearl

The alternate way to access the terminological entries in WIPO Pearl is through the concept map search option. This feature displays a bubble chart representing the subject fields, which have different colors to better differentiate them, and which are larger or smaller depending on the number of terminological entries available in each (**Figure 44**).

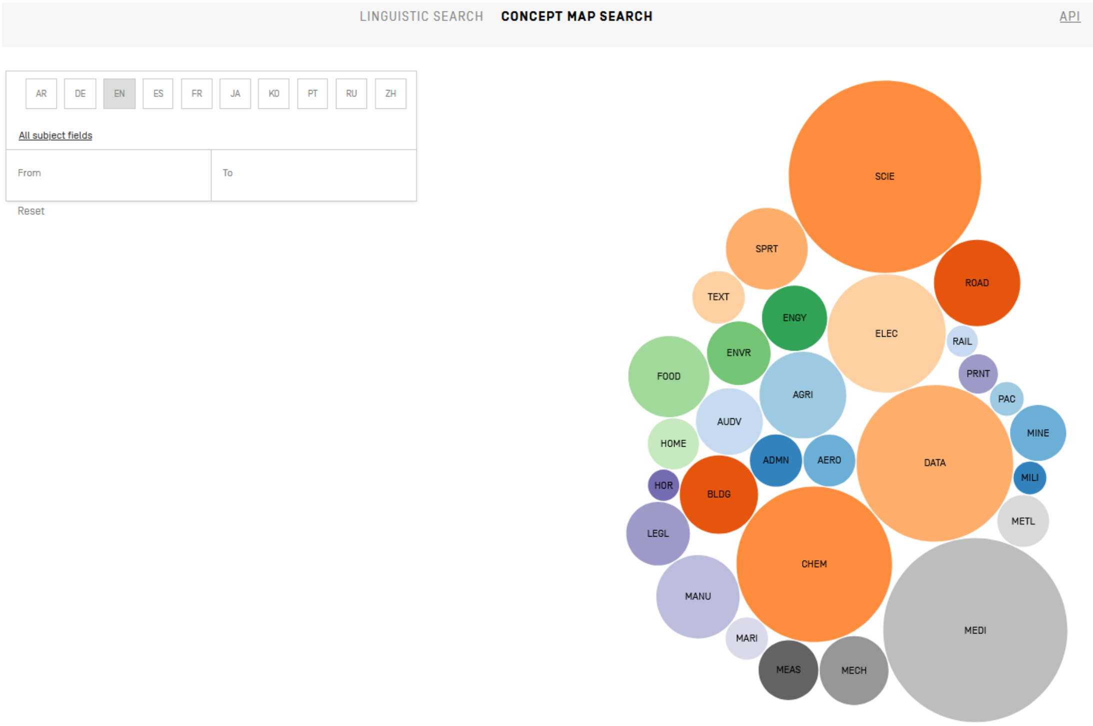


Figure 44. Concept map search in WIPO Pearl

Once a subject field is selected, a graphical representation of the conceptual systems appears (**Figure 45**). These conceptual systems form concept clouds associated by two types of conceptual relation: (i) hierarchical relations (i.e., generic or partitive), shown in blue; and (ii) associative relations, shown with dashed lines. The concepts in red belong to a different subject field, but are related to the concepts of the present subject field in some way. Despite being in the same category as meronymy, hyponymy is explicitly represented in this resource thanks to the visualization of the hierarchical relations in blue. These conceptual systems are created by specifying in each terminological entry if there is a *related concept broader* and a *related concept narrower*.

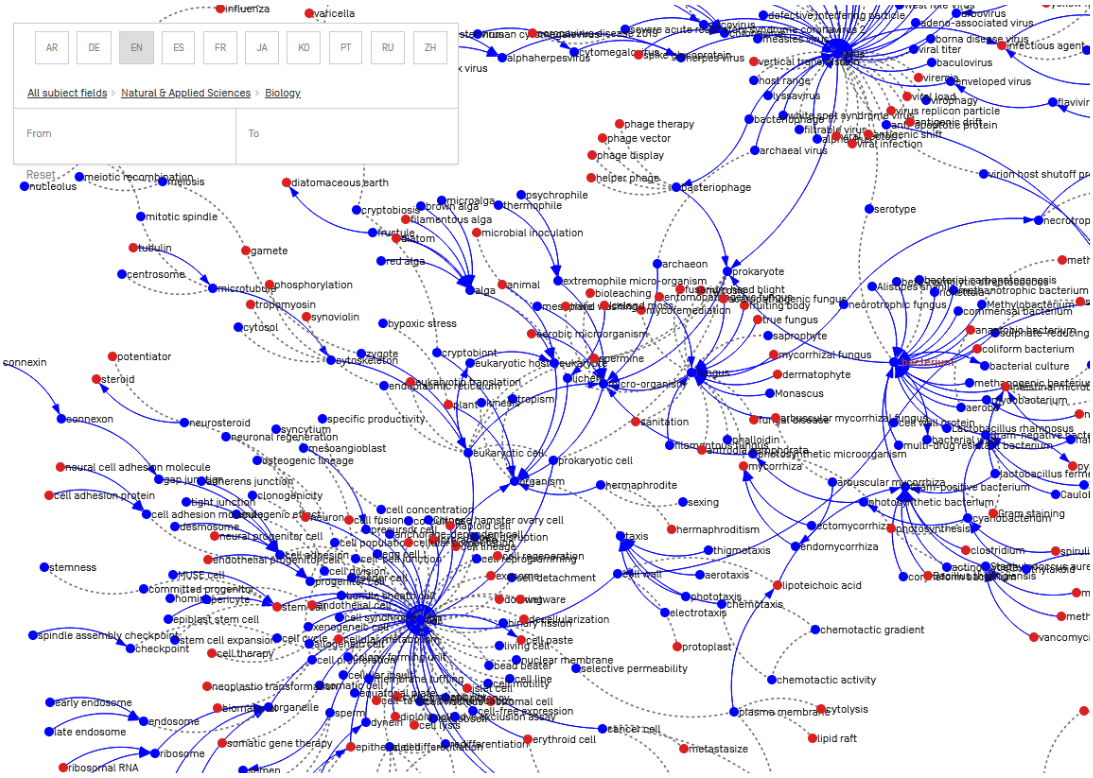


Figure 45. Conceptual systems and conceptual relations in WIPO Pearl

When a concept is selected in the concept map, the corresponding term entry is accessed (this entry can also be accessed from the results of the linguistic search, seen in **Figure 43**). The following information is available in a WIPO Pearl term entry: (i) term; (ii) term type; (iii) usage label; (iv) term reliability; (v) last modification date; (vi) context (which is preferably a defining context); and (vii) source (which is preferably a patent available through PATENTSCOPE). In addition, it is possible to view a segment of the concept map, focusing on the concept of the term entry so that the user can see the closest related concepts and distinguish the types of relation. **Figure 46** shows the entry for BACTERIUM in WIPO Pearl, along with a segment of the concept map with conceptual relations.

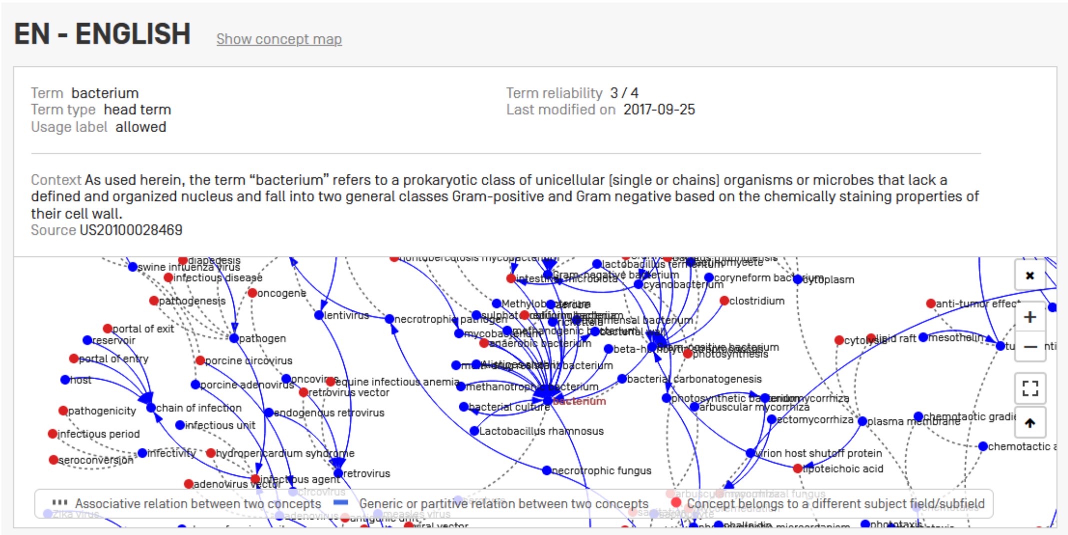


Figure 46. Terminological entry with conceptual relations in WIPO Pearl

Thanks to its cognitive approach, WIPO Pearl provides an explicit representation of hyponymy in its conceptual systems and conceptual relations. Nonetheless, the representation of conceptual relations, namely hyponymy, is open to improvement.

For instance, grouping both hyponymy and meronymy in the same category leads to confusion, because the user cannot tell which type of relation exists between two concepts. Furthermore, the fact that only two types of relation (i.e., hierarchical and associative) are represented leads to information overload because of a lack of filtering. This issue could be successfully addressed if relations were divided into different subtypes (e.g., hyponymy, meronymy, causality, functionality, etc.). Needless to say, this would greatly enhance the representation of hyponymy.

3.3.2.4. Hyponymy in EcoLexicon

Since EcoLexicon and its treatment of conceptual relations have been discussed in sections **§2.1.3** and **§2.2.3.2**, here the focus is on options for representing hyponymic relations in (cloud-like) conceptual systems and (tree-like) conceptual hierarchies.

Conceptual systems in EcoLexicon are cloud-like representations of the interrelations between a given concept and related concepts. Since EcoLexicon can represent up to 18 different types of conceptual relation, the visual map offers the possibility of filtering the relationships so that only those of interest are displayed with a view to reducing information overload. **Figure 47** shows the conceptual system of BACTERIA in EcoLexicon represented with a single hyponymic relation (*type of*) as follows.

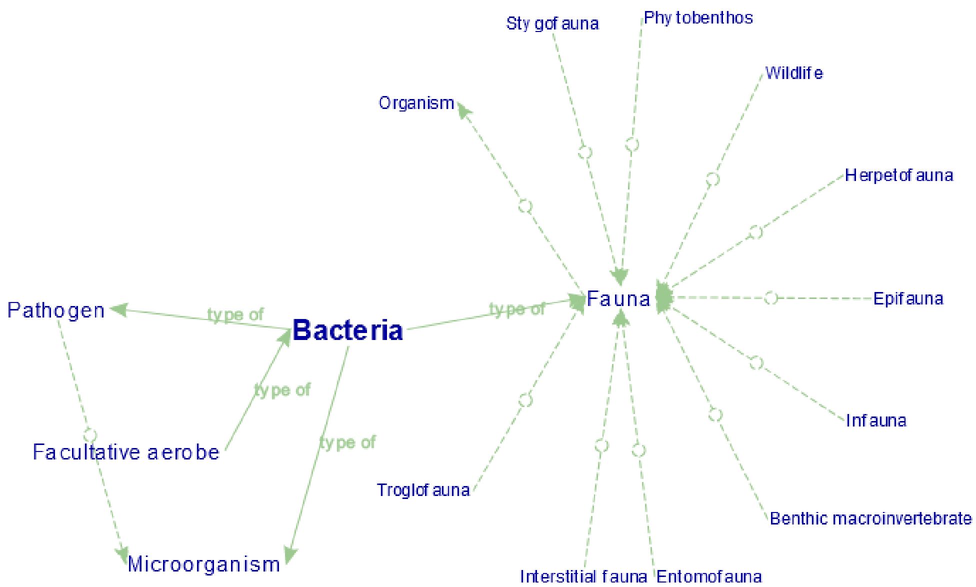


Figure 47. Conceptual system in EcoLexicon with a single hyponymic relation (*type of*)

However, these conceptual systems in cloud-like networks can be confusing when a large number of related concepts are involved. For this reason, EcoLexicon also offers the possibility of representing this information in conceptual hierarchies or tree-like representations with different levels of the hypernyms and hyponyms of a given concept. **Figure 48** shows the conceptual system of ROCK in EcoLexicon represented with a single hyponymic relation (*type of*).

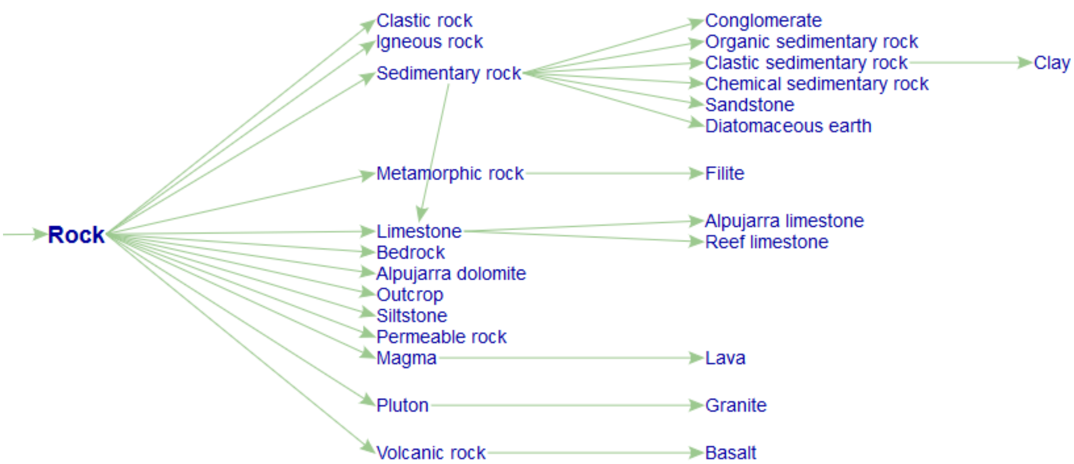


Figure 48. Conceptual hierarchy in EcoLexicon with a single hyponymic relation (*type of*)

Filtering conceptual relations so that only hyponymic relations are selected and generating tree-like hierarchies with this information is an excellent way of representing hyponymy in a terminological resource. Therefore, of all the resources in this section, EcoLexicon has the most effective hyponymy representation.

However, there is still room for improvement. In this regard, Gil-Berrozpe (2016) and Gil-Berrozpe *et al*. (2018) detected a number of problems such as the visualization of dimensions of co-hyponyms at the same level without any distinction, noise, information overload, redundancy, and transitivity problems regarding property inheritance. This led to a hyponymy refinement process and the subsequent corpus-based study (Gil-Berrozpe 2016, 2017) (see **§3.2.3**).

One of the proposals of Gil-Berrozpe (2016) and Gil-Berrozpe *et al*. (2018) not only included dividing hyponymy into subtypes, but also representing the conceptual hierarchies with different colors depending on hyponymy subtype (see **Figure 49** for an example).

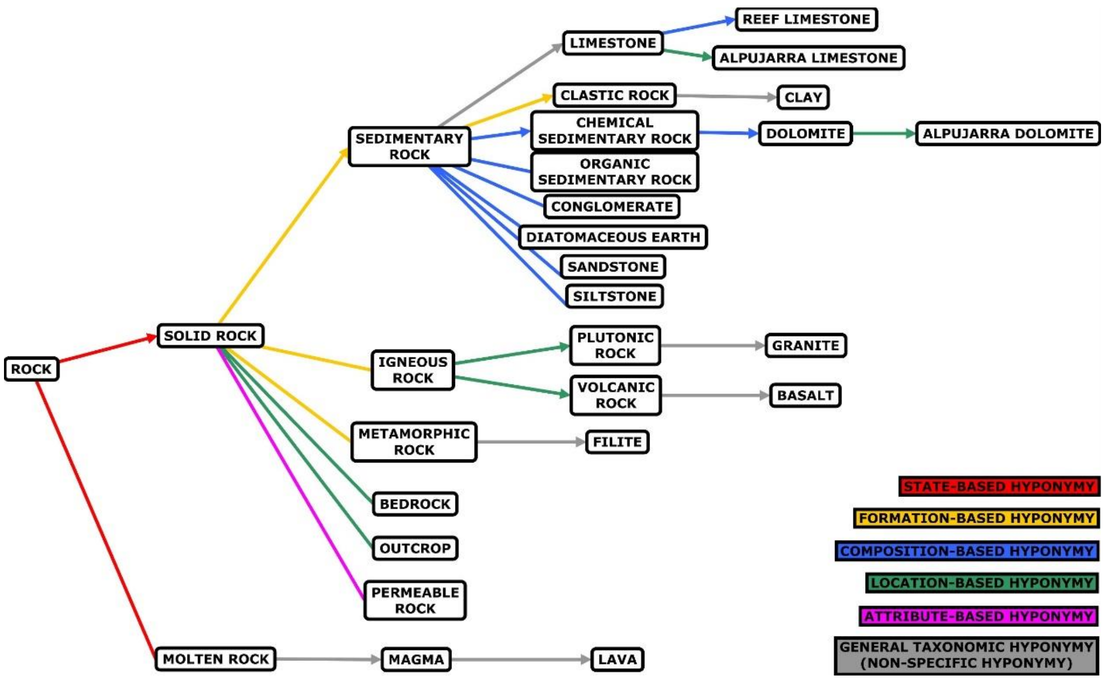


Figure 49. Conceptual hierarchy in EcoLexicon with multiple hyponymy subtypes

Nevertheless, this visualization proposal could not be implemented in EcoLexicon because the corpus-based study (Gil-Berrozpe 2017, Gil-Berrozpe *et al*. 2017) detected up to 32 possible hyponymy subtypes for environmental concepts. Despite having only one hyponymic relation (*type of*), EcoLexicon already has in total 18 conceptual relations which, in some conceptual systems, generate information overload. Adding 32 more conceptual relations would have made the networks too complex. For this reason, it was necessary to find an alternative way to accurately represent hyponymy.

3.3.3. Towards a Comprehensive Representation of Hyponymy

As reflected in the case studies, hyponymy is not explicit in dictionaries and encyclopedias. It is only implicitly alluded to in intensional definitions or references to broader or narrower concepts in extensional definitions. In addition, hyponymy is treated much the same way in term banks.

TKBs, however, have a more cognitive approach, and represent hyponymy in conceptual networks that link concepts with relations. However, such conceptual systems are not problem-free. On the one hand, when hyponymy and meronymy are combined in a single category, it is not clear which relations are generic-specific. On the other hand, when too many hyponymy subtypes are specified, this produces information overload.

Therefore, we propose a new way of representing hyponymy that is based on terminological entries with a structure and content focused on the description and categorization of generic-specific relations. A comprehensive representation of hyponymy should meet the following criteria:

- Concepts should be hierarchically structured with different levels of specificity so as to highlight the transitivity from hypernyms to direct hyponyms rather than to indirect hyponyms.
- Term entries should contain intensional definitions of concepts, based on *genus* and *differentiae*, in order to display property inheritance from hypernyms to hyponyms and to clearly distinguish between co-hyponyms.
- Concept description should be based on conceptual categories so as to better understand the ontological changes that occur in the transition from hypernyms to hyponyms.

- Hyponymy should be categorized in subtypes by identifying the nuance that differentiates a hyponym from its hypernym, and co-hyponyms should be classified according to the hyponymy subtype involved.
- Hyponymic contexts should be reflected in order to understand how hyponymy is encoded in KPs and expressed in specialized language.

The following chapters present the methodology and the results of the representation of generic-specific relations in twelve hyponymy-based terminological entries. This set of entries is the foundation of HypoLexicon, a terminological resource focused on the description, categorization, and representation of environmental terminology.

4. MATERIALS AND METHODS

The materials used in this research include four specialized environmental subcorpora from the EcoLexicon English Corpus and a selection of specialized terminological resources. The software used was the EcoLexicon internal application, Sketch Engine, and Lexonomy. The methodology involved corpus extraction and compilation, corpus analysis (including all hypernym and hyponym extraction, identification, and selection), creation of the conceptual hierarchies, and design of the terminological template for the hyponymy-based terminological entries.

4.1. MATERIALS

As a study based on the theoretical and practical premises of FBT, which is highly influenced by Corpus Linguistics, the main materials were four specialized environmental subcorpora taken from the EcoLexicon English Corpus. However, a selection of specialized terminological resources was also necessary to help build the terminological definitions and improve the conceptual hierarchies. Furthermore, the software used to carry out each process of the methodology (i.e., the EcoLexicon internal application, Sketch Engine, and Lexonomy) are also described in this section.

4.1.1. ECOLEXICON ENGLISH CORPUS

Sager (1990: 55) states that texts are the basic material of terminologists and that the terminological resources reflect the usage of terms in natural language contexts. This is only possible by basing resource content on corpora.

Many authors defined a corpus (Sinclair 1991, 1996b; Atkins *et al.* 1991; Pearson 1998; Bourigault & Slodzian 1999). According to these definitions, a corpus is a large collection of real samples of a language in the form of texts that have been selected according to specific criteria and which are to be used as a representation of language. Since the compilation process of the corpora must respond to the needs of the research to be carried out, different aspects should be taken into account (e.g., quantity, domain, level of expertise, period, etc.). Pearson's (1998: 43) definition is the following:

> (…) a corpus is an artefact; it is selected, chosen or assembled according to explicit criteria. It is stored in electronic form. It consists of pieces of naturally occurring language. In this context, we understand naturally occurring to

> mean that the pieces of language have not been tampered with or edited. The corpus may, however, be annotated during or after the compilation process; grammatical tags or SGML markups (e.g., indicating text origin, authorship) may be added to facilitate information retrieval. A corpus may be used as a "sample of the language" (Sinclair) or because it is "representative of a given language" (Francis). A corpus may be a collection of transcribed spoken and/or written pieces of language, contrary to what the use of the word text might suggest.

Furthermore, other authors highlight how corpora also reflect extra-linguistic elements through knowledge-rich contexts (Meyer 2001), such as semantic or conceptual relations (Condamines 2002). This is particularly relevant for this research, which was based on four specialized environmental subcorpora extracted from the EcoLexicon English Corpus.

The EcoLexicon English Corpus (EEC) is a 23.1-million-word corpus of contemporary environmental texts compiled by the LexiCon Research Group (León-Araúz *et al.* 2018, 2019; San Martín *et al.* 2020). It is a tool designed to be exploited by terminologists, translators or even technical experts for modeling, comprehension, and production tasks. For the purpose of allowing its general use, the EEC is publicly available in two ways: through the *Concordance search* function within EcoLexicon, and through the corpus tool Sketch Engine as part of the Sketch Engine Open Corpora (**Figure 50**). The EEC was precisely made available as an open corpus in Sketch Engine so that any user could freely access it.

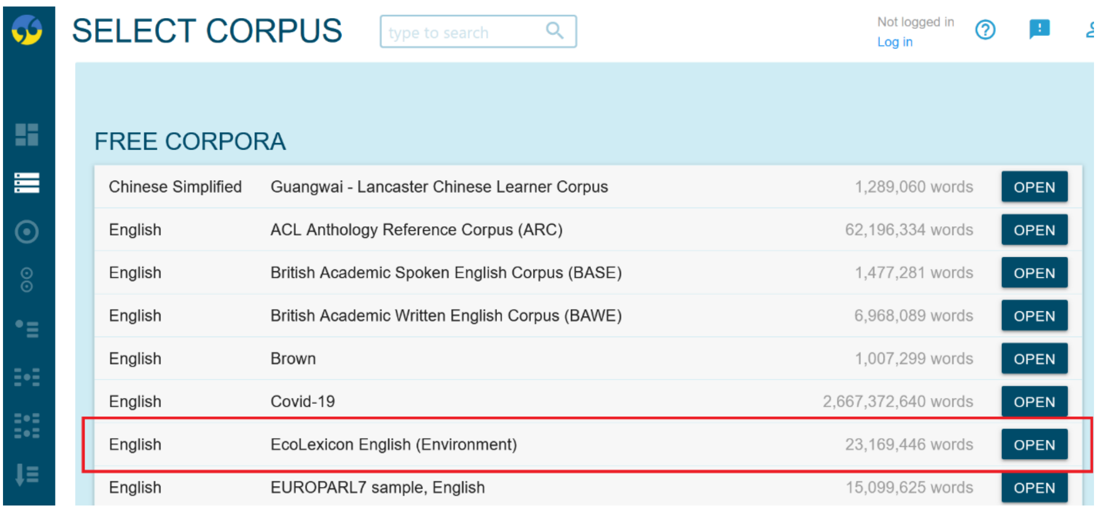


Figure 50. The EcoLexicon English Corpus among Sketch Engine Open Corpora

The texts in the EEC are tagged with a set of XML-based metadata (i.e., domain, user, geographical variant, genre, editor, year, and country). These metadata allow users to perform queries based on pragmatic factors, such as environmental domains and target user. This allows them to compare, for instance, the use of the same term in different contexts (León-Araúz *et al.* 2018).

The EEC was processed and compiled in an internal application of the research group, but it was also recompiled in Sketch Engine with the Penn Treebank tagset (TreeTagger version 3.3) and with the EcoLexicon Semantic Sketch Grammar (ESSG) (León-Araúz *et al.* 2016; León-Araúz & San Martín 2018). As previously mentioned, the ESSG is a customized sketch grammar that extracts semantic word sketches based on the most common conceptual relations (i.e., hyponymic, meronymic, locative, causal, functional). The hyponymic or generic-specific word sketches were used in our research.

For the sake of delimiting the scope of the study, and of analyzing and comparing hyponymy across microdomains, four subcorpora were extracted from the EEC: a Biology, a Chemistry, a Civil Engineering, and a Geology subcorpus.

4.1.1.1. BIOLOGY SUBCORPUS

The Biology (BIO) subcorpus contains 6,217,032 words according to the statistics in Sketch Engine (**Figure 51**). It is composed of specialized texts in the domains of general Biology, Biological Oceanography, Botany, Zoology, Microbiology, Molecular Biology, and Biochemistry. In total, it comprises 241 documents for experts, semi-experts, and the general public. The documents in this subcorpus include scientific articles and journal papers, specialized books, lexicographic material, educational texts, and newspaper articles.

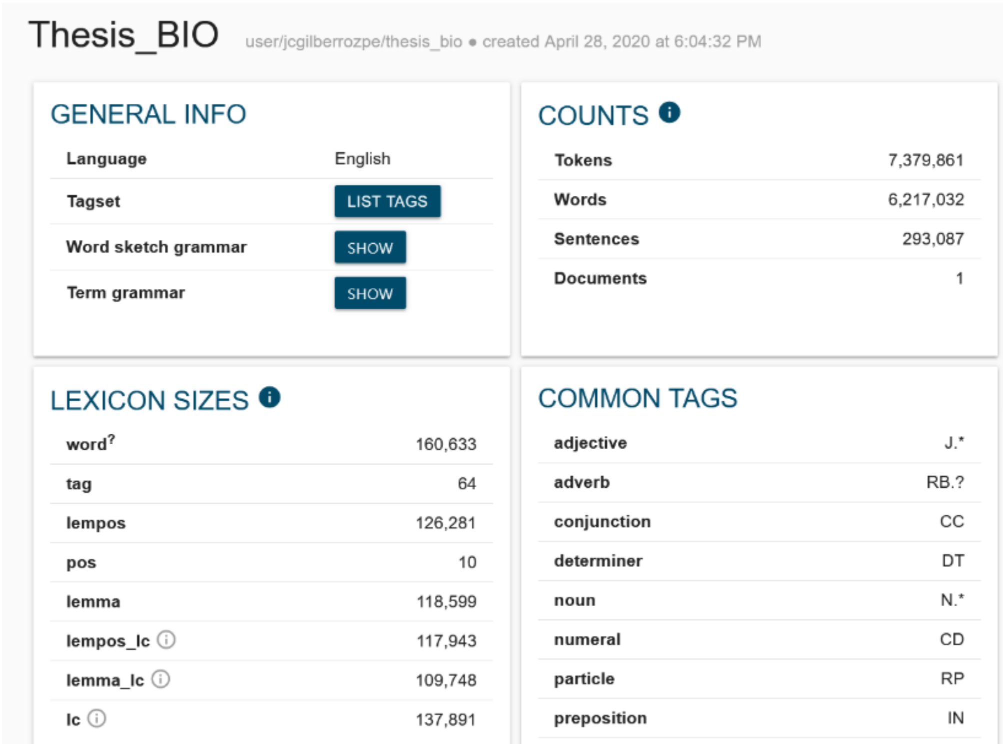


Figure 51. BIO subcorpus information in Sketch Engine

4.1.1.2. CHEMISTRY SUBCORPUS

The Chemistry (CHEM) subcorpus contains 2,984,197 words according to the statistics in Sketch Engine (**Figure 52**). It is a collection of specialized texts in the domains of general Chemistry, Geochemistry, Biochemistry, and Chemical Oceanography. In total, it comprises 54 documents for experts and semi-experts. The documents in this subcorpus are scientific articles or journal papers, specialized books, and lexicographic material.

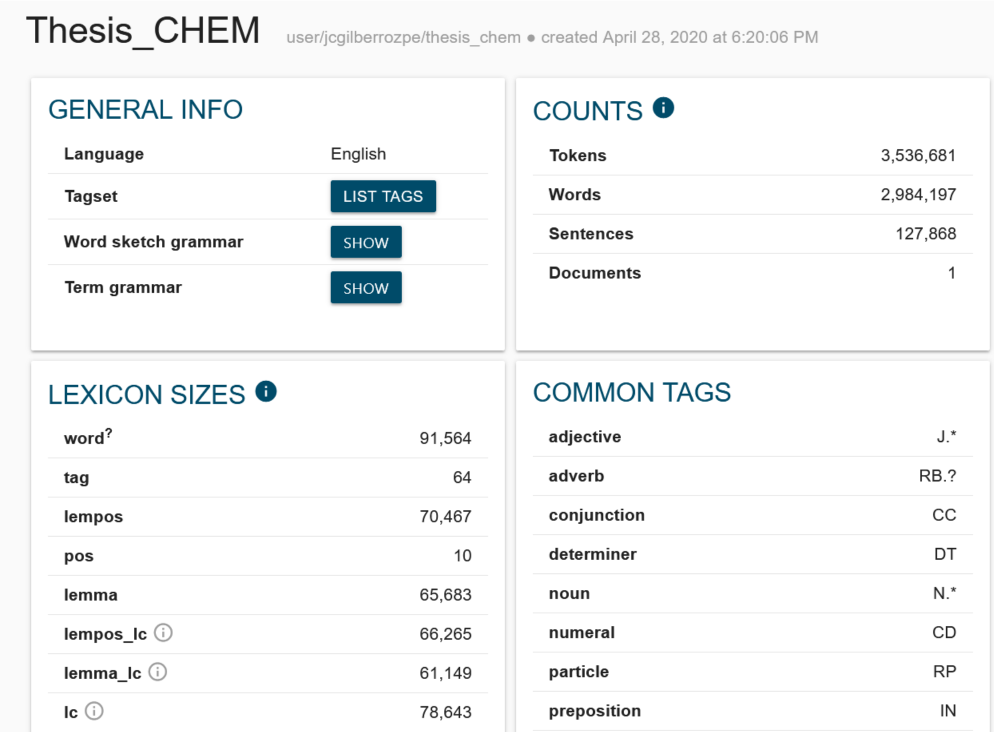


Figure 52. CHEM subcorpus information in Sketch Engine

4.1.1.3. Civil Engineering Subcorpus

The Civil Engineering (CIV) subcorpus contains 4,491,909 words according to the statistics in Sketch Engine (**Figure 53**). It is composed of specialized texts in the domains of Transport and Infrastructure Engineering, Hydraulic Engineering, Mining Engineering, Waste Management, Water Treatment and Supply, Air Quality Management, and Soil Quality Management. In total, it comprises 221 documents for experts, semi-experts, and the general public. The documents in this subcorpus include scientific articles or journal papers, specialized books, lexicographic material, official reports, and educational material.

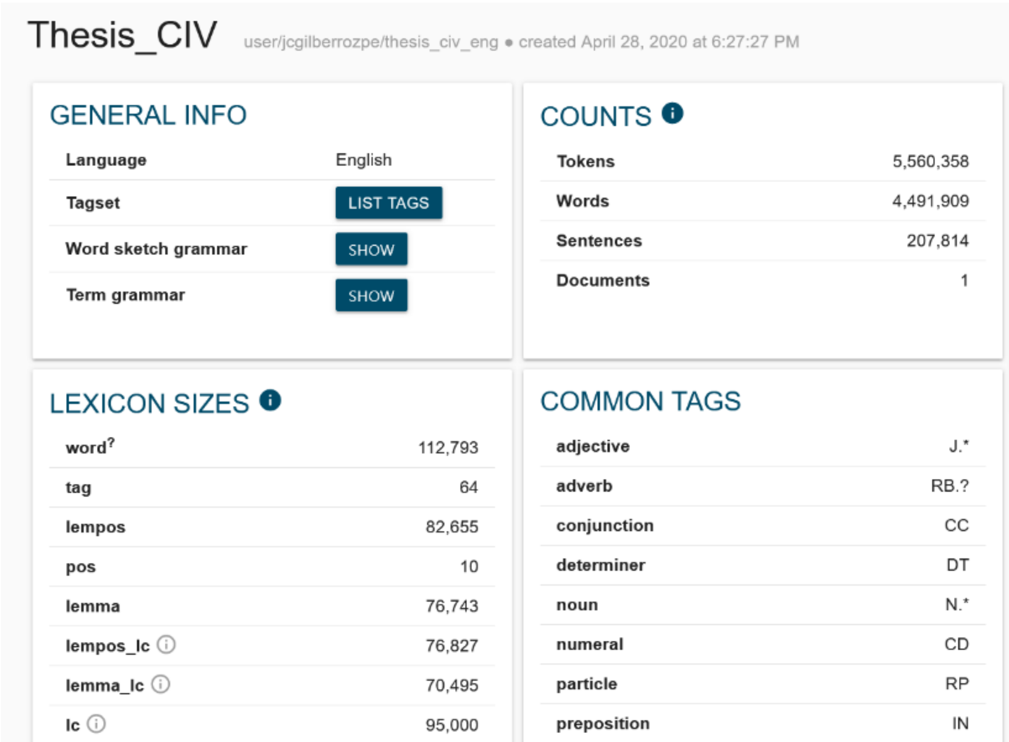


Figure 53. CIV subcorpus information in Sketch Engine

4.1.1.4. GEOLOGY SUBCORPUS

The Geology (GEO) subcorpus contains 3,975,045 words according to the statistics in Sketch Engine (**Figure 54**). It is a collection of specialized texts in the domains of general Geology, Hydrogeology, Geophysics, Geochemistry, Geological Oceanography, and Geomorphology. It has 107 documents for experts, semi-experts, and the general public. The documents in this subcorpus are scientific articles or journal papers, specialized books, lexicographic material, educational texts, and newspaper articles.

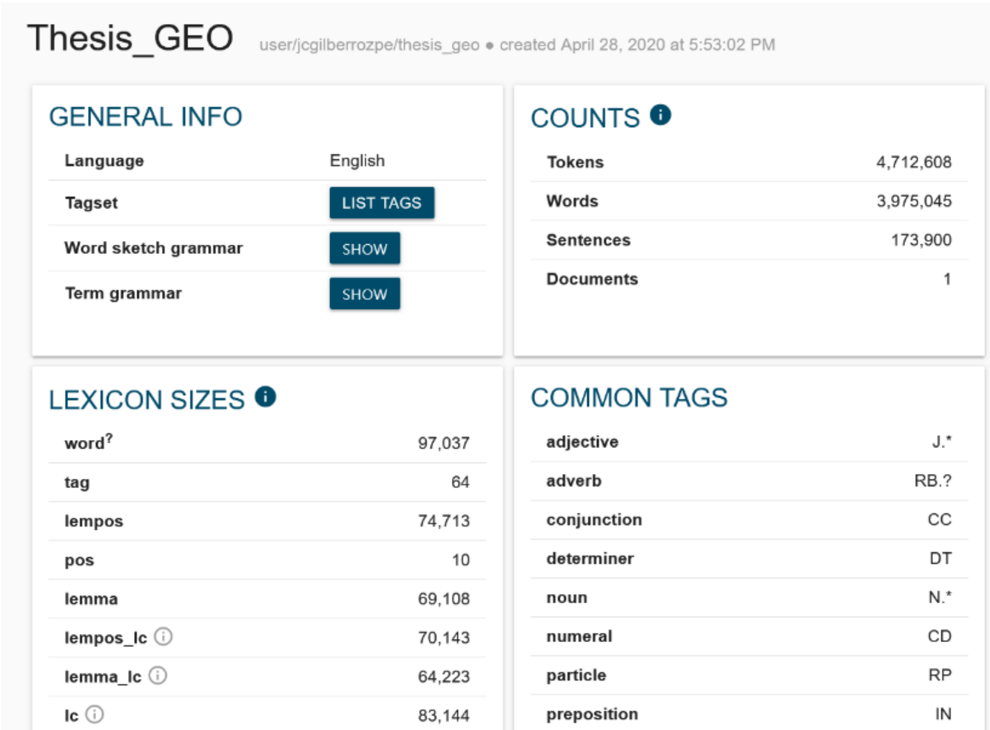


Figure 54. GEO subcorpus information in Sketch Engine

4.1.2. Specialized Terminological Resources

The four environmental subcorpora (BIO, CHEM, CIV, and GEO) were the basis for the corpus analysis of the research. However, they were not the only source of linguistic information. To create the conceptual hierarchies, build the terminological definitions, and hierarchically structure the entries, additional sources were needed. A selection of terminological resources was thus used to obtain further definitional and relational information about the concepts involved.

EcoLexicon was at the core of this selection, and its terminological contents and conceptual hierarchies were taken into account – in particular, all elements related to the hypernyms and hyponyms extracted, identified, and selected from the environmental subcorpora. In other words, all of the original content in EcoLexicon, associated with the terms involved, was further expanded with the new information obtained through corpus analysis.

Furthermore, to describe these hypernyms and hyponyms in as much detail as possible, a series of additional specialized lexicographic resources (i.e., dictionaries and encyclopedias) were also selected. These resources were used to extract more linguistic information (namely, definitions, synonyms, and context regarding the structure of the entries), and to expand and validate the data obtained from corpus analysis and the original conceptual hierarchies in EcoLexicon. The 25 environmental lexicographic resources used were the following:

- *A Comprehensive Dictionary of Chemistry* (Willie 2010)
- *A Dictionary of Biology* (Hine 2019)
- *A Dictionary of Chemistry* (Daintith 2008)
- *A Dictionary of Construction, Surveying and Civil Engineering* (Gorse *et al*. 2012)
- *A Dictionary of Geology and Earth Sciences* (Allaby 2013)
- *A Dictionary of Mechanical Engineering* (Atkins & Escudier 2019)
- *A Dictionary of Science* (Martin 2010)
- *Dictionary of Chemistry* (McGraw-Hill 2003a)
- *Dictionary of Civil Engineering* (Kurtz 2004)
- *Dictionary of Geology & Mineralogy* (McGraw-Hill 2003b)
- *Dictionary of Microbiology and Molecular Biology* (Singleton & Sainsbury 2006)
- *Encyclopedia of Biological Chemistry* (Lennarz & Lane 2013)
- *Encyclopedia of Biology* (Rittner & McCabe 2004)
- *Encyclopedia of Cell Biology* (Bradshaw & Stahl 2016)
- *Encyclopedia of Chemistry* (Rittner & Bailey 2005)
- *Encyclopedia of Engineering Geology* (Bobrowsky & Marker 2018)
- *Encyclopedia of Environmental Science and Engineering* (Pfafflin & Ziegler 2012)
- *Encyclopedia of Geology* (Selley *et al*. 2005)
- *Encyclopedia of Molecular Biology* (Creighton 1999)
- *Encyclopedia of Molecular Cell Biology and Molecular Medicine* (Meyers 2004)
- *Environmental Geology* (LaMoreaux 2019)
- *Henderson's Dictionary of Biology* (Lawrence 2008)
- *The Dictionary of Cell and Molecular Biology* (Lackie 2013)
- *The Wiley Dictionary of Civil Engineering and Construction* (Webster 1997)
- *Ullmann's Encyclopedia of Industrial Chemistry* (Wiley-VCH 2011)

4.1.3. SOFTWARE

The software applications used in the research were the following: (i) the EcoLexicon internal application, used for compiling the environmental subcorpora; (ii) Sketch Engine, used for processing the subcorpora and extracting all the hyponymic information; and (iii) Lexonomy, used for designing the hyponymy-based terminological template.

4.1.3.1. EcoLexicon Internal Application

The EcoLexicon internal application[6] (León-Araúz *et al*. 2018) is a private tool only available to the members of the LexiCon Research Group. It has two main functionalities: entry management and corpus management. Entries can be managed by adding, modifying, and deleting information regarding concepts, terms, definitions, conceptual relations, and other linguistic information. It is also possible to query the EEC and to add, modify, and delete information associated with the texts in the corpus.

Within the EcoLexicon internal application, the EEC can also be queried in regard to pragmatic restrictions such as author, date of publication, target reader, contextual domain, and keywords. This is possible because each text in the EEC is tagged with the following XML-based metadata: (i) domain; (ii) user; (iii) geographical variant; (iv) genre; (v) publisher; (vi) year; and (vii) country (Faber 2022).

Figure 55 depicts the main menu in the EcoLexicon internal application. From here, members of the Lexicon Research Group can search for a particular entry or directly click a button to add a new concept or term. There is also an option to access the corpus management section, which is independent of the main view.

Figure 55. Main menu in the EcoLexicon internal application

[6] Available at: http://ecolexicon.ugr.es/en/private.htm

4.1.3.2. SKETCH ENGINE

Sketch Engine[7] (Kilgarriff *et al*. 2004, 2014) is a web-based corpus management and analysis software. It can be used to identify the grammatical relations in which a term participates by means of *word sketches*. Sketch Engine has a wide variety of modules and features, but this description is only focused on those used in this research. **Figure 56** shows the main menu in Sketch Engine, where all features available are displayed.

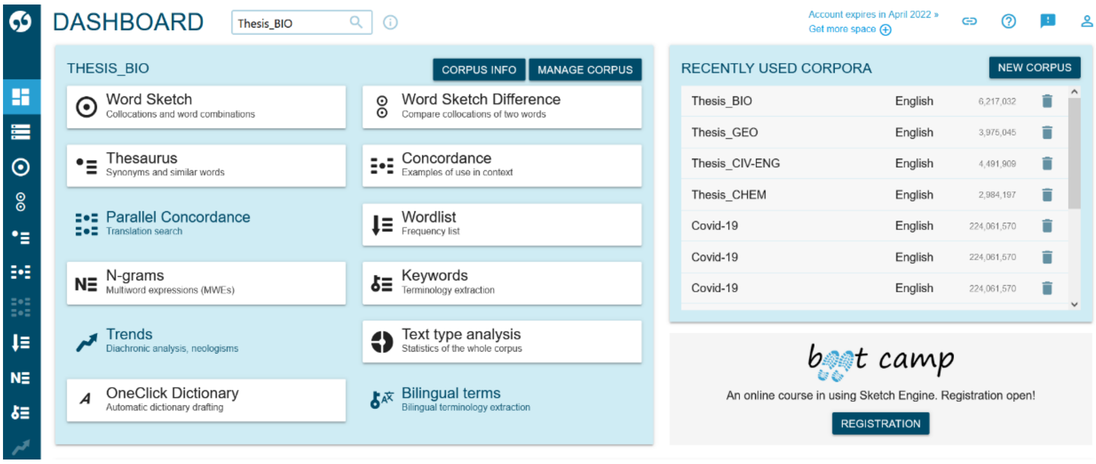


Figure 56. Main menu in Sketch Engine

Sketch Engine contains a wide range of pre-loaded corpora. It also offers the possibility to automatically create a corpus with texts from the web by using the *WebBootCat* tool. For this research, the four environmental subcorpora extracted from the EEC (**§4.1.1**) were uploaded to Sketch Engine. Once a corpus is fully uploaded, the user can compile the corpus, which includes applying a part-of-speech (PoS) tagger, a term grammar, and a sketch grammar. The term grammar used was the *English (TreeTagger – PennTB) for term extraction 2.3*, and the sketch grammar was a customized template based on the ESSG, described in **§3.2.3.2**. Apart from these settings, we also used the following features: (i) *Concordance*; (ii) *Keywords*; and (iii) *Word Sketch*.

The *Concordance* feature allows users to perform searches within a corpus and obtain concordance lines for word or a combination of words. However, it is also possible to perform advanced queries by using CQL to search for morphosyntactic tags, lemmas, and regular expressions (RegEx). After showing the concordances, the

[7] Available at: https://www.sketchengine.eu/

user can also apply different filtering and sorting parameters, analyze frequencies, obtain collocations, see the distribution of hits within the corpus, show results as keywords-in-context (KWIC) or full sentences, and download the concordance results in different formats.

The *Keywords* feature allows users to extract the most relevant terms from a corpus. Lists of terms from a focus corpus are thus extracted and compared with a reference corpus. The results are classified according to keyness score, which determines the specialization of a term with a *simple maths* formula. Both single-word terms (SWTs) and multi-word terms (MWTs) can be identified with the *Keywords* feature.

The *Word Sketch* feature allows for generating a one-page summary of a word's grammatical and collocational behavior based on corpus analysis (Kilgarriff *et al*. 2014). This is known as a word sketch (WS). Each WS contains columns with lists of words that co-occur with the word inserted in the query and which follow a predefined pattern set in the sketch grammar used to compile the corpus. There is a default sketch grammar available, but customized sketch grammars can also be uploaded and used when compiling any corpus. Some of the most basic patterns in the default sketch grammar are the following:

- *object_of* – verbs to which the noun acts as an object
- *subject_of* – verbs to which the noun acts as a subject
- *modifier* – words (e.g., adjectives, nouns, verbs) modifying the noun
- *modifies* – words (e.g., adjectives, nouns, verbs) modified by the noun
- *and/or* – another noun co-occurring with the noun as part of a listing

4.1.3.3. LEXONOMY

Lexonomy[8] (Měchura 2017; Rambousek *et al*. 2021) is a cloud-based, open-source platform for writing and publishing dictionaries. It is an easy-to-use tool for small to medium-sized dictionary projects. Creating a dictionary involves naming the resource, designing an arbitrary XML structure for the entries, editing its entries, and eventually making the dictionary publicly available on the Lexonomy website. In comparison to other tools, Lexonomy offers the following advantages: (i) since it is cloud-based, it requires no previous installation or set-up; (ii) its intuitive interface

[8] Available at: https://www.lexonomy.eu/

makes it possible to create and edit the entries without any knowledge of coding or programming; and (iii) it is free. **Figure 57** shows the homepage of Lexonomy, where users can log in to access their dictionaries.

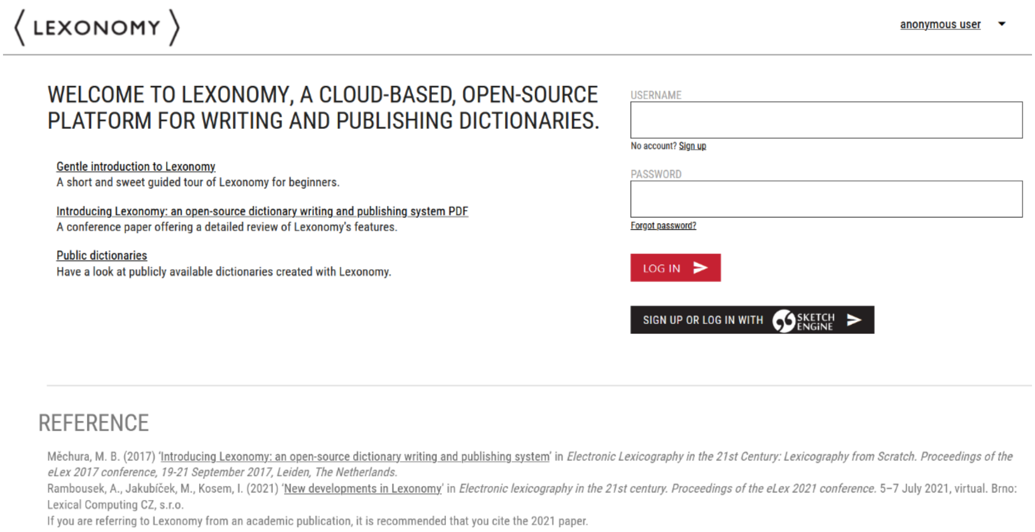


Figure 57. Homepage of Lexonomy

Terminological entries in Lexonomy are stored as XML documents and their structure is defined by a schema which is exclusive to each dictionary in the platform. When creating a new dictionary, users can use a predefined schema (e.g., monolingual dictionary, bilingual dictionary) and customize it afterwards, or directly start from scratch by designing the schema themselves.

According to Měchura (2017: 663), a Lexonomy schema resembles a Document Type Definition (DTD). It lists the XML elements which can appear in the entries and specifies how they can be nested, how many of them must or may be there, which attributes they may or must have, what their values may be, and so on. As an added advantage, Lexonomy offers a visual schema editor where users can define the structure of their entries without having to hand-code anything.

4.2. METHODS

The methodology of this research was based on the previously described materials and was carried out in four states: (i) extracting and compiling the four environmental subcorpora; (ii) analyzing the subcorpora; (iii) creating the conceptual hierarchies; and (iv) designing the terminological template for the description and representation of all hyponymic information.

4.2.1. Subcorpora Extraction and Compilation

With the EEC as the starting point, four specialized subcorpora were extracted: BIO, CHEM, CIV, and GEO. For this purpose, the EcoLexicon internal application was used. Inside the EcoLexicon internal application, the corpus management section was accessed, specifically a subsection designed for filtering and searching the EEC (**Figure 58**).

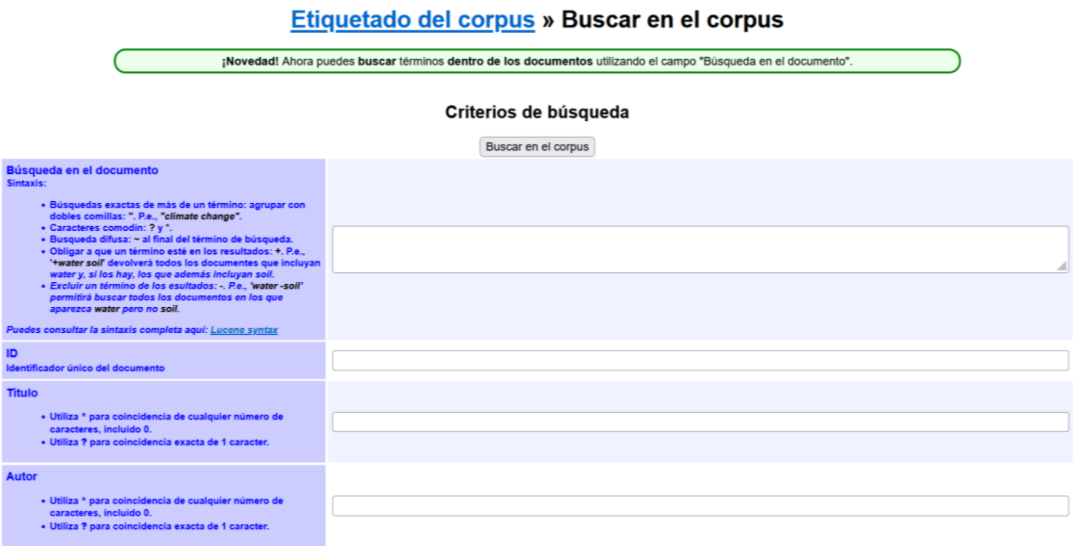


Figure 58. Corpus management section in the EcoLexicon internal application

Four independent queries were then performed so as to obtain each specialized subcorpus. For instance, **Figure 59** shows the results of the corpus extraction process to obtain the texts belonging to the CIV subcorpus. In this case, the search criteria were the following: "language = English; domains = Transport and Infrastructure Engineering, Hydraulic Engineering, Mining Engineering, Waste Management, Water Treatment and Supply, Air Quality Management, Soil Quality Management; quarantine = excluded". In this way, 221 documents were obtained. The same process was followed to extract the text files for the BIO, CHEM, and GEO subcorpora.

Etiquetado del corpus » Buscar » Resultados búsqueda

Criterios de búsqueda

lengua = inglés, dominio1 = 3.2.1 Transport and Infrastructure Engineering, dominio8 = 3.2.2 Hydraulic Engineering, dominio9 = 3.2.4 Mining Engineering, dominio10 = 3.2.5.1 Waste Management, dominio11 = 3.2.5.2 Water Treatment and Supply, dominio12 = 3.2.5.3 Air Quality Management, dominio13 = 3.2.5.4 Soil Quality Management, cuarentena = excluidos

Cambiar criterios

Se han encontrado **221 documentos** correspondientes a los criterios de búsqueda, que hacen un total de **28.79 MB** y **4 560 334 palabras**

« anterior 1 2 3 4 5 6 7 8 9 10 siguiente »

Seleccionar todos Deseleccionar todos 0 documentos seleccionados.

ID	Título del documento	Autor	Editor	Dominios	Lengua	Género	Usuario/lector	Subido por	Tamaño	Documento	Ref. externa	¿Descargar?
225	An improved calculation model for the wave-induced pore pressure distribution in a rubble-mound breakwater core	Dieter Vanneste, Peter Troch	Académico/ Investigador/ Educativo	3.2.2 Hydraulic Engineering, 3.2.3 Coastal Engineering	inglés	artículo	Experto	Pamela	62.4KB	XML · TXT	link	☐
556	Encyclopedia of Global Warming	Steven I. Dutch	Académico/ Investigador/ Educativo	2.7.2 Climatology, 3.2.5.3 Air Quality Management	inglés, Inglés norteamericano	material lexicográfico	Público general/lego	Pamela	4.23MB	XML · TXT	link	☐
625	Highway Meteorology	A.H.Perry and L.J.Symons	Empresas y ONG	2.7.1 Meteorology, 3.2.1 Transport and Infrastructure Engineering	inglés, Inglés británico	libro/monografía	Experto	Pamela	351.74KB	XML · TXT	link	☐
630	WHERE THE WEATHER MEETS THE ROAD: A RESEARCH AGENDA FOR IMPROVING ROAD WEATHER SERVICES	dk	Académico/ Investigador/ Educativo	2.7.1 Meteorology, 3.2.1 Transport and Infrastructure Engineering	inglés, Inglés norteamericano	libro/monografía	Semi-experto	Pamela	339.28KB	XML · TXT	link	☐
654	Modeling the effects of permeable and reflective structures on waves and nearshore flows	Inigo J. Losada	Académico/ Investigador/ Educativo	3.1 Marine Engineering, 3.2.2 Hydraulic Engineering, 3.2.3 Coastal Engineering	inglés	artículo	Experto	Pamela	58.09KB	XML · TXT	link	☐
664	Modeling shore platforms: present status and future developments	Alan S. Trenhaile	Académico/ Investigador/ Educativo	3.1 Marine Engineering, 3.2.2 Hydraulic Engineering, 3.2.3 Coastal Engineering	inglés	artículo	Experto	Pamela	49.19KB	XML · TXT	link	☐
684	Advances in water quality modeling in the coastal environment	Mark S. Dortch	Académico/ Investigador/ Educativo	1 Environmental protection, 2.6 Chemistry, 2.10 Oceanography, 3.2.3 Coastal Engineering, 3.2.5.2 Water Treatment and Supply	inglés	artículo	Experto	Pamela	37.52KB	XML · TXT	link	☐
688	An ecological perspective on the deployment and design of low-crested and other hard coastal defence structures	L. Airoldi, M. Abbiati, M.W. Beck, S.J. Hawkins, P.R. Jonsson, D. Martin, P.S. Moschellad, A. Sundelo, R.C. Thompson,P. Aberg	Académico/ Investigador/ Educativo	3.2.2 Hydraulic Engineering, 3.2.3 Coastal Engineering	inglés	artículo	Experto	Pamela	57.92KB	XML · TXT	link	☐

Figure 59. Corpus extraction in the EcoLexicon internal application

The four environmental subcorpora were uploaded to Sketch Engine. This software was then used to compile the subcorpora by applying the *English (TreeTagger – PennTB) for term extraction 2.3* term grammar and a customized sketch grammar based on the ESSG. This sketch grammar followed the FBT methodology for knowledge extraction based on KPs and KRCs. Moreover, this customized sketch grammar with hyponymic KP-based rules was extremely important for the analysis of the subcorpora, particularly for hyponym extraction and identification.

4.2.2. Subcorpora Analysis

Terminology work is based on corpus analysis for the extraction of conceptual information because most of the knowledge shared by experts is expressed in real texts (Bourigault & Slodzian 1999). Corpus analysis is thus commonly used by terminologists to find terms in large corpora and extract their syntactic and semantic information. Corpus analysis also identifies conceptual information that can afterwards be classified and analyzed so as to characterize concepts within their activation frame (León-Araúz *et al*. 2012).

In the past, the only way to analyze corpus information for terminological work was to manually read concordance lines. This was inefficient because a single term could generate thousands of concordance lines, many without any useful information. León-Araúz *et al*. (2016: 73) state that this time-consuming task led to the development of new corpus-based methods and applications to analyze and extract linguistic information.

With modern corpus analysis tools such as Sketch Engine, it is possible to perform CQL queries or apply KP-based sketch grammars to identify KRCs. Authors such as Barbero (2022) have explored the application of CQL grammars for lexical and semantic information extraction. By identifying KRCs it is possible to extract term-oriented knowledge about a concept from a corpus and then use that information to provide a starting point for any terminological purpose (Bielinskiene *et al*. 2012).

The corpus analysis in this research consisted of three steps: (i) hypernym extraction, identification, and selection; (ii) hyponym extraction and identification; and (iii) hyponym selection. With the four environmental subcorpora compiled in Sketch Engine, it was possible to obtain all the relevant information for this study.

4.2.2.1. HYPERNYM EXTRACTION, IDENTIFICATION, AND SELECTION

Hypernym extraction, identification, and selection was based on the *Keywords* function in Sketch Engine, which extracts the most relevant SWTs and MWTs from a corpus. We thus identified the three most representative terms of each subcorpora, which became the candidate hypernyms for each domain (BIO, CHEM, CIV, and GEO). Three was considered the optimal number because the objective was to create twelve terminological entries with sufficient conceptual, relational, and contextual information for the ontological categories in each environmental domain.

The advanced query of the *Keywords* function was used (**Figure 60**). This option allowed us to apply different criteria (e.g., focus corpus, reference corpus, rarity, minimum frequency, maximum frequency, etc.) to refine the query.

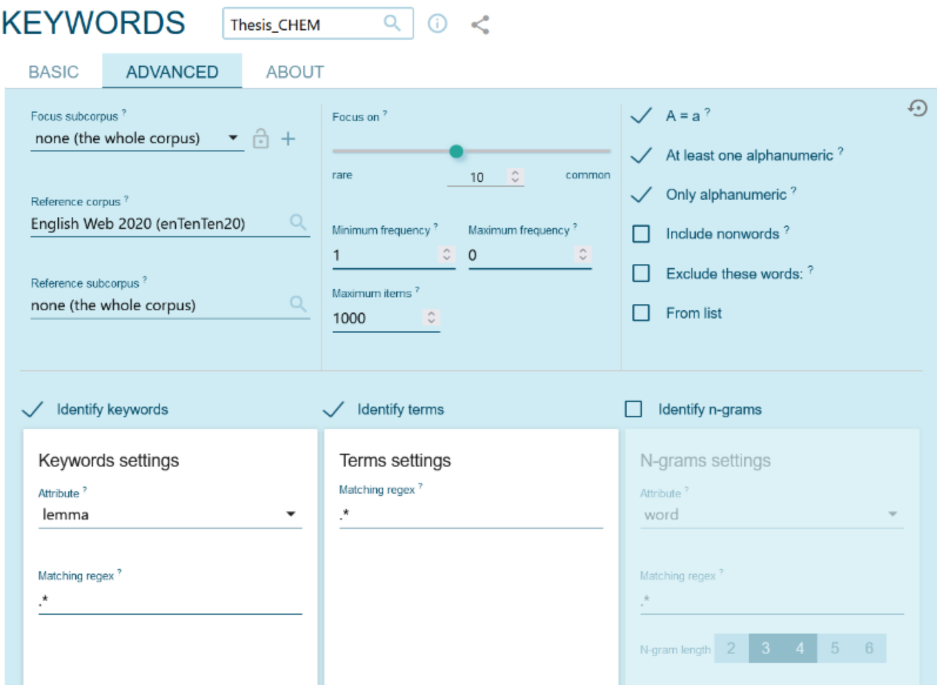
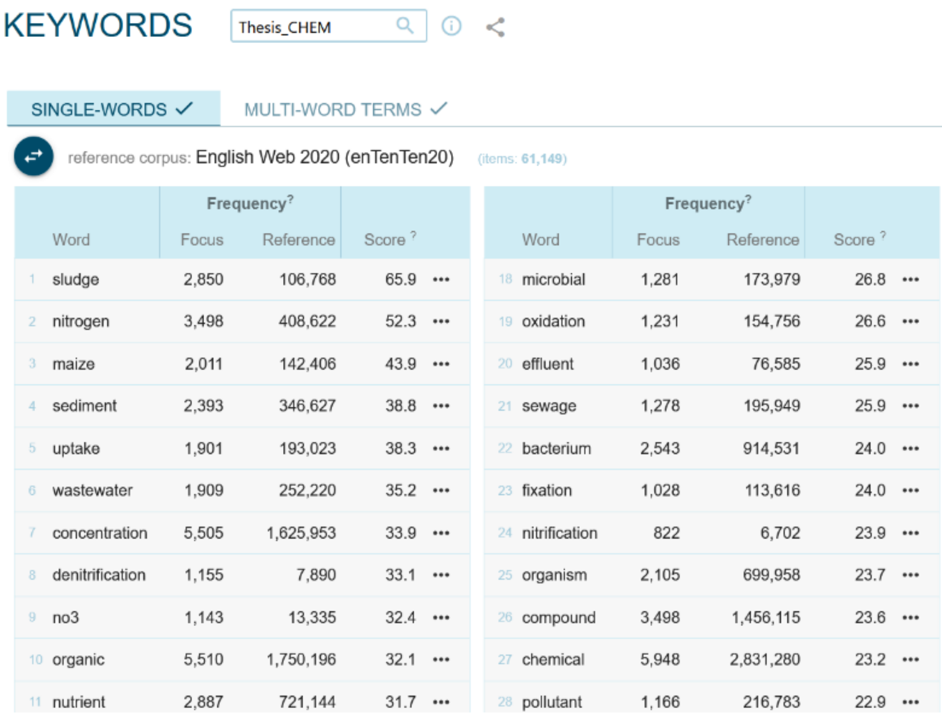


Figure 60. Keywords function in Sketch Engine: query

The results are then given in two tabs (SWTs and MWTs) in the form of columns where it is possible to filter the information (e.g., hits in the focus corpus, hits in the reference corpus, keyness score, etc.) (**Figure 61**).



	Word	Frequency Focus	Frequency Reference	Score		Word	Frequency Focus	Frequency Reference	Score
1	sludge	2,850	106,768	65.9	18	microbial	1,281	173,979	26.8
2	nitrogen	3,498	408,622	52.3	19	oxidation	1,231	154,756	26.6
3	maize	2,011	142,406	43.9	20	effluent	1,036	76,585	25.9
4	sediment	2,393	346,627	38.8	21	sewage	1,278	195,949	25.9
5	uptake	1,901	193,023	38.3	22	bacterium	2,543	914,531	24.0
6	wastewater	1,909	252,220	35.2	23	fixation	1,028	113,616	24.0
7	concentration	5,505	1,625,953	33.9	24	nitrification	822	6,702	23.9
8	denitrification	1,155	7,890	33.1	25	organism	2,105	699,958	23.7
9	no3	1,143	13,335	32.4	26	compound	3,498	1,456,115	23.6
10	organic	5,510	1,750,196	32.1	27	chemical	5,948	2,831,280	23.2
11	nutrient	2,887	721,144	31.7	28	pollutant	1,166	216,783	22.9

Figure 61. Keywords function in Sketch Engine: results

The four environmental subcorpora were processed by comparing them with the English Web 2020 (enTenTen20) general corpus in Sketch Engine. Moreover, empty words and non-terms were excluded. The four queries were performed according to different rarity levels (BIO: 50; CHEM: 10; CIV: 10; GEO: 100) so as to obtain the three best hypernyms because the texts in each subcorpus had different levels of specificity.

Once the four subcorpora were processed, the results were sorted by keyness score, which follows a *simple maths* formula to determine the specificity of a term. In other words, the higher its keyness score is, the more frequently a term occurs in the focus corpus in comparison with the reference corpus. In addition to the empty words and non-terms that were previously omitted, repetitions, acronyms, and generic hypernyms also had to be manually discarded. The results of each query were then identified.

Finally, the three candidate terms of each subcorpus with the highest keyness score were selected as the three hypernyms. They were BACTERIUM, REEF, and CELL in the BIO subcorpus (**Table 11**); SLUDGE, NITROGEN, and MAIZE in the CHEM subcorpus (**Table 12**); WASTEWATER, BREAKWATER, and POLLUTANT in the CIV subcorpus (**Table 13**); and EARTHQUAKE, SEDIMENT, and SOIL in the GEO subcorpus (**Table 14**).

	TERM	FOCUS	REFERENCE	KEYNESS SCORE
1	**bacterium**	8,210	914,531	16.5
2	**reef**	6,001	586,941	13.7
3	**cell**	17,853	7,040,791	12.0
4	gene	8,154	2,143,942	11.8
5	habitat	6,190	1,169,695	11.7
6	plant	18,068	7,917,054	11.1
7	protein	7,924	2,375,619	10.9
8	rice	6,010	1,322,354	10.9
9	DNA	5,870	1,299,505	10.7
10	fish	9,311	3,933,526	9.5

Table 11. Hypernym identification and selection in the BIO subcorpus. [Focus corpus: Thesis_BIO | Reference corpus: English Web 2020 | Rarity: 50 | Sorted by keyness score]

	TERM	FOCUS	REFERENCE	KEYNESS SCORE
1	**sludge**	2,850	106,768	65.9
2	**nitrogen**	3,498	408,622	52.3
3	**maize**	2,011	142,406	43.9
4	uptake	1,901	193,023	38.3
5	concentration	5,505	1,625,953	33.9
6	denitrification	1,155	7,890	33.1
7	nutrient	2,887	721,144	31.7
8	rhizosphere	1,011	5,504	29.2
9	oxygen	2,821	832,065	28.3
10	phytoplankton	1,049	42,174	28.0

Table 12. Hypernym identification and selection in the CHEM subcorpus. [Focus corpus: Thesis_CHEM Ι Reference corpus: English Web 2020 Ι Rarity: 10 Ι Sorted by keyness score]

	TERM	FOCUS	REFERENCE	KEYNESS SCORE
1	**wastewater**	7,204	252,220	83.6
2	**breakwater**	2,584	29,710	44.5
3	**pollutant**	2,351	216,783	29.2
4	adsorption	1,679	54,650	27.8
5	deposition	2,316	265,816	26.8
6	nitrogen	2,644	408,622	25.4
7	emission	6,390	1,642,319	24.9
8	wave	10,456	2,975,621	24.8
9	aeration	1,420	31,974	24.8
10	oxygen	3,662	832,065	23.5

Table 13. Hypernym identification and selection in the CIV subcorpus. [Focus corpus: Thesis_CIV Ι Reference corpus: English Web 2020 Ι Rarity: 10 Ι Sorted by keyness score]

	TERM	FOCUS	REFERENCE	KEYNESS SCORE
1	**earthquake**	6,292	734,916	12.3
2	**sediment**	4,698	346,627	10.2
3	**soil**	6,659	2,469,621	9.8
4	wave	5,901	2,975,621	8.1
5	earth	7,226	5,210,790	7.6
6	surface	7,220	5,472,604	7.4
7	rock	6,424	5,011,777	6.9
8	water	14,184	17,935,266	6.2
9	ecosystem	3,034	937,880	6.2
10	velocity	2,776	591,949	6.1

Table 14. Hypernym identification and selection in the GEO subcorpus. [Focus corpus: Thesis_GEO Ι Reference corpus: English Web 2020 Ι Rarity: 100 Ι Sorted by keyness score]

4.2.2.2. HYPONYM EXTRACTION AND IDENTIFICATION

Hyponym extraction and identification was based on the *Word Sketch* (WS) function in Sketch Engine, which provides summaries of a term's grammatical and collocational behavior. This selection was validated and expanded by CQL queries performed with the *Concordance* function.

The advanced query of the WS feature was used (**Figure 62**). This option was used to select the part of speech of the word and to apply different criteria (e.g., maximum frequency, minimum score) so as to obtain better results.

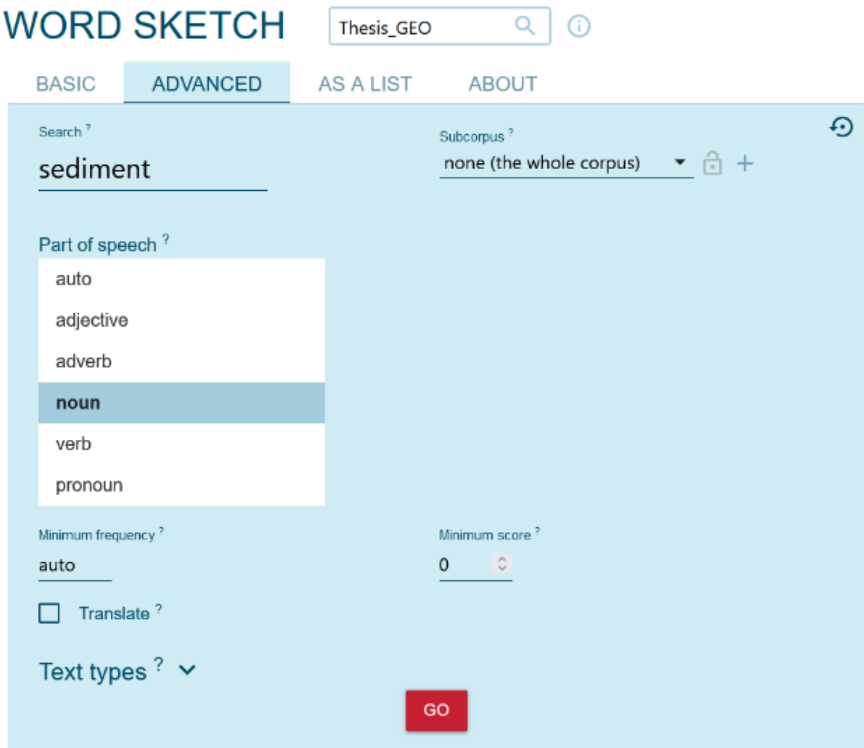


Figure 62. WS function in Sketch Engine: query

Figure 63 shows the results of the WS. Independent sheets show the collocates of the search term obtained with different KP-based rules (e.g., *"X" is the generic of*, *MWterms_Head*, *modifier*). As previously stated, the customized sketch grammar (León-Araúz *et al.* 2016) was used to obtain valuable results regarding hyponymic information.

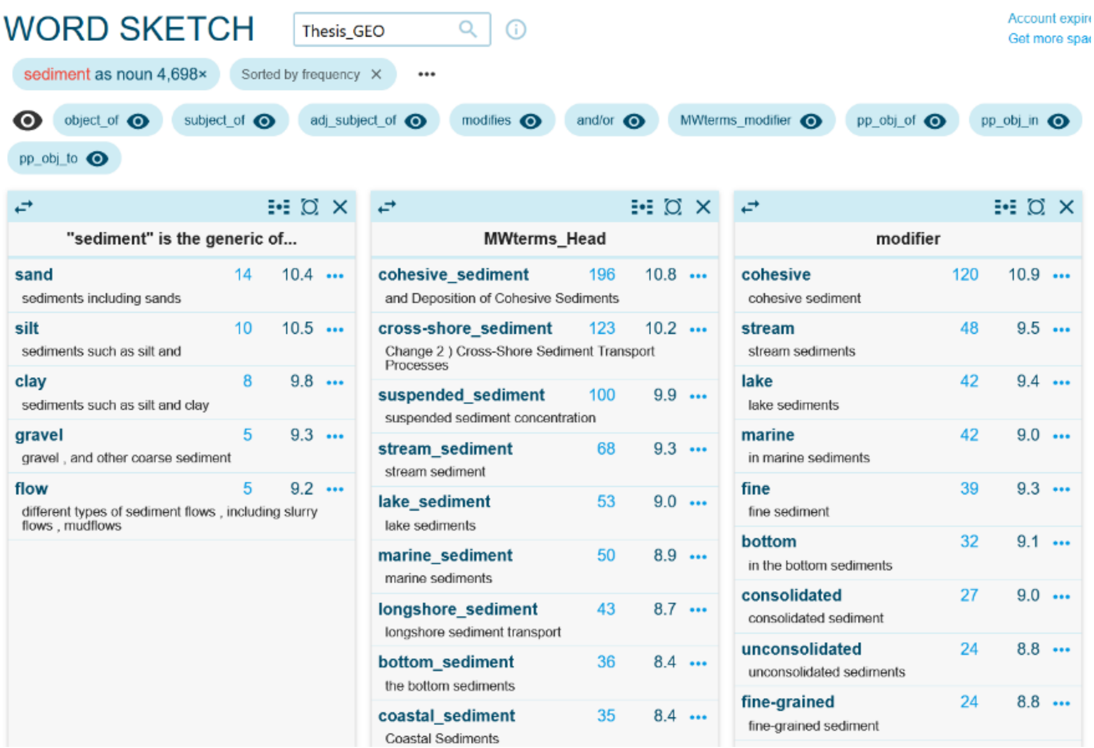


Figure 63. WS function in Sketch Engine: results

The WS queries were performed for each of the twelve hypernyms extracted in the previous step of the corpus analysis (BIO: BACTERIUM, REEF, and CELL; CHEM: SLUDGE, NITROGEN, and MAIZE; CIV: WASTEWATER, BREAKWATER, and POLLUTANT; GEO: EARTHQUAKE, SEDIMENT, and SOIL). The results of the WSs were sorted by typicality score (LogDice) – the higher the score, the stronger the collocation indicated in the WS.

On the one hand, the *"X" is the generic of* WS was used to extract hyponymic SWTs and certain hyponymic MWTs. Since the hyponyms obtained with this WS were not numerous, they were all extracted for each hypernym. On the other hand, the *MWterms_Head* WS (MWT WS) and the *modifier* WS were used to extract hyponymic MWTs. In this case, the hyponyms extracted with these two WSs were limited to the first thirty results regardless of their frequency so as to limit the retrieval to the most representative examples.

Furthermore, a customized CQL search using the *Concordance* function was employed to validate and expand the hyponym lists extracted with the WSs. This search focused on obtaining hyponymic MWTs as well, and was based on Cabezas-García's (2019, 2020) custom CQL for MWTs with one or more modifiers:

[tag="N.*|JJ.*|RB.*|VVN.*|VVG.*"]{1,}[lemma="**HYPERNYM**"]

Accordingly, this CQL query searches for any lemmatized hypernym ([lemma="hypernym"]) preceded one or more times ({1,}) by any noun (N.*), adjective (JJ.*), adverb (RB.*), verb in past participle form (VVN.*), or verb in gerund or present participle form (VVG.*). In order to perform this search, the advanced query of the *Concordance* feature was used (**Figure 64**). This option allows users to carry out different types of queries (e.g., simple, lemma, phrase, word, character), including CQL queries, where the possibility of building the CQL through an accessible and intuitive interface is given.

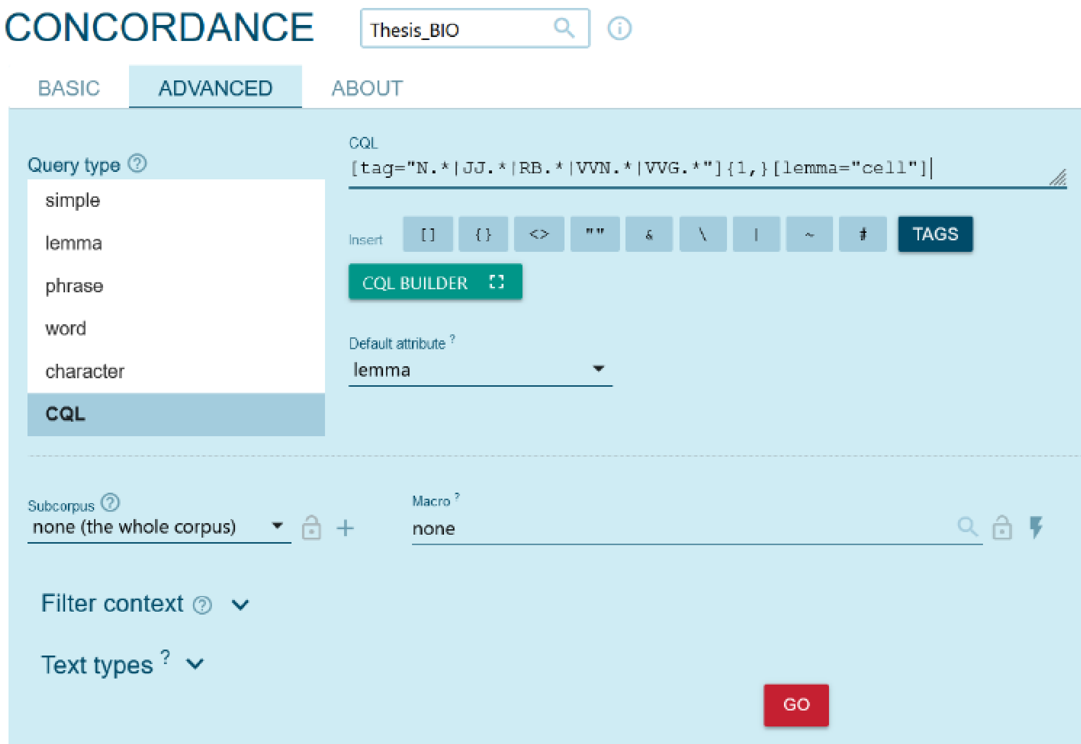


Figure 64. CQL concordance search in Sketch Engine: query

The results are then displayed as concordance lines (**Figure 65**).

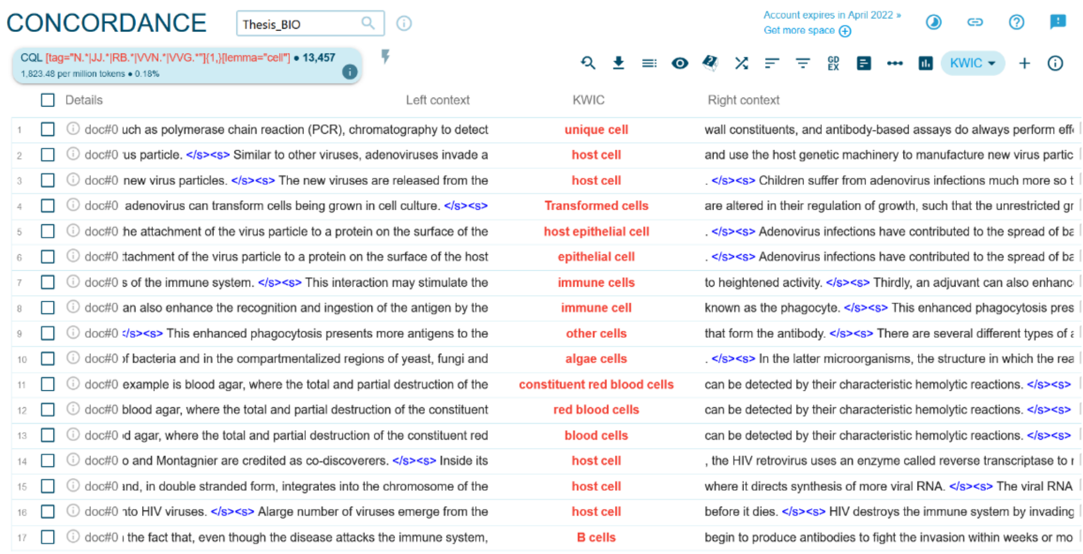


Figure 65. CQL concordance search in Sketch Engine: results

Nevertheless, there are different options to filter them. In this case, we chose the frequency option for sorting the results by lemma in the form of KWIC (**Figure 66**). Finally, the first thirty results independent of their frequency were retrieved for each of the twelve hypernyms.

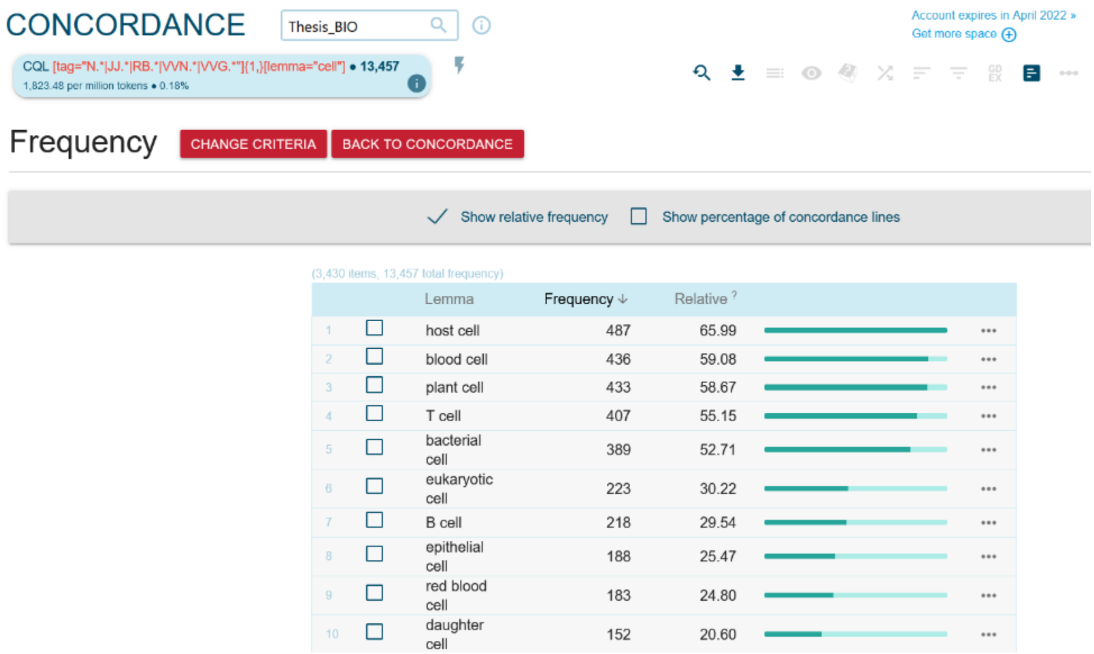


Figure 66. CQL concordance search in Sketch Engine: filtering by frequency

While performing these processes, it was necessary to manually discard non-terms, duplicates, repetitions, and false positives. In this way, it was possible to identify all the relevant hyponyms according to the criteria established.

As an illustration of the logic behind the manual discard process, an example of false positive was a combination of words composed of a generic non-defining modifier and the term involved; for instance, STRONG EARTHQUAKE or DIFFERENT BACTERIUM. Another typical case of false positives involved modifiers referring to specific locations, as in SAN FRANCISCO EARTHQUAKE and ALASKAN EARTHQUAKE.

4.2.2.3. HYPONYM SELECTION

Hyponym selection involved classifying and comparing all the hyponymic information retrieved in the previous extraction and selection process. The data from the WS queries were validated and further expanded with the data from the customized CQL queries.

In this subsection, the results of this process are shown for each hypernym. They are displayed in a table divided in two sections: (i) a section showing the selection of SWTs and certain MWTs, using the *"X" is the generic of* WS, where hyponyms are listed by frequency and typicality score; and (ii) a section showing the MWTs selected by using a combination of the MWT WS, the *modifier* WS and the MWT CQL query. Here, hyponyms are also listed by frequency and typicality score (along with the method used to identify them). Also provided is a comparison with the hyponyms in the already existing hierarchies in EcoLexicon.

4.2.2.3.1. HYPONYMS OF BACTERIUM (BIO-1)

Table 15 shows the hyponyms of the hypernym BACTERIUM, one of the concepts with the highest number of hyponyms. *BACTERIUM is the generic of* WS produced eleven hyponyms (e.g., ESCHERICHIA COLI, STREPTOCOCCUS, BACILLUS). The combination of the MWT WS, the *modifier* WS, and the MWT CQL query resulted in 21 hyponyms (e.g., GRAM-NEGATIVE BACTERIUM, HETEROTROPHIC BACTERIUM, GRAM-POSITIVE BACTERIUM).

	BACTERIUM **hyponyms** (*is the generic of* WS) [total frequency = 8,210]		
	TERM	**FREQUENCY**	**SCORE**
1	Escherichia coli (E. coli)	24	9.7
2	Streptococcus	18	9.3
3	Bacillus	15	9.1
4	Pseudomonas	14	9.0
5	Salmonella	10	8.5
6	Cyanobacterium	9	8.4
7	Mycobacterium	8	8.2
8	Haemophilus	6	7.8
9	Chromatium	6	7.8
10	Bacillus subtilis (B. subtilis)	6	7.8
11	Staphylococcus	6	7.8

	BACTERIUM **hyponyms** (MWT WS, modifier WS & MWT CQL) [total frequency = 8,210]			**FOUND WITH**		
	TERM	**FREQUENCY**	**SCORE**	**MWT WS**	**MODIFIER WS**	**MWT CQL**
1	gram-negative bacterium	197	10.2	X	X	X
2	heterotrophic bacterium	116	9.4	X	X	X
3	gram-positive bacterium	101	9.2	X	X	X
4	anaerobic bacterium	93	9.1	X	X	X
5	pathogenic bacterium	80	8.9	X	X	X
6	marine bacterium	74	8.8	X	X	X
7	nitrifying bacterium	70	8.7	X		X
8	enteric bacterium	66	8.6	X	X	X
9	soil bacterium	52	8.3	X	X	X
10	sulfur bacterium	51	8.3	X	X	X
11	photosynthetic bacterium	45	8.1	X	X	X
12	symbiotic bacterium	42	8.0	X	X	X
13	methanogenic bacterium	37	7.8	X	X	X
14	nitrogen-fixing bacterium	37	7.8	X	X	X
15	coliform bacterium	34	7.7	X	X	X
16	aerobic bacterium	31	7.6	X	X	X
17	oral bacterium	28	7.4	X	X	X
18	filamentous bacterium	27	7.4	X	X	X
19	lactic acid bacterium	25	7.3	X	X	X
20	AAnP bacterium	25	7.3	X	X	X
21	autotrophic bacterium	18	7.2		X	

Table 15. Hyponym selection of BACTERIUM (BIO-1)

In EcoLexicon, the original conceptual hierarchy of BACTERIUM (**Figure 67**) was not sufficiently developed. The concept BACTERIUM was associated with three hypernyms (i.e., FAUNA, MICROORGANISM, PATHOGEN), but there was only one hyponym (i.e., FACULTATIVE AEROBE) linked to it.

Figure 67. Original conceptual hierarchy of BACTERIUM in EcoLexicon

4.2.2.3.2. *HYPONYMS OF REEF (BIO-2)*

Table 16 shows the hyponym selection of the hypernym REEF. No hyponym could be extracted by using the *REEF is the generic of* WS. However, based on the combination of the MWT WS, the *modifier* WS, and the MWT CQL query, 15 hyponyms (e.g., CORAL REEF, FRINGING REEF, PATCH REEF) were selected.

REEF **hyponyms** (*is the generic of* **WS**) **[total frequency = 5,601]**			
	TERM	**FREQUENCY**	**SCORE**
-	NONE	-	-

REEF **hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 5,601]**				**FOUND WITH**		
	TERM	**FREQUENCY**	**SCORE**	**MWT WS**	**MODIFIER WS**	**MWT CQL**
1	coral reef	1,363	12.8	X	X	X
2	fringing reef	121	9.7	X		X
3	patch reef	117	9.6	X	X	X
4	barrier reef	47	8.3	X	X	X
5	outer reef	34	7.9	X	X	X
6	tropical reef	34	7.9	X	X	X
7	artificial reef	33	7.8	X	X	X
8	inner reef	21	7.2	X	X	X
9	shallow reef	21	7.2	X	X	X
10	coastal reef	16	6.8	X	X	X
11	rocky reef	16	6.8	X	X	X
12	natal reef	13	6.5	X	X	X
13	seaward reef	12	6.4	X		
14	offshore reef	9	6.9		X	
15	submerged reef	8	6.8		X	

Table 16. Hyponym selection of REEF (BIO-2)

In EcoLexicon, the original conceptual hierarchy of REEF (**Figure 68**) was well populated with different hyponymy levels. Eight hyponyms (e.g., COASTAL REEF, BARRIER REEF, FRINGING REEF) were located at the first level, three hyponyms (i.e., REEF PATCH, ATOLL, MESOAMERICAN BARRIER REEF SYSTEM) were at the second level, and one hyponym (i.e., COMPLETELY LAND RINGED ATOLL) was at the third level of specificity.

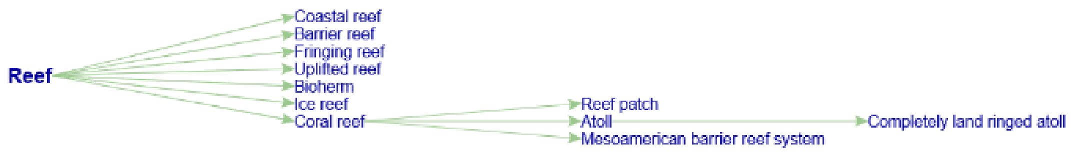


Figure 68. Original conceptual hierarchy of REEF in EcoLexicon

4.2.2.3.3. HYPONYMS OF CELL (BIO-3)

Table 17 shows the hyponym selection of the hypernym CELL. In accordance with the *CELL is the generic of* WS extraction and identification, only two hyponyms (i.e., MACROPHAGE, LYMPHOCYTE) could be selected. The combination of the MWT WS, the *modifier* WS, and the MWT CQL query resulted in twenty hyponyms (e.g., HOST CELL, BLOOD CELL, PLANT CELL).

CELL hyponyms (*is the generic of* WS) [total frequency = 17,853]			
	TERM	FREQUENCY	SCORE
1	macrophage	16	9.5
2	lymphocyte	13	9.2

CELL hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 17,853]				FOUND WITH		
	TERM	FREQUENCY	SCORE	MWT WS	MODIFIER WS	MWT CQL
1	host cell	490	10.0	X	X	X
2	blood cell	442	9.9	X	X	X
3	plant cell	435	9.8	X	X	X
4	T cell	407	9.7	X	X	X
5	eukaryotic cell	249	9.0	X	X	X
6	B cell	218	8.9	X	X	X
7	red blood cell	192	8.7	X	X	X
8	epithelial cell	188	8.7	X	X	X
9	daughter cell	152	8.3	X	X	X
10	white blood cell	150	8.3	X	X	X
11	nerve cell	111	7.9	X	X	X
12	target cell	101	7.8	X	X	X
13	mother cell	98	7.7	X	X	X
14	animal cell	91	7.6	X		X
15	epidermal cell	90	7.6	X	X	X
16	cancer cell	90	7.6	X	X	X
17	stem cell	85	7.5	X	X	X
18	muscle cell	73	7.3	X	X	X
19	granulosa cell	62	7.1	X		X
20	yeast cell	59	7.9			X

Table 17. Hyponym selection of CELL (BIO-3)

In EcoLexicon, the original conceptual hierarchy of CELL (**Figure 69**) did not have much content. The concept CELL was associated with three hyponyms (i.e., ZYGOTE, CHROMATOPHORE, EOSINOPHIL) at the same level of specificity.

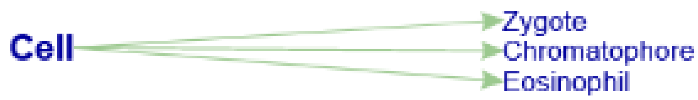


Figure 69. Original conceptual hierarchy of CELL in EcoLexicon

4.2.2.3.4. HYPONYMS OF SLUDGE (CHEM-1)

Table 18 shows the hyponym selection of the hypernym SLUDGE. No hyponym could be extracted by using the *SLUDGE is the generic of* WS. However, the combination of the MWT WS, the *modifier* WS, and the MWT CQL query produced 28 hyponyms (e.g., ACTIVATED SLUDGE, SEWAGE SLUDGE, RAW SLUDGE).

SLUDGE hyponyms (*is the generic of* WS) [total frequency = 2,850]			
	TERM	FREQUENCY	SCORE
-	NONE	-	-

SLUDGE hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 2,850]				FOUND WITH		
	TERM	FREQUENCY	SCORE	MWT WS	MODIFIER WS	MWT CQL
1	activated sludge	587	12.7	X		X
2	sewage sludge	251	11.6	X	X	X
3	raw sludge	43	9.2	X	X	X
4	digested sludge	36	9.0	X		X
5	primary sludge	36	9.0	X	X	X
6	conventional activated sludge	30	8.7	X		X
7	secondary sludge	25	8.4	X	X	X
8	liquid sludge	22	8.3	X	X	X
9	dewatered sludge	21	8.2	X	X	X
10	settled sludge	15	7.7	X		X
11	returned sludge	13	7.5	X		X
12	excess sludge	13	7.5	X	X	X
13	undigested sludge	12	7.4	X	X	X
14	treated sludge	12	7.4	X		X
15	bulking sludge	12	7.4	X		X
16	residual sludge	9	7.0	X	X	X
17	waste sludge	9	7.0	X	X	X
18	dry sludge	9	2.54			X
19	oily sludge	8	8.4		X	X
20	municipal sludge	8	8.0		X	
21	API separator sludge	8	6.8	X	X	X
22	chemical sludge	8	6.8	X		
23	wet sludge	7	7.9		X	
24	surplus activated sludge	7	6.6	X		
25	stabilized sludge	7	1.98			X
26	wastewater sludge	6	7.2		X	
27	biological sludge	6	6.4		X	
28	untreated sludge	5	7.6		X	

Table 18. Hyponym selection of SLUDGE (CHEM-1)

In EcoLexicon, the original conceptual hierarchy of SLUDGE (**Figure 70**) was also lacking in content. The concept SLUDGE was linked to two hypernyms (i.e., MUD, SEDIMENT) and to four hyponyms (e.g., PRIMARY SLUDGE, RAW SLUDGE, BIOLOGICAL SLUDGE) at the same level of subdivision.

Figure 70. Original conceptual hierarchy of SLUDGE in EcoLexicon

4.2.2.3.5. HYPONYMS OF NITROGEN (CHEM-2)

Table 19 shows the hyponym selection of the hypernym NITROGEN. No hyponym could be extracted by using the *NITROGEN is the generic of* WS. However, the combination of the MWT WS, the *modifier* WS, and the MWT CQL query resulted in 21 hyponyms (e.g., ORGANIC NITROGEN, FIXED NITROGEN, MARINE NITROGEN).

NITROGEN hyponyms (*is the generic of* WS) [total frequency = 3,498]			
	TERM	FREQUENCY	SCORE
-	NONE	-	-

NITROGEN hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 3,498]				FOUND WITH		
	TERM	FREQUENCY	SCORE	MWT WS	MODIFIER WS	MWT CQL
1	organic nitrogen	160	11.2	X	X	X
2	fixed nitrogen	87	10.4	X		X
3	marine nitrogen	83	10.3	X	X	X
4	inorganic nitrogen	78	10.2	X	X	X
5	dissolved organic nitrogen	41	9.3	X		X
6	combined nitrogen	37	9.2	X	X	X
7	tissue nitrogen	25	8.6	X	X	X
8	oceanic nitrogen	25	8.6	X		X
9	dissolved inorganic nitrogen	24	8.6	X		X
10	riverine nitrogen	16	8.0	X		X
11	atmospheric nitrogen	16	8.0	X	X	X
12	ammoniacal nitrogen	15	7.9	X	X	X
13	particulate nitrogen	13	7.7	X	X	X
14	anthropogenic nitrogen	13	7.7	X		X
15	ocean nitrogen	13	7.7	X	X	X
16	gaseous nitrogen	13	7.7	X	X	X
17	reactive nitrogen	12	7.6	X	X	X
18	oxidized nitrogen	11	7.5	X		X

19	molecular nitrogen	11	7.5	X	X	X
20	mineral nitrogen	10	7.3	X	X	X
21	liquid nitrogen	8	7.8		X	X

Table 19. Hyponym selection of NITROGEN (CHEM-2)

In EcoLexicon, the original conceptual hierarchy of NITROGEN (**Figure 71**) was empty in terms of hyponymic information. The concept NITROGEN was only associated with its hypernym, CHEMICAL ELEMENT, which was also linked to a more general hypernym, SUBSTANCE.

Figure 71. Original conceptual hierarchy of NITROGEN in EcoLexicon

4.2.2.3.6. *HYPONYMS OF MAIZE (CHEM-3)*

Table 20 shows the hyponym selection of the hypernym MAIZE, which was one of the concepts with the least hyponymic content. No hyponym could be extracted by using the *MAIZE is the generic of* WS. However, the combination of the MWT WS, the *modifier* WS, and the MWT CQL query produced eleven hyponyms (e.g., TRANSGENIC MAIZE, QUALITY PROTEIN MAIZE, FERTILE TRANSGENIC MAIZE).

MAIZE hyponyms (*is the generic of* WS) [total frequency = 1,732]			
	TERM	FREQUENCY	SCORE
-	NONE	-	-

MAIZE hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 1,732]				FOUND WITH		
	TERM	FREQUENCY	SCORE	MWT WS	MODIFIER WS	MWT CQL
1	transgenic maize	115	12.4	X	X	X
2	quality protein maize	12	9.4	X	X	X
3	fertile transgenic maize	7	8.6	X		X
4	high oil maize	6	8.4	X	X	X
5	inbred maize	6	8.4	X		X
6	opaque-2 maize	5	8.1	X	X	X
7	Bt maize	5	8.1	X	X	X
8	tropical maize	4	7.8	X		X
9	commercial maize	4	7.8	X		X
10	high lysine maize	3	0.85			X
11	hybrid maize	3	0.85			X

Table 20. Hyponym selection of MAIZE (CHEM-3)

The concept MAIZE did not exist in EcoLexicon. That is the reason why there is no comparison with the original conceptual hierarchy.

4.2.2.3.7. HYPONYMS OF WASTEWATER (CIV-1)

Table 21 shows the hyponym selection of the hypernym WASTEWATER. In accordance with the *WASTEWATER is the generic of* WS, three hyponyms (i.e., SEWAGE, LEACHATE, EFFLUENT) were selected. The combination of the MWT WS, the *modifier* WS, and the MWT CQL query resulted in 17 hyponyms (e.g., INDUSTRIAL WASTEWATER, DOMESTIC WASTEWATER, MUNICIPAL WASTEWATER).

WASTEWATER hyponyms (*is the generic of* WS) [total frequency = 6,660]			
	TERM	FREQUENCY	SCORE
1	sewage	10	9.4
2	leachate	9	9.4
3	effluent	8	9.1

WASTEWATER hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 6,660]				FOUND WITH		
	TERM	FREQUENCY	SCORE	MWT WS	MODIFIER WS	MWT CQL
1	industrial wastewater	338	11.2	X	X	X
2	domestic wastewater	212	10.6	X	X	X
3	municipal wastewater	134	9.9	X	X	X
4	influent wastewater	107	9.6	X	X	X
5	treated wastewater	85	9.3	X		X
6	raw wastewater	50	8.5	X	X	X
7	settled wastewater	45	8.4	X		X
8	olive mill wastewater	43	8.3	X	X	X
9	food-processing wastewater	40	8.2	X	X	X
10	textile wastewater	36	8.1	X	X	X
11	untreated wastewater	29	7.8	X	X	X
12	petrochemical wastewater	25	7.5	X	X	X
13	hydraulic fracturing wastewater	21	7.3	X		X
14	PTA wastewater	18	7.1	X	X	X
15	molasses wastewater	15	7.8		X	
16	dairy wastewater	14	7.7		X	
17	sugar beet wastewater	12	7.5		X	

Table 21. Hyponym selection of WASTEWATER (CIV-1)

In EcoLexicon, the conceptual hierarchy of WASTEWATER (**Figure 72**) was well populated with different hyponymy levels. On the one hand, WASTEWATER was linked to two hypernyms (i.e., WATER, LIQUID WASTE). On the other hand, six hyponyms (e.g., EFFLUENT, INFLUENT, URBAN WASTEWATER) were located at the first level of specificity, three hyponyms (i.e., BLACKWATER, GRAYWATER, SURFACE RUNOFF) at the second level, and one hyponym (i.e., SEPTIC WATERS) at the third level.

Figure 72. Original conceptual hierarchy of WASTEWATER in EcoLexicon

4.2.2.3.8. HYPONYMS OF BREAKWATER (CIV-2)

Table 22 shows the hyponym selection of the hypernym BREAKWATER. No hyponym could be extracted by using the *BREAKWATER is the generic of* WS. However, the combination of the MWT WS, the *modifier* WS, and the MWT CQL query produced 21 hyponyms (e.g., VERTICAL BREAKWATER, SUBMERGED BREAKWATER, FLOATING BREAKWATER).

BREAKWATER hyponyms (*is the generic of* WS) [total frequency = 2,565]			
	TERM	**FREQUENCY**	**SCORE**
-	NONE	-	-

BREAKWATER hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 2,565]				**FOUND WITH**		
	TERM	**FREQUENCY**	**SCORE**	**MWT WS**	**MODIFIER WS**	**MWT CQL**
1	vertical breakwater	168	11.3	X	X	X
2	submerged breakwater	143	11.1	X	X	X
3	floating breakwater	94	10.6	X	X	X
4	rubble-mound breakwater	81	10.4	X	X	X
5	caisson breakwater	46	9.6	X	X	X
6	sloping breakwater	43	9.5	X	X	X
7	semicircular breakwater	34	9.1	X	X	X
8	berm breakwater	31	9.0	X	X	X
9	detached breakwater	30	9.0	X	X	X
10	offshore breakwater	27	8.8	X	X	X
11	composite breakwater	26	8.8	X	X	X

12	porous breakwater	22	8.5	X	X	X
13	low-crested breakwater	20	8.4	X	X	X
14	perforated-wall caisson breakwater	17	8.1	X	X	X
15	box-type breakwater	14	7.9	X	X	X
16	coastal breakwater	13	7.8	X	X	X
17	skirt breakwater	13	7.8	X	X	X
18	permeable breakwater	12	7.6	X	X	X
19	shore-parallel breakwater	11	7.5	X	X	X
20	dual cylindrical caisson breakwater	8	7.1	X	X	
21	high-crested breakwater	8	7.6		X	

Table 22. Hyponym selection of BREAKWATER (CIV-2)

In EcoLexicon, the original conceptual hierarchy of BREAKWATER (**Figure 73**) had some generic-specific content. The concept BREAKWATER was associated with two hypernyms (i.e., CONSTRUCTION, SHORE-PARALLEL STRUCTURE) and it was also linked to eight hyponyms (e.g., BERM BREAKWATER, REEF BREAKWATER, OFFSHORE BREAKWATER), all at the same level.

Figure 73. Original conceptual hierarchy of BREAKWATER in EcoLexicon

4.2.2.3.9. HYPONYMS OF POLLUTANT (CIV-3)

Table 23 shows the hyponym selection of the hypernym POLLUTANT. In accordance with the *POLLUTANT is the generic of* WS extraction and identification, six hyponyms (e.g., NITROGEN OXIDE, SULFUR OXIDE, BENZENE) were selected. The combination of the MWT WS, the *modifier* WS, and the MWT CQL query resulted in 14 hyponyms (e.g., AIR POLLUTANT, ORGANIC POLLUTANT, PERSISTENT ORGANIC POLLUTANT).

POLLUTANT hyponyms (*is the generic of* WS) [total frequency = 2,351]			
	TERM	FREQUENCY	SCORE
1	NOx (nitrogen oxide)	7	9.2
2	SO2 (sulfur dioxide)	7	9.1
3	benzene	7	9.0
4	ozone	7	8.9
5	CO2 (carbon dioxide)	7	8.4
6	NO2 (nitrogen dioxide)	1	9.4

POLLUTANT hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 2,351]				FOUND WITH		
	TERM	FREQUENCY	SCORE	MWT WS	MODIFIER WS	MWT CQL
1	air pollutant	320	12.0	X	X	X
2	organic pollutant	140	11.0	X	X	X
3	persistent organic pollutant	60	9.8	X	X	X
4	priority pollutant	54	9.6	X	X	X
5	gaseous pollutant	19	8.2	X	X	X
6	atmospheric pollutant	15	7.8	X	X	X
7	anthropogenic pollutant	14	7.9		X	
8	primary pollutant	12	7.5	X	X	X
9	acidifying pollutant	11	7.4	X		X
10	secondary pollutant	11	7.4	X	X	X
11	aromatic pollutant	10	7.6		X	
12	water pollutant	10	7.2	X		
13	inorganic pollutant	10	7.2	X	X	X
14	metal pollutant	9	6.8		X	

Table 23. Hyponym selection of POLLUTANT (CIV-3)

In EcoLexicon, the original conceptual hierarchy of POLLUTANT (**Figure 74**) had a significant amount of hyponymic content, with many hyponymy levels. Nine hyponyms (e.g., PESTICIDE, PERSISTENT ORGANIC POLLUTANT, POLYCHLORINATED BIPHENYL) were located at the first level; eleven hyponyms (e.g., DICHLORODIPHENYLTRICHLOROETHANE, METHYL BROMIDE, AEROSOL) were at the second level; and two hyponyms (i.e., CARBON DIOXIDE, METHANE) were at the third and last level.

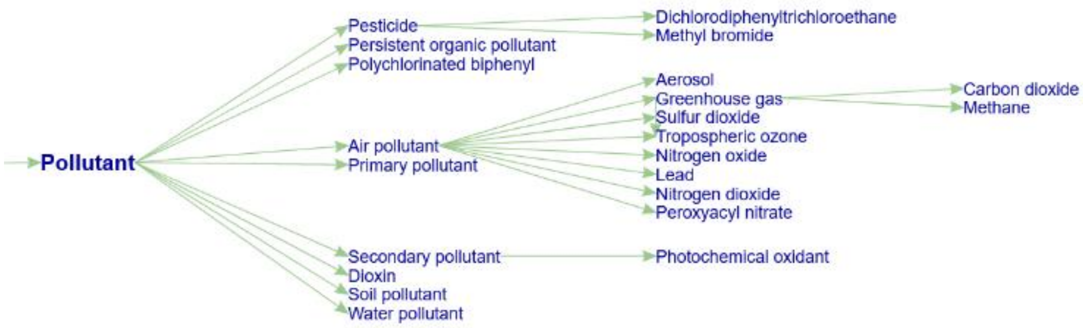


Figure 74. Original conceptual hierarchy of POLLUTANT in EcoLexicon

4.2.2.3.10. HYPONYMS OF EARTHQUAKE (GEO-1)

Table 24 shows the hyponym selection of the hypernym EARTHQUAKE, which was the concept with the least hyponymic content of all. No hyponym could be extracted by using the *EARTHQUAKE is the generic of* WS. However, the combination of the MWT WS, the *modifier* WS, and the MWT CQL query resulted in ten hyponyms (e.g., LOCAL EARTHQUAKE, SHALLOW EARTHQUAKE, CHARACTERISTIC EARTHQUAKE).

EARTHQUAKE hyponyms (*is the generic of* WS) [total frequency = 5,978]			
	TERM	**FREQUENCY**	**SCORE**
-	NONE	-	-

EARTHQUAKE hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 5,978]				FOUND WITH		
	TERM	**FREQUENCY**	**SCORE**	**MWT WS**	**MODIFIER WS**	**MWT CQL**
1	local earthquake	30	8.0	X	X	X
2	shallow earthquake	23	7.6	X	X	X
3	characteristic earthquake	21	7.5	X		X
4	inland crustal earthquake	19	7.7		X	
5	deep earthquake	19	7.4	X	X	X
6	scenario earthquake	19	7.4	X	X	X
7	crustal earthquake	19	7.4	X	X	X
8	subduction earthquake	16	7.5		X	
9	intraplate earthquake	16	7.1	X		
10	shallow crustal earthquake	15	7.0	X	X	

Table 24. Hyponym selection of EARTHQUAKE (GEO-1)

In EcoLexicon, the original conceptual hierarchy of EARTHQUAKE (**Figure 75**) was lacking in content. The concept EARTHQUAKE was associated with a hypernym (i.e., EXTREME EVENT) and also linked to four hyponyms (e.g., TECTONIC EARTHQUAKE, VOLCANIC EARTHQUAKE, TSUNAMIGENIC EARTHQUAKE) at the same level of specificity.

Figure 75. Original conceptual hierarchy of EARTHQUAKE in EcoLexicon

4.2.2.3.11. HYPONYMS OF SEDIMENT (GEO-2)

Table 25 shows the hyponym selection of the hypernym SEDIMENT, which was another one of the concepts with the most hyponymic content. In accordance with the *SEDIMENT is the generic of* WS extraction and identification, four hyponyms (e.g., SAND, SILT, CLAY) were selected. The combination of the MWT WS, the *modifier* WS, and the MWT CQL query produced 27 hyponyms (e.g., COHESIVE SEDIMENT, SUSPENDED SEDIMENT, STREAM SEDIMENT).

SEDIMENT hyponyms (*is the generic of* WS) [total frequency = 4,698]			
	TERM	**FREQUENCY**	**SCORE**
1	sand	14	10.4
2	silt	10	10.5
3	clay	8	9.8
4	gravel	5	9.3

SEDIMENT hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 4,698]				**FOUND WITH**		
	TERM	**FREQUENCY**	**SCORE**	**MWT WS**	**MODIFIER WS**	**MWT CQL**
1	cohesive sediment	196	10.8	X	X	X
2	suspended sediment	100	9.9	X		X
3	stream sediment	68	9.3	X	X	X
4	lake sediment	53	9.0	X	X	X
5	marine sediment	50	8.9	X	X	X
6	bottom sediment	36	8.4	X	X	X
7	coastal sediment	35	8.4	X	X	X
8	fine sediment	33	8.3	X	X	X
9	coarse sediment	23	8.7		X	X

10	fine-grained sediment	22	7.7	X	X	X
11	clastic sediment	21	7.6	X	X	X
12	unconsolidated sediment	20	7.6	X	X	X
13	consolidated sediment	19	7.5	X	X	X
14	fluvial sediment	18	7.4	X	X	X
15	carbonate sediment	16	7.3	X	X	X
16	beach sediment	16	7.3	X	X	X
17	ocean sediment	15	7.4		X	
18	deposited sediment	15	7.2	X		X
19	chemical sediment	15	7.2	X		
20	soft sediment	15	7.2	X	X	X
21	shallow sediment	13	7.5		X	
22	noncohesive sediment	13	7.0	X	X	
23	deep-sea sediment	12	6.8	X	X	
24	pelagic sediment	12	6.8	X	X	
25	biogenic sediment	11	7.7		X	
26	cohesionless sediment	8	7.3		X	
27	muddy sediment	8	7.3		X	

Table 25. Hyponym selection of SEDIMENT (GEO-2)

In EcoLexicon, the original conceptual hierarchy of SEDIMENT (**Figure 76**) also had a significant amount of hyponymic content, displayed in many hyponymy levels. The first level had 13 hyponyms (e.g., CONTINENTAL SEDIMENT, MARINE SEDIMENT, ALLUVIAL SEDIMENT); ten hyponyms (e.g., FLUVIAL SEDIMENT, FINE SAND, COARSE SAND) were at the second level; and one hyponym (i.e., RECESSIONAL MORAINE) was at the third and last level.

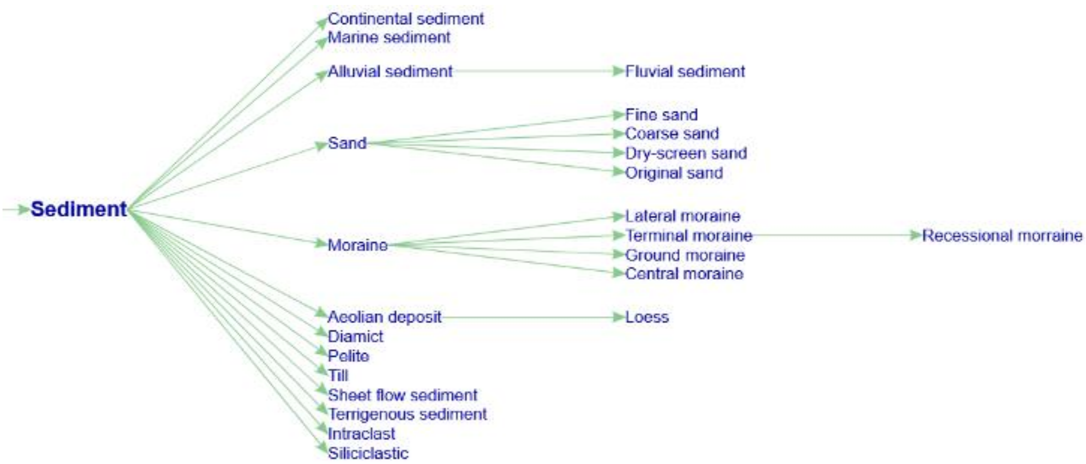


Figure 76. Original conceptual hierarchy of SEDIMENT in EcoLexicon

4.2.2.3.12. Hyponyms of Soil (GEO-3)

Table 26 shows the hyponym selection of the hypernym SOIL. No hyponym could be extracted by using the *SOIL is the generic of* WS. However, the combination of the MWT WS, the *modifier* WS, and the MWT CQL query resulted in 27 hyponyms (e.g., SURFACE SOIL, SANDY SOIL, CONTAMINATED SOIL).

	SOIL hyponyms (*is the generic of* WS) [total frequency = 6,612]		
	Term	**Frequency**	**Score**
-	NONE	-	-

	SOIL hyponyms (MWT WS, modifier WS & MWT CQL) [total frequency = 6,612]			Found with		
	Term	**Frequency**	**Score**	**MWT WS**	**Modifier WS**	**MWT CQL**
1	surface soil	47	9.1	X	X	X
2	sandy soil	46	9.1	X	X	X
3	contaminated soil	41	8.9	X	X	X
4	mineral soil	33	8.6	X	X	X
5	saturated soil	24	8.2	X	X	X
6	expansive soil	23	8.1	X	X	X
7	fertile soil	21	8.0	X	X	X
8	soft soil	20	7.9	X	X	X
9	bulk soil	19	7.8	X	X	X
10	organic soil	18	7.8	X		X
11	urban soil	17	7.7	X	X	X
12	tropical soil	17	7.7	X	X	X
13	clay soil	17	7.7	X	X	X
14	moist soil	14	8.2		X	
15	rich soil	14	8.2		X	
16	wet soil	14	8.2		X	
17	agricultural soil	14	7.4	X	X	X
18	shallow soil	14	7.4	X	X	X
19	forest soil	13	7.7		X	
20	polluted soil	13	7.3	X	X	
21	cohesive soil	13	7.3	X	X	
22	granular soil	11	8.0		X	
23	backfill soil	10	7.9		X	
24	saline soil	10	7.8		X	
25	calcareous soil	9	7.7		X	
26	contractive soil	8	7.6		X	
27	clayey soil	8	7.6		X	

Table 26. Hyponym selection of SOIL (GEO-3)

In EcoLexicon, the original conceptual hierarchy of SOIL (**Figure 77**) had an average quantity of generic-specific content. The concept SOIL (here referred to as GROUND) was associated with one hypernym (i.e., SURFACE) and it was also linked to eleven hyponyms (e.g., CREEP, SUBSOIL, LOAM), all at the same level of specificity.

Figure 77. Original conceptual hierarchy of SOIL (GROUND) in EcoLexicon

4.2.3. CREATION OF THE CONCEPTUAL HIERARCHIES

Once all the data regarding both hypernyms and hyponyms had been selected, the next step was to create the conceptual hierarchies for the hyponymy-based terminological entries. These terminological entries portrayed four elements: (i) terminological definitions according to the FBT (Faber 2022); (ii) conceptual categories (Gil-Berrozpe *et al*. 2019); (iii) hyponymy subtypes in a hierarchical structure (Gil-Berrozpe *et al*. 2017); and (iv) hyponymic contexts (Gil-Berrozpe *et al*. 2017).

For this reason, this process involved the following four steps: (i) building the terminological definitions; (ii) classifying the conceptual categories; (iii) structuring the conceptual hierarchies with hyponymy subtypes; and (iv) extracting and including the hyponymic contexts.

Since the objective of these results was to improve EcoLexicon (**§2.2.3**) by creating a new module, the original conceptual hierarchies of the twelve hypernyms in this TKB were the starting point for creating the new enhanced conceptual hierarchies (except in the case of MAIZE, because this concept initially did not exist in EcoLexicon). This means that, apart from the information filtered and selected by corpus analysis, the hyponyms in the original conceptual hierarchies in EcoLexicon were also taken into account. Nevertheless, they were cross-checked and reviewed according to the criteria of this research.

For instance, in the original conceptual hierarchy of REEF, there were some hyponyms (e.g., BIOHERM, ATOLL) included in the new proposal, whilst other hyponyms (e.g., MESOAMERICAN BARRIER REEF SYSTEM) were not regarded as terms based on the criteria established.

After including the extra concepts from EcoLexicon, the twelve terminological entries contained a total of 309 concepts (12 general hypernyms and 297 hyponyms).

4.2.3.1. BUILDING THE TERMINOLOGICAL DEFINITIONS

Previous sections described the terminological definition as well as its relevance on the representation of hyponymy. Concept definitions are the most elementary aspect of any terminological resource because they express specialized knowledge about both simple and complex ideas. The terminological definition defines the concept designated by a specialized lexical unit or term, and identifies the necessary aspects of such concept within the limits of a given specialized domain (Sager & Ndi-Kimbi 1995).

Since a definition also has a non-linguistic nature and takes into account the ontological and relational aspects of concepts, it was regarded as the best starting point when creating the hyponymy-based terminological entries. Moreover, the template of these definitions follows a structure based on the explicitation of the *genus* (i.e., the closest hypernym) and the *differentiae* (i.e., the characteristics that differentiate each co-hyponym). This makes this type of definition optimal for describing generic-specific relations and hierarchically structuring concepts.

The definitions were constructed according to the following protocol:

1. If the concept had originally been present in EcoLexicon, its definition was directly retrieved from this TKB. However, these definitions were reviewed, cross-checked, and enhanced with additional information provided by the definitional and encyclopedic information consulted in the selection of specialized terminological resources, as well as by the definitional and contextual information obtained from KRCs.
2. If the concept was not present in EcoLexicon, its definition was retrieved from the selection of specialized terminological resources, and later enhanced with additional definitional and contextual information provided by the corpus analysis.
3. If a concept was neither present in EcoLexicon nor in the specialized resources used as sources, its definition was created from the definitional and contextual information provided by the corpus analysis.

In this way, terminological definitions were created, which reflected property inheritance through the *genus* and specified the *differentiae* that make each co-hyponym different from each other. Terminological entries also had to be concise and describe all the elements about the concept in a single sentence. Longer descriptions of more than one sentence are more typical of encyclopedic definitions.

Interestingly, as definitions were created, many synonyms for different terms were found both in the selection of specialized resources and in corpus analysis. Since they were regarded as relevant linguistic information, they were also included in the final version of the hyponymy-based terminological entries.

Table 27 shows the terminological definition building process for the hyponym ATOLL. The most repeated and salient characteristics are underlined so as to visualize the elements finally included in terminological definition. Additionally, property inheritance was corrected in order to show ATOLL as a hyponym of CORAL REEF, not as a direct hyponym of REEF.

	ATOLL
Original definition (EcoLexicon)	Ring-shaped reef composed largely of coral, mainly found in the tropical waters of the Pacific Ocean.
Additional definition (Martin 2010)	A circular or elliptical coral reef that encloses a shallow central lagoon. It may be continuous or, more often, broken into closely spaced islets. The water outside the reef is deep. Atolls range in size from a few kilometres to more than 100 km across and are most often found in the Pacific Ocean. They represent the craters of volcanic islands that have sunk as coral grew on or around the rim.
Additional definition (Lawrence 2008)	Coral reef surrounding a central lagoon.
Corpus information	An atoll is a ring of reef, and often islands or sand cays, surrounding a central lagoon. The vast majority of atolls occur in the Indo-West Pacific region, that is, the tropical Indian and western Pacific oceans. Atolls are rare in the Caribbean and the rest of the tropical Atlantic Ocean.
Corpus information	Atolls are rings of reef, with steep outer slopes, that enclose a shallow lagoon.
Corpus information	Darwin reasoned that atolls are formed by reef growth on a subsiding island. The atoll gets its start when a deep-sea volcano erupts to build an island or seamount. Corals soon colonize the shores of the new island, and a fringing reef develops.

Final definition	**Ring-shaped coral reef with steep outer slopes that encloses a shallow central lagoon and that is typically found in the Pacific Ocean.**

Table 27. Example of a terminological definition building (concept: ATOLL)

In this way, a total of 309 terminological definitions reflecting property inheritance were built for the 309 concepts in the hyponymy-based terminological entries. This made it easier to structure the conceptual hierarchies.

Moreover, thanks to the synonyms identified during this process (e.g., the concept CYANOBACTERIUM was also found to be named BLUE-GREEN BACTERIUM and BLUE-GREEN ALGA), the number of terms associated with those concepts increased. They were originally 309, one term per concept, and after adding all the synonyms the number increased to 465.

4.2.3.2. CLASSIFYING THE CONCEPTUAL CATEGORIES

The second elements to be included in the terminological entries were the conceptual categories of each concept. As previously explained, conceptual categories are semantic classes or tags that identify a concept and show its interrelation with similar concepts. They determine the degree of specificity (Murphy & Lassaline 1997) of a given concept and its ontological similarity (Hahn & Chater 1997) to related concepts.

Gil-Berrozpe *et al*. (2019) classified all concepts in EcoLexicon with top-down categories based on their definition and the contextual information in the EcoLexicon corpus. The criteria followed in this study were also applied to this classification.

Thus, all 309 concepts (both hypernyms and hyponyms) were classified in the conceptual categories in EcoLexicon. Depending on the nature of the concept, more than one conceptual category could be attributed to the same concept. In this process, there were 22 different categories: *chemical substance*, *defense structure*, *deposit*, *earth/soil movement*, *fluid matter*, *gas*, *landform*, *layer*, *microorganism*, *mineral*, *model*, *particle*, *part of lifeform*, *part of animal*, *part of fungus*, *part of plant*, *plant*, *rock*, *soil*, *solid matter*, *structure*, and *water*.

Most of the concepts were members of a single conceptual category that tended to be the same as that of their direct hypernym, as in the case of BOX-TYPE BREAKWATER (**Table 28**). Other concepts were members of more than one category, such as ANTHROPOGENIC POLLUTANT, which was simultaneously categorized as *chemical substance*, *fluid matter*, and *solid matter*.

CONCEPT	CONCEPTUAL CATEGORY
breakwater Coastal defense structure, generally parallel to the coastline, made of wood, concrete or stone, to protect the coast from the impact of the wave and to provide shelter for ports and harbors.	E-1.3.2: *Defense structure*
floating breakwater Breakwater consisting of a moored assembly of floating objects with a limited range of movement to protect vessels riding at anchor.	E-1.3.2: *Defense structure*
box-type breakwater Floating breakwater with reinforced concrete modules that are either empty inside or have a core of light material.	E-1.3.2: *Defense structure*

Table 28. Example of a conceptual category classification (hierarchy: BREAKWATER)

Given the importance of conceptual categories in the hyponymy-based terminological entries, the full conceptual category hierarchy in EcoLexicon by Gil-Berrozpe *et al*. (2019) is provided in **Annex I**.

4.2.3.3. STRUCTURING THE CONCEPTUAL HIERARCHIES WITH HYPONYMY SUBTYPES

The third elements to be included in the terminological entries were the hyponymy subtype of each hyponym. As explained in section **§3.2.3.1**, hyponymy subtypes can be used to filter and obtain a more fine-grained vision of conceptual hierarchies. They are based on the main characteristic that defines a hyponym and distinguishes it from its hypernym. They are also dependent on the conceptual nature of the hyponym.

Gil-Berrozpe *et al*. (2017, 2018) identified the hyponymy subtypes typical of environmental concepts. This research study used this inventory as a reference and followed the same criteria to identify the hyponymy subtypes in the generic-specific relations of the conceptual hierarchies.

Given that hyponymy subtypes can only be attributed to hyponyms, 297 hyponyms were classified with hyponymy subtypes. In this process, 19 subtypes were applied: *ability-based*, *color-based*, *composition-based*, *degree-based*, *denomination-based*, *effect-based*, *function-based*, *height-based*, *location-based*, *method-based*, *moisture-based*, *movement-based*, *origin-based*, *relation-based*, *shape-based*, *size-based*, *state-based*, *status-based*, and *technology-based* hyponymy. Since the concepts had been previously defined by showing property inheritance, and they had also been ontologically classified, this facilitated the specification of the new conceptual hierarchies.

Because of the information in these enhanced conceptual hierarchies, they did not display the information in alphabetical order, like traditional resources. Instead, the conceptual hierarchies grouped hyponyms by subtype and level (from general to specific). Then, within each classification, information was organized alphabetically.

Table 29 displays a segment of the hyponymy subtype classification for the concept REEF. As can be seen, BIOHERM and CORAL REEF are both first-level hyponyms of REEF based on their composition; UPLIFTED REEF is a second-level hyponym of CORAL REEF based on its height; and BARRIER REEF, COASTAL REEF, FRINGING REEF, and OUTER REEF are also second-level hyponyms of CORAL REEF, but based on location.

CONCEPT	HYPONYMY SUBTYPE
reef Ridge or mound-like structure made of rock or other sedimentary material lying just below the surface of the sea and found in the tidal zone along a coastline.	
	Composition-based hyponymy (first level)
bioherm Reef made of sedentary organisms such as marine invertebrates (corals, echinoderms, gastropods, mollusks, etc.) and enclosed or surrounded by rock of different origin. **coral reef** Reef made of coral consolidated into limestone and that is generally found below the ocean surface in shallow warm tropical waters.	
	Height-based hyponymy (second level)
uplifted reef Coral reef that is above water level.	
	Location-based hyponymy (second level)
barrier reef Coral reef roughly parallel to a shore and separated from it by a lagoon or other body of water that is too deep for coral to proliferate. **coastal reef** Coral reef occurring near and parallel to a coastline. **fringing reef** Coral reef directly connected to the coast that can be separated from it by a barrier lagoon or a canal. **outer reef** Coral reef that is located on the point where the coral system meets the ocean.	

Table 29. Example of a hyponymy subtype identification (hierarchy: REEF)

Given the importance of hyponymy subtypes in the hyponymy-based terminological entries, the inventory created by Gil-Berrozpe *et al*. (2017) is given in **Annex II**. As a matter of fact, all the hyponymy subtypes in the terminological entries coincided with those in this inventory.

4.2.3.4. EXTRACTING AND INCLUDING THE HYPONYMIC CONTEXTS

The fourth and last elements in the terminological entries were the hyponymic contexts (**§3.2.3.2**) of certain concepts. Hyponymic contexts are a type of KRC that were extracted from hyponymic KP-based queries in corpus. Many authors have described methods to extract semantic relations from corpora by searching for specific KRCs, such as Lafourcade & Ramadier (2016) or Rojas-García & Cabezas-García (2019). In this line, Gil-Berrozpe *et al*. (2017) identified a set of KP-based hyponymic contexts typical of the EEC.

In this case, the hyponymic contexts were extracted with the customized sketch grammar developed by León-Araúz *et al*. (2016) and by applying the WS function in Sketch Engine. The KP-based rule used was the *"X" is the type of* WS. **Table 30** shows a segment with examples of the hyponymic contexts retrieved by using this method for BENZENE.

[…] individual aromatics such as **BENZENE**, toluene, and naphthalene are toxic at concentrations about 100 times lower.
[...] It is usually acquired by exposure to certain drugs, to toxins such as **BENZENE**, or to ionizing radiation.
Common contaminants include volatile hydrocarbons –such as **BENZENE**, toluene, ethylene, and xylene (BTEX compounds)– found in fuels; [...]
The greater hazards come from radon, lead, indoor air pollution, and fumes from chemicals such as **BENZENE** and formaldehyde.
There are traces of aromatic compounds such as **BENZENE** present at the percent level in most crude oils.
BENZENE and other volatile organic compounds have been detected in bottled water despite manufacturers' claims of purity.
Used oil may also contain toxic substances such as lead, **BENZENE**, zinc, and cadmium.
Benzene ring compounds such as **BENZENE**, toluene, phenol, and pentachlorophenol all inhibit methanogens.
[...] Of particular interest in this class of pollutants are **BENZENE**, toluene, ethylbenzene, and xylene isomers (BTEX).
Although nearly all organic compounds are oxidised by this procedure, some aromatic compounds, such as **BENZENE**, pyridine, and toluene are either unaffected or only partially oxidised during the test.
[...] a variety of substances that are not soluble in polar solvents such as water (recall that oil and water do not mix) but will dissolve in nonpolar solvents such as **BENZENE** and chloroform.

Table 30. Example of a hyponymic context retrieval (concept: BENZENE)

Of the 309 concepts, 107 had a hyponymic context. In cases where a concept was associated with more than a term, the hyponymic context was only extracted for one of the synonyms and not for all of them so as to avoid information overload. In those cases where many hyponymic contexts were retrieved for the same concept, only one was selected to be included in the terminological entries.

Given the importance of hyponymy contexts in the hyponymy-based terminological entries, the inventory of hyponymic KPs identified by Gil-Berrozpe *et al*. (2017) is given in **Annex III**. In fact, many of the hyponymic contexts in the terminological entries followed the structure of the hyponymic KPs listed in this inventory.

4.2.4. DESIGN OF THE TERMINOLOGICAL TEMPLATE

The final step was the design of the terminological template for the twelve hyponymy-based terminological entries. The software Lexonomy was used for this purpose. Since the terminological entries in Lexonomy are written in XML, they can be designed from scratch to meet the needs of the terminological resource in question.

The elements included in the design of the terminological template for the hyponymy-based terminological entries were the following:

- Parent or superordinate concept (represented by a hypernym)
- Child or subordinate concepts (represented by hyponyms)
- Up to six hyponymy levels
- Terminological definitions
- Conceptual categories
- Hyponymy subtypes
- Hyponymic contexts

As required by Lexonomy, the entry structure of this terminological template was designed in XML (**Table 31**), following a hierarchical structure for the representation of the hyponymic relations. Parent concepts and hypernyms, child concepts and hyponyms, definitions, hyponymy subtypes, and hyponymic contexts were introduced as elements (represented between angle brackets, < >), whilst conceptual categories were introduced as attributes (preceded by @).

```
<parentconcept>
        <hypernym>
                @conceptualcategory
        <definition_hyper>
        <hyponymiccontext_hyper>
        <childconcept-1>
                <hyponymysubtype_hypo1>
                <hyponym-1>
                        @conceptualcategory
                <definition_hypo1>
                <hyponymiccontext_hypo1>
                <childconcept-2>
                        <hyponymysubtype_hypo2>
                        <hyponym-2>
                                @conceptualcategory
                        <definition_hypo2>
                        <hyponymiccontext_hypo2>
                        <childconcept-3>
                                <hyponymysubtype_hypo3>
                                <hyponym-3>
                                        @conceptualcategory
                                <definition_hypo3>
                                <hyponymiccontext_hypo3>
                                <childconcept-4>
                                        <hyponymysubtype_hypo4>
                                        <hyponym-4>
                                                @conceptualcategory
                                        <definition_hypo4>
                                        <hyponymiccontext_hypo4>
                                        <childconcept-5>
                                                <hyponymysubtype_hypo5>
                                                <hyponym-5>
                                                        @conceptualcategory
                                                <definition_hypo5>
                                                <hyponymiccontext_hypo5>
                                                <childconcept-6>
                                                        <hyponymysubtype_hypo6>
                                                        <hyponym-6>
                                                                @conceptualcategory
                                                        <definition_hypo6>
                                                        <hyponymiccontext_hypo6>
```

Table 31. Entry format of the hyponymy-based terminological template in XML

The design of this entry structure also permits the modification of individual concept entries. For instance, if a child concept has synonyms, more than one hyponym can be added, since they share the same hyponymy subtype, conceptual category, and definition. Likewise, if there are several child concepts of the same hypernym, the entry structure allows the creation of *sibling* entries at the same level, either before or after the selected element.

5. RESULTS AND DISCUSSION

This chapter presents and analyzes the final results of this doctoral thesis. Firstly, the twelve hyponymy-based terminological entries are shown, described, and analyzed. Secondly, we present HypoLexicon, the practical application of this research. Finally, a statistical summary of our results is given.

5.1. HYPONYMY-BASED TERMINOLOGICAL ENTRIES

The twelve hyponymy-based terminological entries correspond to the conceptual hierarchies of the three BIO hypernyms (i.e., BACTERIUM, REEF, and CELL), the three CHEM hypernyms (i.e., SLUDGE, NITROGEN, and MAIZE), the three CIV hypernyms (i.e., WASTEWATER, BREAKWATER, and POLLUTANT), and the three GEO hypernyms (i.e., EARTHQUAKE, SEDIMENT, and SOIL) of the corpus analysis performed as part of this research.

The structure and format of the terminological entries facilitates the representation of their hyponymic content. The information is displayed in conceptual hierarchies which are alphabetically ordered as follows: (i) hyponymy subtypes (with possible multiple hyponymy levels); and (ii) hyponyms and co-hyponyms related to each hyponymy subtype.

Each terminological entry includes: (i) terms (i.e., hypernyms, hyponyms, and synonyms) that designate the concepts, in boldface and red; (ii) conceptual categories, in blue and with a light-blue background; (iii) definitions, in black; (iv) hyponymic contexts, in black, in italics, and with a bullet point; and (v) hyponymy subtypes, in black, in boldface, and with a yellow background dividing sections per subtype to facilitate their visualization.

This schema applies to all hyponymy-based terminological entries, shown below. Furthermore, the structure and format are the same as in the terminological entries in HypoLexicon seen with the entry view.

5.1.1. BACTERIUM

The hyponymy-based terminological entry of BACTERIUM is shown in **Table 32**.

bacterium E-7.4: Microorganism

Microscopic unicellular prokaryotic organism characterized by having a cell wall with DNA and a primitive nucleus, and by lacking visible chromosomes and membrane.

- *[…] the interior of microorganisms such as BACTERIA is almost entirely comprised of water.*

ABILITY-BASED HYPONYMY

aerobic bacterium E-7.4: Microorganism
aerobe

Bacterium that survives and grows only in the presence of oxygen.

- *When oxygen becomes limiting, some microorganisms, mainly AEROBIC BACTERIA, have the ability to switch to the use of the nitrogen oxides […].*

ABILITY-BASED HYPONYMY

facultative aerobic bacterium E-7.4: Microorganism
facultative aerobe

Aerobic bacterium that usually lives in the presence of oxygen, but does not require it.

anaerobic bacterium E-7.4: Microorganism
anaerobe

Bacterium that can survive in a partial or complete absence of air.

- *[…] studies have shown that ANAEROBES, including B. fragilis, can produce compounds such as succinic acid and short-chain fatty acids.*

ABILITY-BASED HYPONYMY

methanogenic bacterium E-7.4: Microorganism
methanogen

Anaerobic bacterium that is capable of producing methane to aid respiration.

- *[…] in association with H2-consuming microorganisms, such as the METHANOGENS, the H2 partial pressure is maintained at levels low enough […].*

autotrophic bacterium E-7.4: Microorganism
autotroph

Bacterium that is capable of synthesizing its own food by performing several reactions involving light energy and chemicals to derive energy for its biological sustainability.

- *Some AUTOTROPHIC BACTERIA, such as the cyanobacteria previously mentioned, are photosynthetic or photoautotrophic.*

ABILITY-BASED HYPONYMY

photosynthetic bacterium E-7.4: Microorganism

Autotrophic bacterium that is capable of converting light energy into chemical energy through light absorbing pigments and reaction centers.

- *Diatoms and bacteria, including PHOTOSYNTHETIC BACTERIA, actually account for most of the primary production on mudflats.*

COLOR-BASED HYPONYMY

Cyanobacterium E-7.4: Microorganism
blue-green bacterium
blue-green alga
Photosynthetic bacterium of a phylum generally being blue-green in color and widespread in marine and freshwater environments, even capable of nitrogen fixation.

- *CYANOBACTERIA, once known as BLUE-GREEN ALGAE, are a group of photosynthetic bacteria.*

METHOD-BASED HYPONYMY

AAnP bacterium E-7.4: Microorganism
aerobic anoxygenic phototrophic bacterium
Photosynthetic aerobic bacterium that captures energy from light by anoxygenic photosynthesis.

heterotrophic bacterium E-7.4: Microorganism
heterotroph
Bacterium that requires organic compounds of nitrogen and carbon as a source of energy, which is obtained through metabolic synthesis.

- *A HETEROTROPH is an organism that must obtain its carbon in an organic form.*

EFFECT-BASED HYPONYMY

nitrogen-fixing bacterium E-7.4: Microorganism
Bacterium that transforms nitrogen gas from the atmosphere into fixed nitrogen compounds, such as ammonia.

- *[…] transgenic plants have a limited impact on their plant-associated microorganisms, including NITROGEN-FIXING BACTERIA, mycorrhizal fungi, and endophytic microbiota.*

pathogenic bacterium E-7.4: Microorganism
Bacterium that causes disease.

- *Capsules are formed by a few PATHOGENIC BACTERIA, such as Streptococcus pneumoniae.*

sulfur bacterium E-7.4: Microorganism
Bacterium that obtains energy by oxidizing inorganic sulfur compounds.

- *An example of a colorless SULFUR BACTERIA is the genus Thiothrix.*

LOCATION-BASED HYPONYMY

marine bacterium E-7.4: Microorganism
Bacterium that lives in the saltwater of a sea or ocean.

oral bacterium E-7.4: Microorganism
Bacterium that lives in the mouth of an organism.

- *[…] certain ORAL BACTERIA, such as Streptococcus sanguis, S. gordonii, S. mitis, S. oralis, and Actinomyces species, have higher affinity […].*

soil bacterium E-7.4: Microorganism
Bacterium that develops microcolonies in soil.

EFFECT-BASED HYPONYMY

nitrifying bacterium **E-7.4: Microorganism**

Soil bacterium that obtains energy either by oxidizing ammonium compounds into nitrites or by oxidizing nitrites into nitrates, thus playing an important role in the nitrogen cycle.

METHOD-BASED HYPONYMY

gram-negative bacterium **E-7.4: Microorganism**
Gram-negative bacterium

Bacterium that does not retain the crystal violet stain used in the gram-staining method of bacterial differentiation, generally resistant to the effects of antibiotics or the actions of the body's immune cells, and containing thin cell walls.

- *[...] the microbial populations usually found in systems treated with recommended doses of biocide (GRAM-NEGATIVE BACTERIA, especially Pseudomonas species).*

DENOMINATION-BASED HYPONYMY

Chromatium **E-7.4: Microorganism**

Gram-negative, photoautotrophic, ovoid to rod-shaped bacterium of a genus which oxidizes sulfide to produce the sulfur that is deposited in intracellular granules of the cytoplasm.

Haemophilus **E-7.4: Microorganism**

Gram-negative, rod-shaped, non-motile bacterium of a genus including the pathogenic agents of several human diseases (such as meningitis, pneumonia, and conjunctivitis).

- *The LPS of mucosal pathogens such as Neisseria and HAEMOPHILUS spp. is smaller than S-LPS [...].*

Pseudomonas **E-7.4: Microorganism**

Gram-negative, rod-shaped, motile, non-sporogenous bacterium of a genus occurring in soil and detritus, and including saprophytes and plant or animal pathogens.

- *Other bacteria such as Bacillus, PSEUDOMONAS, Spirillum, and Thiobacillus can carry out this denitrification process to completion.*

LOCATION-BASED HYPONYMY

enteric bacterium **E-7.4: Microorganism**

Gram-negative, rod-shaped bacterium that occurs normally or pathogenically in the intestines of organisms.

- *When E. coli and other ENTERIC BACTERIA are nitrogen limited, the synthesis of a number of proteins is dramatically induced.*

DENOMINATION-BASED HYPONYMY

Salmonella **E-7.4: Microorganism**

Enteric, rod-shaped, usually motile and non-sporogenous bacterium of a genus including pathogens for warm-blooded organisms that cause food poisoning, gastrointestinal inflammation, or septicemia.

- *SALMONELLA is a bacterium that is prevalent in the intestines of birds, mammals, and reptiles.*

SHAPE-BASED HYPONYMY

coliform bacterium E-7.4: Microorganism

Rod-shaped enteric bacterium that can ferment lactose with the production of acid and gas when incubated at 35–37°C.

- *Other gas-forming microbes such as COLIFORM BACTERIA, certain Clostridium species, heterofermentative lactic acid bacteria, and wild yeasts […].*

DENOMINATION-BASED HYPONYMY

Escherichia coli E-7.4: Microorganism
E. coli

Rod-shaped coliform bacterium of the genus Escherichia.

- *In a typical bacterium such as ESCHERICHIA COLI there are about 5000 genes.*

gram-positive bacterium E-7.4: Microorganism
Gram-positive bacterium

Bacterium that retains the crystal violet stain used in the gram-staining method of bacterial differentiation, generally sensitive to the effects of antibiotics or the actions of the body's immune cells, and containing thick cell walls.

DENOMINATION-BASED HYPONYMY

Bacillus E-7.4: Microorganism

Gram-positive, rod-shaped, usually aerobic bacterium of a genus producing endospores and including many saprophytes and some parasites.

- *Other bacteria such as BACILLUS, Pseudomonas, Spirillum, and Thiobacillus can carry out this denitrification process to completion.*

DENOMINATION-BASED HYPONYMY

Bacillus subtilis E-7.4: Microorganism
B. subtilis
hay bacillus
grass bacillus

Bacillus which is catalase-positive and is found in soil and the gastrointestinal tract of humans, ruminants and marine sponges.

- *BACILLUS SUBTILIS is a bacterium that has been widely used as antagonistic for soil-borne pathogens.*

Mycobacterium E-7.4: Microorganism

Gram-positive, rod-shaped bacterium of a genus including the pathogenic agents of tuberculosis and leprosy.

- *Infections with bacteria such as MYCOBACTERIUM are particularly problematic and severe.*

Staphylococcus E-7.4: Microorganism

Gram-positive, non-motile bacterium of a genus including pathogenic agents of various diseases (such as skin infections, food poisoning, and endocarditis).

- *Such selection by a virus for its host is useful in typing some bacteria, primarily STAPHYLOCOCCUS and Salmonella.*

Streptococcus E-7.4: Microorganism

Gram-positive, spherical or ovoid bacterium of a genus including the agents of souring of milk and dental decay, and haemolytic pathogens causing infections such as scarlet fever and pneumonia.

- *Within 8 to 12 hours after delivery, the newborn typically has been colonized by bacteria such as STREPTOCOCCI, staphylococci, and lactobacilli, acquired primarily from its mother.*

EFFECT-BASED HYPONYMY

lactic acid bacterium E-7.4: Microorganism

Gram-positive, non-motile, non-sporogenous bacterium that produces the fermentation of carbohydrate materials to form lactic acid.

- *Diets that are high in sugars are especially harmful to teeth because LACTIC ACID BACTERIA (especially Streptococcus sanguis and S. mutans) ferment the sugars to lactic acid [...].*

RELATION-BASED HYPONYMY

symbiotic bacterium E-7.4: Microorganism

Bacterium that lives in symbiosis with another organism or bacterium, forming an association in which both are benefited.

- *Some biologists even classify mitochondria and chloroplasts as SYMBIOTIC BACTERIA rather than organelles.*

SHAPE-BASED HYPONYMY

filamentous bacterium E-7.4: Microorganism

Bacterium that takes the form of branching filaments.

- *[...] a variety of substrates including zoogloeal bacteria, fungal hyphae, insect debris, and the larger FILAMENTOUS BACTERIA.*

Table 32. Full hyponymy-based terminological entry of BACTERIUM

BACTERIUM, the first BIO hyponymy-based terminological entry, has the following information: 34 concepts with their definitions; 49 terms that designate those concepts; one conceptual category (i.e., *microorganism*); eight hyponymy subtypes (i.e., *ability-based*, *color-based*, *denomination-based*, *effect-based*, *location-based*, *method-based*, *relation-based*, and *shape-based* hyponymy); up to four hyponymy levels; and 27 concepts with hyponymic contexts. This is the richest of all the BIO entries insofar as concepts and variety of hyponymy subtypes, but evidently not in relation to conceptual categories. Interestingly, this is also the entry with the highest number of hyponymic contexts.

As previously mentioned, *microorganism* is the only conceptual category. Since this is a BIO entry for a type of organism, it is specified through hyponymy in a form similar to that of a taxonomy, in that an organism cannot change species (e.g., a type of BACTERIUM will always be a BACTERIUM from a taxonomical perspective).

The hyponymy subtypes in this entry are *denomination-based*, *ability-based*, and *effect-based* hyponymy. The *denomination-based* hyponyms (e.g., BACILLUS SUBTILIS, ESCHERICHIA COLI, SALMONELLA) have a terminological designation that is very different from that of the hypernym, BACTERIUM. This is because their designations are their scientific names in Latin, something that is very typical of specialized disciplines within BIO. Moreover, in this entry, *ability-based* hyponyms (e.g., AUTOTROPHIC BACTERIUM, FACULTATIVE AEROBIC BACTERIUM, PHOTOSYNTHETIC BACTERIUM) specify the capacity of bacteria to synthesize chemical substances that generate other elements or to live in certain media. Finally, the *effect-based* hyponyms (e.g., LACTIC ACID BACTERIUM, NITROGEN-FIXING BACTERIUM, PATHOGENIC BACTERIUM) focus on the results of processes carried out by bacteria, such as transforming nitrogen gas from the atmosphere or causing disease.

Furthermore, the hyponym ESCHERICHIA COLI has four hyponymy levels: ESCHERICHIA COLI is a *denomination-based* type of COLIFORM BACTERIUM, which is a *shape-based* type of ENTERIC BACTERIUM, which is a *location-based* type of GRAM-NEGATIVE BACTERIUM, which is a *method-based* type of BACTERIUM. This sequence is quite rich in terms of hyponymic nuances because all four hyponyms are specified with a different subtype. Since the category (i.e., *microorganism*) remains the same, this taxonomy reveals more conceptual information about the concepts than conventional ones.

In relation to hyponymic contexts, it is possible to find a wide variety of KPs that express hyponymy in different ways (e.g., "gas-forming microbes such as coliform bacteria"; "some microorganisms, mainly aerobic bacteria"; "plant-associated microorganisms, including nitrogen-fixing bacteria"; etc.). However, this research only collected these contexts as a reference to more hyponymic information. Since it did not focus on their phraseology, but rather on conceptual elements, this aspect is not discussed in this or subsequent terminological entries.

Table 33 summarizes the information in the hyponymy-based terminological entry of BACTERIUM. The following elements are specified: (i) number of concepts, terms, and terminological definitions; (ii) conceptual categories and concepts per conceptual category; (iii) hyponymy subtypes and hyponyms per hyponymy subtype; (iv) hyponyms per hyponymy level; and (v) concepts with hyponymic contexts. Tables with the same elements are also provided for each of the other hyponymy-based terminological entries.

BACTERIUM entry

Concepts, terms, and terminological definitions	
No. of concepts	34
No. of terms	49
No. of terminological definitions	34

Conceptual categories	
No. of conceptual categories	1
Conceptual categories	*Microorganism*
No. of *microorganism* concepts	34
***Microorganism* concepts**	AANP BACTERIUM ǀ AEROBIC BACTERIUM ǀ ANAEROBIC BACTERIUM ǀ AUTOTROPHIC BACTERIUM ǀ BACILLUS ǀ BACILLUS SUBTILIS ǀ BACTERIUM ǀ CHROMATIUM ǀ COLIFORM BACTERIUM ǀ CYANOBACTERIUM ǀ ENTERIC BACTERIUM ǀ ESCHERICHIA COLI ǀ FACULTATIVE AEROBIC BACTERIUM ǀ FILAMENTOUS BACTERIUM ǀ GRAM-NEGATIVE BACTERIUM ǀ GRAM-POSITIVE BACTERIUM ǀ HAEMOPHILUS ǀ HETEROTROPHIC BACTERIUM ǀ LACTIC ACID BACTERIUM ǀ MARINE BACTERIUM ǀ METHANOGENIC BACTERIUM ǀ MYCOBACTERIUM ǀ NITRIFYING BACTERIUM ǀ NITROGEN-FIXING BACTERIUM ǀ ORAL BACTERIUM ǀ PATHOGENIC BACTERIUM ǀ PHOTOSYNTHETIC BACTERIUM ǀ PSEUDOMONAS ǀ SALMONELLA ǀ SOIL BACTERIUM ǀ STAPHYLOCOCCUS ǀ STREPTOCOCCUS ǀ SULFUR BACTERIUM ǀ SYMBIOTIC BACTERIUM

Hyponymy subtypes	
No. of hyponymy subtypes	8
Hyponymy subtypes	*ability-based* ǀ *color-based* ǀ *denomination-based* ǀ *effect-based* ǀ *location-based* ǀ *method-based* ǀ *relation-based* ǀ *shape-based*
No. of *ability-based* hyponyms	7
***Ability-based* hyponyms**	AEROBIC BACTERIUM ǀ ANAEROBIC BACTERIUM ǀ AUTOTROPHIC BACTERIUM ǀ FACULTATIVE AEROBIC BACTERIUM ǀ HETEROTROPHIC BACTERIUM ǀ METHANOGENIC BACTERIUM ǀ PHOTOSYNTHETIC BACTERIUM
No. of *color-based* hyponyms	1
***Color-based* hyponyms**	CYANOBACTERIUM
No. of *denomination-based* hyponyms	10
***Denomination-based* hyponyms**	BACILLUS ǀ BACILLUS SUBTILIS ǀ CHROMATIUM ǀ ESCHERICHIA COLI ǀ HAEMOPHILUS ǀ MYCOBACTERIUM ǀ PSEUDOMONAS ǀ SALMONELLA ǀ STAPHYLOCOCCUS ǀ STREPTOCOCCUS
No. of *effect-based* hyponyms	5
***Effect-based* hyponyms**	LACTIC ACID BACTERIUM ǀ NITRIFYING BACTERIUM ǀ NITROGEN-FIXING BACTERIUM ǀ PATHOGENIC BACTERIUM ǀ SULFUR BACTERIUM
No. of *location-based* hyponyms	4
***Location-based* hyponyms**	ENTERIC BACTERIUM ǀ MARINE BACTERIUM ǀ ORAL BACTERIUM ǀ SOIL BACTERIUM

No. of *method-based* hyponyms	3
***Method-based* hyponyms**	AANP BACTERIUM | GRAM-NEGATIVE BACTERIUM | GRAM-POSITIVE BACTERIUM
No. of *relation-based* hyponyms	1
***Relation-based* hyponyms**	SYMBIOTIC BACTERIUM
No. of *shape-based* hyponyms	2
***Shape-based* hyponyms**	COLIFORM BACTERIUM, FILAMENTOUS BACTERIUM

Hyponymy levels	
No. of hyponymy levels	4
No. of first-level hyponyms	14
First-level hyponyms	AEROBIC BACTERIUM | ANAEROBIC BACTERIUM | AUTOTROPHIC BACTERIUM | FILAMENTOUS BACTERIUM | GRAM-NEGATIVE BACTERIUM | GRAM-POSITIVE BACTERIUM | HETEROTROPHIC BACTERIUM | MARINE BACTERIUM | NITROGEN-FIXING BACTERIUM | ORAL BACTERIUM | PATHOGENIC BACTERIUM | SOIL BACTERIUM | SULFUR BACTERIUM | SYMBIOTIC BACTERIUM
No. of second-level hyponyms	13
Second-level hyponyms	BACILLUS | CHROMATIUM | ENTERIC BACTERIUM | FACULTATIVE AEROBIC BACTERIUM | HAEMOPHILUS | LACTIC ACID BACTERIUM | METHANOGENIC BACTERIUM | MYCOBACTERIUM | NITRIFYING BACTERIUM | PHOTOSYNTHETIC BACTERIUM | PSEUDOMONAS | STAPHYLOCOCCUS | STREPTOCOCCUS
No. of third-level hyponyms	5
Third-level hyponyms	AANP BACTERIUM | BACILLUS SUBTILIS | COLIFORM BACTERIUM | CYANOBACTERIUM | SALMONELLA
No. of fourth-level hyponyms	1
Fourth-level hyponyms	ESCHERICHIA COLI

Hyponymic contexts	
No. of concepts with hyponymic contexts	27
Concepts with hyponymic contexts	AEROBIC BACTERIUM | ANAEROBIC BACTERIUM | AUTOTROPHIC BACTERIUM | BACILLUS | BACILLUS SUBTILIS | BACTERIUM | COLIFORM BACTERIUM | CYANOBACTERIUM | ENTERIC BACTERIUM | ESCHERICHIA COLI | FILAMENTOUS BACTERIUM | GRAM-NEGATIVE BACTERIUM | HAEMOPHILUS | HETEROTROPHIC BACTERIUM | LACTIC ACID BACTERIUM | METHANOGENIC BACTERIUM | MYCOBACTERIUM | NITROGEN-FIXING BACTERIUM | ORAL BACTERIUM | PATHOGENIC BACTERIUM | PHOTOSYNTHETIC BACTERIUM | PSEUDOMONAS | SALMONELLA | STAPHYLOCOCCUS | STREPTOCOCCUS | SULFUR BACTERIUM | SYMBIOTIC BACTERIUM

Table 33. Summary of the information in the terminological entry of BACTERIUM

5.1.2. REEF

The hyponymy-based terminological entry of REEF is shown in **Table 34**.

reef E-4.2.1: Landform

Ridge or mound-like structure made of rock or other sedimentary material lying just below the surface of the sea and found in the tidal zone along a coastline.

- *[...] further reducing the resiliency of REEFS and other marine ecosystems to human impacts.*

COMPOSITION-BASED HYPONYMY

bioherm E-4.2.1: Landform
reef knoll

Reef made of sedentary organisms such as marine invertebrates (corals, echinoderms, gastropods, mollusks, etc.) and enclosed or surrounded by rock of different origin.

coral reef E-4.2.1: Landform

Reef made of coral consolidated into limestone and that is generally found below the ocean surface in shallow warm tropical waters.

- *[...] ecologically important and diverse habitats including CORAL REEFS, mangrove forests, seagrass beds and coastal forests.*

HEIGHT-BASED HYPONYMY

uplifted reef E-4.2.1: Landform

Coral reef that is above water level.

LOCATION-BASED HYPONYMY

barrier reef E-4.2.1: Landform

Coral reef roughly parallel to a shore and separated from it by a lagoon or other body of water that is too deep for coral to proliferate.

coastal reef E-4.2.1: Landform

Coral reef occurring near and parallel to a coastline.

fringing reef E-4.2.1: Landform
inner reef

Coral reef directly connected to the coast that can be separated from it by a barrier lagoon or a canal.

outer reef E-4.2.1: Landform

Coral reef that is located on the point where the coral system meets the ocean.

ORIGIN-BASED HYPONYMY

reef patch E-4.2.1: Landform
patch reef

Coral reef formed independently on a shelf at depths less than 70 meters in the lagoon of a barrier reef or of an atoll.

SHAPE-BASED HYPONYMY

atoll **E-4.2.1: Landform**
Ring-shaped coral reef with steep outer slopes that encloses a shallow central lagoon and that is typically found in the Pacific Ocean.

SHAPE-BASED HYPONYMY

completely land ringed atoll **E-4.2.1: Landform**
Atoll with a lagoon area which is completely land-locked from the open ocean by a continuous string of reef islets.

rocky reef **E-4.2.1: Landform**
Reef made of rock outcrops with varying relief, creating refuges for juvenile and smaller fish and surface area for colonization of algae and invertebrates.

HEIGHT-BASED HYPONYMY

shallow reef **E-4.2.1: Landform**
Reef located at a depth which is close to the surface of water.

submerged reef **E-4.2.1: Landform**
Reef located in water depths ranging from 25 to 2000 m, as a consequence of being formed during periods of lower sea level.

LOCATION-BASED HYPONYMY

natal reef **E-4.2.1: Landform**
Reef where a certain species or organism has been given birth.

offshore reef **E-4.2.1: Landform**
Reef located at some distance from the shore.

seaward reef **E-4.2.1: Landform**
Reef that is facing the sea.

tropical reef **E-4.2.1: Landform**
Reef found in tropical regions between the Tropic of Cancer and the Tropic of Capricorn.

ORIGIN-BASED HYPONYMY

artificial reef **E-1.3: Structure**
Reef of human origin typically built to promote marine life in areas with a generally featureless bottom or to control erosion.

- *The development of other alternatives including aquaculture, reef restoration and ARTIFICIAL REEFS are also being explored.*

ice reef **E-4.2.1: Landform**
Ridge formed in an ice sheet by the action of external pressure.

Table 34. Full hyponymy-based terminological entry of REEF

REEF, the second BIO hyponymy-based terminological entry, has the following information: 20 concepts with their definitions; 23 terms that designate those concepts; two conceptual categories (i.e., *landform* and *structure*); five hyponymy subtypes (i.e., *composition-based*, *height-based*, *location-based*, *origin-based*, and *shape-based* hyponymy); up to three hyponymy levels; and three concepts with hyponymic contexts. This is an average entry in terms of content, compared to the other BIO entries.

There are two types of conceptual category in this terminological entry: *landform* and *structure*. In fact, these categories are diametrically opposed since *landform* belongs to a hierarchy of natural categories, and *structure* to a hierarchy of artificial categories. Not surprisingly, most of the hyponyms of REEF (e.g., BIOHERM, COASTAL REEF, ROCKY REEF) are classified in the *landform* category, and only the human-made type of REEF (i.e., ARTIFICIAL REEF) inherits the *structure* category.

The most outstanding hyponymy subtype is, by far, *location-based* hyponymy. As might be expected, most of the hyponyms in this conceptual hierarchy (e.g., COASTAL REEF, FRINGING REEF, TROPICAL REEF) are characterized by the place they occupy in relation to the coast or the sea. The remaining hyponymy subtypes (i.e., *composition-based*, *height-based*, *origin-based*, and *shape-based* hyponymy) have fewer cases, which reveals that the locative hyponymic nuance is in fact the determining feature of REEF hyponyms.

COMPLETELY LAND RINGED ATOLL is specified and classified in the following three hyponymy levels: COMPLETELY LAND RINGED ATOLL is a *shape-based* type of ATOLL, which is a *shape-based* type of CORAL REEF, which is a *composition-based* type of REEF. The most decisive hyponymic nuance in this sequence is related to the shape of the concepts (i.e., ATOLL differs from its CORAL REEF for being ring-shaped, and then COMPLETELY LAND RINGED ATOLL differs from ATOLL for having a continuous string of reef islets, which gives it a more characteristic shape).

As for hyponymic contexts, this entry has examples of hyponymic KRCs extracted from the corpus (e.g., "reefs and other marine ecosystems"; "ecologically important and diverse habitats including coral reefs"; "the development of other alternatives including aquaculture, reef restoration and artificial reefs"). Finally, **Table 35** summarizes the information in the hyponymy-based terminological entry of REEF.

REEF entry

Concepts, terms, and terminological definitions	
No. of concepts	20
No. of terms	23
No. of terminological definitions	20

Conceptual categories	
No. of conceptual categories	2
Conceptual categories	*landform* ǀ *structure*
No. of *landform* concepts	19
***Landform* concepts**	ATOLL ǀ BARRIER REEF ǀ BIOHERM ǀ COASTAL REEF ǀ COMPLETELY LAND RINGED ATOLL ǀ CORAL REEF ǀ FRINGING REEF ǀ ICE REEF ǀ NATAL REEF ǀ OFFSHORE REEF ǀ OUTER REEF ǀ REEF ǀ REEF PATCH ǀ ROCKY REEF ǀ SEAWARD REEF ǀ SHALLOW REEF ǀ SUBMERGED REEF ǀ TROPICAL REEF ǀ UPLIFTED REEF
No. of *structure* concepts	1
***Structure* concepts**	ARTIFICIAL REEF

Hyponymy subtypes	
No. of hyponymy subtypes	5
Hyponymy subtypes	*composition-based* ǀ *height-based* ǀ *location-based* ǀ *origin-based* ǀ *shape-based*
No. of *composition-based* hyponyms	3
***Composition-based* hyponyms**	BIOHERM ǀ CORAL REEF ǀ ROCKY REEF
No. of *height-based* hyponyms	3
***Height-based* hyponyms**	SHALLOW REEF ǀ SUBMERGED REEF ǀ UPLIFTED REEF
No. of *location-based* hyponyms	8
***Location-based* hyponyms**	BARRIER REEF ǀ COASTAL REEF ǀ FRINGING REEF ǀ NATAL REEF ǀ OFFSHORE REEF ǀ OUTER REEF ǀ SEAWARD REEF ǀ TROPICAL REEF
No. of *origin-based* hyponyms	3
***Origin-based* hyponyms**	ARTIFICIAL REEF ǀ ICE REEF ǀ REEF PATCH
No. of *shape-based* hyponyms	2
***Shape-based* hyponyms**	ATOLL ǀ COMPLETELY LAND RINGED ATOLL

Hyponymy levels	
No. of hyponymy levels	3
No. of first-level hyponyms	11
First-level hyponyms	ARTIFICIAL REEF ǀ BIOHERM ǀ CORAL REEF ǀ ICE REEF ǀ NATAL REEF ǀ OFFSHORE REEF ǀ ROCKY REEF ǀ SEAWARD REEF ǀ SHALLOW REEF ǀ SUBMERGED REEF ǀ TROPICAL REEF
No. of second-level hyponyms	7
Second-level hyponyms	ATOLL ǀ BARRIER REEF ǀ COASTAL REEF ǀ FRINGING REEF ǀ OUTER REEF ǀ REEF PATCH ǀ UPLIFTED REEF
No. of third-level hyponyms	1
Third-level hyponyms	COMPLETELY LAND RINGED ATOLL

Hyponymic contexts	
No. of concepts with hyponymic contexts	3
Concepts with hyponymic contexts	ARTIFICIAL REEF ǀ CORAL REEF ǀ REEF

Table 35. Summary of the information in the terminological entry of REEF

5.1.3. CELL

The hyponymy-based terminological entry of CELL is shown in **Table 36**.

cell E-9.3: Part of lifeform

The smallest structural unit of an organism that is capable of independent functioning, and that is usually composed of cytoplasm, a nucleus, and various organelles, all surrounded by a semipermeable cell membrane.

- *Viruses are known to parasitize all types of CELLS, including bacteria, algae, fungi, protozoa, animals, and plants.*

ABILITY-BASED HYPONYMY

host cell E-9.3: Part of lifeform

Cell that is invaded by or capable of being invaded by an infectious agent or vector.

target cell E-9.3: Part of lifeform

Cell that has a specific receptor for an antigen, antibody, hormone or drug.

COMPOSITION-BASED HYPONYMY

eukaryotic cell E-9.3: Part of lifeform

eucaryotic cell

Cell with a clearly defined nucleus and various organelles that mainly composes multicellular organisms.

- *Animal cells and other EUKARYOTIC CELLS possess, in addition to the plasma membrane, numerous intracellular membranes which form the organelles that perform specialized metabolic functions.*

RELATION-BASED HYPONYMY

animal cell E-9.3.1: Part of animal

Eukaryotic cell that is the basic structural and functional unit of animal tissues and organs.

- *ANIMAL CELLS and other eukaryotic cells possess, in addition to the plasma membrane, numerous intracellular membranes which form the organelles that perform specialized metabolic functions.*

ABILITY-BASED HYPONYMY

chromatophore E-9.3.1: Part of animal

Pigment-bearing animal cell of integument that is capable of causing color changes by expanding or contracting.

EFFECT-BASED HYPONYMY

cancer cell E-9.3.1: Part of animal

Animal cell that divides continually, forming solid tumors or flooding the blood with abnormal cells.

- *Chemotherapy is the treatment of a disease or condition with chemicals that have a specific effect on its cause, such as a microorganism or CANCER CELL.*

FUNCTION-BASED HYPONYMY

stem cell E-9.3.1: Part of animal

Animal cell with no specialization that can give rise to one or more different types of specialized cells, such as blood cells and nerve cells.

LOCATION-BASED HYPONYMY

blood cell E-9.3.1: Part of animal

Animal cell that is present and contained in blood.

- *The glomerulus retains BLOOD CELLS, proteins, and other useful large molecules in the blood [...].*

COLOR-BASED HYPONYMY

red blood cell E-9.3.1: Part of animal
erythrocyte
red cell
red corpuscle

Blood cell with hemoglobin that transports oxygen and carbon dioxide to and from the tissues.

white blood cell E-9.3.1: Part of animal
leukocyte
white cell
white corpuscle

Blood cell that lacks hemoglobin and functions in the immune system to protect against infection and diseases.

- *Like blood, it also transports numerous WHITE BLOOD CELLS (especially lymphocytes) and miscellaneous materials [...].*

FUNCTION-BASED HYPONYMY

eosinophil E-9.3.1: Part of animal

White blood cell containing granules that is readily stained by eosin and that is present at sites of allergic reactions and parasitic infections.

- *Because inflammation would enlist blood cells such as EOSINOPHILS and lymphocytes to the site of infection [...].*

lymphocyte E-9.3.1: Part of animal

White blood cell formed in lymphoid tissue that recognizes and deactivates specific foreign substances called antigens.

- *LYMPHOCYTES, a type of white blood cell, fight viruses, bacteria, and other antigens by producing antibodies.*

FUNCTION-BASED HYPONYMY

B cell E-9.3.1: Part of animal
B lymphocyte

Lymphocyte that develops from stem cells in the bone marrow and is found in the blood, lymph nodes and tissues, mainly responsible for manufacturing antibodies.

- *Two distinct types of lymphocytes are B CELLS and T cells, both of which are central to the immune response.*

T cell E-9.3.1: Part of animal
T lymphocyte
Lymphocyte that forms in bone marrow and matures in the thymus, activating when receptors on the cell surface recognize specific antigens and regulating the immune system's response to infected or malignant cells.

- *Two distinct types of lymphocytes are B cells and T CELLS, both of which are central to the immune response.*

macrophage E-9.3.1: Part of animal
White blood cell occurring mainly in connective tissue and in the bloodstream that ingests foreign particles and infectious microorganisms by phagocytosis, also stimulating the action of other immune system cells.

- *Because cells such as MACROPHAGES are also negatively charged, repulsive forces will further discourage interaction of macrophages with the capsular material.*

epidermal cell E-9.3.1: Part of animal
Animal cell that is present at and makes up the epidermis of an organism.

epithelial cell E-9.3.1: Part of animal
Animal cell that is present at and forms a thin tissue barrier called an epithelium, which can be on the surface or on the inside of any tissue.

granulosa cell E-9.3.1: Part of animal
follicular cell
Animal cell that is present at and forms the membrana granulosa lining the vesicular ovarian follicle and is associated with the developing female gamete in the ovary of mammals.

muscle cell E-9.3.1: Part of animal
myocyte
Animal cell that is present at and forms the muscles of the body and that is capable of shorten its length using a series of motor proteins.

nerve cell E-9.3.1: Part of animal
neurocyte
neurone
Animal cell that is present at and constitutes the nervous system, consisting of the nerve cell body, dendrites, and axon.

ORIGIN-BASED HYPONYMY

zygote E-9.3.1: Part of animal
Animal cell resulting from the union of two gametes (an ovum and a spermatozoon), so it is a fertilized egg cell.

plant cell E-9.3.3: Part of plant
Eukaryotic cell that is a structural and functional unit of a plant, capable of performing photosynthesis.

yeast cell E-9.3.2: Part of fungus
Eukaryotic cell that reproduces vegetatively by budding or germinates to produce a mycelium.

RELATION-BASED HYPONYMY

daughter cell E-9.3: Part of lifeform
Cell formed in the division of a mother cell.

mother cell E-9.3: Part of lifeform
parent cell
Cell that divides to produce two or more daughter cells.

Table 36. Full hyponymy-based terminological entry of CELL

CELL, the third and last BIO hyponymy-based terminological entry, has the following information: 26 concepts with their definitions; 39 terms designating those concepts; four conceptual categories (i.e., *part of animal*, *part of fungus*, *part of lifeform*, and *part of plant*); eight hyponymy subtypes (i.e., *ability-based*, *color-based*, *composition-based*, *effect-based*, *function-based*, *location-based*, *origin-based*, and *relation-based* hyponymy); up to six hyponymy levels; and eleven concepts with hyponymic contexts. This entry stands out because it is the only one of the twelve entries that contains up to six hyponymy levels in the conceptual hierarchy.

Regarding conceptual categories, all of them are partitive: *part of animal*, *part of fungus*, *part of lifeform*, and *part of plant*. This is not surprising since organisms are composed of cells. The most frequent category is *part of animal* (e.g., BLOOD CELL, LYMPHOCYTE, T CELL). Since *part of lifeform* is a broader conceptual category than the others, it is applicable to those cases where CELL can be related to animals, fungi, and plants without any other distinction (e.g., EUKARYOTIC CELL, HOST CELL, MOTHER CELL).

The most important hyponymy subtypes in this entry are *function-based*, *location-based*, and *relation-based* hyponymy. The *function-based* hyponyms (e.g., B CELL, LYMPHOCYTE, MACROPHAGE) are characterized in terms of their usefulness or their role within an organism to protect the system or produce certain substances. *Location-based* hyponyms (e.g., BLOOD CELL, EPITHELIAL CELL, NERVE CELL) are identified by the part of the organism where they are located. Finally, *relation-based* hyponyms (e.g., ANIMAL CELL, MOTHER CELL, YEAST CELL) highlight their association with another concept that they affect or depend on in some way.

Regarding hyponymy levels, both co-hyponyms B CELL and T CELL are classified in six levels, following this schema: B CELL / T CELL is a *function-based* type of LYMPHOCYTE, which is a *function-based* type of WHITE BLOOD CELL, which is a *color-based* type of BLOOD CELL, which is a *location-based* type of ANIMAL CELL, which is a *relation-based* type of EUKARYOTIC CELL, which is a *composition-based* type of CELL. In fact, these are the most complex hyponymic hierarchies of all the entries, since this is the only occasion in which there are six hyponymy levels. Moreover, the conceptual complexity behind them can be appreciated, since five different typologies of hyponymy subtypes are traversed to reach the last hyponyms.

Moreover, the hyponymic contexts identified have KPs expressing generic-specific relations (e.g., "animal cells and other eukaryotic cells"; "blood cells, proteins, and other useful large molecules in the blood"; "blood cells such as eosinophils and lymphocytes"; etc.). Finally, **Table 37** summarizes the information in the hyponymy-based terminological entry of CELL.

CELL entry

Concepts, terms, and terminological definitions	
No. of concepts	26
No. of terms	39
No. of terminological definitions	26

Conceptual categories	
No. of conceptual categories	4
Conceptual categories	*part of animal* ǀ *part of fungus* ǀ *part of lifeform* ǀ *part of plant*
No. of *part of animal* concepts	18
***Part of animal* concepts**	ANIMAL CELL ǀ B CELL ǀ BLOOD CELL ǀ CANCER CELL ǀ CHROMATOPHORE ǀ EOSINOPHIL ǀ EPIDERMAL CELL ǀ EPITHELIAL CELL ǀ GRANULOSA CELL ǀ LYMPHOCYTE ǀ MACROPHAGE ǀ MUSCLE CELL ǀ NERVE CELL ǀ RED BLOOD CELL ǀ STEM CELL ǀ T CELL ǀ WHITE BLOOD CELL ǀ ZYGOTE
No. of *part of fungus* concepts	1
***Part of fungus* concepts**	YEAST CELL
No. of *part of lifeform* concepts	6
***Part of lifeform* concepts**	CELL ǀ DAUGHTER CELL ǀ EUKARYOTIC CELL ǀ HOST CELL ǀ MOTHER CELL ǀ TARGET CELL
No. of *part of plant* concepts	1
***Part of plant* concepts**	PLANT CELL

Hyponymy subtypes	
No. of hyponymy subtypes	8
Hyponymy subtypes	*ability-based* ǀ *color-based* ǀ *composition-based* ǀ *effect-based* ǀ *function-based* ǀ *location-based* ǀ *origin-based* ǀ *relation-based*

No. of *ability-based* hyponyms	3
***Ability-based* hyponyms**	CHROMATOPHORE ǀ HOST CELL ǀ TARGET CELL
No. of *color-based* hyponyms	2
***Color-based* hyponyms**	RED BLOOD CELL ǀ WHITE BLOOD CELL
No. of *composition-based* hyponyms	1
***Composition-based* hyponyms**	EUKARYOTIC CELL
No. of *effect-based* hyponyms	1
***Effect-based* hyponyms**	CANCER CELL
No. of *function-based* hyponyms	6
***Function-based* hyponyms**	B CELL ǀ EOSINOPHIL ǀ LYMPHOCYTE ǀ MACROPHAGE ǀ STEM CELL ǀ T CELL
No. of *location-based* hyponyms	6
***Location-based* hyponyms**	BLOOD CELL ǀ EPIDERMAL CELL ǀ EPITHELIAL CELL ǀ GRANULOSA CELL ǀ MUSCLE CELL ǀ NERVE CELL
No. of *origin-based* hyponyms	1
***Origin-based* hyponyms**	ZYGOTE
No. of *relation-based* hyponyms	5
***Relation-based* hyponyms**	ANIMAL CELL ǀ DAUGHTER CELL ǀ MOTHER CELL ǀ PLANT CELL ǀ YEAST CELL

Hyponymy levels	
No. of hyponymy levels	6
No. of first-level hyponyms	5
First-level hyponyms	DAUGHTER CELL ǀ EUKARYOTIC CELL ǀ HOST CELL ǀ MOTHER CELL ǀ TARGET CELL
No. of second-level hyponyms	3
Second-level hyponyms	ANIMAL CELL ǀ PLANT CELL ǀ YEAST CELL
No. of third-level hyponyms	10
Third-level hyponyms	BLOOD CELL ǀ CANCER CELL ǀ CHROMATOPHORE ǀ EPIDERMAL CELL ǀ EPITHELIAL CELL ǀ GRANULOSA CELL ǀ MUSCLE CELL ǀ NERVE CELL ǀ STEM CELL ǀ ZYGOTE
No. of fourth-level hyponyms	2
Fourth-level hyponyms	RED BLOOD CELL ǀ WHITE BLOOD CELL
No. of fifth-level hyponyms	3
Fifth-level hyponyms	EOSINOPHIL ǀ LYMPHOCYTE ǀ MACROPHAGE
No. of sixth-level hyponyms	2
Sixth-level hyponyms	B CELL ǀ T CELL

Hyponymic contexts	
No. of concepts with hyponymic contexts	11
Concepts with hyponymic contexts	ANIMAL CELL ǀ B CELL ǀ BLOOD CELL ǀ CANCER CELL ǀ CELL ǀ EOSINOPHIL ǀ EUKARYOTIC CELL ǀ LYMPHOCYTE ǀ MACROPHAGE ǀ T CELL ǀ WHITE BLOOD CELL

Table 37. Summary of the information in the terminological entry of CELL

5.1.4. SLUDGE

The hyponymy-based terminological entry of SLUDGE is shown in **Table 38**.

sludge E-8.2: Fluid matter | E-8.4: Solid matter
Semi-solid mud or sediment that is usually produced from a range of industrial processes.
- *Incinerators can be designed to accept wastes of any physical form, including gases, liquids, solids, SLUDGES, and slurries.*

MOISTURE-BASED HYPONYMY

wet sludge E-8.2: Fluid matter
Sludge with a high quantity of water.

ORIGIN-BASED HYPONYMY

oily sludge E-8.4: Solid matter
oil sludge
bottom sludge
tank bottom sludge
Solid sludge produced by the petroleum industry, which contains emulsified petroleum hydrocarbons, heavy metals, and solid particles.

sewage sludge E-8.2: Fluid matter | E-8.4: Solid matter
wastewater sludge
waste sludge
residual sludge
Sludge produced as a by-product of wastewater treatment processes.

MOISTURE-BASED HYPONYMY

dewatered sludge E-8.4: Solid matter
dry sludge
dried sludge
sludge cake
dewatered sludge cake
Sewage sludge which has had water removed from it before disposal.
- *This ensures that all parts of the compost mixture, especially the SLUDGE CAKE, spends an adequate period in the centre of the pile at 55°C.*

liquid sludge E-8.2: Fluid matter
non-dewatered sludge
thin sludge
Sewage sludge that has not been dewatered before disposal.

ORIGIN-BASED HYPONYMY

API separator sludge E-8.2: Fluid matter | E-8.4: Solid matter
API sludge
Sewage sludge obtained from an API separator which is used for treating wastewater from the cleaning of oil storage tanks, and which mainly consists of heavy organic compounds and combustible matter.

municipal sludge E-8.2: Fluid matter
municipal sewage sludge
Sewage sludge from the municipal wastewater treatment process.

STATUS-BASED HYPONYMY

bulking sludge E-8.2: Fluid matter | E-8.4: Solid matter
Sewage sludge with extremely bad settling and thickening characteristics.

digested sludge E-8.2: Fluid matter
Sewage sludge that has undergone a digestion process.

raw sludge E-8.2: Fluid matter | E-8.4: Solid matter
untreated sludge
Sewage sludge that has not been subject to any treatment, digestion or stabilization process.

stabilized sludge E-8.2: Fluid matter
Sewage sludge that has undergone a stabilization process.

treated sludge E-8.2: Fluid matter
Sewage sludge that has undergone biological, chemical or heat treatment, long-term storage or any other appropriate process so as to significantly reduce its fermentability and the health hazards resulting from its use.

METHOD-BASED HYPONYMY

primary sludge E-8.2: Fluid matter
Treated sludge generated during primary treatment; resulting from chemical precipitation, sedimentation or other primary process; taken from the bottom of the primary settling tanks; and composed of settleable raw solids.

METHOD-BASED HYPONYMY

chemical sludge E-8.2: Fluid matter
Primary sludge that results from the use of chemicals to remove constituents through precipitation.

secondary sludge E-8.2: Fluid matter
biological sludge
settled sludge
Treated sludge generated during secondary treatment; resulting from biological treatments which involve removing organic matter from sewage by saturating it with air and microbial organisms grown in a settling tank; and mainly consisting of activated waste biomass.

COMPOSITION-BASED HYPONYMY

activated sludge E-8.2: Fluid matter
active sludge
conventional activated sludge
Secondary sludge that contains a flocculent culture of organisms developed in aeration tanks under controlled conditions.

ORIGIN-BASED HYPONYMY

excess sludge **E-8.2: Fluid matter**
waste activated sludge
WAS
surplus activated sludge
SAS
Activated sludge that results from decanting off the cleaner water and is removed from the treatment process to keep the ratio of biomass to food supplied in the wastewater in balance.

returned sludge **E-8.2: Fluid matter**
return activated sludge
RAS
return sludge
Activated sludge that results from decanting off the cleaner water and is transported back to the aeration basin.

undigested sludge **E-8.2: Fluid matter | E-8.4: Solid matter**
Sewage sludge that has not undergone a digestion process.

Table 38. Full hyponymy-based terminological entry of SLUDGE

SLUDGE, the first CHEM hyponymy-based terminological entry, has the following information: 20 concepts with their definitions; 46 terms that designate those concepts; two conceptual categories (i.e., *fluid matter* and *solid matter*); five hyponymy subtypes (i.e., *composition-based*, *method-based*, *moisture-based*, *origin-based*, and *status-based* hyponymy); up to five hyponymy levels; and two concepts with hyponymic contexts. This is the richest entry of all CHEM entries with regard to concepts, terms, and hyponymy levels, but not in relation to the variety of conceptual categories and hyponymy subtypes.

There are two conceptual categories in this entry: *fluid matter* and *solid matter*. Most concepts are classified as *fluid matter* (e.g., ACTIVATED SLUDGE, CHEMICAL SLUDGE, EXCESS SLUDGE) though a few are *solid matter* (e.g., DEWATERED SLUDGE, OILY SLUDGE) and there are even many concepts which are members of both categories (e.g., API SEPARATOR SLUDGE, RAW SLUDGE, SEWAGE SLUDGE), due to the fact that they can adopt both states of matter without changing their intrinsic nature.

In this entry, the most important hyponymy subtypes are *origin-based* and *status-based* hyponymy. *Origin-based* hyponyms (e.g., EXCESS SLUDGE, MUNICIPAL SLUDGE, RETURNED SLUDGE) are determined by the phase of the treatment process or by their source, whereas *status-based* hyponyms (e.g., DIGESTED SLUDGE, STABILIZED

SLUDGE, TREATED SLUDGE) are characterized by the properties that they have or retain after the treatment process.

This entry also has co-hyponyms, such as EXCESS SLUDGE and RETURNED SLUDGE, which reach the fifth and last hyponymy level in this hierarchy. They follow the following schema: EXCESS SLUDGE / RETURNED SLUDGE is an *origin-based* type of ACTIVATED SLUDGE, which is a *composition-based* type of SECONDARY SLUDGE, which is a *method-based* type of BULKING SLUDGE, which is a *status-based* type of SEWAGE SLUDGE, which is an *origin-based* type of SLUDGE.

In this entry, only a couple of hyponymic KRCs are identified (i.e., "wastes of any physical form, including gases, liquids, solids, sludges, and slurries"; "all parts of the compost mixture, especially the sludge cake"). Finally, **Table 39** summarizes the information in the hyponymy-based terminological entry of SLUDGE.

SLUDGE entry

Concepts, terms, and terminological definitions	
No. of concepts	20
No. of terms	46
No. of terminological definitions	20

Conceptual categories	
No. of conceptual categories	2
Conceptual categories	*fluid matter* ǀ *solid matter*
No. of *fluid matter* concepts	18
***Fluid matter* concepts**	ACTIVATED SLUDGE ǀ API SEPARATOR SLUDGE ǀ BULKING SLUDGE ǀ CHEMICAL SLUDGE ǀ DIGESTED SLUDGE ǀ EXCESS SLUDGE ǀ LIQUID SLUDGE ǀ MUNICIPAL SLUDGE ǀ PRIMARY SLUDGE ǀ RAW SLUDGE ǀ RETURNED SLUDGE ǀ SECONDARY SLUDGE ǀ SEWAGE SLUDGE ǀ SLUDGE ǀ STABILIZED SLUDGE ǀ TREATED SLUDGE ǀ UNDIGESTED SLUDGE ǀ WET SLUDGE
No. of *solid matter* concepts	8
***Solid matter* concepts**	API SEPARATOR SLUDGE ǀ BULKING SLUDGE ǀ DEWATERED SLUDGE ǀ OILY SLUDGE ǀ RAW SLUDGE ǀ SEWAGE SLUDGE ǀ SLUDGE ǀ UNDIGESTED SLUDGE

Hyponymy subtypes	
No. of hyponymy subtypes	5
Hyponymy subtypes	*composition-based* ǀ *method-based* ǀ *moisture-based* ǀ *origin-based* ǀ *status-based*
No. of *composition-based* hyponyms	1
***Composition-based* hyponyms**	ACTIVATED SLUDGE
No. of *method-based* hyponyms	3
***Method-based* hyponyms**	CHEMICAL SLUDGE ǀ PRIMARY SLUDGE ǀ SECONDARY SLUDGE

No. of *moisture-based* hyponyms	3
Moisture-based hyponyms	DEWATERED SLUDGE ǀ LIQUID SLUDGE ǀ WET SLUDGE
No. of *origin-based* hyponyms	6
Origin-based hyponyms	API SEPARATOR SLUDGE ǀ EXCESS SLUDGE ǀ MUNICIPAL SLUDGE ǀ OILY SLUDGE ǀ RETURNED SLUDGE ǀ SEWAGE SLUDGE
No. of *status-based* hyponyms	6
Status-based hyponyms	BULKING SLUDGE ǀ DIGESTED SLUDGE ǀ RAW SLUDGE ǀ STABILIZED SLUDGE ǀ TREATED SLUDGE ǀ UNDIGESTED SLUDGE

Hyponymy levels	
No. of hyponymy levels	5
No. of first-level hyponyms	3
First-level hyponyms	OILY SLUDGE ǀ SEWAGE SLUDGE ǀ WET SLUDGE
No. of second-level hyponyms	10
Second-level hyponyms	API SEPARATOR SLUDGE ǀ BULKING SLUDGE ǀ DEWATERED SLUDGE ǀ DIGESTED SLUDGE ǀ LIQUID SLUDGE ǀ MUNICIPAL SLUDGE ǀ RAW SLUDGE ǀ STABILIZED SLUDGE ǀ TREATED SLUDGE ǀ UNDIGESTED SLUDGE
No. of third-level hyponyms	2
Third-level hyponyms	PRIMARY SLUDGE ǀ SECONDARY SLUDGE
No. of fourth-level hyponyms	2
Fourth-level hyponyms	ACTIVATED SLUDGE ǀ CHEMICAL SLUDGE
No. of fifth-level hyponyms	2
Fifth-level hyponyms	EXCESS SLUDGE ǀ RETURNED SLUDGE

Hyponymic contexts	
No. of concepts with hyponymic contexts	2
Concepts with hyponymic contexts	DEWATERED SLUDGE ǀ SLUDGE

Table 39. Summary of the information in the terminological entry of SLUDGE

5.1.5. NITROGEN

The hyponymy-based terminological entry of NITROGEN is shown in **Table 40**.

nitrogen E-8.1: Chemical substance
N
Non-metallic chemical element that under standard conditions is a colorless, odorless, tasteless gas constituting 78% of the Earth's atmosphere, represented by the symbol N and with atomic number 7.
■ *Atoms of other elements, typically oxygen, NITROGEN and sulphur, are incorporated into the basic hydrocarbon structures.*

EFFECT-BASED HYPONYMY

inorganic nitrogen E-8.1: Chemical substance
Nitrogen that forms compounds originating in inorganic materia (i.e., not containing carbon and hydrogen).

STATUS-BASED HYPONYMY

dissolved inorganic nitrogen E-8.1: Chemical substance
DIN
Inorganic nitrogen composed of nitrate, nitrite, and ammonium, which is readily available to phytoplankton and often controls the formation of blooms.

organic nitrogen E-8.1: Chemical substance
Nitrogen that forms compounds originating in organic materia (i.e., containing carbon and hydrogen).

STATUS-BASED HYPONYMY

dissolved organic nitrogen E-8.1: Chemical substance
DON
Organic nitrogen related to dissolved organic matter that is found in soils, sediments, seawater, and freshwater, and is a mixture of diverse molecules.

- *DON, such as urea and amino acids, is generally considered to be excreted at lower, but substantial levels compared to ammonia.*

LOCATION-BASED HYPONYMY

atmospheric nitrogen E-8.1: Chemical substance
Nitrogen that is located in the atmosphere.

marine nitrogen E-8.1: Chemical substance
oceanic nitrogen
ocean nitrogen
Nitrogen that is located in the sea or ocean.

riverine nitrogen E-8.1: Chemical substance
Nitrogen that is located in a river.

tissue nitrogen E-8.1: Chemical substance
Nitrogen that is located in tissues of organisms.

ORIGIN-BASED HYPONYMY

anthropogenic nitrogen E-8.1: Chemical substance
Nitrogen that is caused by human activity.

RELATION-BASED HYPONYMY

combined nitrogen E-8.1: Chemical substance
Nitrogen covalently bonded to one or more elements other than nitrogen.

COMPOSITION-BASED HYPONYMY

ammoniacal nitrogen E-8.1: Chemical substance
ammonical nitrogen
NH3-N

Combined nitrogen that is used as a measure for the amount of ammonia in landfill leachate and in waste products, or as a measure of the health of water in rivers, lakes or reservoirs.

- *Many of the soluble nutrients in sewage, mainly potassium salts and AMMONICAL NITROGEN derived from urine, are lost in the final effluent.*

mineral nitrogen E-8.1: Chemical substance

Combined nitrogen in the form of ammonium or nitrate salts which represents the small fraction of the total nitrogen that will be available to plants during the growing season.

oxidized nitrogen E-8.1: Chemical substance
oxidised nitrogen

Combined nitrogen consisting of nitrates and nitrites.

reactive nitrogen E-8.1: Chemical substance
Nr

Combined nitrogen that supports growth directly or indirectly, including gases such as nitrogen oxides, ammonia, nitrous oxide, and anion nitrate.

particulate nitrogen E-8.1: Chemical substance
PN

Nitrogen that is isolated by filtration with a filter capable of withstanding the combustion needed to convert the N on the filter to N2 gas so that a sample can be analyzed isotopically.

- *PN is the N that is retained by the filter.*

STATE-BASED HYPONYMY

gaseous nitrogen E-8.1: Chemical substance | E-8.2.2: Gas
molecular nitrogen
N_2

Nitrogen compound referred to the diatomic molecules of nitrogen formed when individual atoms bond, resulting in a colorless, odorless, and inert gas that is suitable for use in industrial applications.

EFFECT-BASED HYPONYMY

fixed nitrogen E-8.1: Chemical substance | E-8.2.2: Gas

Gaseous nitrogen that has been converted to ammonia, an ammonium ion, nitrate or another nitrogen oxide so that it can be used as a nutrient by living organisms.

STATE-BASED HYPONYMY

liquid nitrogen E-8.1: Chemical substance | E-8.2: Fluid matter
LN_2

Gaseous nitrogen in a liquid state at very low temperature, produced industrially by fractional distillation of liquid air, and widely used as a coolant.

Table 40. Full hyponymy-based terminological entry of NITROGEN

NITROGEN, the second CHEM hyponymy-based terminological entry, has the following information: 19 concepts with their definitions; 32 terms that designate those concepts; three conceptual categories (i.e., *chemical substance*, *fluid matter*, and *gas*); seven hyponymy subtypes (i.e., *composition-based*, *effect-based*, *location-based*, *origin-based*, *relation-based*, *state-based*, and *status-based* hyponymy); up to two hyponymy levels; and four concepts with hyponymic contexts. Of all CHEM entries, this is the entry with the most conceptual categories and hyponymy subtypes. However, it only has a few hyponymy levels.

The conceptual categories in the NITROGEN entry are *chemical substance*, *fluid matter*, and *gas*. Despite the fact that all concepts are classified as *chemical substance* (e.g., AMMONIACAL NITROGEN, COMBINED NITROGEN, MARINE NITROGEN), LIQUID NITROGEN is also classified as *fluid matter,* whereas FIXED NITROGEN and GASEOUS NITROGEN are also classified as *gas*. Even though these concepts are chemical substances, some of them are also determined by the type of matter.

The most important hyponymy subtypes in this entry are *composition-based*, *effect-based*, and *location-based* hyponymy. The *composition-based* hyponyms (e.g., AMMONIACAL NITROGEN, MINERAL NITROGEN, OXIDIZED NITROGEN) designate a component within their chemical structure that makes them a more specific type of nitrogen. The *effect-based* hyponyms (e.g., FIXED NITROGEN, INORGANIC NITROGEN, ORGANIC NITROGEN) are based on the material affected when they form compounds or on their use by living organisms to obtain nutrients. Finally, the *location-based* hyponyms (e.g., ATMOSPHERIC NITROGEN, MARINE NITROGEN, RIVERINE NITROGEN) are based on the environmental medium where they are found.

In this entry, there are only two hyponymy levels (e.g., DISSOLVED INORGANIC NITROGEN is a *status-based* type of INORGANIC NITROGEN, which is an *effect-based* type of NITROGEN; OXIDIZED NITROGEN is a *composition-based* type of COMBINED NITROGEN, which is a *relation-based* type of NITROGEN). This conceptual hierarchy is thus more horizontal than vertical. Its horizontal dimension highlights how co-hyponyms differ from each other. Moreover, *status-based* hyponymy is related to an additional category (e.g., GASEOUS NITROGEN is also *gas*, and LIQUID NITROGEN is also *fluid matter*).

In relation to hyponymic contexts, certain KRCs are detected (e.g., "atoms of other elements, typically oxygen, nitrogen and sulphur"; "DON, such as urea and amino acids"; "soluble nutrients in sewage, mainly potassium salts and ammonical nitrogen"). Finally, **Table 41** summarizes the information in the hyponymy-based terminological entry of NITROGEN.

NITROGEN entry

Concepts, terms, and terminological definitions	
No. of concepts	19
No. of terms	32
No. of terminological definitions	19

Conceptual categories	
No. of conceptual categories	3
Conceptual categories	*chemical substance* ǀ *fluid matter* ǀ *gas*
No. of *chemical substance* concepts	19
***Chemical substance* concepts**	AMMONIACAL NITROGEN ǀ ANTHROPOGENIC NITROGEN ǀ ATMOSPHERIC NITROGEN ǀ COMBINED NITROGEN ǀ DISSOLVED INORGANIC NITROGEN ǀ DISSOLVED ORGANIC NITROGEN ǀ FIXED NITROGEN ǀ GASEOUS NITROGEN ǀ INORGANIC NITROGEN ǀ LIQUID NITROGEN ǀ MARINE NITROGEN ǀ MINERAL NITROGEN ǀ NITROGEN ǀ ORGANIC NITROGEN ǀ OXIDIZED NITROGEN ǀ PARTICULATE NITROGEN ǀ REACTIVE NITROGEN ǀ RIVERINE NITROGEN ǀ TISSUE NITROGEN
No. of *fluid matter* concepts	1
***Fluid matter* concepts**	LIQUID NITROGEN
No. of *gas* concepts	2
***Gas* concepts**	FIXED NITROGEN ǀ GASEOUS NITROGEN

Hyponymy subtypes	
No. of hyponymy subtypes	7
Hyponymy subtypes	*composition-based* ǀ *effect-based* ǀ *location-based* ǀ *origin-based* ǀ *relation-based* ǀ *state-based* ǀ *status-based*
No. of *composition-based* hyponyms	4
***Composition-based* hyponyms**	AMMONIACAL NITROGEN ǀ MINERAL NITROGEN ǀ OXIDIZED NITROGEN ǀ REACTIVE NITROGEN
No. of *effect-based* hyponyms	3
***Effect-based* hyponyms**	FIXED NITROGEN ǀ INORGANIC NITROGEN ǀ ORGANIC NITROGEN
No. of *location-based* hyponyms	4
***Location-based* hyponyms**	ATMOSPHERIC NITROGEN ǀ MARINE NITROGEN ǀ RIVERINE NITROGEN ǀ TISSUE NITROGEN
No. of *origin-based* hyponyms	1
***Origin-based* hyponyms**	ANTHROPOGENIC NITROGEN
No. of *relation-based* hyponyms	2
***Relation-based* hyponyms**	COMBINED NITROGEN ǀ PARTICULATE NITROGEN
No. of *state-based* hyponyms	2
***State-based* hyponyms**	GASEOUS NITROGEN ǀ LIQUID NITROGEN
No. of *status-based* hyponyms	2
***Status-based* hyponyms**	DISSOLVED INORGANIC NITROGEN ǀ DISSOLVED ORGANIC NITROGEN

Hyponymy levels	
No. of hyponymy levels	2
No. of first-level hyponyms	10

First-level hyponyms	INORGANIC NITROGEN ǀ ORGANIC NITROGEN ǀ ATMOSPHERIC NITROGEN ǀ MARINE NITROGEN ǀ RIVERINE NITROGEN ǀ TISSUE NITROGEN ǀ ANTHROPOGENIC NITROGEN ǀ COMBINED NITROGEN ǀ PARTICULATE NITROGEN ǀ GASEOUS NITROGEN
No. of second-level hyponyms	8
Second-level hyponyms	AMMONIACAL NITROGEN ǀ DISSOLVED INORGANIC NITROGEN ǀ DISSOLVED ORGANIC NITROGEN ǀ FIXED NITROGEN ǀ LIQUID NITROGEN ǀ MINERAL NITROGEN ǀ OXIDIZED NITROGEN ǀ REACTIVE NITROGEN

Hyponymic contexts	
No. of concepts with hyponymic contexts	4
Concepts with hyponymic contexts	AMMONIACAL NITROGEN ǀ DISSOLVED ORGANIC NITROGEN ǀ NITROGEN ǀ PARTICULATE NITROGEN

Table 41. Summary of the information in the terminological entry of NITROGEN

5.1.6. MAIZE

The hyponymy-based terminological entry of MAIZE is shown in **Table 42**.

maize E-7.5: Plant
Tall annual grass (Zea mays) cultivated for its yellow edible grains, which develop on a spike.
- *[...] the Bar gene has been used to develop herbicide-resistant crops including canola, MAIZE, cotton, rice, and soybean [...].*

COMPOSITION-BASED HYPONYMY

high oil maize E-7.5: Plant
high lysine maize
Maize that has more protein, lysine and carotenoids than conventional maize.
- *HIGH LYSINE MAIZE is an example of a nutritionally enhanced crop that improves performance in animal studies.*

quality protein maize E-7.5: Plant
QPM
Maize that contains nearly twice as much lysine and tryptophan and is an example of biofortification.

FUNCTION-BASED HYPONYMY

commercial maize E-7.5: Plant
Maize that is suitable, adequate or prepared for commerce and human use.

LOCATION-BASED HYPONYMY

tropical maize E-7.5: Plant
Maize that has been cultivated in tropical regions between the Tropic of Cancer and the Tropic of Capricorn.

ORIGIN-BASED HYPONYMY

hybrid maize E-7.5: Plant
Maize that is produced by cross pollinating different inbred lines of maize.

inbred maize E-7.5: Plant
Maize that is produced as a result of inbreeding (i.e., breeding organisms that are closely related genetically).

transgenic maize E-7.5: Plant
genetically modified maize
GM maize
Maize that is genetically engineered to express agriculturally-desirable traits, such as resistance to pests and to herbicides.

ABILITY-BASED HYPONYMY

fertile transgenic maize E-7.5: Plant
Transgenic maize that is capable of producing offspring.

COMPOSITION-BASED HYPONYMY

opaque-2 maize E-7.5: Plant
o2 maize
Transgenic maize with an increased level of free amino acids in the mature endosperm, doubling the concentration of lysine and improving the protein quality of the grain.

EFFECT-BASED HYPONYMY

Bt maize E-7.5: Plant
Transgenic maize that expresses insecticidal proteins from the bacterium Bacillus thuringiensis, including Delta endotoxins.

Table 42. Full hyponymy-based terminological entry of MAIZE

MAIZE, the third and last CHEM hyponymy-based terminological entry, has the following information: eleven concepts with their definitions; 16 terms designating those concepts; one conceptual category (i.e., *plant*); six hyponymy subtypes (i.e., *ability-based*, *composition-based*, *effect-based*, *function-based*, *location-based*, and *origin-based* hyponymy); up to two hyponymy levels; and two concepts with hyponymic contexts. This is the poorest entry with regard to general content not only of the CHEM entries, but also of all entries. Despite the fact that it only has a few concepts, it has a variety of hyponymy subtypes.

Like BACTERIUM, this entry only has one conceptual category, *plant*, and the same reasoning applies. MAIZE is a plant, and both animal and plant species are hyponymically structured in similar way to a taxonomy. For this reason, since MAIZE cannot ever be or become another class of entity or species, it remains in the same conceptual category.

The most prevalent hyponymy subtypes in this entry are *composition-based* and *origin-based* hyponymy. *Composition-based* hyponyms (e.g., HIGH OIL MAIZE, OPAQUE-2 MAIZE, QUALITY PROTEIN MAIZE) are those that contain special chemical elements or nutrients that transform these types of MAIZE into more specific entities. On the other hand, the *origin-based* hyponyms (e.g., HYBRID MAIZE, INBRED MAIZE, TRANSGENIC MAIZE) are determined by the growing or farming of MAIZE, which determines their nature and characteristics.

As for hyponymy levels, since this terminological entry has less content in comparison to others, its conceptual hierarchy only has two levels. However, these hyponymic sequences all relate to the concept TRANSGENIC MAIZE by activating different dimensions or microsenses in the hyponyms. Accordingly, FERTILE TRANSGENIC MAIZE is an *ability-based* type of TRANSGENIC MAIZE; OPAQUE-2 MAIZE is a *composition-based* type of TRANSGENIC MAIZE; and BT MAIZE is an *effect-based* type of TRANSGENIC MAIZE. Afterwards, TRANSGENIC MAIZE is an *origin-based* type of MAIZE. This branching demonstrates the validity of hyponymy subtypes to intuitively specify which nuance characterizes each generic-specific relation, whilst also distinguishing the microsenses of co-hyponyms from each other.

The hyponymic contexts in this terminological entry are also limited, and there are only two examples with hyponymic KPs (i.e., "herbicide-resistant crops including canola, maize, cotton, rice, and soybean"; "high lysine maize is an example of a nutritionally enhanced crop"). Finally, **Table 43** summarizes the information in the hyponymy-based terminological entry of MAIZE.

MAIZE entry

Concepts, terms, and terminological definitions	
No. of concepts	11
No. of terms	16
No. of terminological definitions	11

Conceptual categories	
No. of conceptual categories	1
Conceptual categories	*plant*
No. of *plant* concepts	11
***Plant* concepts**	BT MAIZE ǀ COMMERCIAL MAIZE ǀ FERTILE TRANSGENIC MAIZE ǀ HIGH OIL MAIZE ǀ HYBRID MAIZE ǀ INBRED MAIZE ǀ MAIZE ǀ OPAQUE-2 MAIZE ǀ QUALITY PROTEIN MAIZE ǀ TRANSGENIC MAIZE ǀ TROPICAL MAIZE

Hyponymy subtypes	
No. of hyponymy subtypes	6
Hyponymy subtypes	*ability-based* ǀ *composition-based* ǀ *effect-based* ǀ *function-based* ǀ *location-based* ǀ *origin-based*
No. of *ability-based* hyponyms	1
***Ability-based* hyponyms**	FERTILE TRANSGENIC MAIZE
No. of *composition-based* hyponyms	3
***Composition-based* hyponyms**	HIGH OIL MAIZE ǀ OPAQUE-2 MAIZE ǀ QUALITY PROTEIN MAIZE
No. of *effect-based* hyponyms	1
***Effect-based* hyponyms**	BT MAIZE
No. of *function-based* hyponyms	1
***Function-based* hyponyms**	COMMERCIAL MAIZE
No. of *location-based* hyponyms	1
***Location-based* hyponyms**	TROPICAL MAIZE
No. of *origin-based* hyponyms	3
***Origin-based* hyponyms**	HYBRID MAIZE ǀ INBRED MAIZE ǀ TRANSGENIC MAIZE

Hyponymy levels	
No. of hyponymy levels	2
No. of first-level hyponyms	7
First-level hyponyms	COMMERCIAL MAIZE ǀ HIGH OIL MAIZE ǀ HYBRID MAIZE ǀ INBRED MAIZE ǀ QUALITY PROTEIN MAIZE ǀ TRANSGENIC MAIZE ǀ TROPICAL MAIZE
No. of second-level hyponyms	3
Second-level hyponyms	BT MAIZE ǀ FERTILE TRANSGENIC MAIZE ǀ OPAQUE-2 MAIZE

Hyponymic contexts	
No. of concepts with hyponymic contexts	2
Concepts with hyponymic contexts	HIGH OIL MAIZE ǀ MAIZE

Table 43. Summary of the information in the terminological entry of MAIZE

5.1.7. WASTEWATER

The hyponymy-based terminological entry of WASTEWATER is shown in **Table 44**.

wastewater E-8.2.3: Water
Water containing waste, usually referring to liquid matter discharged as useless from a manufacturing process, pumped through sewers, and generally treated at a wastewater treatment plant.
■ *[...] other waters such as WASTEWATER and river water may offer only limited ability to dilute these constituents by blending.*

MOVEMENT-BASED HYPONYMY

effluent E-8.2.3: Water
Wastewater or other liquid, partially treated or treated, that flows out from an industrial installation or a treatment plant.

influent E-8.2.3: Water
influent wastewater

Wastewater or other liquid, raw or partially treated, that flows into a reservoir, basin or treatment plant.

leachate E-8.2.3: Water
Wastewater that has percolated through soil and dissolved some soil materials in the process.

ORIGIN-BASED HYPONYMY

industrial wastewater E-8.2.3: Water
Wastewater from an industrial production, transformation or manipulation process in which water is used.

ORIGIN-BASED HYPONYMY

abattoir wastewater E-8.2.3: Water
slaughterhouse wastewater
meat processing wastewater
Industrial wastewater of a reddish color that contains animal blood, dissolved soils, oil and grease, gut contents, and urine, since it is the result of cleaning up slaughtered cattle, slaughter equipment, floors and personnel.

food processing wastewater E-8.2.3: Water
food-processing wastewater
Industrial wastewater generated during food production processes.

- *The use of a lagoon or pond covered with a geomembrane to treat high BOD wastewaters such as cow manure or FOOD PROCESSING WASTEWATERS.*

ORIGIN-BASED HYPONYMY

dairy wastewater E-8.2.3: Water
Food processing wastewater generated during dairy and milk production processes.

- *[...] two different types of wastewaters, namely DAIRY WASTEWATER and landfill leachate, consisting of a large range of substrates.*

olive mill wastewater E-8.2.3: Water
OMW
OMWW
Food processing wastewater composed of olive washing water, olive pulp water, water added to olive paste in the centrifugation step, and water coming from washing extraction plants.

sugar beet wastewater E-8.2.3: Water
Food processing wastewater discharged from sugar beet and generated during sugar production processes.

ORIGIN-BASED HYPONYMY

molasses wastewater E-8.2.3: Water
Sugar beet wastewater originating from producing molasses, which is a brown viscid syrup prepared from raw sugar.

hydraulic fracturing wastewater E-8.2.3: Water
Industrial wastewater originating from hydraulically fractured oil and gas wells.

petrochemical wastewater E-8.2.3: Water
Industrial wastewater generated during oil-related processes and originating from oilfield production, crude oil refinery plants, olefin process plants, refrigeration, and energy units.

PTA wastewater E-8.2.3: Water
PTA plant wastewater
purified terephthalic acid wastewater
Industrial wastewater generated during PTA manufacturing processes.

textile wastewater E-8.2.3: Water
Industrial wastewater generated during textile processes, containing dyes mixed with various contaminants.

- *Anaerobic biological treatment and activated carbon adsorption are often integrated for the treatment of high-strength, refractory, or toxic wastewaters such as phenolic or TEXTILE WASTEWATERS.*

surface runoff E-8.2.3: Water
overland flow
Wastewater resulting from excess precipitation on the ground surface that can no longer sufficiently rapidly infiltrate in the soil.

ORIGIN-BASED HYPONYMY

stormwater E-8.2.3: Water
rainwater
Surface runoff due to rainfall collected from roofs, impervious surfaces, and drainage systems.

- *Separation of different streams, such as STORMWATER, cooling water, process water, sanitary, sewage, etc.*

urban wastewater E-8.2.3: Water
municipal wastewater
Wastewater from urban population centers that is the result of domestic, public or industrial use, containing sanitary sewage and sometimes combining with stormwater or surface runoff.

COLOR-BASED HYPONYMY

blackwater E-8.2.3: Water
black water
sanitary wastewater
lavatory wastewater
Urban wastewater containing urine, feces, food residues and other waste resulting from animal and human use.

STATUS-BASED HYPONYMY

septic water E-8.2.3: Water
Blackwater in decomposition under anaerobic conditions.

greywater E-8.2.3: Water
grey water

graywater
gray water
domestic wastewater
Urban wastewater coming from domestic processes (e.g., washing dishes, laundry, bathing) that does not contain serious contaminants, so it can be recycled specially for use in gardening or for flushing toilets.

LOCATION-BASED HYPONYMY

sewage E-8.2.3: Water
sewerage
Urban wastewater that is carried off by sewers.
- *Anaerobic bacteria are common in SEWAGE and other wastewaters.*

STATUS-BASED HYPONYMY

raw wastewater E-8.2.3: Water
untreated wastewater
Wastewater that has not undergone any sort of physical, chemical or biological treatment process.

settled wastewater E-8.2.3: Water
Wastewater that remains in a tank before undergoing a treatment process.

treated wastewater E-8.2.3: Water
Wastewater that has been processed through a physical, chemical and biological procedure to reduce its organic matter content and meet the quality criteria to be discharged into the environment.

Table 44. Full hyponymy-based terminological entry of WASTEWATER

WASTEWATER, the first CIV hyponymy-based terminological entry, contains the following information: 25 concepts with their definitions; 44 terms designating those concepts; one conceptual category (i.e., *water*); five hyponymy subtypes (i.e., *color-based*, *location-based*, *movement-based*, *origin-based*, and *status-based* hyponymy); four hyponymy levels; and six concepts with hyponymic contexts. Interestingly, this entry is similar (e.g., number of concepts, conceptual categories, hyponymy subtypes) to the next CIV entry, BREAKWATER, although the WASTEWATER entry possesses more terms. However, in comparison to the POLLUTANT entry, both are smaller in volume.

In this entry, all hyponyms belong to the *water* category (e.g., FOOD PROCESSING WASTEWATER, INFLUENT, SEPTIC WATERS), regardless of the nuances of the hyponymy subtypes. In regard to hyponymy subtypes, there is a predominance of *origin-based* hyponymy in this terminological entry. Generic-specific relations in this conceptual hierarchy are thus mainly determined by this hyponymic type. The *origin-based* hyponyms (e.g., DAIRY WASTEWATER, INDUSTRIAL WASTEWATER, STORMWATER) are all characterized by the process that generated them, and are either related to natural phenomena (e.g., STORMWATER, SURFACE RUNOFF) or to artificial processes in industries (e.g., HYDRAULIC FRACTURING WASTEWATER, OLIVE MILL WASTEWATER, TEXTILE WASTEWATER).

As for hyponymy levels, MOLASSES WASTEWATER reaches the fourth level of the hierarchy as follows: MOLASSES WASTEWATER is an *origin-based* type of SUGAR BEET WASTEWATER, which is an *origin-based* type of FOOD-PROCESSING WASTEWATER, which is an *origin-based* type of INDUSTRIAL WASTEWATER, which is an *origin-based* type of WASTEWATER. All the hyponymy levels of this sequence depend on *origin-based* hyponymy, which confirms that this hyponymy subtype is the one marks the verticality of WASTEWATER concepts.

Furthermore, this terminological entry has hyponymic contexts showing KPs expressing hyponymy in different ways (e.g., "other waters such as wastewater and river water"; "toxic wastewater such as phenolic or textile wastewaters"; "anaerobic bacteria are common in sewage and other wastewaters"). Finally, **Table 45** summarizes the information in the hyponymy-based terminological entry of WASTEWATER.

WASTEWATER entry

Concepts, terms, and terminological definitions	
No. of concepts	25
No. of terms	44
No. of terminological definitions	25

Conceptual categories	
No. of conceptual categories	1
Conceptual categories	*water*
No. of *water* concepts	25
***Water* concepts**	ABATTOIR WASTEWATER ǀ BLACKWATER ǀ DAIRY WASTEWATER ǀ EFFLUENT ǀ FOOD PROCESSING WASTEWATER ǀ GREYWATER ǀ HYDRAULIC FRACTURING WASTEWATER ǀ INDUSTRIAL WASTEWATER ǀ INFLUENT ǀ LEACHATE ǀ MOLASSES WASTEWATER ǀ OLIVE MILL WASTEWATER ǀ PETROCHEMICAL WASTEWATER ǀ PTA WASTEWATER ǀ RAW WASTEWATER ǀ SEPTIC WATERS ǀ SETTLED WASTEWATER ǀ SEWAGE ǀ STORMWATER ǀ SUGAR BEET WASTEWATER ǀ SURFACE RUNOFF ǀ TEXTILE WASTEWATER ǀ TREATED WASTEWATER ǀ URBAN WASTEWATER ǀ WASTEWATER
Hyponymy subtypes	
No. of hyponymy subtypes	5

Hyponymy subtypes	*color-based* ǀ *location-based* ǀ *movement-based* ǀ *origin-based* ǀ *status-based*
No. of *color-based* **hyponyms**	2
Color-based **hyponyms**	BLACKWATER ǀ GREYWATER
No. of *location-based* **hyponyms**	1
Location-based **hyponyms**	SEWAGE
No. of *movement-based* **hyponyms**	3
Movement-based **hyponyms**	EFFLUENT ǀ INFLUENT ǀ LEACHATE
No. of *origin-based* **hyponyms**	14
Origin-based **hyponyms**	ABATTOIR WASTEWATER ǀ DAIRY WASTEWATER ǀ FOOD PROCESSING WASTEWATER ǀ HYDRAULIC FRACTURING WASTEWATER ǀ INDUSTRIAL WASTEWATER ǀ MOLASSES WASTEWATER ǀ OLIVE MILL WASTEWATER ǀ PETROCHEMICAL WASTEWATER ǀ PTA WASTEWATER ǀ STORMWATER ǀ SUGAR BEET WASTEWATER ǀ SURFACE RUNOFF ǀ TEXTILE WASTEWATER ǀ URBAN WASTEWATER
No. of *status-based* **hyponyms**	4
Status-based **hyponyms**	RAW WASTEWATER ǀ SEPTIC WATERS ǀ SETTLED WASTEWATER ǀ TREATED WASTEWATER

Hyponymy levels	
No. of hyponymy levels	4
No. of first-level hyponyms	9
First-level hyponyms	EFFLUENT ǀ INDUSTRIAL WASTEWATER ǀ INFLUENT ǀ LEACHATE ǀ RAW WASTEWATER ǀ SETTLED WASTEWATER ǀ SURFACE RUNOFF ǀ TREATED WASTEWATER ǀ URBAN WASTEWATER
No. of second-level hyponyms	10
Second-level hyponyms	ABATTOIR WASTEWATER ǀ BLACKWATER ǀ FOOD PROCESSING WASTEWATER ǀ GREYWATER ǀ HYDRAULIC FRACTURING WASTEWATER ǀ PETROCHEMICAL WASTEWATER ǀ PTA WASTEWATER ǀ SEWAGE ǀ STORMWATER ǀ TEXTILE WASTEWATER
No. of third-level hyponyms	4
Third-level hyponyms	DAIRY WASTEWATER ǀ OLIVE MILL WASTEWATER ǀ SEPTIC WATER ǀ SUGAR BEET WASTEWATER
No. of fourth-level hyponyms	1
Fourth-level hyponyms	MOLASSES WASTEWATER

Hyponymic contexts	
No. of concepts with hyponymic contexts	6
Concepts with hyponymic contexts	DAIRY WASTEWATER ǀ FOOD PROCESSING WASTEWATER ǀ SEWAGE ǀ STORMWATER ǀ TEXTILE WASTEWATER ǀ WASTEWATER

Table 45. Summary of the information in the terminological entry of WASTEWATER

5.1.8. BREAKWATER

The hyponymy-based terminological entry of BREAKWATER is shown in **Table 46**.

breakwater E-1.3.2: Defense structure

Coastal defense structure, generally parallel to the coastline, made of wood, concrete or stone, to protect the coast from the impact of the wave and to provide shelter for ports and harbors.

- *New shore protection structures such as seawalls, groins, BREAKWATERS, revetments and artificial reefs are increasingly being developed.*

COMPOSITION-BASED HYPONYMY

floating breakwater E-1.3.2: Defense structure

Breakwater consisting of a moored assembly of floating objects with a limited range of movement to protect vessels riding at anchor.

TECHNOLOGY-BASED HYPONYMY

box-type breakwater E-1.3.2: Defense structure

Floating breakwater with reinforced concrete modules that are either empty inside or have a core of light material.

rubble-mound breakwater E-1.3.2: Defense structure
rubble mound breakwater

Breakwater composed of a mound of non-selectively formed and placed stones which are protected with a covering layer of selected stones or of specially shaped concrete armored elements.

TECHNOLOGY-BASED HYPONYMY

berm breakwater E-1.3.2: Defense structure

Rubble-mound breakwater with a horizontal berm of armor stones at about sea-side water level, which is progressively modelled and shaped by the waves until it becomes stable.

permeable breakwater E-1.3.2: Defense structure

Rubble-mound breakwater with permeable sections to enable efficient seawater exchange in enclosed basins without investing additional energy on the seawater pumping.

reef breakwater E-1.3.2: Defense structure

Rubble-mound breakwater of single-sized stone with a crest at or below sea level, in such a way that it can be shaped by the waves.

sloping breakwater E-1.3.2: Defense structure

Rubble-mound breakwater protected by a layer in the form of concrete or large rocks shaped like a slope, so it is more suitable for use in soft soil conditions.

s-slope breakwater E-1.3.2: Defense structure

Rubble-mound breakwater with a gentle slope around still-water level and steeper slopes above and below, thus giving it the shape of an "s".

vertical breakwater E-1.3.2: Defense structure
wall breakwater
upright breakwater
vertical faced breakwater

Rubble-mound breakwater with a vertical and seaward face superstructure of masonry or concrete blocks to reflect waves so that they will not break.

TECHNOLOGY-BASED HYPONYMY

composite breakwater **E-1.3.2: Defense structure**
vertical composite breakwater
vertically composite breakwater
Vertical breakwater composed of an upright caisson structure and rubble mound which consists of core and revetments.

TECHNOLOGY-BASED HYPONYMY

caisson breakwater **E-1.3.2: Defense structure**
Composite breakwater made of a rubble-mound foundation, a vertical superstructure, and a watertight retaining substructure in the shape of a large box.

TECHNOLOGY-BASED HYPONYMY

dual cylindrical caisson breakwater **E-1.3.2: Defense structure**
DCBW
Caisson breakwater with two cylindrical walls – the outer cylinder has a hole on the upper portion, forming a wave chamber of a round shape surrounded by the inner impermeable cylinder.

perforated-wall caisson breakwater **E-1.3.2: Defense structure**
Caisson breakwater widely used in harbors and port areas to dissipate incident wave energy, allowing for safe navigation conditions during sea storms.

semicircular breakwater **E-1.3.2: Defense structure**
Composite breakwater composed of a precast reinforced concrete structure built with a semicircular vault and a bottom slab.

skirt breakwater **E-1.3.2: Defense structure**
curtainwall pile breakwater
Vertical breakwater consisting of a row of solid wall projecting from a certain depth of water to above water surface but not reaching down to the sea bottom, leaving a significant gap below it.

HEIGHT-BASED HYPONYMY

submerged breakwater **E-1.3.2: Defense structure**
Breakwater crowned at or below the still water level, protecting the coasts against erosion and port channels from sand deposition.

LOCATION-BASED HYPONYMY

coastal breakwater **E-1.3.2: Defense structure**

Breakwater located at the shoreline and commonly used to protect it against wave impact and erosion.

HEIGHT-BASED HYPONYMY

high-crested breakwater **E-1.3.2: Defense structure**
Coastal breakwater with its crest far above water level.

low-crested breakwater **E-1.3.2: Defense structure**
Coastal breakwater with its crest below water level or slightly appearing on the surface.

LOCATION-BASED HYPONYMY

headland breakwater **E-1.3.2: Defense structure**
shore-connected breakwater
Coastal breakwater that is connected to the shore or coastline.

offshore breakwater **E-1.3.2: Defense structure**
detached breakwater
exempt breakwater
nearshore breakwater
Coastal breakwater that is not connected to the shore or coastline.

LOCATION-BASED HYPONYMY

shore-parallel breakwater **E-1.3.2: Defense structure**
Offshore breakwater located parallel to the shoreline, usually used to stabilize the coast or to provide swimming areas that are sheltered from the incoming waves.

- *Wave overtopping nearshore coastal structures, such as SHORE-PARALLEL BREAKWATERS, can significantly alter the current circulation and sediment transport patterns around the structures [...].*

STATUS-BASED HYPONYMY

porous breakwater **E-1.3.2: Defense structure**
Breakwater characterized by its porosity and its depth of submergence that creates a tranquil water basin for the berthing of vessels.

Table 46. Full hyponymy-based terminological entry of BREAKWATER

BREAKWATER, the second CIV hyponymy-based terminological entry, has the following information: 24 concepts with their definitions; 35 terms that designate those concepts; one conceptual category (i.e., *defense structure*); five hyponymy subtypes (i.e., *composition-based*, *height-based*, *location-based*, *status-based*, and *technology-based* hyponymy); up to five hyponymy levels; and two concepts with hyponymic contexts. As previously mentioned, the BREAKWATER entry and WASTEWATER entry are similar in content though there are slight differences in their volume. However, since BREAKWATER is an artificial concept, and WASTEWATER is a natural one, their conceptualization differs.

This is another case of a terminological entry with only one conceptual category: *defense structure*. All BREAKWATER hyponyms (e.g., CAISSON BREAKWATER, HIGH-CRESTED BREAKWATER, RUBBLE-MOUND BREAKWATER) are artifacts in the form of defense structures, and their conceptual category does not vary, regardless of the number of hyponymy subtypes and levels in the conceptual hierarchy.

Of the many hyponymy subtypes in the entry, the one that predominates is *technology-based* hyponymy. These hyponyms (e.g., BOX-TYPE BREAKWATER, PERFORATED WALL-CAISSON BREAKWATER, SKIRT BREAKWATER) are characterized by elements or attributes (such as components or shape) that reflect their construction technology. This hyponymy subtype is thus typical of artificial entities such as inventions and structures, which are human-made.

Regarding hyponymy levels, the two co-hyponyms at the fifth and last level. are DUAL CYLINDRICAL CAISSON BREAKWATER and PERFORATED-WALL CAISSON BREAKWATER, expressed as follows: DUAL CYLINDRICAL CAISSON BREAKWATER / PERFORATED-WALL CAISSON BREAKWATER is a *technology-based* type of CAISSON BREAKWATER, which is a *technology-based* type of COMPOSITE BREAKWATER, which is a *technology-based* type of VERTICAL BREAKWATER, which is a *technology-based* type of RUBBLE-MOUND BREAKWATER, which is a *composition-based* type of BREAKWATER. The verticality in the generic-specific relations of BREAKWATER concepts is thus based on the technology that specifies their meaning. However, the first level is determined by the composition of BREAKWATER, since a RUBBLE-MOUND BREAKWATER is characterized by a mound composed of non-selectively formed and placed stones.

This entry only has two hyponymic contexts (i.e., "new shore protection structures such as seawalls, groins, breakwaters, revetments and artificial reefs"; "nearshore coastal structures, such as shore-parallel breakwaters"). The first context mentions ARTIFICIAL REEF and associates it with protection structures. This same context is also included in the REEF entry, where REEF is categorized as a *structure* rather than as a *landform.* Finally, **Table 47** summarizes the information in the hyponymy-based terminological entry of BREAKWATER.

BREAKWATER entry

Concepts, terms, and terminological definitions	
No. of concepts	24
No. of terms	35
No. of terminological definitions	24

Conceptual categories	
No. of conceptual categories	1
Conceptual categories	*defense structure*
No. of *defense structure* concepts	24
***Defense structure* concepts**	BERM BREAKWATER | BOX-TYPE BREAKWATER | BREAKWATER | CAISSON BREAKWATER | COASTAL BREAKWATER | COMPOSITE BREAKWATER | DUAL CYLINDRICAL CAISSON BREAKWATER | FLOATING BREAKWATER | HEADLAND BREAKWATER | HIGH-CRESTED BREAKWATER | LOW-CRESTED BREAKWATER | OFFSHORE BREAKWATER | PERFORATED-WALL CAISSON BREAKWATER | PERMEABLE BREAKWATER | POROUS BREAKWATER | REEF BREAKWATER | RUBBLE-MOUND BREAKWATER | SEMICIRCULAR BREAKWATER | SHORE-PARALLEL BREAKWATER | SKIRT BREAKWATER | SLOPING BREAKWATER | S-SLOPE BREAKWATER | SUBMERGED BREAKWATER | VERTICAL BREAKWATER

Hyponymy subtypes	
No. of hyponymy subtypes	5
Hyponymy subtypes	*composition-based* | *height-based* | *location-based* | *status-based* | *technology-based*
No. of *composition-based* hyponyms	2
***Composition-based* hyponyms**	FLOATING BREAKWATER | RUBBLE-MOUND BREAKWATER
No. of *height-based* hyponyms	3
***Height-based* hyponyms**	HIGH-CRESTED BREAKWATER | LOW-CRESTED BREAKWATER | SUBMERGED BREAKWATER
No. of *location-based* hyponyms	4
***Location-based* hyponyms**	COASTAL BREAKWATER | HEADLAND BREAKWATER | OFFSHORE BREAKWATER | SHORE-PARALLEL BREAKWATER
No. of *status-based* hyponyms	1
***Status-based* hyponyms**	POROUS BREAKWATER
No. of *technology-based* hyponyms	13
***Technology-based* hyponyms**	BERM BREAKWATER | BOX-TYPE BREAKWATER | CAISSON BREAKWATER | COMPOSITE BREAKWATER | DUAL CYLINDRICAL CAISSON BREAKWATER | PERFORATED WALL-CAISSON BREAKWATER | PERMEABLE BREAKWATER | REEF BREAKWATER | SEMICIRCULAR BREAKWATER | SKIRT BREAKWATER | SLOPING BREAKWATER | S-SLOPE BREAKWATER | VERTICAL BREAKWATER

Hyponymy levels	
No. of hyponymy levels	5
No. of first-level hyponyms	5

First-level hyponyms	COASTAL BREAKWATER ǀ FLOATING BREAKWATER ǀ POROUS BREAKWATER ǀ RUBBLE-MOUND BREAKWATER ǀ SUBMERGED BREAKWATER
No. of second-level hyponyms	11
Second-level hyponyms	BERM BREAKWATER ǀ BOX-TYPE BREAKWATER ǀ HEADLAND BREAKWATER ǀ HIGH-CRESTED BREAKWATER ǀ LOW-CRESTED BREAKWATER ǀ OFFSHORE BREAKWATER ǀ PERMEABLE BREAKWATER ǀ REEF BREAKWATER ǀ SLOPING BREAKWATER ǀ S-SLOPE BREAKWATER ǀ VERTICAL BREAKWATER
No. of third-level hyponyms	3
Third-level hyponyms	COMPOSITE BREAKWATER ǀ SHORE-PARALLEL BREAKWATER ǀ SKIRT BREAKWATER
No. of fourth-level hyponyms	2
Fourth-level hyponyms	CAISSON BREAKWATER ǀ SEMICIRCULAR BREAKWATER
No. of fifth-level hyponyms	2
Fifth-level hyponyms	DUAL CYLINDRICAL CAISSON BREAKWATER ǀ PERFORATED-WALL CAISSON BREAKWATER

Hyponymic contexts	
No. of concepts with hyponymic contexts	2
Concepts with hyponymic contexts	BREAKWATER ǀ SHORE-PARALLEL BREAKWATER

Table 47. Summary of the information in the terminological entry of BREAKWATER

5.1.9. POLLUTANT

The hyponymy-based terminological entry of POLLUTANT is shown in **Table 48**.

pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter

Physical, chemical, biological or radiological substance that has an adverse effect on air, water or soil.

- *More harmful POLLUTANTS, such as sulfur and nitrogen oxides, were ignored, because people were unaware of their existence.*

ABILITY-BASED HYPONYMY

metal pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter

Pollutant that is a chemical element characterized by its metallic luster, its capacity to lose electrons and form a positive ion, and the ability to conduct heat and electricity.

DENOMINATION-BASED HYPONYMY

lead E-8.1: Chemical substance | E-8.4: Solid matter

Pb

Metal pollutant which is a chemical element with symbol Pb and atomic number 82.

- *Pollution elements such as PB, Zn, V and Sb showed their lowest concentration, further indicating that the cleaner air was from the sea.*

COMPOSITION-BASED HYPONYMY

inorganic pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter

Pollutant that does not contain carbon elements in its composition.

organic pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter
Pollutant that contains carbon elements in its composition.

- *[...] higher concentrations of ORGANIC POLLUTANTS such as PCBs, DDE, dieldrin, and chlordane [...].*

EFFECT-BASED HYPONYMY

persistent organic pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter
POP
Organic pollutant resulting from a manufacturing process and that is capable of bioaccumulating in human and animal tissue, with potential significant impacts on health and the environment.

- *PCBs are categorised as a POP.*

DENOMINATION-BASED HYPONYMY

priority pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter
Pollutant in a list of 126 specific pollutants that includes heavy metals and specific organic chemicals which are frequently found in wastewater.

- *For most of the adsorbable PRIORITY POLLUTANTS such as benzene, chlorobenzenes, phenol, chlorophenols, and nitrophenols, removal efficiencies exceeded 99%.*

FUNCTION-BASED HYPONYMY

pesticide E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter
Pollutant used to kill or control any type of pest (e.g., rodents, insects, weeds).

- [...] *the artificial marsh and solar aquatic systems [...] have even been shown capable of removing pollutants such as heavy metals, PESTICIDES, and industrial toxins.*

COMPOSITION-BASED HYPONYMY

dichlorodiphenyltrichloroethane E-8.1: Chemical substance | E-8.2: Fluid matter
DDT
Pesticide with formula (ClC6H4)2CH(CCl3) used mainly as an insecticide.

- *Wastewaters should be checked for pesticides such as DDT and PCP and for metals such as mercury, arsenic, and copper.*

methyl bromide E-8.1: Chemical substance | E-8.2: Fluid matter
bromomethane
CH3Br
Pesticide with formula CH3Br used mainly as a fumigant, but also as a solvent, refrigerant, and in organic synthesis.

- *Certain industrial chemicals, such as refrigerants, halons, and METHYL BROMIDE (a deadly pesticide used on crops), destroy Earth's ozone layer.*

LOCATION-BASED HYPONYMY

air pollutant E-8.1: Chemical substance | E-8.2: Fluid matter
atmospheric pollutant
Pollutant in air that, at a sufficiently high concentration, can cause disease to humans and animals, kill vegetation, and damage structures.

■ *Once released, AIR POLLUTANTS, such as SOx and NOx, are transported over long distances.*

COMPOSITION-BASED HYPONYMY

aromatic pollutant E-8.1: Chemical substance | E-8.2: Fluid matter
Air pollutant possessing one or more benzene rings.

COMPOSITION-BASED HYPONYMY

benzene E-8.1: Chemical substance | E-8.2: Fluid matter
benzol
C6H6
Aromatic pollutant with formula C6H6 derived from petroleum and used in or to manufacture a wide variety of chemical products, including detergents, insecticides, and motor fuels.

■ *[...] individual aromatics such as BENZENE, toluene, and naphthalene are toxic at concentrations about 100 times lower.*

polychlorinated biphenyl E-8.1: Chemical substance | E-8.2: Fluid matter
PCB
Aromatic pollutant used in industry with electrical insulators and manufacturing plastics that can be toxic for organisms.

■ *Both pathogens and micro-pollutants, such as PCB and PAH, have a high affinity for particulate matter.*

ORIGIN-BASED HYPONYMY

aerosol E-8.1: Chemical substance | E-8.2: Fluid matter
Air pollutant dispensed from a small metal container by a propellant under pressure.

primary pollutant E-8.1: Chemical substance | E-8.2: Fluid matter
Air pollutant emitted from a source directly into the atmosphere.

■ *[...] PRIMARY POLLUTANTS such as sulfur dioxides and nitrogen oxides [...].*

secondary pollutant E-8.1: Chemical substance | E-8.2: Fluid matter
Air pollutant produced by the reaction of a primary pollutant with some other component in the air.

STATE-BASED HYPONYMY

gaseous pollutant E-8.1: Chemical substance | E-8.2.2: Gas
Air pollutant manifested in the form of gas.

COMPOSITION-BASED HYPONYMY

dioxin E-8.1: Chemical substance | E-8.2.2: Gas
Gaseous pollutant found in a number of chemical products as lipophilic contaminant which originates from industrial processes (e.g., pesticide manufacture, papermaking) and waste incineration.

■ *Chemical dehalogenation can be an effective process for removing halogens from hazardous organic compounds such as DIOXINS, furans, PCBs, and chlorinated pesticides.*

nitrogen dioxide E-8.1: Chemical substance | E-8.2.2: Gas

NO2

Gaseous pollutant with formula NO2 often found in smog and automobile exhaust fumes, and synthesized for use as a nitrating agent, a catalyst and an oxidizing agent.

- *Flue gas after it has left the chimney that includes suspended dust, grit, fly ash […], and rarely a coloured gas such as NITROGEN DIOXIDE.*

nitrogen oxide E-8.1: Chemical substance | E-8.2.2: Gas

NOx

Gaseous pollutant composed of nitrogen and oxygen which is produced in combustion.

- *Because ozone is a secondary pollutant, its formation is controlled by the concentrations of other pollutants, namely NITROGEN OXIDES and hydrocarbons (VOCs).*

sulfur dioxide E-8.1: Chemical substance | E-8.2.2: Gas

sulphur dioxide

SO2

Gaseous pollutant with formula SO2 produced from volcanic eruptions, ocean spray, organic decomposition and the burning of fossil fuels, which is a component in the creation of acid precipitation.

- *The troposphere also contains trace amounts of many other gases, such as methane, various nitrogen oxides, ammonia, SULFUR DIOXIDE, and ozone, and these come from both natural and anthropogenic sources.*

EFFECT-BASED HYPONYMY

greenhouse gas E-8.1: Chemical substance | E-8.2.2: Gas

GHG

Gaseous pollutant composed of molecules that absorb and re-radiate infrared electromagnetic radiation, thus heating the Earth's atmosphere.

- *GREENHOUSE GASES (GHGs) in the atmosphere –such as carbon dioxide (CO2), methane, nitrous oxide, water vapor, and ozone– […].*

COMPOSITION-BASED HYPONYMY

carbon dioxide E-8.1: Chemical substance | E-8.2.2: Gas

CO2

Greenhouse gas with formula CO2 produced by respiration and combustion of carbon-containing fuels, needed by plants for photosynthesis, and used in food refrigeration, carbonated drinks, and fire extinguishers.

- *Atmospheric gases, such as CARBON DIOXIDE, water vapor, and methane, that trap heat radiation from Earth's surface by absorbing it and reemitting it.*

methane E-8.1: Chemical substance | E-8.2.2: Gas

CH4

Greenhouse gas with formula CH4 which is the main constituent of natural gas and is used as a fuel.

- *GHGs, such as carbon dioxide (CO2) and METHANE, are often trapped within frozen ground.*

ozone E-8.1: Chemical substance | E-8.2.2: Gas

O3

Greenhouse gas with formula O3 formed naturally in the ozone layer from atmospheric oxygen by electric discharge or exposure to ultraviolet radiation and that is used to deodorize air, purify water, and treat industrial wastes.

- *[...] sulfur dioxide may not be the primary cause of plant injury, and other pollutants such as OZONE may have a greater impact.*

LOCATION-BASED HYPONYMY

tropospheric ozone E-8.1: Chemical substance | E-8.2.2: Gas
ground-level ozone
Ozone located at the troposphere (i.e., the lowest level of the Earth's atmosphere).

- *[...] benzene in the lower atmosphere will react with other atmospheric components, contributing to the formation of GROUND-LEVEL OZONE and other air pollutants [...].*

photochemical oxidant E-8.1: Chemical substance | E-8.2.2: Gas
Gaseous pollutant which enters into oxidation reactions in the presence of light or other radiant energy.

COMPOSITION-BASED HYPONYMY

peroxyacyl nitrate E-8.1: Chemical substance | E-8.2.2: Gas
PAN
acyl peroxy nitrate
APN
Photochemical oxidant that is a powerful respiratory and eye irritant present in photochemical smog.

- *Any strong oxidising agent in photochemical smog, mainly ozone, nitrogen dioxide or PAN.*

soil pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter
Pollutant in the soil that negatively affects the edaphic biota, plants, animal life and human health.

- *This reagent is a good oxidizer for herbicides and other SOIL POLLUTANTS such as hexadecane or Dieldrin.*

water pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter
Pollutant in any body of water that alters its quality and negatively affects human health, aquatic ecosystems and associated terrestrial ecosystems.

ORIGIN-BASED HYPONYMY

anthropogenic pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter
Pollutant resulting from human activities, including polluting residuals from consumption and production activity.

- *Deposition can include a wide variety of natural and ANTHROPOGENIC POLLUTANTS, including inorganic elements and compounds [...] and organic compounds [...].*

EFFECT-BASED HYPONYMY

acidifying pollutant E-8.1: Chemical substance | E-8.2: Fluid matter | E-8.4: Solid matter
Anthropogenic pollutant resulting from emissions of sulfur dioxide, nitrogen oxides and ammonia, emitted under the combustion of fossil fuel in electricity generating power stations, industrial plants, residential heating, commercial, and service sectors.

Table 48. Full hyponymy-based terminological entry of POLLUTANT

POLLUTANT, the third and last CIV hyponymy-based terminological entry, contains the following information: 33 concepts with their definitions; 55 terms designating those concepts; four conceptual categories (i.e., *chemical substance*, *fluid matter*, *gas*, and *solid matter*); eight hyponymy subtypes (i.e., *ability-based*, *composition-based*, *denomination-based*, *effect-based*, *function-based*, *location-based*, *origin-based*, and *state-based* hyponymy); up to 5 hyponymy levels; and 24 concepts with hyponymic contexts. This is the richest CIV entry in terms of number of concepts, terms, types of conceptual categories, and variety of hyponymy subtypes.

This terminological entry has various conceptual categories: *chemical substance*, *fluid matter*, *gas*, and *solid matter*. All concepts from the hypernym to the last hyponym are classified as *chemical substance* (e.g., AIR POLLUTANT, CARBON DIOXIDE, TROPOSPHERIC OZONE). However, since the hierarchy has various levels, the hyponyms acquire characteristics that often involve a change of state. This generates additional categories such as *fluid matter* (e.g., AEROSOL, ORGANIC POLLUTANT, WATER POLLUTANT), *gas* (e.g., GREENHOUSE GAS, METHANE, OZONE), and *solid matter* (e.g., ANTHROPOGENIC POLLUTANT, LEAD, SOIL POLLUTANT). Moreover, certain hyponyms belong to more than one category, depending on how broad their conceptualization is (e.g., ORGANIC POLLUTANT, PRIORITY POLLUTANT, SOIL POLLUTANT).

Furthermore, the predominant hyponymy subtype in this terminological entry is clearly *composition-based* hyponymy. Its hyponyms (e.g., AROMATIC POLLUTANT, CARBON DIOXIDE, PEROXYACYL NITRATE) are determined by their chemical elements or compounds. Given that these concepts belong to the category of *chemical substance*, it is not surprising that this hierarchy hinges on composition.

The only hyponym reaching the fifth hyponymy level is TROPOSPHERIC OZONE, which has the following schema: TROPOSPHERIC OZONE is a *location-based* type of OZONE, which is a *composition-based* type of GREENHOUSE GAS, which is an *effect-based* type of GASEOUS POLLUTANT, which is a *state-based* type of AIR POLLUTANT, which is a *location-based* type of POLLUTANT. Therefore, although *composition-based* hyponymy is the most prevalent among POLLUTANT concepts, it does not determine verticality because it is conveyed by a combination of nuances specified by different hyponymy subtypes. POLLUTANT hyponyms thus possess a complex conceptualization that takes into account very different characteristics when their meaning is specified at further hyponymy levels.

In relation to hyponymic contexts, this terminological entry has a wide variety of examples showing different KPs involving various terms (e.g., "more harmful pollutants, such as sulfur and nitrogen oxides"; "most of the adsorbable priority pollutants such as benzene, chlorobenzenes, phenol, chlorophenols, and nitrophenols"; "ground-level ozone and other air pollutants"). Finally, **Table 49** summarizes the information in the hyponymy-based terminological entry of POLLUTANT.

POLLUTANT entry

Concepts, terms, and terminological definitions	
No. of concepts	33
No. of terms	55
No. of terminological definitions	33

Conceptual categories	
No. of conceptual categories	4
Conceptual categories	*chemical substance* I *fluid matter* I *gas* I *solid matter*
No. of *chemical substance* concepts	33
***Chemical substance* concepts**	ACIDIFYING POLLUTANT I AEROSOL I AIR POLLUTANT I ANTHROPOGENIC POLLUTANT I AROMATIC POLLUTANT I BENZENE I CARBON DIOXIDE I DICHLORODIPHENYLTRICHLOROETHANE I DIOXIN I GASEOUS POLLUTANT I GREENHOUSE GAS I INORGANIC POLLUTANT I LEAD I METAL POLLUTANT I METHANE I METHYL BROMIDE I NITROGEN DIOXIDE I NITROGEN OXIDE I ORGANIC POLLUTANT I OZONE I PEROXYACYL NITRATE I PERSISTENT ORGANIC POLLUTANT I PESTICIDE I PHOTOCHEMICAL OXIDANT I POLLUTANT I POLYCHLORINATED BIPHENYL I PRIMARY POLLUTANT I PRIORITY POLLUTANT I SECONDARY POLLUTANT I SOIL POLLUTANT I SULFUR DIOXIDE I TROPOSPHERIC OZONE I WATER POLLUTANT
No. of *fluid matter* concepts	20
***Fluid matter* concepts**	ACIDIFYING POLLUTANT I AEROSOL I AIR POLLUTANT I ANTHROPOGENIC POLLUTANT I AROMATIC POLLUTANT I BENZENE I DICHLORODIPHENYLTRICHLOROETHANE I INORGANIC POLLUTANT I METAL POLLUTANT I METHYL BROMIDE I ORGANIC POLLUTANT I PERSISTENT ORGANIC POLLUTANT I PESTICIDE I POLLUTANT I POLYCHLORINATED BIPHENYL I PRIMARY POLLUTANT I PRIORITY POLLUTANT I SECONDARY POLLUTANT I SOIL POLLUTANT I WATER POLLUTANT
No. of *gas* concepts	12
***Gas* concepts**	CARBON DIOXIDE I DIOXIN I GASEOUS POLLUTANT I GREENHOUSE GAS I METHANE I NITROGEN DIOXIDE I NITROGEN OXIDE I OZONE I PEROXYACYL NITRATE I PHOTOCHEMICAL OXIDANT I SULFUR DIOXIDE I TROPOSPHERIC OZONE
No. of *solid matter* concepts	12
***Solid matter* concepts**	ACIDIFYING POLLUTANT I ANTHROPOGENIC POLLUTANT I INORGANIC POLLUTANT I LEAD I METAL POLLUTANT I ORGANIC POLLUTANT I PERSISTENT ORGANIC POLLUTANT I PESTICIDE I POLLUTANT I PRIORITY POLLUTANT I SOIL POLLUTANT I WATER POLLUTANT

Hyponymy subtypes	
No. of hyponymy subtypes	8
Hyponymy subtypes	*ability-based* ǀ *composition-based* ǀ *denomination-based* ǀ *effect-based* ǀ *function-based* ǀ *location-based* ǀ *origin-based* ǀ *state-based*
No. of *ability-based* hyponyms	1
***Ability-based* hyponyms**	METAL POLLUTANT
No. of *composition-based* hyponyms	15
***Composition-based* hyponyms**	AROMATIC POLLUTANT ǀ BENZENE ǀ CARBON DIOXIDE ǀ DICHLORODIPHENYLTRICHLOROETHANE ǀ DIOXIN ǀ INORGANIC POLLUTANT ǀ METHANE ǀ METHYL BROMIDE ǀ NITROGEN DIOXIDE ǀ NITROGEN OXIDE ǀ ORGANIC POLLUTANT ǀ OZONE ǀ PEROXYACYL NITRATE ǀ POLYCHLORINATED BIPHENYL ǀ SULFUR DIOXIDE
No. of *denomination-based* hyponyms	2
***Denomination-based* hyponyms**	LEAD ǀ PRIORITY POLLUTANT
No. of *effect-based* hyponyms	4
***Effect-based* hyponyms**	ACIDIFYING POLLUTANT ǀ GREENHOUSE GAS ǀ PERSISTENT ORGANIC POLLUTANT ǀ PHOTOCHEMICAL OXIDANT
No. of *function-based* hyponyms	1
***Function-based* hyponyms**	PESTICIDE
No. of *location-based* hyponyms	4
***Location-based* hyponyms**	AIR POLLUTANT ǀ SOIL POLLUTANT ǀ TROPOSPHERIC OZONE ǀ WATER POLLUTANT
No. of *origin-based* hyponyms	4
***Origin-based* hyponyms**	AEROSOL ǀ ANTHROPOGENIC POLLUTANT ǀ PRIMARY POLLUTANT ǀ SECONDARY POLLUTANT
No. of *state-based* hyponyms	1
***State-based* hyponyms**	GASEOUS POLLUTANT

Hyponymy levels	
No. of hyponymy levels	5
No. of first-level hyponyms	9
First-level hyponyms	AIR POLLUTANT ǀ ANTHROPOGENIC POLLUTANT ǀ INORGANIC POLLUTANT ǀ METAL POLLUTANT ǀ ORGANIC POLLUTANT ǀ PESTICIDE ǀ PRIORITY POLLUTANT ǀ SOIL POLLUTANT ǀ WATER POLLUTANT
No. of second-level hyponyms	10
Second-level hyponyms	ACIDIFYING POLLUTANT ǀ AEROSOL ǀ AROMATIC POLLUTANT ǀ DICHLORODIPHENYLTRICHLOROETHANE ǀ GASEOUS POLLUTANT ǀ LEAD ǀ METHYL BROMIDE ǀ PERSISTENT ORGANIC POLLUTANT ǀ PRIMARY POLLUTANT ǀ SECONDARY POLLUTANT
No. of third-level hyponyms	8
Third-level hyponyms	BENZENE ǀ DIOXIN ǀ GREENHOUSE GAS ǀ NITROGEN DIOXIDE ǀ NITROGEN OXIDE ǀ PHOTOCHEMICAL OXIDANT ǀ POLYCHLORINATED BIPHENYL ǀ SULFUR DIOXIDE
No. of fourth-level hyponyms	4
Fourth-level hyponyms	CARBON DIOXIDE ǀ METHANE ǀ OZONE ǀ PEROXYACYL NITRATE
No. of fifth-level hyponyms	1
Fifth-level hyponyms	TROPOSPHERIC OZONE

Hyponymic contexts	
No. of concepts with hyponymic contexts	24
Concepts with hyponymic contexts	AIR POLLUTANT | ANTHROPOGENIC POLLUTANT | BENZENE | CARBON DIOXIDE | DICHLORODIPHENYLTRICHLOROETHANE | DIOXIN | GREENHOUSE GAS | LEAD | METHANE | METHYL BROMIDE | NITROGEN DIOXIDE | NITROGEN OXIDE | ORGANIC POLLUTANT | OZONE | PEROXYACYL NITRATE | PERSISTENT ORGANIC POLLUTANT | PESTICIDE | POLLUTANT | POLYCHLORINATED BIPHENYL | PRIMARY POLLUTANT | PRIORITY POLLUTANT | SOIL POLLUTANT | SULFUR DIOXIDE | TROPOSPHERIC OZONE

Table 49. Summary of the information in the terminological entry of POLLUTANT

5.1.10. EARTHQUAKE

The hyponymy-based terminological entry of EARTHQUAKE is shown in **Table 50**.

earthquake P-11.1: Earth/soil movement

Geologic phenomenon involving a sudden movement of the Earth's crust or upper mantle caused by the release of stress accumulated along geologic faults or by volcanic activity and resulting in the generation of seismic waves which can be destructive.

- *One example of a strong EARTHQUAKE is the 1755 Lisbon earthquake.*

DEGREE-BASED HYPONYMY

aftershock P-11.1: Earth/soil movement

Earthquake of lesser magnitude that follows a more forceful one in the same area.

- *The procedure for removing dependent earthquakes, such as foreshocks and AFTERSHOCKS.*

EFFECT-BASED HYPONYMY

tsunamigenic earthquake P-11.1: Earth/soil movement

Earthquake that produces a large size tsunami relative to the value of its surface wave magnitude.

FUNCTION-BASED HYPONYMY

scenario earthquake P-11.1: Earth/soil movement | E-6.5.5: Model

Earthquake referring to a specific magnitude and source.

FUNCTION-BASED HYPONYMY

characteristic earthquake P-11.1: Earth/soil movement | E-6.5.5: Model

Scenario earthquake referring to the maximum considered values of an earthquake of a specific fault.

- *A CHARACTERISTIC EARTHQUAKE is the largest earthquake that ruptures the same fault segment with similar magnitude during repeated seismic cycles.*

HEIGHT-BASED HYPONYMY

deep earthquake P-11.1: Earth/soil movement

deep-focus earthquake
Earthquake with a hypocenter depth of over 300 kilometers from the Earth's surface.

shallow earthquake P-11.1: Earth/soil movement
shallow-focus earthquake
Earthquake whose hypocenter is located within 70 kilometers of the Earth's surface.

LOCATION-BASED HYPONYMY

local earthquake P-11.1: Earth/soil movement
Earthquake of small magnitude that is local to only a geographical region, such as a volcano group.

- *[...] data from different types of seismic sources including natural sources such as teleseismic events and LOCAL EARTHQUAKES [...].*

ORIGIN-BASED HYPONYMY

tectonic earthquake P-11.1: Earth/soil movement
Earthquake originating when the Earth's crust breaks due to geological forces on rocks and adjoining plates that cause physical and chemical changes.

LOCATION-BASED HYPONYMY

intraplate earthquake P-11.1: Earth/soil movement
Tectonic earthquake occurring within the interior of a tectonic plate.

subduction earthquake P-11.1: Earth/soil movement
subduction zone earthquake
Tectonic earthquake occurring in the subduction zones (i.e., the regions where portions of the Earth's tectonic plates are diving beneath other plates and into the Earth's interior).

ORIGIN-BASED HYPONYMY

crustal earthquake P-11.1: Earth/soil movement
Tectonic earthquake originating from a fault movement in the crust (i.e., a relatively thin layer on the Earth's surface).

HEIGHT-BASED HYPONYMY

shallow crustal earthquake P-11.1: Earth/soil movement
Crustal earthquake whose hypocenter is located close to the Earth's surface.

LOCATION-BASED HYPONYMY

inland crustal earthquake P-11.1: Earth/soil movement
Crustal earthquake occurring in the interior part of a region, away from the shoreline.

volcanic earthquake P-11.1: Earth/soil movement
Earthquake originating from the direct action of volcanic activity in conjunction with tectonic forces.

Table 50. Full hyponymy-based terminological entry of EARTHQUAKE

EARTHQUAKE, the first GEO hyponymy-based terminological entry, has the following information: 15 concepts with their definitions; 18 terms designating those concepts; two conceptual categories (i.e., *earth/soil movement* and *model*); six hyponymy subtypes (i.e., *degree-based*, *effect-based*, *function-based*, *height-based*, *location-based* and *origin-based* hyponymy); up to three hyponymy levels; and four concepts with hyponymic contexts. In comparison with the other two GEO entries, this is the one with the least content. However, it features some unique characteristics, such as exclusive conceptual categories (i.e., *earth/soil movement* and *model*) and an exclusive hyponymy subtype (i.e., *degree-based* hyponymy), probably because this is the only terminological entry related to a process concept.

In this entry, there is one main conceptual category: *earth/soil movement* (e.g., AFTERSHOCK, INTRAPLATE EARTHQUAKE, TECTONIC EARTHQUAKE). In fact, EARTHQUAKE is the only general hypernym of all twelve terminological entries which refers to a natural process. This is the reason why all of its concepts are classified as *earth/soil movement*. However, there are two concepts (i.e., CHARACTERISTIC EARTHQUAKE and SCENARIO EARTHQUAKE) which not only refer to an *earth/soil movement*, but also to a mathematical model that expresses a specific magnitude or a source. They thus have double category membership and are either a process or entity, depending on the dimension focused.

The most important hyponymy subtypes in this entry are *location-based*, *height-based*, and *origin-based* hyponymy. The *location-based* hyponyms (e.g., INLAND CRUSTAL EARTHQUAKE, INTRAPLATE EARTHQUAKE, SUBDUCTION EARTHQUAKE) are distinguished by the location within the Earth's layers or plates where they take place. The *height-based* hyponyms (e.g., DEEP EARTHQUAKE, SHALLOW CRUSTAL EARTHQUAKE, SHALLOW EARTHQUAKE) are characterized by the height with respect to sea level of their occurrence. Finally, the *origin-based* hyponyms (e.g., CRUSTAL EARTHQUAKE, TECTONIC EARTHQUAKE, VOLCANIC EARTHQUAKE) are determined by the force or source causing these types of earthquakes.

Regarding hyponymy levels, the most specific hyponyms are two co-hyponyms at the third level, namely, INLAND CRUSTAL EARTHQUAKE and SHALLOW CRUSTAL EARTHQUAKE. However, they are related to their direct hypernym with a different hyponymy subtype. They have the following sequence: SHALLOW CRUSTAL EARTHQUAKE is a *height-based* type of CRUSTAL EARTHQUAKE whilst INLAND CRUSTAL EARTHQUAKE is a *location-based* type of CRUSTAL EARTHQUAKE, which is then an *origin-based* type of TECTONIC EARTHQUAKE, which is an *origin-based* type of EARTHQUAKE.

Therefore, in these sequences, nuances of earthquake origin are used for the first branching, after which the conceptual hierarchy is further specified by other attributes such as height and location.

Regarding hyponymic contexts, the entry also identifies hyponymic KPs in the examples provided (e.g., "dependent earthquakes, such as foreshocks and aftershocks"; "a characteristic earthquake is the largest earthquake"; "natural sources such as teleseismic events and local earthquakes"). Finally, **Table 51** summarizes the information in the hyponymy-based terminological entry of EARTHQUAKE.

EARTHQUAKE entry

Concepts, terms, and terminological definitions	
No. of concepts	15
No. of terms	18
No. of terminological definitions	15

Conceptual categories	
No. of conceptual categories	2
Conceptual categories	*earth/soil movement* ǀ *model*
No. of *earth/soil movement* concepts	15
***Earth/soil movement* concepts**	AFTERSHOCK ǀ CHARACTERISTIC EARTHQUAKE ǀ CRUSTAL EARTHQUAKE ǀ DEEP EARTHQUAKE ǀ EARTHQUAKE ǀ INLAND CRUSTAL EARTHQUAKE ǀ INTRAPLATE EARTHQUAKE ǀ LOCAL EARTHQUAKE ǀ SCENARIO EARTHQUAKE ǀ SHALLOW CRUSTAL EARTHQUAKE ǀ SHALLOW EARTHQUAKE ǀ SUBDUCTION EARTHQUAKE ǀ TECTONIC EARTHQUAKE ǀ TSUNAMIGENIC EARTHQUAKE ǀ VOLCANIC EARTHQUAKE
No. of *model* concepts	2
***Model* concepts**	CHARACTERISTIC EARTHQUAKE ǀ SCENARIO EARTHQUAKE

Hyponymy subtypes	
No. of hyponymy subtypes	6
Hyponymy subtypes	*degree-based* ǀ *effect-based* ǀ *function-based* ǀ *height-based* ǀ *location-based* ǀ *origin-based*
No. of *degree-based* hyponyms	1
***Degree-based* hyponyms**	AFTERSHOCK
No. of *effect-based* hyponyms	1
***Effect-based* hyponyms**	TSUNAMIGENIC EARTHQUAKE
No. of *function-based* hyponyms	2
***Function-based* hyponyms**	CHARACTERISTIC EARTHQUAKE ǀ SCENARIO EARTHQUAKE
No. of *height-based* hyponyms	3
***Height-based* hyponyms**	DEEP EARTHQUAKE ǀ SHALLOW CRUSTAL EARTHQUAKE ǀ SHALLOW EARTHQUAKE
No. of *location-based* hyponyms	4

Location-based **hyponyms**	INLAND CRUSTAL EARTHQUAKE ǀ INTRAPLATE EARTHQUAKE ǀ LOCAL EARTHQUAKE ǀ SUBDUCTION EARTHQUAKE
No. of *origin-based* **hyponyms**	3
Origin-based **hyponyms**	CRUSTAL EARTHQUAKE ǀ TECTONIC EARTHQUAKE ǀ VOLCANIC EARTHQUAKE

Hyponymy levels	
No. of hyponymy levels	3
No. of first-level hyponyms	8
First-level hyponyms	AFTERSHOCK ǀ DEEP EARTHQUAKE ǀ LOCAL EARTHQUAKE ǀ SCENARIO EARTHQUAKE ǀ SHALLOW EARTHQUAKE ǀ TECTONIC EARTHQUAKE ǀ TSUNAMIGENIC EARTHQUAKE ǀ VOLCANIC EARTHQUAKE
No. of second-level hyponyms	4
Second-level hyponyms	CHARACTERISTIC EARTHQUAKE ǀ CRUSTAL EARTHQUAKE ǀ INTRAPLATE EARTHQUAKE ǀ SUBDUCTION EARTHQUAKE
No. of third-level hyponyms	2
Third-level hyponyms	INLAND CRUSTAL EARTHQUAKE ǀ SHALLOW CRUSTAL EARTHQUAKE

Hyponymic contexts	
No. of concepts with hyponymic contexts	4
Concepts with hyponymic contexts	AFTERSHOCK ǀ CHARACTERISTIC EARTHQUAKE ǀ EARTHQUAKE ǀ LOCAL EARTHQUAKE

Table 51. Summary of the information in the terminological entry of EARTHQUAKE

5.1.11. SEDIMENT

The hyponymy-based terminological entry of SEDIMENT is shown in **Table 52**.

sediment E-8.4.1: Deposit

Solid unconsolidated rock and mineral fragment that comes from the weathering of rocks and is transported by water, air, or ice and forms layers on the Earth's surface.

- *This 'classical' definition […] would exclude most speciation studies on solid materials, such as soils and SEDIMENTS […].*

COMPOSITION-BASED HYPONYMY

aeolian deposit E-8.4.1: Deposit

Sediment composed of windblown grains of sand or dust.

COMPOSITION-BASED HYPONYMY

loess E-8.4.1: Deposit | E-4.2.1: Landform

Aeolian deposit composed largely of silt-sized quartz particles and showing little or no stratification.

chemical sediment E-8.4.1: Deposit
Sediment composed of previously dissolved minerals that have either precipitated from evaporated water or been extracted from water by living organisms and deposited when the organisms died or discarded their shells.

clastic sediment E-8.4.1: Deposit
Sediment made of clastic materials transported by mechanical agents.

cohesive sediment E-8.4.1: Deposit
Sediment with a significant proportion of clays, whose electromagnetic properties cause the sediment to bind together.

gravel E-8.4.1: Deposit
Sediment consisting in aggregate or unconsolidated sediments composed of rock fragments coarser than sand, with a diameter of more than 2 mm.
- *Some unusual types of GRAVEL and conglomerate include tillites [...] and diamictite [...].*

intraclast E-8.4.1: Deposit | E-8.3: Particle
Sediment composed of calcium carbonate from local penecontemporary erosion of the soil in the sedimentary basin.

noncohesive sediment E-8.4.1: Deposit
non-cohesive sediment
cohesionless sediment
Sediment composed of sand and gravel-sized material that is not bounded in nature.

pelite E-8.4.1: Deposit | E-8.4.2.2: Rock
pelyte
Sediment and sedimentary rock composed of fine fragments, such as clay or mud.
- *Other shallow shelf metasedimentary rocks, including quartzites, limestones, and PELITES, were metamorphosed to granulite facies at 2.7 billion years ago.*

sand E-8.4.1: Deposit | E-8.4.2.1: Mineral | E-8.3: Particle
Sediment consisting of rock or mineral grains that is finer than a granule and coarser than silt, with grains between 0.05 and 2 mm in diameter.
- *[...] braided stream deposits show less order and are characteristically dominated by bed load material such as gravel and SAND.*

FUNCTION-BASED HYPONYMY

dry-screen sand E-8.4.1: Deposit | E-8.4.2.1: Mineral
Sand used in construction that has been sifted to eliminate over-size particles.

ORIGIN-BASED HYPONYMY

original sand E-8.4.1: Deposit | E-8.4.2.1: Mineral
native sand
Sand found in a beach before it is replenished by sand from another source.

SIZE-BASED HYPONYMY

coarse sand E-8.4.1: Deposit | E-8.4.2.1: Mineral | E-8.3: Particle
Sand with grains between 0.5 and 1 mm in diameter.

fine sand E-8.4.1: Deposit | E-8.4.2.1: Mineral | E-8.3: Particle
Sand with grains between 0.125 and 0.25 mm in diameter.

soft sediment E-8.4.1: Deposit
muddy sediment
Sediment composed of a mixture of soil with sufficient water to make it soft.

- *This is generally associated with SOFT SEDIMENT, such as silt and clay deposits.*

EFFECT-BASED HYPONYMY

moraine E-8.4.1: Deposit | E-4.2.1: Landform
Sediment composed of unsorted fragmental material that is transported or deposited by glaciers, and results in a landform like a mound, ridge or other prominence on the terrain.

- *These include famous landforms such as the Matterhorn and other horns, aretes, U-shaped valleys, erratics, and MORAINES.*

LOCATION-BASED HYPONYMY

ground moraine E-8.4.1: Deposit | E-4.2.1: Landform
bottom moraine
subglacial moraine
Moraine formed at the base of a glacier.

lateral moraine E-8.4.1: Deposit | E-4.2.1: Landform
Moraine formed along the sides of a glacier, commonly found on glaciers occupying a valley.

terminal moraine E-8.4.1: Deposit | E-4.2.1: Landform
end moraine
marginal moraine
Moraine formed at the end of a glacier, marking the maximum advance of the glacier.

ORIGIN-BASED HYPONYMY

recessional moraine E-8.4.1: Deposit | E-4.2.1: Landform
stadial moraine
Terminal moraine formed during a temporary halt in the final retreat of a glacier or during a minor readvance of the ice front during a period of glacial recession.

central moraine E-8.4.1: Deposit | E-4.2.1: Landform
Moraine formed by the union of the lateral moraines of two glaciers to form a larger one.

LOCATION-BASED HYPONYMY

continental sediment E-8.4.1: Deposit
Sediment deposited in a non-marine environment.

LOCATION-BASED HYPONYMY

alluvial sediment E-8.4.1: Deposit
alluvium
Continental sediment carried by rushing streams and deposited in a river bed, flood plain, or delta.

- *[...] geologic conditions can range from hard rock such as granite to soft, unconsolidated geologic formation such as ALLUVIAL SEDIMENTS.*

LOCATION-BASED HYPONYMY

deposited sediment E-8.4.1: Deposit
bed sediment
Alluvial sediment made of inorganic particles below 2 mm in diameter that has been deposited on the bed of a river or stream.

fluvial sediment E-8.4.1: Deposit
Alluvial sediment that has been deposited by the water of a river.

ORIGIN-BASED HYPONYMY

stream sediment E-8.4.1: Deposit
Alluvial sediment derived from the erosion and transport of soil, rock debris and other materials within the catchment basin upstream of the sampling site.

- *The chemical composition of secondary minerals in STREAM SEDIMENTS, soils, and other near-surface materials is dominantly contributed by local bedrock [...].*

lake sediment E-8.4.1: Deposit
lacustrine sediment
Continental deposited in lakes or lacustrine environments.

- *Methods of studying global change include examination of the geologic record from LAKE SEDIMENTS, glacial ice, and other Earth materials.*

marine sediment E-8.4.1: Deposit
ocean sediment
Sediment deposited in a marine environment, such as a sea or an ocean, and that is finer and smoother than continental sediment.

COMPOSITION-BASED HYPONYMY

carbonate sediment E-8.4.1: Deposit
Marine sediment made of skeletons of organisms like foraminifera, scleractinian corals and calcareous algae.

LOCATION-BASED HYPONYMY

coastal sediment E-8.4.1: Deposit
Marine sediment deposited in coasts or coastal environments.

- *Sulfur can be reduced to sulfide or to other trace sulfur gases in anaerobic environments such as wetlands and COASTAL SEDIMENTS.*

LOCATION-BASED HYPONYMY

beach sediment E-8.4.1: Deposit
Coastal sediment deposited in a beach.

deep-sea sediment E-8.4.1: Deposit
pelagic sediment
Marine sediment that deposits slowly in the abyssal ocean, beyond the continental margins.
- *Many ophiolites are overlain by DEEP-SEA SEDIMENTS, including chert, red clay, in some cases carbonates, or sulfide layers.*

shallow sediment E-8.4.1: Deposit
shallow marine sediment
Marine sediment deposited in shallow marine areas and composed of solid particle skeleton and pores which are full of water.

COMPOSITION-BASED HYPONYMY

terrigenous sediment E-8.4.1: Deposit
detrital sediment
Shallow sediment composed of eroded terrestrial material.

COMPOSITION-BASED HYPONYMY

diamictite E-8.4.1: Deposit | E-8.4.2.2: Rock
mixtite
Terrigenous sediment of calcareous nature that is not sorted or poorly sorted and contains particles of many sizes.
- *The base of the Morondova basin [...] has spectacular glacial deposits including DIAMICTITES, tillites, and glacial outwash gravels.*

MOVEMENT-BASED HYPONYMY

sheet flow sediment E-8.4.1: Deposit
Sediment under high sheer stress moving as a layer that extends from the bed surface to some distance below, being transported in the direction of fluid flow.

ORIGIN-BASED HYPONYMY

biogenic sediment E-8.4.1: Deposit
Sediment resulting from the physiological activities of organisms.
- *The benthic environment also hosts BIOGENIC SEDIMENTS, including calcareous and siliceous oozes that form by the gradual rain of shells of deceased planktonic organisms.*

bottom sediment E-8.4.1: Deposit
basic sediment and water
BS&W
bottom settling
Sediment consisting in oil, water and foreign matter collected in the bottom of petroleum storage tanks.

clay E-8.4.1: Deposit | E-8.4.2.2: Rock
Sediment and detrital sedimentary rock of fine grain, formed from marine and lacustrine sediments of particles of 0.002 mm in diameter composed, among others, of clay minerals and quartz.

siliciclastic E-8.4.1: Deposit
siliclastic
Non-carbon silica-based sediment which is fragmented from pre-existing rocks, transported to another location, and redeposited before forming another rock.

silt E-8.4.1: Deposit | E-8.4.2.2: Rock
Sediment and sedimentary rock composed mainly of calcium carbonate, formed by the break up and deposition of shells, coral and other marine organisms by wave-action and ocean currents.

till E-8.4.1: Deposit
Unsorted heterogeneous sediment deposited directly by glacial ice and showing no stratification.

SIZE-BASED HYPONYMY

coarse sediment E-8.4.1: Deposit
Sediment with grains above 2 mm in diameter.
- *In mountainous regions where there is a lot of sand and gravel, a river must thread its way around bars of sand, gravel, and other COARSE SEDIMENT.*

fine sediment E-8.4.1: Deposit | E-8.3: Particle
fine-grained sediment
Sediment with grains below 2 mm in diameter.
- *FINE SEDIMENT (such as silt) is a type of water pollution because it reduces soil resources […].*

STATUS-BASED HYPONYMY

consolidated sediment E-8.4.1: Deposit | E-8.4.2.2: Rock
Sediment that has been converted into rock by compaction, deposition of cement in pore spaces, or by physical and chemical changes in the constituents.
- *All results determined by using this technique for various types of CONSOLIDATED SEDIMENT (including mudstone, till, and lacustrine clay) […].*

suspended sediment E-8.4.1: Deposit
Sediment transported by a fluid that it is fine enough for turbulent eddies to outweigh settling of the particles through the fluid.

unconsolidated sediment E-8.4.1: Deposit
Sediment that is not compacted and lithified, formed by secondary sedimentation of previously weathered rocks and redeposition of their fragments, or by chemical and biochemical precipitation from solutions.
- *The normal mode in a surface layer with very low shear velocity, such as UNCONSOLIDATED SEDIMENTS, that may be affected significantly by gravity at long periods such as tsunami in the ocean.*

Table 52. Full hyponymy-based terminological entry of SEDIMENT

SEDIMENT, the second GEO hyponymy-based terminological entry has the following information: 48 concepts with their definitions; 71 terms designating those concepts; five conceptual categories (i.e., *deposit*, *landform*, *mineral*, *particle*, and *rock*); eight hyponymy subtypes (i.e., *composition-based*, *effect-based*, *function-based*, *location-based*, *movement-based*, *origin-based*, *size-based* and *status-based* hyponym); up to four hyponymy levels; and 17 concepts with hyponymic contexts. This is the richest entry not only of the GEO entries, but of all entries in relation to the number of concepts and terms. Therefore, this generates a wider variety of conceptual categories and hyponymy subtypes.

Of the conceptual categories in this entry, the predominant one is *deposit*, which is present in all concepts (e.g., COHESIVE SEDIMENT, FINE SEDIMENT, TERRIGENOUS SEDIMENT). However, as concepts are further specified in the conceptual hierarchy through hyponymy subtypes, this also generates more specific categories, such as *landform* (e.g., CENTRAL MORAINE, LATERAL MORAINE, LOESS), *mineral* (e.g., DRY-SCREEN SAND, FINE SAND, ORIGINAL SAND), *particle* (e.g., COARSE SAND, FINE SEDIMENT, INTRACLAST), and *rock* (CLAY, DIAMICTITE, SILT). Therefore, the *deposit* category, which is more general, also acquires more nuances as the conceptual hierarchy progresses.

On the other hand, the most relevant hyponymy subtypes in this entry are *composition-based*, *location-based*, and *origin-based* hyponymy. The *composition-based* hyponyms (e.g., CARBONATE SEDIMENT, INTRACLAST, TERRIGENOUS SEDIMENT) are differentiated from their hypernyms because of the materials they are made of. The *location-based* hyponyms (e.g., ALLUVIAL SEDIMENT, GROUND MORAINE, SHALLOW SEDIMENT) are determined by the place where the sediment is deposited or where it is typically formed. Finally, the *origin-based* hyponyms (e.g., BIOGENIC SEDIMENT, ORIGINAL SAND, RECESSIONAL MORAINE) are characterized by the process that creates or originates them.

Interestingly, the only concept at the fourth and last hyponymy level of this entry is DIAMICTITE, which has the following schema: DIAMICTITE is a *composition-based* type of TERRIGENOUS SEDIMENT, which is a *composition-based* type of SHALLOW SEDIMENT, which is a *location-based* type of MARINE SEDIMENT, which is a *location-based* type of SEDIMENT. In this sequence, *location-based* hyponymy is at the first and second levels, whereas *composition-based* hyponymy is at the third and fourth levels. However, *composition-based* hyponymy is at the highest hyponymy levels of other sequences (e.g., those of *chemical sediment*, *cohesive sediment*, and *intraclast*), and *location-based* is at the lower hyponymy levels of other sequences (e.g., those of DEPOSITED SEDIMENT, FLUVIAL SEDIMENT, and BEACH SEDIMENT). It is thus difficult to establish a common pattern regarding which levels are characteristic of certain subtypes.

In relation to hyponymic contexts, there are many examples in this entry that show different ways of expressing hyponymic KPs (e.g., "unusual types of gravel and conglomerate include tillites [...] and diamictite"; "bed load material such as gravel and sand"; "stream sediments, soils, and other near-surface materials"). Finally, **Table 53** summarizes the information in the hyponymy-based terminological entry of SEDIMENT.

SEDIMENT entry

Concepts, terms, and terminological definitions	
No. of concepts	48
No. of terms	71
No. of terminological definitions	48

Conceptual categories	
No. of conceptual categories	5
Conceptual categories	*deposit* ǀ *landform* ǀ *mineral* ǀ *particle* ǀ *rock*
No. of *deposit* concepts	48
***Deposit* concepts**	AEOLIAN DEPOSIT ǀ ALLUVIAL SEDIMENT ǀ BEACH SEDIMENT ǀ BIOGENIC SEDIMENT ǀ BOTTOM SEDIMENT ǀ CARBONATE SEDIMENT ǀ CENTRAL MORAINE ǀ CHEMICAL SEDIMENT ǀ CLASTIC SEDIMENT ǀ CLAY ǀ COARSE SAND ǀ COARSE SEDIMENT ǀ COASTAL SEDIMENT ǀ COHESIVE SEDIMENT ǀ CONSOLIDATED SEDIMENT ǀ CONTINENTAL SEDIMENT ǀ DEEP-SEA SEDIMENT ǀ DEPOSITED SEDIMENT ǀ DIAMICTITE ǀ DRY-SCREEN SAND ǀ FINE SAND ǀ FINE SEDIMENT ǀ FLUVIAL SEDIMENT ǀ GRAVEL ǀ GROUND MORAINE ǀ INTRACLAST ǀ LAKE SEDIMENT ǀ LATERAL MORAINE ǀ LOESS ǀ MARINE SEDIMENT ǀ MORAINE ǀ NONCOHESIVE SEDIMENT ǀ ORIGINAL SAND ǀ PELITE ǀ RECESSIONAL MORAINE ǀ SAND ǀ SEDIMENT ǀ SHALLOW SEDIMENT ǀ SHEET FLOW SEDIMENT ǀ SILICICLASTIC ǀ SILT ǀ SOFT SEDIMENT ǀ STREAM SEDIMENT ǀ SUSPENDED SEDIMENT ǀ TERMINAL MORAINE ǀ TERRIGENOUS SEDIMENT ǀ TILL ǀ UNCONSOLIDATED SEDIMENT
No. of *landform* concepts	7
***Landform* concepts**	CENTRAL MORAINE ǀ GROUND MORAINE ǀ LATERAL MORAINE ǀ LOESS ǀ MORAINE ǀ RECESSIONAL MORAINE ǀ TERMINAL MORAINE
No. of *mineral* concepts	5
***Mineral* concepts**	COARSE SAND ǀ DRY-SCREEN SAND ǀ FINE SAND ǀ ORIGINAL SAND ǀ SAND
No. of *particle* concepts	5
***Particle* concepts**	COARSE SAND ǀ FINE SAND ǀ FINE SEDIMENT ǀ INTRACLAST ǀ SAND
No. of *rock* concepts	5

***Rock* concepts**	CLAY | CONSOLIDATED SEDIMENT | DIAMICTITE | PELITE | SILT
Hyponymy subtypes	
No. of hyponymy subtypes	8
Hyponymy subtypes	*composition-based* | *effect-based* | *function-based* | *location-based* | *movement-based* | *origin-based* | *size-based* | *status-based*
No. of *composition-based* hyponyms	14
***Composition-based* hyponyms**	AEOLIAN DEPOSIT | CARBONATE SEDIMENT | CHEMICAL SEDIMENT | CLASTIC SEDIMENT | COHESIVE SEDIMENT | DIAMICTITE | GRAVEL | INTRACLAST | LOESS | NONCOHESIVE SEDIMENT | PELITE | SAND | SOFT SEDIMENT | TERRIGENOUS SEDIMENT
No. of *effect-based* hyponyms	1
***Effect-based* hyponyms**	MORAINE
No. of *function-based* hyponyms	1
***Function-based* hyponyms**	DRY-SCREEN SAND
No. of *location-based* hyponyms	14
***Location-based* hyponyms**	ALLUVIAL SEDIMENT | BEACH SEDIMENT | CENTRAL MORAINE | COASTAL SEDIMENT | CONTINENTAL SEDIMENT | DEEP-SEA SEDIMENT | DEPOSITED SEDIMENT | FLUVIAL SEDIMENT | GROUND MORAINE | LAKE SEDIMENT | LATERAL MORAINE | MARINE SEDIMENT | SHALLOW SEDIMENT | TERMINAL MORAINE
No. of *movement-based* hyponyms	1
***Movement-based* hyponyms**	SHEET FLOW SEDIMENT
No. of *origin-based* hyponyms	9
***Origin-based* hyponyms**	BIOGENIC SEDIMENT | BOTTOM SEDIMENT | CLAY | ORIGINAL SAND | RECESSIONAL MORAINE | SILICICLASTIC | SILT | STREAM SEDIMENT | TILL
No. of *size-based* hyponyms	4
***Size-based* hyponyms**	COARSE SAND | COARSE SEDIMENT | FINE SAND | FINE SEDIMENT
No. of *status-based* hyponyms	3
***Status-based* hyponyms**	CONSOLIDATED SEDIMENT | SUSPENDED SEDIMENT | UNCONSOLIDATED SEDIMENT

Hyponymy levels	
No. of hyponymy levels	4
No. of first-level hyponyms	25
First-level hyponyms	AEOLIAN DEPOSIT | BIOGENIC SEDIMENT | BOTTOM SEDIMENT | CHEMICAL SEDIMENT | CLASTIC SEDIMENT | CLAY | COARSE SEDIMENT | COHESIVE SEDIMENT | CONSOLIDATED SEDIMENT | CONTINENTAL SEDIMENT | FINE SEDIMENT | GRAVEL | INTRACLAST | MARINE SEDIMENT | MORAINE | NONCOHESIVE SEDIMENT | PELITE | SAND | SHEET FLOW SEDIMENT | SILICICLASTIC | SILT | SOFT SEDIMENT |SUSPENDED SEDIMENT | TILL | UNCONSOLIDATED SEDIMENT
No. of second-level hyponyms	15
Second-level hyponyms	ALLUVIAL SEDIMENT | CARBONATE SEDIMENT | CENTRAL MORAINE | COARSE SAND | COASTAL SEDIMENT | DEEP-SEA SEDIMENT | DRY-SCREEN SAND | FINE SAND | GROUND MORAINE | LAKE SEDIMENT | LATERAL

	MORAINE | LOESS | ORIGINAL SAND | SHALLOW SEDIMENT | TERMINAL MORAINE
No. of third-level hyponyms	6
Third-level hyponyms	BEACH SEDIMENT | DEPOSITED SEDIMENT | FLUVIAL SEDIMENT | RECESSIONAL MORAINE | STREAM SEDIMENT | TERRIGENOUS SEDIMENT
No. of fourth-level hyponyms	1
Fourth-level hyponyms	DIAMICTITE

Hyponymic contexts	
No. of concepts with hyponymic contexts	17
Concepts with hyponymic contexts	ALLUVIAL SEDIMENT | BIOGENIC SEDIMENT | COARSE SEDIMENT | COASTAL SEDIMENT |CONSOLIDATED SEDIMENT | DEEP-SEA SEDIMENT | DIAMICTITE | FINE SEDIMENT | GRAVEL | LAKE SEDIMENT | MORAINE | PELITE | SAND | SEDIMENT | SOFT SEDIMENT | STREAM SEDIMENT | UNCONSOLIDATED SEDIMENT

Table 53. Summary of the information in the terminological entry of SEDIMENT

5.1.12. SOIL

The hyponymy-based terminological entry of SOIL is shown in **Table 54**.

soil E-8.4.2.3: Soil | E-12.2: Layer

Top layer of the Earth's surface, consisting in weathered, unconsolidated material, dead and living organic matter, air space, and the soil solution on top of bed rock, which is capable of supporting plant growth.

■ *SOIL, till, and other near-surface materials make up a relatively thin (<1.5-m thick) layer that overlies bedrock.*

ABILITY-BASED HYPONYMY

fertile soil E-8.4.2.3: Soil

rich soil

Soil capable of sustaining agricultural plant growth, providing plant habitat and resulting in sustained and consistent yields of high quality.

COMPOSITION-BASED HYPONYMY

loam E-8.4.2.3: Soil

Fertile soil that contains a roughly equal mixture of clay, sand and silt, and which is good for growing most crops.

FUNCTION-BASED HYPONYMY

agricultural soil E-8.4.2.3: Soil

Fertile soil which is used to produce high quality plants for human use, providing the populations with food.

LOCATION-BASED HYPONYMY

forest soil E-8.4.2.3: Soil
Fertile soil influenced by forest vegetation and generally characterized by deeply rooted trees, recycling of organic matter and nutrients, including wood, and wide varieties of soil-dwelling organisms.

COMPOSITION-BASED HYPONYMY

calcareous soil E-8.4.2.3: Soil
Soil that has calcium carbonate (CaCO3) in abundance.

cohesive soil E-8.4.2.3: Soil
Soil composed of particles that adhere to each other by means of adhesive and cohesive forces.

COMPOSITION-BASED HYPONYMY

clay soil E-8.4.2.3: Soil
clayey soil
Cohesive soil mostly composed of clay, which forms hard lumps when dry and becomes sticky when wet.

- *Centrifugation has been widely applied to the extraction of pore waters from various materials including sediments, chalks, sandstones and CLAYEY SOILS.*

STATUS-BASED HYPONYMY

soft soil E-8.4.2.3: Soil
Clay soil that is geologically young and has not undergone significant secondary or delayed consolidation since its formation.

granular soil E-8.4.2.3: Soil
Soil composed of crumb with open structures that allow water and air to penetrate through the earth.

EFFECT-BASED HYPONYMY

contractive soil E-8.4.2.3: Soil
Granular soil that decreases or tends to decrease in volume during large shear deformation.

mineral soil E-8.4.2.3: Soil
Soil composed of mineral or rock derivatives, with little organic matter.

organic soil E-8.4.2.3: Soil
Soil consisting chiefly of, or containing at least, 30% of organic matter.

saline soil E-8.4.2.3: Soil
Soil containing enough soluble salts to interfere with the ability of plants to take up water.

sandy soil E-8.4.2.3: Soil
Soil with a high percentage of sand or large soil particles, mainly consisting of rock particles such as limestone, shale, granite and quartz.

DENOMINATION-BASED HYPONYMY

cambisol E-8.4.2.3: Soil
Soil pertaining to one of the orders of the World Reference Base for Soil Resources that is characterized by a layer in the subsoil that looks more like soil than rock, and also accumulation of materials such as clay, calcium carbonate, iron and manganese.

leptosol E-8.4.2.3: Soil
Soil pertaining to one of the orders of the World Reference Base for Soil Resources that is characterized by a very shallow layer over hard rock or a deeper layer that is extremely gravelly or stony.

lithosol E-8.4.2.3: Soil
Soil pertaining to one of the orders of the FAO85 classification that results from erosive phenomena and is formed on hard bedrock, consisting chiefly of unweathered or partly weathered rock fragments.

mollisol E-8.4.2.3: Soil
Soil pertaining to one of the orders of the U.S. Soil Taxonomy that is characterized by a significant accumulation of humus in the surface horizon and that is almost always formed under native grass vegetation.

- *Prairie soils, mainly MOLLISOLS, are rich and black with a thick, fertile organic layer composed of fine roots, microorganisms, and soil.*

regosol E-8.4.2.3: Soil
Soil pertaining to one of the orders of the World Reference Base for Soil Resources that consists of unconsolidated material derived from freshly deposited alluvium or sands.

EFFECT-BASED HYPONYMY

expansive soil E-8.4.2.3: Soil
Soil that, upon wetting and drying, alternately expands and contracts, causing problems for foundations of buildings and other structures.

FUNCTION-BASED HYPONYMY

backfill soil E-8.4.2.3: Soil
Soil used to refill an excavated hole or trench, composed of a mixture of soil, sand, gravel, and even commercial products.

HEIGHT-BASED HYPONYMY

shallow soil E-8.4.2.3: Soil | E-12.2: Layer
Soil that is not deep (less than 50 cm depth of solum) and has little room for water storage.

surface soil E-8.4.2.3: Soil | E-12.2: Layer
Soil that extends 13 to 20 cm below the surface.

LOCATION-BASED HYPONYMY

bulk soil E-8.4.2.3: Soil | E-12.2: Layer
Soil that is located outside the rhizosphere and that is not penetrated by plant roots.

subsoil E-8.4.2.3: Soil | E-12.2: Layer
Soil that is located between the top soil and bedrock, which is usually less fertile and of poorer texture.

tropical soil E-8.4.2.3: Soil
Soil found in tropical regions between the Tropic of Cancer and the Tropic of Capricorn, and characterized due to regional natural factors, such as the ocean, wildlife, and decomposing minerals found around the Equator.

urban soil E-8.4.2.3: Soil
Soil typically found in urban areas, which has been extensively influenced by human activities (i.e., construction, contamination).

MOISTURE-BASED HYPONYMY

metastable soil E-8.4.2.3: Soil
Soil susceptible of undergoing substantial volumetric change on variation of its moisture content.

saturated soil E-8.4.2.3: Soil
Soil whose pores are temporarily or permanently filled with water.

MOISTURE-BASED HYPONYMY

moist soil E-8.4.2.3: Soil
Saturated soil whose pores are partially filled with water.

wet soil E-8.4.2.3: Soil
Saturated soil whose pores are completely filled with water.

ORIGIN-BASED HYPONYMY

residual soil E-8.4.2.3: Soil
Soil formed in place by weathering of the underlying mineral materials and which remains after the soluble elements have been dissolved.

- *Sediments sorted into fine- and coarse-grained layers [...] are generally more susceptible to liquefaction than unsorted sediments, such as RESIDUAL SOILS and glacial tills.*

STATUS-BASED HYPONYMY

contaminated soil E-8.4.2.3: Soil
polluted soil
Soil with a concentration of pollutants at a high level, posing a potential health or ecological risk.

- *CONTAMINATED SOILS and other materials were to be buried in new waste landfills.*

Table 54. Full hyponymy-based terminological entry of SOIL

SOIL, the third and last GEO hyponymy-based terminological entry, has the following information: 34 concepts with their definitions; 37 terms that designate those concepts; two conceptual categories (i.e., *layer* and *soil*); ten hyponymy subtypes (i.e., *ability-based*, *composition-based*, *denomination-based*, *effect-based*, *function-based*, *height-based*, *location-based*, *moisture-based*, *origin-based*, and *status-based* hyponymy); up to three hyponymy levels; and five concepts with hyponymic contexts. This is the entry with the greatest variety of hyponymy subtypes, which reveals the conceptual complexity in the categorization of its hypernym-hyponym pairs within the hierarchy.

Regarding the conceptual categories in this terminological entry, all concepts are classified in one conceptual category, *soil* (e.g., AGRICULTURAL SOIL, LEPTOSOL, ORGANIC SOIL). However, certain hyponyms also possess an additional category, *layer* (e.g., BULK SOIL, SUBSOIL, SURFACE SOIL). This means that though the concepts all belong to the *soil* category, not all of them are *layers*, since this would mean that they referred to a specific part of the Earth's surface.

As previously mentioned, this is the entry with the greatest variety of hyponymy subtypes, the most relevant of which are *composition-based*, *denomination-based*, *location-based*, and *moisture-based* hyponymy. The *composition-based* hyponyms (e.g., CLAY SOIL, GRANULAR SOIL, SALINE SOIL) designate more independent concepts because of the materials that they are made of. The *denomination-based* hyponyms (e.g., CAMBISOL, LITHOSOL, REGOSOL) are characterized by elements with scientific names, and are also part of the official taxonomy of soil denominations. The *location-based* hyponyms (e.g., FOREST SOIL, TROPICAL SOIL, URBAN SOIL) are determined by their environment. Finally, the *moisture-based* hyponyms (e.g., METASTABLE SOIL, SATURATED SOIL, WET SOIL) are differentiated by the amount of water that they contain.

The hyponymy levels in this entry are limited since the only hyponym at the third level is SOFT SOIL, which has the following sequence: SOFT SOIL is a *status-based* type of CLAY SOIL, which is a *composition-based* type of COHESIVE SOIL, which is a *composition-based* type of SOIL. Given the limited number of hyponymy levels and the significant number of hyponymy subtypes, in this hierarchy, verticality is not as important as horizontality. SOIL hyponyms reveal different dimensions or microsenses of co-hyponyms. For example, there are ten hyponymy subtypes activated at the first hyponymy level, which means that SOIL acts as a direct hypernym for ten different hyponymic nuances: *ability* (e.g., FERTILE SOIL); *composition* (e.g., CALCAREOUS SOIL); *denomination* (e.g., CAMBISOL); *effect* (e.g., EXPANSIVE SOIL); *function* (e.g., BACKFILL SOIL); *height* (e.g., SHALLOW SOIL); *location* (e.g., BULK SOIL); *moisture* (e.g., METASTABLE SOIL); *origin* (e.g., RESIDUAL SOIL); and *status* (e.g., CONTAMINATED SOIL).

As for the hyponymic contexts in this entry, there are examples that identify certain hyponymic KPs (e.g., "soil, till, and other near-surface materials"; "various materials including sediments, chalks, sandstones and clayey soils"; "prairie soils, mainly mollisols"). Finally, **Table 55** summarizes the information in the hyponymy-based terminological entry of SOIL.

SOIL entry

Concepts, terms, and terminological definitions	
No. of concepts	34
No. of terms	37
No. of terminological definitions	34

Conceptual categories	
No. of conceptual categories	2
Conceptual categories	*layer* ǀ *soil*
No. of *layer* concepts	5
***Layer* concepts**	BULK SOIL ǀ SHALLOW SOIL ǀ SOIL ǀ SUBSOIL ǀ SURFACE SOIL
No. of *soil* concepts	34
***Soil* concepts**	AGRICULTURAL SOIL ǀ BACKFILL SOIL ǀ BULK SOIL ǀ CALCAREOUS SOIL ǀ CAMBISOL ǀ CLAY SOIL ǀ COHESIVE SOIL ǀ CONTAMINATED SOIL ǀ CONTRACTIVE SOIL ǀ EXPANSIVE SOIL ǀ FERTILE SOIL ǀ FOREST SOIL ǀ GRANULAR SOIL ǀ LEPTOSOL ǀ LITHOSOL ǀ LOAM ǀ METASTABLE SOIL ǀ MINERAL SOIL ǀ MOIST SOIL ǀ MOLLISOL ǀ ORGANIC SOIL ǀ REGOSOL ǀ RESIDUAL SOIL ǀ SALINE SOIL ǀ SANDY SOIL ǀ SATURATED SOIL ǀ SHALLOW SOIL ǀ SOFT SOIL ǀ SOIL ǀ SUBSOIL ǀ SURFACE SOIL ǀ TROPICAL SOIL ǀ URBAN SOIL ǀ WET SOIL

Hyponymy subtypes	
No. of hyponymy subtypes	10
Hyponymy subtypes	*ability-based* ǀ *composition-based* ǀ *denomination-based* ǀ *effect-based* ǀ *function-based* ǀ *height-based* ǀ *location-based* ǀ *moisture-based* ǀ *origin-based* ǀ *status-based*
No. of *ability-based* hyponyms	1
***Ability-based* hyponyms**	FERTILE SOIL
No. of *composition-based* hyponyms	9
***Composition-based* hyponyms**	CALCAREOUS SOIL ǀ CLAY SOIL ǀ COHESIVE SOIL ǀ GRANULAR SOIL ǀ LOAM ǀ MINERAL SOIL ǀ ORGANIC SOIL ǀ SALINE SOIL ǀ SANDY SOIL
No. of *denomination-based* hyponyms	5
***Denomination-based* hyponyms**	CAMBISOL ǀ LEPTOSOL ǀ LITHOSOL ǀ MOLLISOL ǀ REGOSOL

No. of *effect-based* hyponyms	2
***Effect-based* hyponyms**	CONTRACTIVE SOIL ǀ EXPANSIVE SOIL
No. of *function-based* hyponyms	2
***Function-based* hyponyms**	AGRICULTURAL SOIL ǀ BACKFILL SOIL
No. of *height-based* hyponyms	2
***Height-based* hyponyms**	SHALLOW SOIL ǀ SURFACE SOIL
No. of *location-based* hyponyms	5
***Location-based* hyponyms**	BULK SOIL ǀ FOREST SOIL ǀ SUBSOIL ǀ TROPICAL SOIL ǀ URBAN SOIL
No. of *moisture-based* hyponyms	4
***Moisture-based* hyponyms**	METASTABLE SOIL ǀ MOIST SOIL ǀ SATURATED SOIL ǀ WET SOIL
No. of *origin-based* hyponyms	1
***Origin-based* hyponyms**	RESIDUAL SOIL
No. of *status-based* hyponyms	2
***Status-based* hyponyms**	CONTAMINATED SOIL ǀ SOFT SOIL

Hyponymy levels	
No. of hyponymy levels	3
No. of first-level hyponyms	25
First-level hyponyms	BACKFILL SOIL ǀ BULK SOIL ǀ CALCAREOUS SOIL ǀ CAMBISOL ǀ COHESIVE SOIL ǀ CONTAMINATED SOIL ǀ EXPANSIVE SOIL ǀ FERTILE SOIL ǀ GRANULAR SOIL ǀ LEPTOSOL ǀ LITHOSOL ǀ METASTABLE SOIL ǀ MINERAL SOIL ǀ MOLLISOL ǀ ORGANIC SOIL ǀ REGOSOL ǀ RESIDUAL SOIL ǀ SALINE SOIL ǀ SANDY SOIL ǀ SATURATED SOIL ǀ SHALLOW SOIL ǀ SUBSOIL ǀ SURFACE SOIL ǀ TROPICAL SOIL ǀ URBAN SOIL
No. of second-level hyponyms	7
Second-level hyponyms	AGRICULTURAL SOIL ǀ CLAY SOIL ǀ CONTRACTIVE SOIL ǀ FOREST SOIL ǀ LOAM ǀ MOIST SOIL ǀ WET SOIL
No. of third-level hyponyms	1
Third-level hyponyms	SOFT SOIL

Hyponymic contexts	
No. of concepts with hyponymic contexts	5
Concepts with hyponymic contexts	CLAY SOIL ǀ CONTAMINATED SOIL ǀ MOLLISOL ǀ RESIDUAL SOIL ǀ SOIL

Table 55. Summary of the information in the terminological entry of SOIL

5.2. HYPOLEXICON: A HYPONYMY-BASED TERMINOLOGICAL RESOURCE

HypoLexicon[9] is a terminological resource focused on the description, categorization, and representation of hyponymy in environmental concepts. It is designed as a stand-alone module for EcoLexicon, since it is also one of its by-products. It includes definitional, relational, ontological, and contextual information about specialized hypernyms and hyponyms of environmental terminology. It is thus the main result and the practical application of this research, because it is the resource in which the hyponymy-based terminological entries were compiled and publicly shared.

HypoLexicon is also the convergence point of four resources: (i) EcoLexicon for the basic structure and information of the terminological entries; (ii) the EEC and the four specialized subcorpora for the population of the terminological entries; (iii) Sketch Engine for the extraction of hyponymic and contextual information through corpus analysis; and (iv) Lexonomy for the design of the terminological template and for the implementation of all data. HypoLexicon is publicly available on the Lexonomy platform and was published using the Creative Commons (CC) Attribution 4.0 International license. There are two ways to access HypoLexicon: through the user interface and through the admin interface.

5.2.1. USER INTERFACE

The user interface of HypoLexicon provides a view of all content for any interested user (e.g., terminologist, linguist, translator, technical expert, lay user). It has a home view and an entry view.

5.2.1.1. HOME VIEW

The home view shows the main menu in HypoLexicon on the Lexonomy platform (**Figure 78**). This section is composed of the following elements: (i) resource title; (ii) resource description; (iii) search bar to perform queries; and (iv) list of random entries.

The search bar and the list of random entries are designed for terminological resources published in Lexonomy with a large number of entries. However, since the number of entries in HypoLexicon is still rather small, these features are less relevant, other than providing direct access to all entries from the list of random entries.

[9] Available at: https://www.lexonomy.eu/hypolexicon

In the upper right corner of the home view, users can log in to Lexonomy if an account on this platform is available (e.g., access as an administrator to manage this resource, access as a contributor to add or modify entries in this resource). The font size can also be increased or reduced for better accessibility.

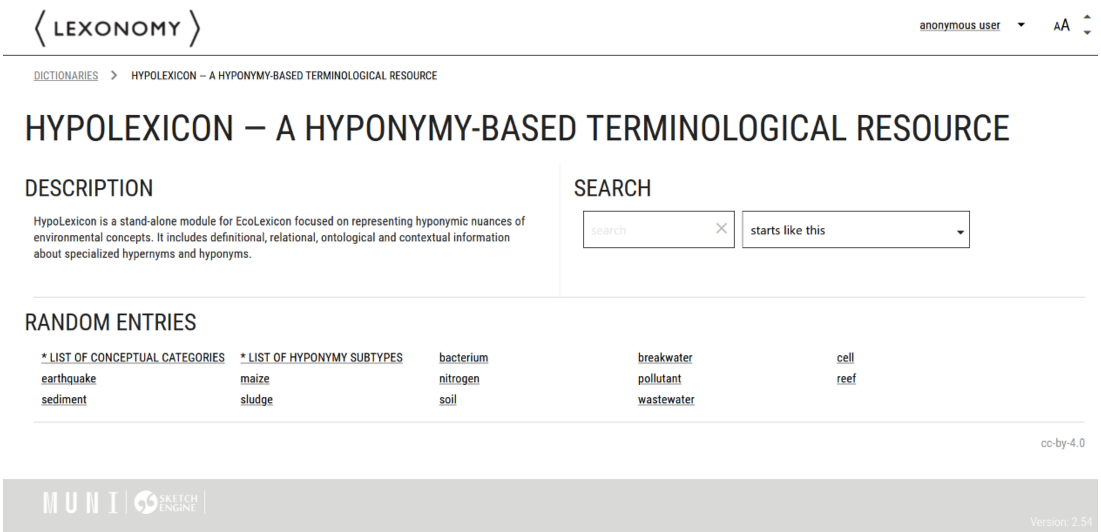


Figure 78. Main menu in HypoLexicon

5.2.1.2. ENTRY VIEW

The entry view shows the contents of HypoLexicon, which are the following: (i) the twelve hyponymy-based terminological entries; (ii) list of conceptual categories; and (iii) list of hyponymy subtypes. Because of the importance of conceptual categories and hyponymy subtypes in HypoLexicon, legends were added so that users could access the complete inventories within the same platform, without searching for the information in EcoLexicon.

Figures 79–90 show segments of the twelve hyponymy-based terminological entries in HypoLexicon. They correspond to the three BIO hypernyms (i.e., BACTERIUM, REEF, and CELL), the three CHEM hypernyms (i.e., SLUDGE, NITROGEN, and MAIZE), the three CIV hypernyms (i.e., WASTEWATER, BREAKWATER, and POLLUTANT), and the three GEO hypernyms (i.e., EARTHQUAKE, SEDIMENT, and SOIL). On the left side of the entry view, the search bar is followed by the full list of entries in HypoLexicon in alphabetical order. On the right side of the entry view, each terminological entry is displayed.

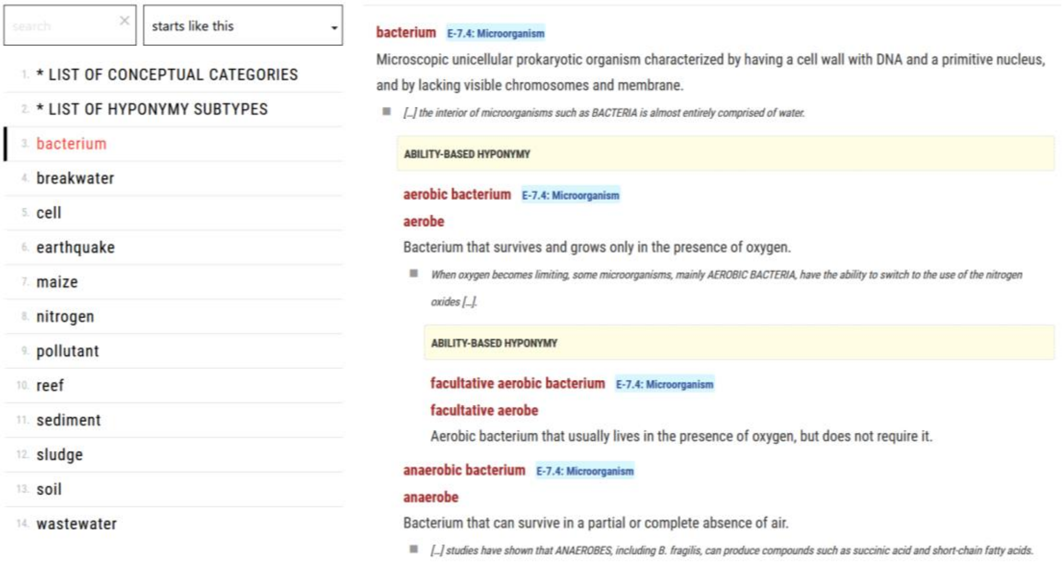


Figure 79. Segment of the BACTERIUM terminological entry in HypoLexicon

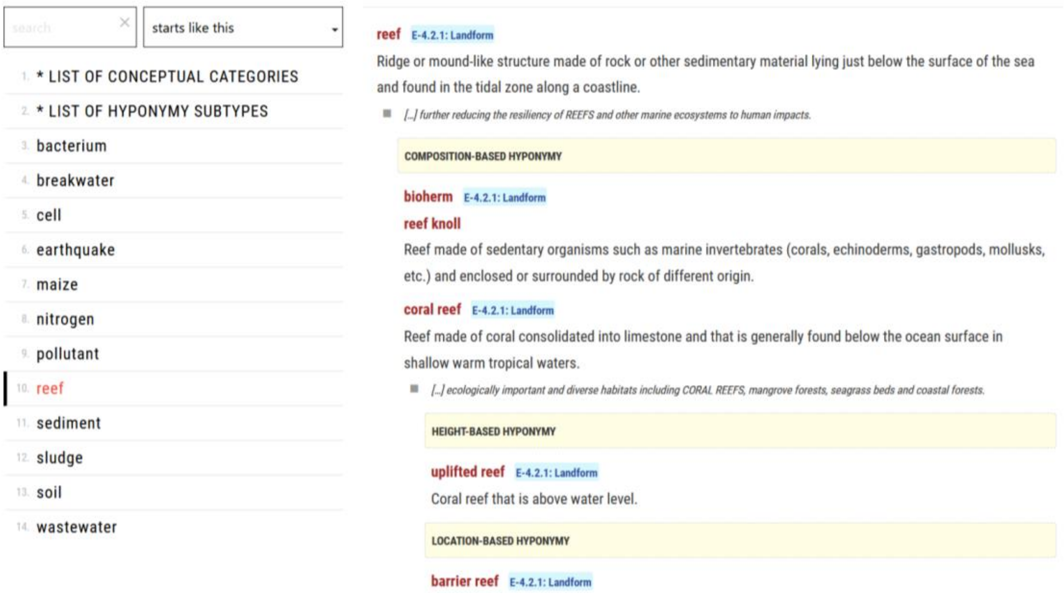


Figure 80. Segment of the REEF terminological entry in HypoLexicon

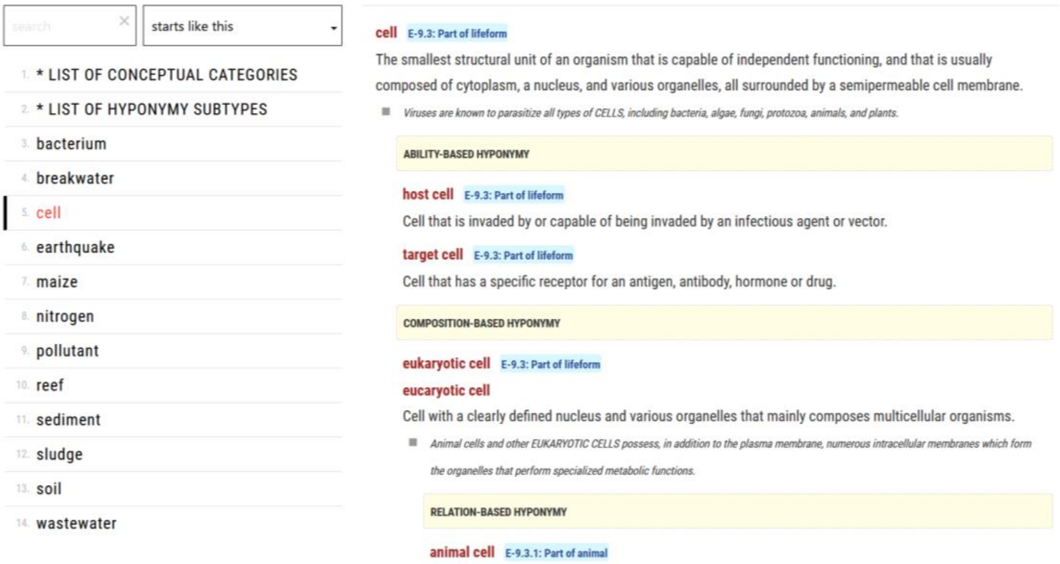


Figure 81. Segment of the CELL terminological entry in HypoLexicon

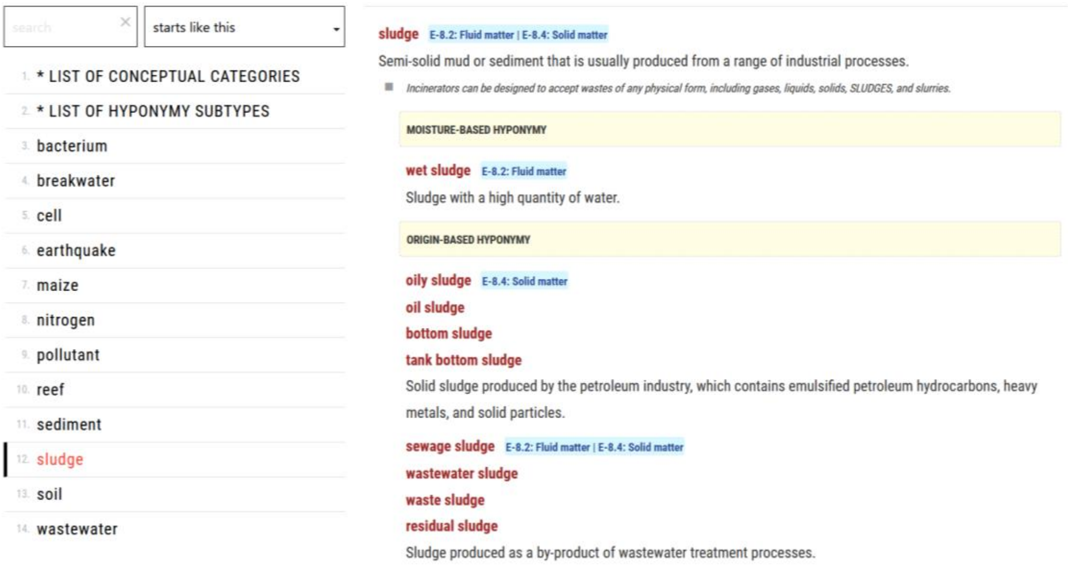


Figure 82. Segment of the SLUDGE terminological entry in HypoLexicon

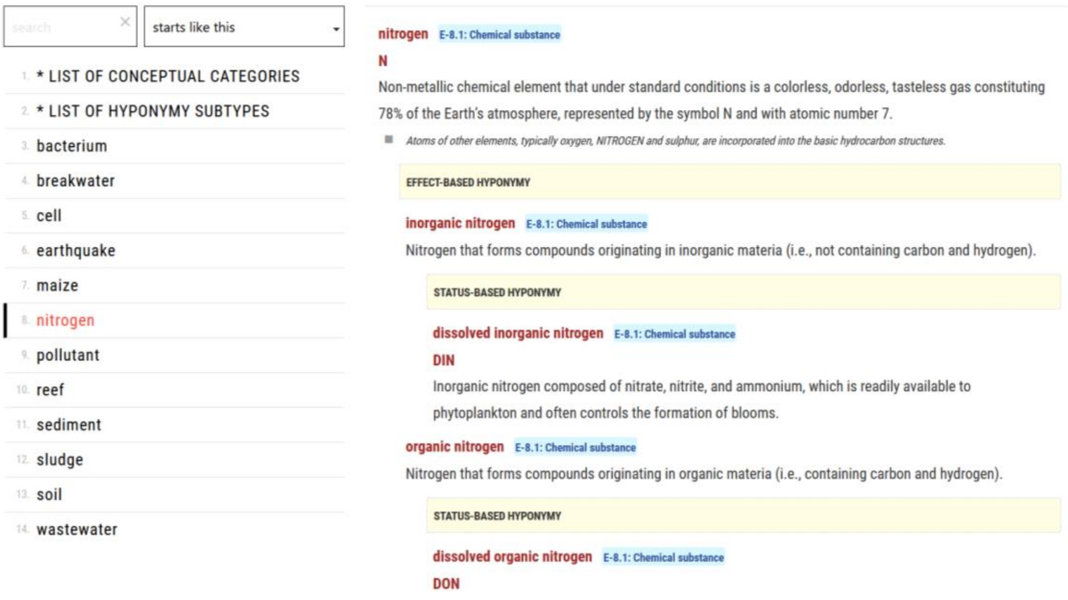


Figure 83. Segment of the NITROGEN terminological entry in HypoLexicon

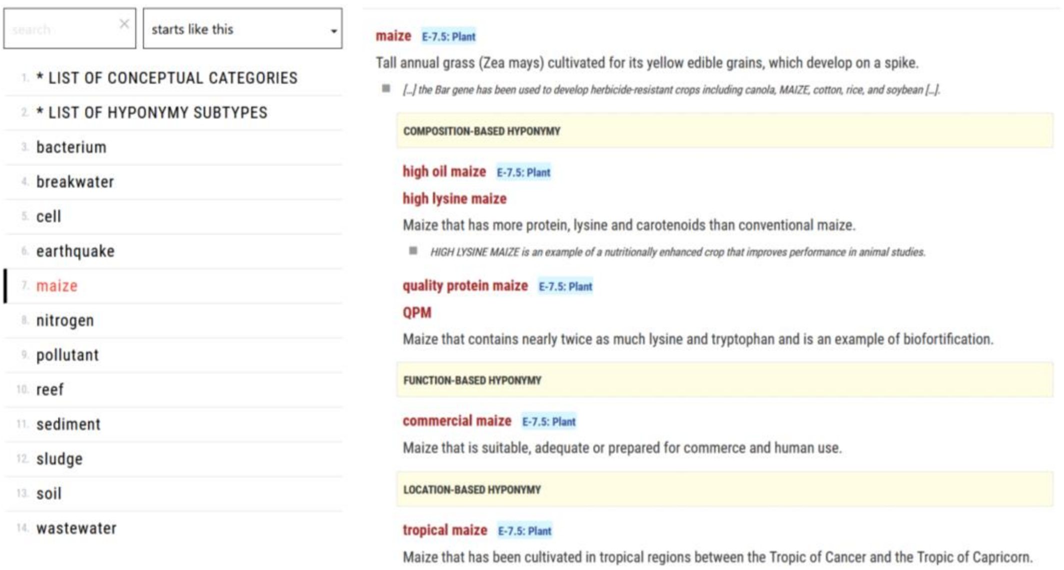


Figure 84. Segment of the MAIZE terminological entry in HypoLexicon

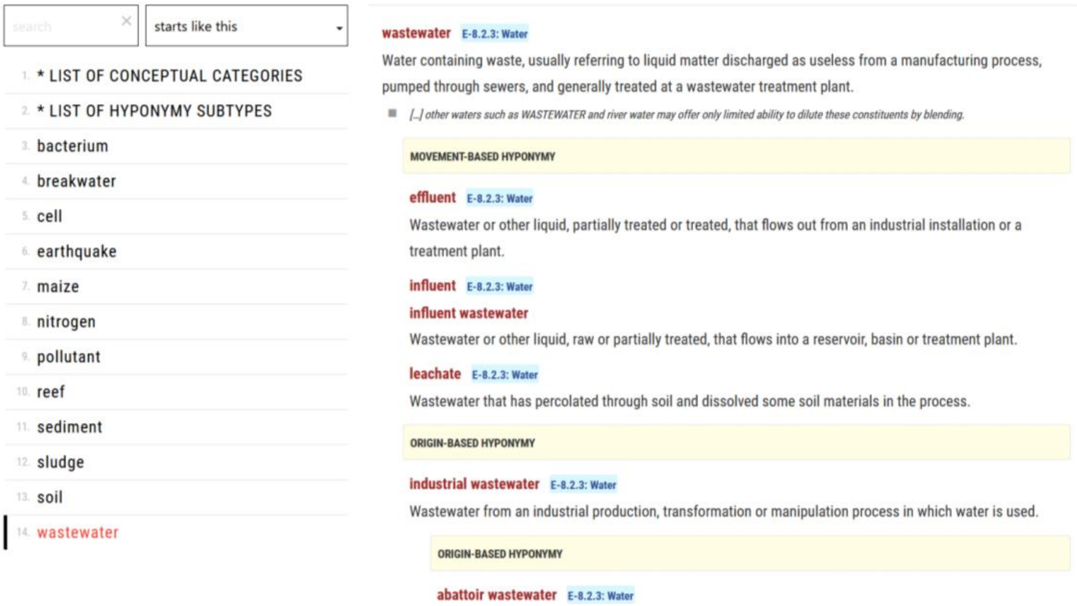


Figure 85. Segment of the WASTEWATER terminological entry in HypoLexicon

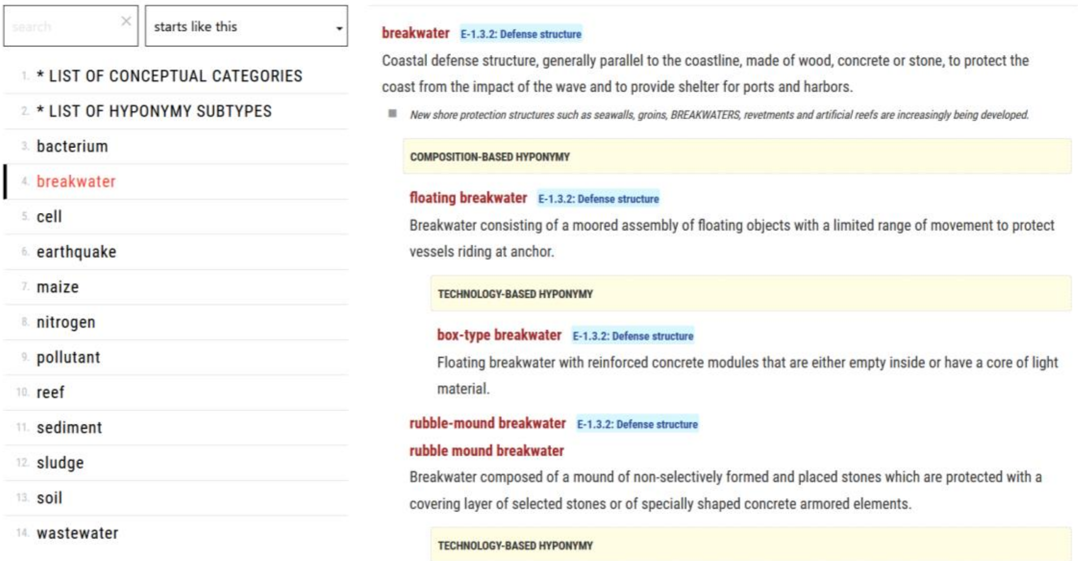


Figure 86. Segment of the BREAKWATER terminological entry in HypoLexicon

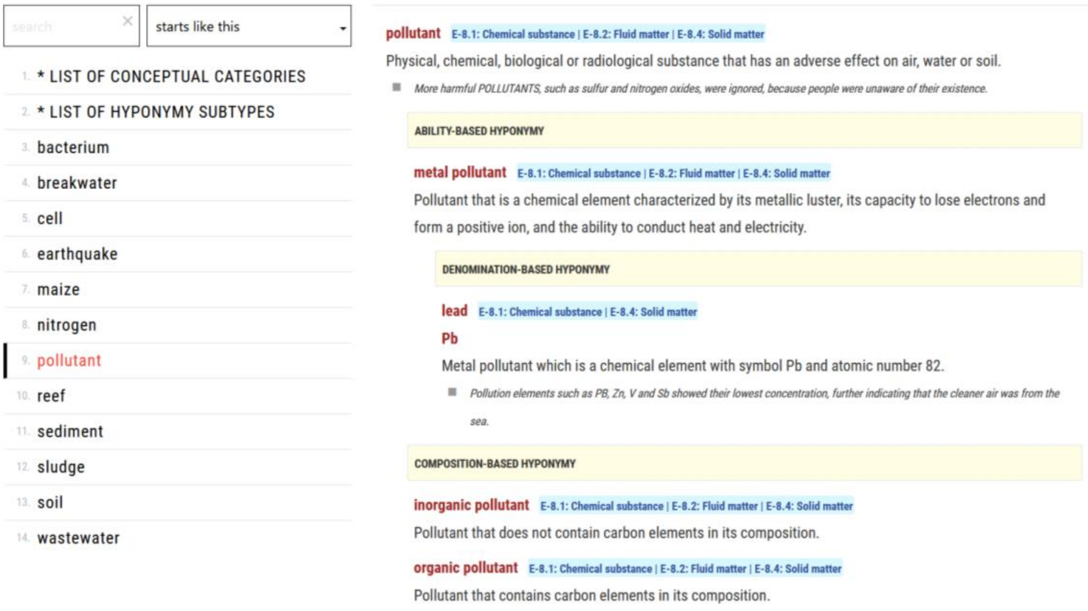


Figure 87. Segment of the POLLUTANT terminological entry in HypoLexicon

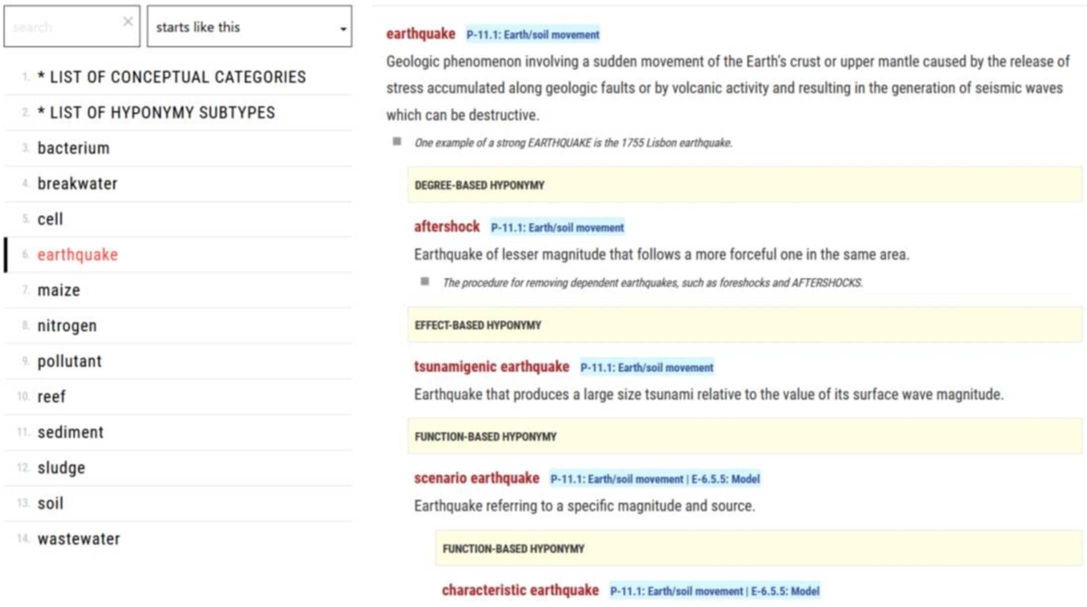


Figure 88. Segment of the EARTHQUAKE terminological entry in HypoLexicon

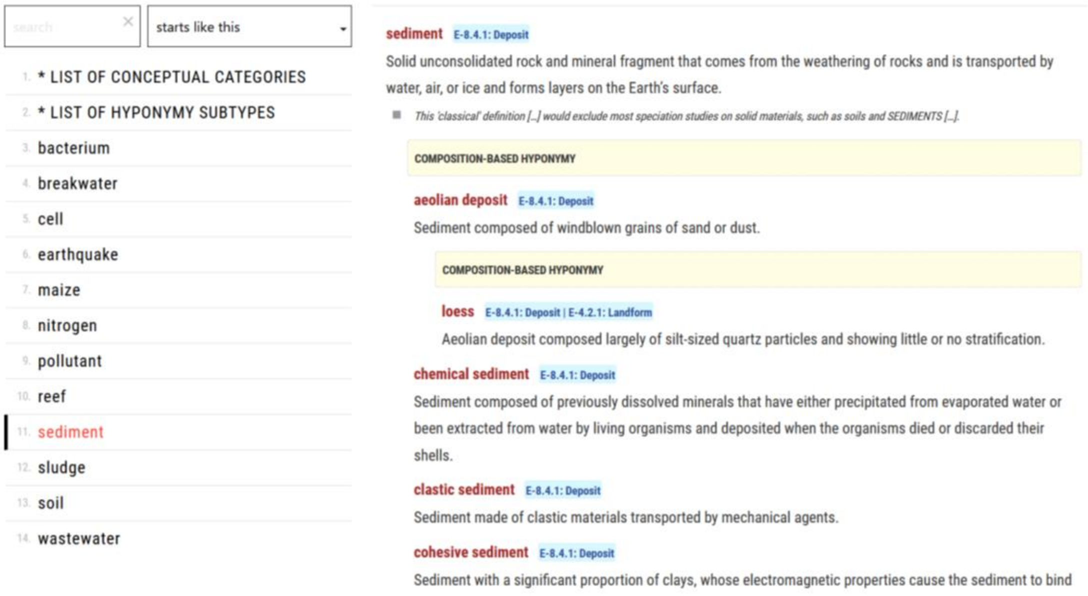


Figure 89. Segment of the SEDIMENT terminological entry in HypoLexicon

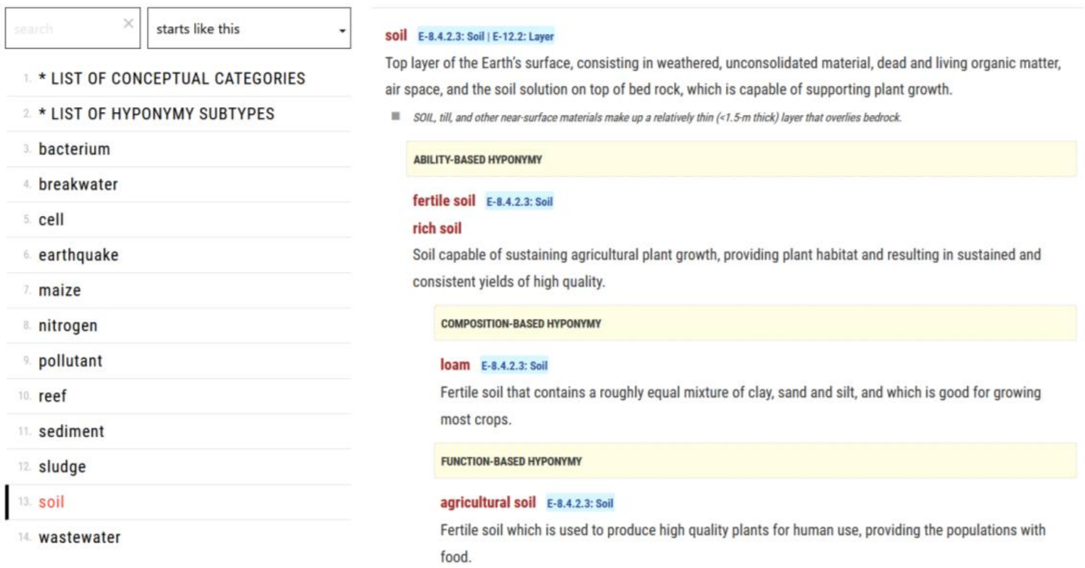


Figure 90. Segment of the SOIL terminological entry in HypoLexicon

Figures 91 and 92 show the views in HypoLexicon of the full list of conceptual categories and hyponymy subtypes, respectively. Unlike the twelve hyponymy-based terminological entries, these have a much simpler structure and a flat format.

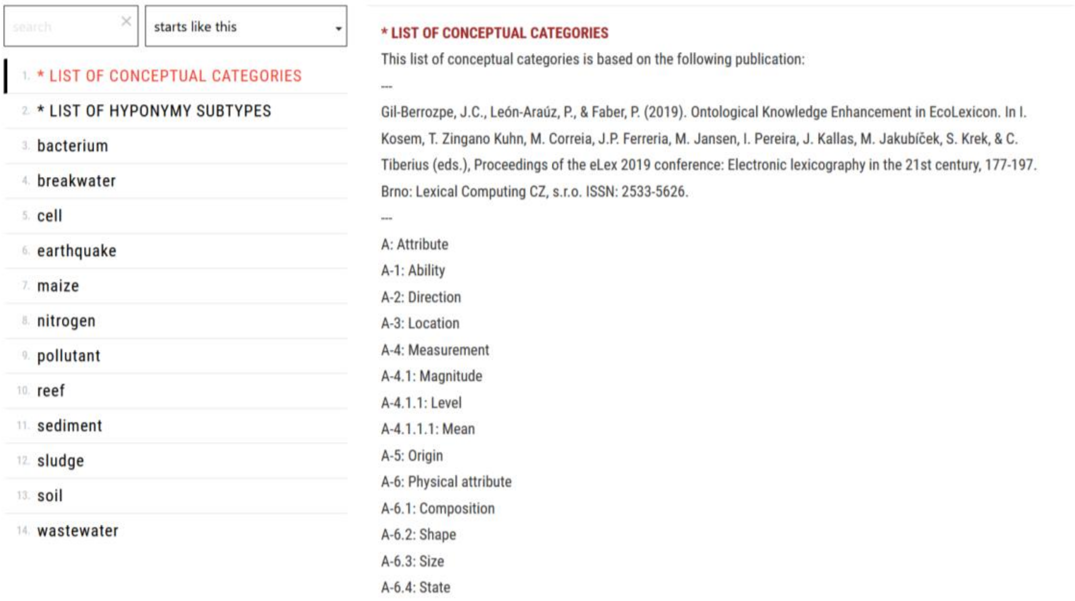


Figure 91. Segment of the list of conceptual categories in HypoLexicon

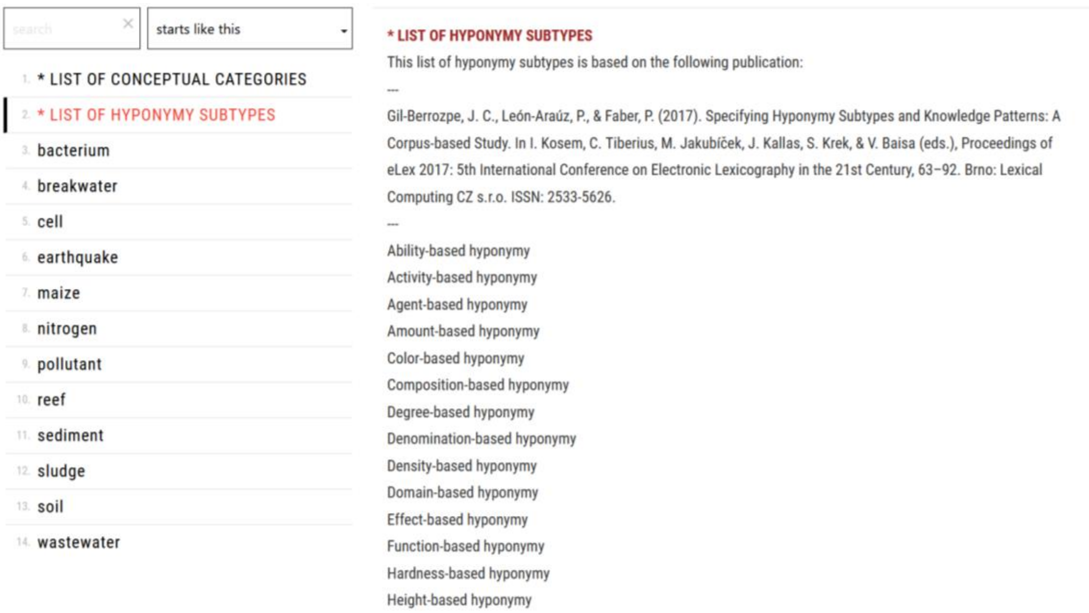


Figure 92. Segment of the list of hyponymy subtypes in HypoLexicon

5.2.2. ADMIN INTERFACE

The admin interface of HypoLexicon grants access to the configuration menu of the resource (**Figure 93**). This menu contains three options: (i) settings for managing the dictionary (i.e., description, users, publishing, change URL); (ii) entry settings (i.e., structure, formatting, headword list, searching); and (iii) expert settings (i.e., entry editor, flags, auto-numbering, linking, download settings, subentries, Sketch Engine, KonText, multimedia API).

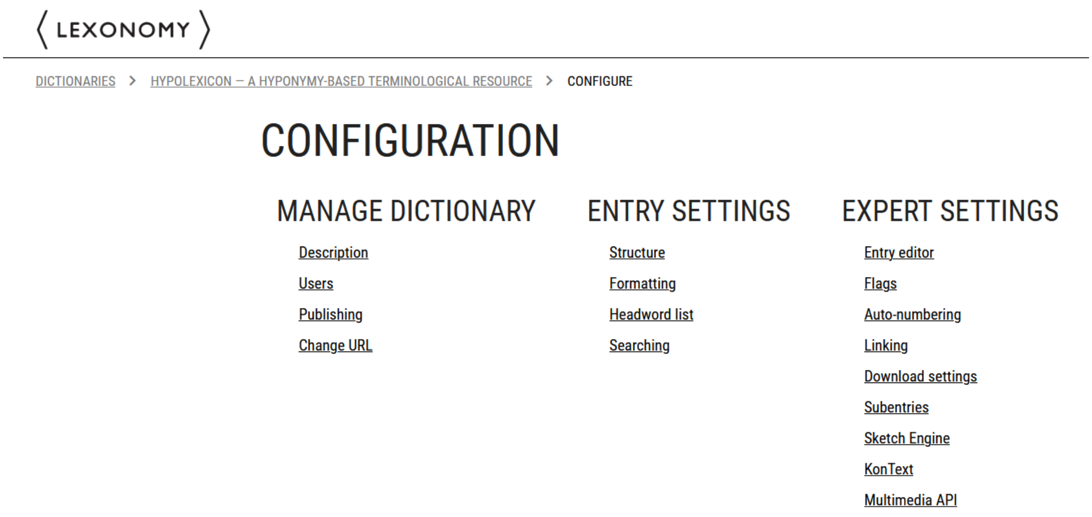


Figure 93. Configuration menu in HypoLexicon

The following subsection describes the most important configuration options in the design of HypoLexicon: (i) the entry settings related to structure and formatting; and (ii) the entry editor.

5.2.2.1. ENTRY SETTINGS: STRUCTURE AND FORMATTING

The entry structure configuration (**Figure 94**) is used to design and modify the structure of HypoLexicon, which follows the hierarchy written in XML language. As many elements as necessary can be created and structured within the hierarchy. After each element is named (e.g., <hyponym-1>), attributes can then be added (e.g., @conceptualcategory), content type must be selected (i.e., child elements, text, text with markup, value from list, empty, or media), and child elements associated with it can be specified, as well as the maximum and minimum number of each one.

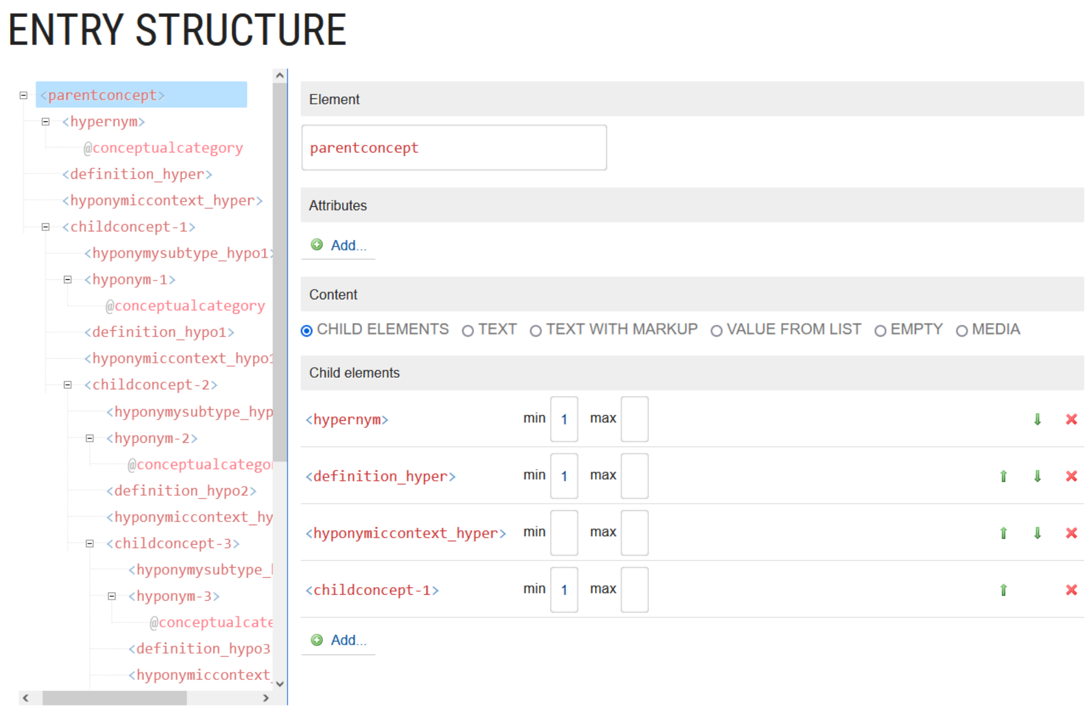


Figure 94. Entry structure configuration in HypoLexicon

The entry formatting configuration (**Figure 95**) allows the editor to design and alter the visual appearance of the terminological entries. When any of the elements is selected, a series of formatting options appear. For example, it is possible to specify whether the element is visible to viewers. The layout option indicates whether there should be a line break before and after or inline. Finally, there are many other appearance options (i.e., separation from other content, indentation and bulleting, box border, background color, outer punctuation, text color, text slant, text weight, text size, inner punctuation, interactivity), with multiple possibilities for each. In this way, each type of element can be designed independently to display information to the users.

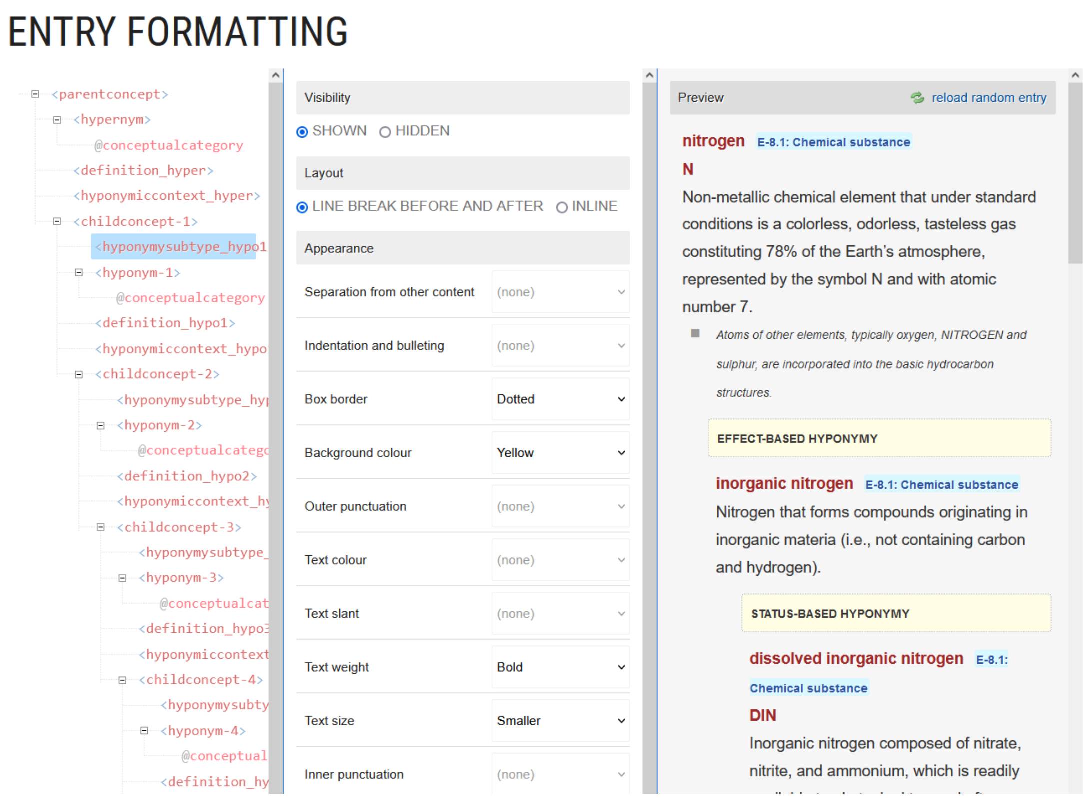


Figure 95. Entry formatting configuration in HypoLexicon

5.2.2.2. Entry Editor

The entry editor is used to add, delete or modify entries. Before entering the edit mode, it is possible to customize it by choosing the *nerd* or *laic* edit mode, and single-line or multi-line edit mode, which allows the administrator to use a smaller and simpler text editor or a more sophisticated one. **Figure 96** shows these options.

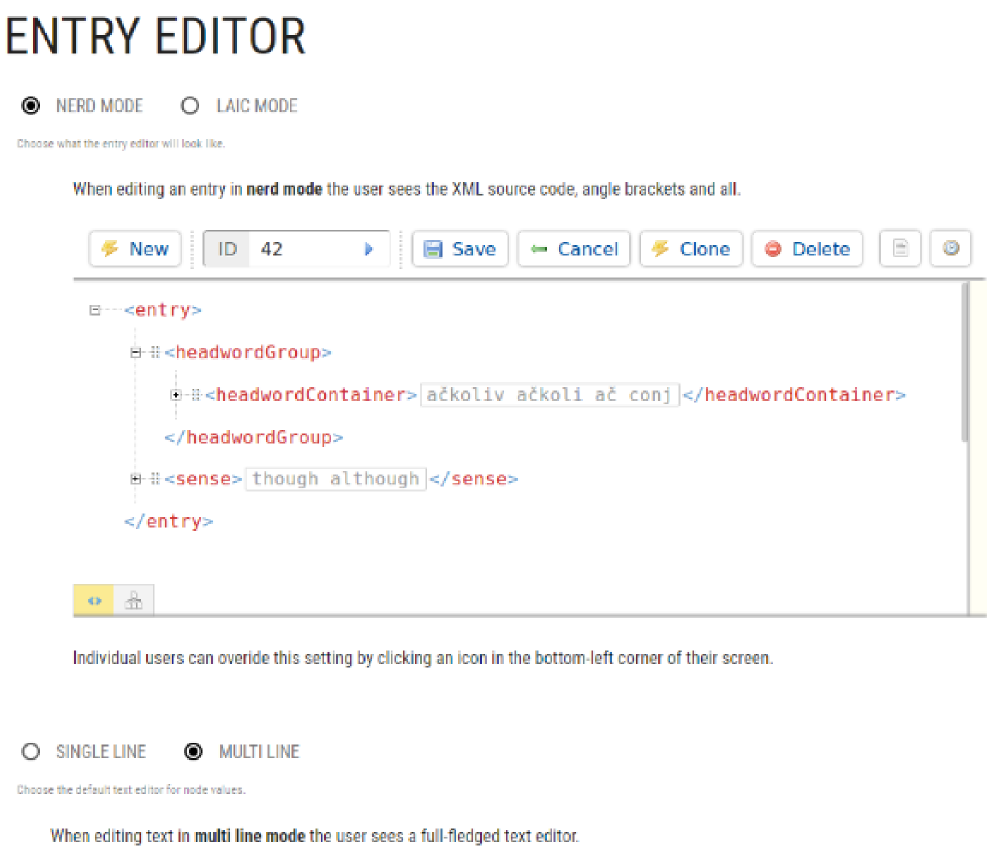


Figure 96. Entry editor mode selection in HypoLexicon

To access the edit mode, an entry should first be selected through the admin interface, where entries have the same appearance as the user interface. The only difference is that now a series of buttons are available above the entry: new, edit, clone, and delete (**Figure 97**).

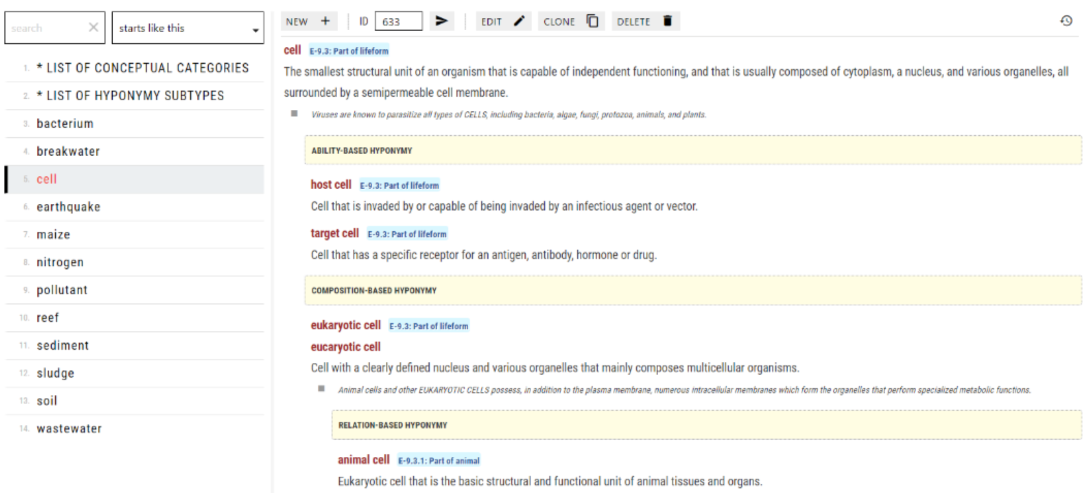


Figure 97. Edit mode in HypoLexicon

The edit mode is activated by clicking on the edit button. Its appearance depends on the type of edit mode selected: *nerd* edit mode (**Figure 98**) allows the administrator to see the XML source code, including the angle brackets (< >), the at sign (@) and the slashes (/); whereas the laic mode (**Figure 99**) hides all the XML source code so that the entry looks like a bulleted list, which can be easier to understand for lay users.

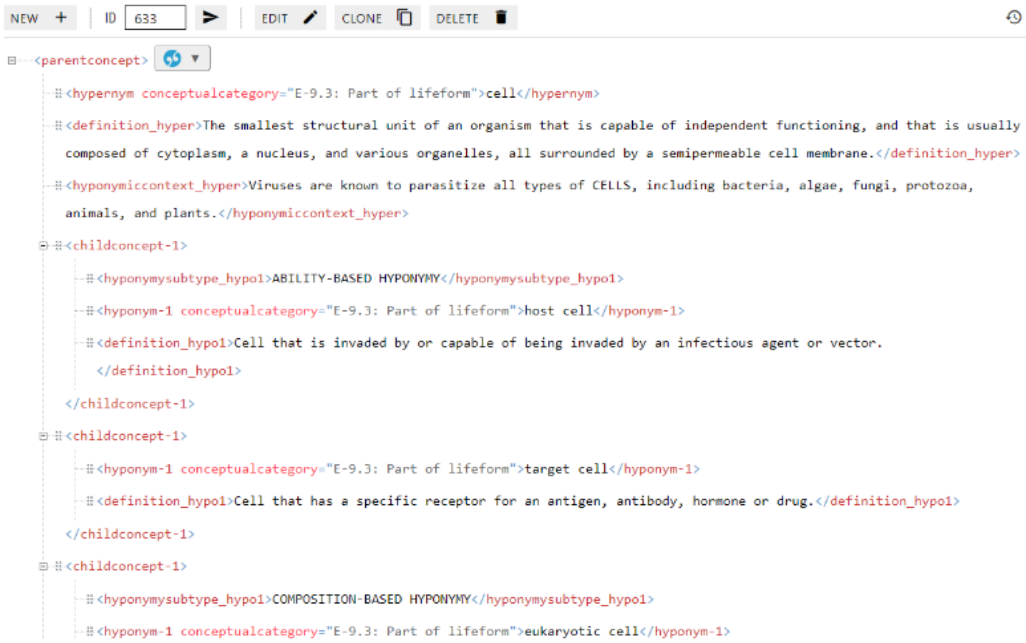


Figure 98. *Nerd* edit mode in HypoLexicon

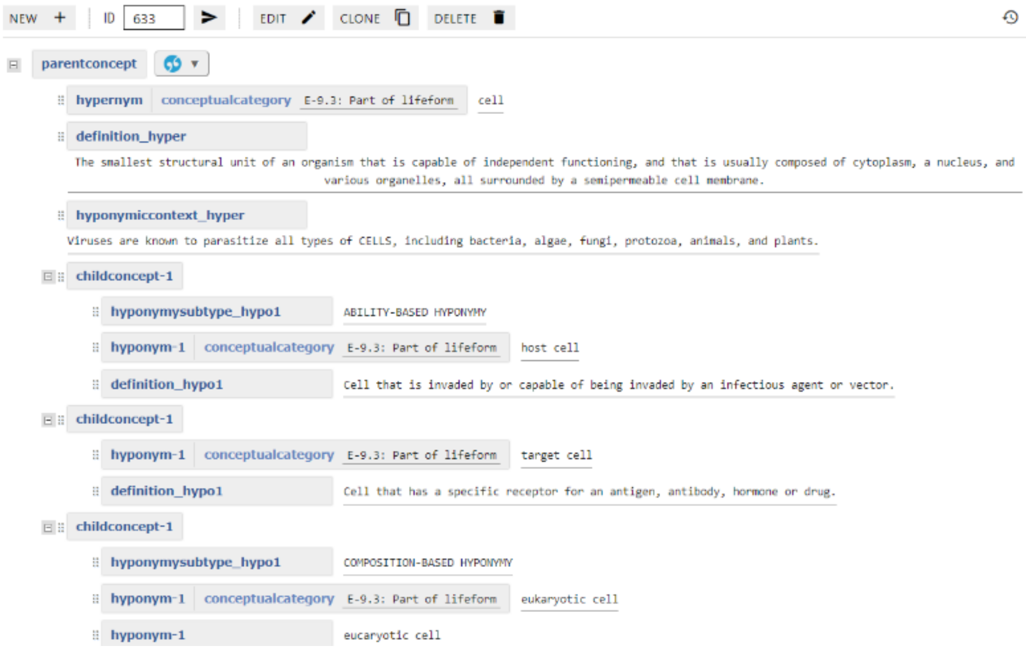


Figure 99. *Laic* edit mode in HypoLexicon

5.3. SUMMARY OF THE RESULTS

The results of this research are the twelve hyponymy-based terminological entries as well as the design and features of HypoLexicon. This section shows a summary of all the information in the terminological entries and uses ring diagrams to statistically summarize their content, including conceptual categories and hyponymy subtypes. In addition, it compares the terminological data in HypoLexicon with the original data in EcoLexicon to highlight the improvements achieved.

5.3.1. SUMMARY OF ALL HYPONYMY-BASED TERMINOLOGICAL ENTRIES

Generally speaking, the data in HypoLexicon consist of 309 concepts (i.e., 12 general hypernyms and 297 hyponyms) with their definitions; 465 terms that designate those concepts; 22 conceptual categories; 19 hyponymy subtypes; up to six hyponymy levels; and 107 concepts with hyponymic contexts.

Table 56 summarizes the information in all hyponymy-based terminological entries in HypoLexicon. These include the following: (i) general volume of concepts, terms, and terminological definitions; (ii) conceptual categories and concepts per conceptual category; (iii) hyponymy subtypes and hyponyms per hyponymy subtype; (iv) hyponyms per hyponymy level; and (v) concepts with hyponymic contexts.

Concepts, terms, and terminological definitions	
No. of concepts	309
No. of terms	465
No. of terminological definitions	309

Conceptual categories	
No. of conceptual categories	22
Conceptual categories	*chemical substance* ǀ *defense structure* ǀ *deposit* ǀ *earth/soil movement* ǀ *fluid matter* ǀ *gas* ǀ *landform* ǀ *layer* ǀ *microorganism* ǀ *mineral* ǀ *model* ǀ *particle* ǀ *part of lifeform* ǀ *part of animal* ǀ *part of fungus* ǀ *part of plant* ǀ *plant* ǀ *rock* ǀ *soil* ǀ *solid matter* ǀ *structure* ǀ *water*
No. of *chemical substance* concepts	52
***Chemical substance* concepts**	ACIDIFYING POLLUTANT ǀ AEROSOL ǀ AIR POLLUTANT ǀ AMMONIACAL NITROGEN ǀ ANTHROPOGENIC NITROGEN ǀ ANTHROPOGENIC POLLUTANT ǀ AROMATIC POLLUTANT ǀ ATMOSPHERIC NITROGEN ǀ BENZENE ǀ CARBON DIOXIDE ǀ COMBINED NITROGEN ǀ DICHLORODIPHENYLTRICHLOROETHANE ǀ DIOXIN ǀ DISSOLVED INORGANIC NITROGEN ǀ DISSOLVED ORGANIC NITROGEN ǀ FIXED NITROGEN ǀ GASEOUS NITROGEN ǀ GASEOUS POLLUTANT ǀ GREENHOUSE GAS ǀ INORGANIC

	NITROGEN ǀ INORGANIC POLLUTANT ǀ LEAD ǀ LIQUID NITROGEN ǀ MARINE NITROGEN ǀ METAL POLLUTANT ǀ METHANE ǀ METHYL BROMIDE ǀ MINERAL NITROGEN ǀ NITROGEN ǀ NITROGEN DIOXIDE ǀ NITROGEN OXIDE ǀ ORGANIC NITROGEN ǀ ORGANIC POLLUTANT ǀ OXIDIZED NITROGEN ǀ OZONE ǀ PARTICULATE NITROGEN ǀ PEROXYACYL NITRATE ǀ PERSISTENT ORGANIC POLLUTANT ǀ PESTICIDE ǀ PHOTOCHEMICAL OXIDANT ǀ POLLUTANT ǀ POLYCHLORINATED BIPHENYL ǀ PRIMARY POLLUTANT ǀ PRIORITY POLLUTANT ǀ REACTIVE NITROGEN ǀ RIVERINE NITROGEN ǀ SECONDARY POLLUTANT ǀ SOIL POLLUTANT ǀ SULFUR DIOXIDE ǀ TISSUE NITROGEN ǀ TROPOSPHERIC OZONE ǀ WATER POLLUTANT
No. of *defense structure* concepts	24
***Defense structure* concepts**	BERM BREAKWATER ǀ BOX-TYPE BREAKWATER ǀ BREAKWATER ǀ CAISSON BREAKWATER ǀ COASTAL BREAKWATER ǀ COMPOSITE BREAKWATER ǀ DUAL CYLINDRICAL CAISSON BREAKWATER ǀ FLOATING BREAKWATER ǀ HEADLAND BREAKWATER ǀ HIGH-CRESTED BREAKWATER ǀ LOW-CRESTED BREAKWATER ǀ OFFSHORE BREAKWATER ǀ PERFORATED-WALL CAISSON BREAKWATER ǀ PERMEABLE BREAKWATER ǀ POROUS BREAKWATER ǀ REEF BREAKWATER ǀ RUBBLE-MOUND BREAKWATER ǀ SEMICIRCULAR BREAKWATER ǀ SHORE-PARALLEL BREAKWATER ǀ SKIRT BREAKWATER ǀ SLOPING BREAKWATER ǀ S-SLOPE BREAKWATER ǀ SUBMERGED BREAKWATER ǀ VERTICAL BREAKWATER
No. of *deposit* concepts	48
***Deposit* concepts**	AEOLIAN DEPOSIT ǀ ALLUVIAL SEDIMENT ǀ BEACH SEDIMENT ǀ BIOGENIC SEDIMENT ǀ BOTTOM SEDIMENT ǀ CARBONATE SEDIMENT ǀ CENTRAL MORAINE ǀ CHEMICAL SEDIMENT ǀ CLASTIC SEDIMENT ǀ CLAY ǀ COARSE SAND ǀ COARSE SEDIMENT ǀ COASTAL SEDIMENT ǀ COHESIVE SEDIMENT ǀ CONSOLIDATED SEDIMENT ǀ CONTINENTAL SEDIMENT ǀ DEEP-SEA SEDIMENT ǀ DEPOSITED SEDIMENT ǀ DIAMICTITE ǀ DRY-SCREEN SAND ǀ FINE SAND ǀ FINE SEDIMENT ǀ FLUVIAL SEDIMENT ǀ GRAVEL ǀ GROUND MORAINE ǀ INTRACLAST ǀ LAKE SEDIMENT ǀ LATERAL MORAINE ǀ LOESS ǀ MARINE SEDIMENT ǀ MORAINE ǀ NONCOHESIVE SEDIMENT ǀ ORIGINAL SAND ǀ PELITE ǀ RECESSIONAL MORAINE ǀ SAND ǀ SEDIMENT ǀ SHALLOW SEDIMENT ǀ SHEET FLOW SEDIMENT ǀ SILICICLASTIC ǀ SILT ǀ SOFT SEDIMENT ǀ STREAM SEDIMENT ǀ SUSPENDED SEDIMENT ǀ TERMINAL MORAINE ǀ TERRIGENOUS SEDIMENT ǀ TILL ǀ UNCONSOLIDATED SEDIMENT
No. of *earth/soil movement* concepts	15
***Earth/soil movement* concepts**	AFTERSHOCK ǀ CHARACTERISTIC EARTHQUAKE ǀ CRUSTAL EARTHQUAKE ǀ DEEP EARTHQUAKE ǀ EARTHQUAKE ǀ INLAND CRUSTAL EARTHQUAKE ǀ INTRAPLATE EARTHQUAKE ǀ LOCAL EARTHQUAKE ǀ SCENARIO EARTHQUAKE ǀ SHALLOW CRUSTAL EARTHQUAKE ǀ SHALLOW EARTHQUAKE ǀ SUBDUCTION

	EARTHQUAKE ǀ TECTONIC EARTHQUAKE ǀ TSUNAMIGENIC EARTHQUAKE ǀ VOLCANIC EARTHQUAKE
No. of *fluid matter* concepts	39
***Fluid matter* concepts**	ACIDIFYING POLLUTANT ǀ ACTIVATED SLUDGE ǀ AEROSOL ǀ AIR POLLUTANT ǀ ANTHROPOGENIC POLLUTANT ǀ API SEPARATOR SLUDGE ǀ AROMATIC POLLUTANT ǀ BENZENE ǀ BULKING SLUDGE ǀ CHEMICAL SLUDGE ǀ DICHLORODIPHENYLTRICHLOROETHANE ǀ DIGESTED SLUDGE ǀ EXCESS SLUDGE ǀ INORGANIC POLLUTANT ǀ LIQUID NITROGEN ǀ LIQUID SLUDGE ǀ METAL POLLUTANT ǀ METHYL BROMIDE ǀ MUNICIPAL SLUDGE ǀ ORGANIC POLLUTANT ǀ PERSISTENT ORGANIC POLLUTANT ǀ PESTICIDE ǀ POLLUTANT ǀ POLYCHLORINATED BIPHENYL ǀ PRIMARY POLLUTANT ǀ PRIMARY SLUDGE ǀ PRIORITY POLLUTANT ǀ RAW SLUDGE ǀ RETURNED SLUDGE ǀ SECONDARY POLLUTANT ǀ SECONDARY SLUDGE ǀ SEWAGE SLUDGE ǀ SLUDGE ǀ SOIL POLLUTANT ǀ STABILIZED SLUDGE ǀ TREATED SLUDGE ǀ UNDIGESTED SLUDGE ǀ WATER POLLUTANT ǀ WET SLUDGE
No. of *gas* concepts	14
***Gas* concepts**	CARBON DIOXIDE ǀ DIOXIN ǀ FIXED NITROGEN ǀ GASEOUS NITROGEN ǀ GASEOUS POLLUTANT ǀ GREENHOUSE GAS ǀ METHANE ǀ NITROGEN DIOXIDE ǀ NITROGEN OXIDE ǀ OZONE ǀ PEROXYACYL NITRATE ǀ PHOTOCHEMICAL OXIDANT ǀ SULFUR DIOXIDE ǀ TROPOSPHERIC OZONE
No. of *landform* concepts	26
***Landform* concepts**	ATOLL ǀ BARRIER REEF ǀ BIOHERM ǀ CENTRAL MORAINE ǀ COASTAL REEF ǀ COMPLETELY LAND RINGED ATOLL ǀ CORAL REEF ǀ FRINGING REEF ǀ GROUND MORAINE ǀ ICE REEF ǀ LATERAL MORAINE ǀ LOESS ǀ MORAINE ǀ NATAL REEF ǀ OFFSHORE REEF ǀ OUTER REEF ǀ RECESSIONAL MORAINE ǀ REEF ǀ REEF PATCH ǀ ROCKY REEF ǀ SEAWARD REEF ǀ SHALLOW REEF ǀ SUBMERGED REEF ǀ TERMINAL MORAINE ǀ TROPICAL REEF ǀ UPLIFTED REEF
No. of *layer* concepts	5
***Layer* concepts**	BULK SOIL ǀ SHALLOW SOIL ǀ SOIL ǀ SUBSOIL ǀ SURFACE SOIL
No. of *microorganism* concepts	34
***Microorganism* concepts**	AANP BACTERIUM ǀ AEROBIC BACTERIUM ǀ ANAEROBIC BACTERIUM ǀ AUTOTROPHIC BACTERIUM ǀ BACILLUS ǀ BACILLUS SUBTILIS ǀ BACTERIUM ǀ CHROMATIUM ǀ COLIFORM BACTERIUM ǀ CYANOBACTERIUM ǀ ENTERIC BACTERIUM ǀ ESCHERICHIA COLI ǀ FACULTATIVE AEROBIC BACTERIUM ǀ FILAMENTOUS BACTERIUM ǀ GRAM-NEGATIVE BACTERIUM ǀ GRAM-POSITIVE BACTERIUM ǀ HAEMOPHILUS ǀ HETEROTROPHIC BACTERIUM ǀ LACTIC ACID BACTERIUM ǀ MARINE BACTERIUM ǀ METHANOGENIC BACTERIUM ǀ MYCOBACTERIUM ǀ NITRIFYING BACTERIUM ǀ NITROGEN-FIXING BACTERIUM ǀ ORAL BACTERIUM ǀ PATHOGENIC BACTERIUM ǀ PHOTOSYNTHETIC BACTERIUM ǀ PSEUDOMONAS ǀ SALMONELLA ǀ SOIL BACTERIUM ǀ

	STAPHYLOCOCCUS ǀ STREPTOCOCCUS ǀ SULFUR BACTERIUM ǀ SYMBIOTIC BACTERIUM
No. of *mineral* concepts	5
***Mineral* concepts**	COARSE SAND ǀ DRY-SCREEN SAND ǀ FINE SAND ǀ ORIGINAL SAND ǀ SAND
No. of *model* concepts	2
***Model* concepts**	CHARACTERISTIC EARTHQUAKE ǀ SCENARIO EARTHQUAKE
No. of *part of animal* concepts	18
***Part of animal* concepts**	ANIMAL CELL ǀ B CELL ǀ BLOOD CELL ǀ CANCER CELL ǀ CHROMATOPHORE ǀ EOSINOPHIL ǀ EPIDERMAL CELL ǀ EPITHELIAL CELL ǀ GRANULOSA CELL ǀ LYMPHOCYTE ǀ MACROPHAGE ǀ MUSCLE CELL ǀ NERVE CELL ǀ RED BLOOD CELL ǀ STEM CELL ǀ T CELL ǀ WHITE BLOOD CELL ǀ ZYGOTE
No. of *part of fungus* concepts	1
***Part of fungus* concepts**	YEAST CELL
No. of *part of lifeform* concepts	6
***Part of lifeform* concepts**	CELL ǀ DAUGHTER CELL ǀ EUKARYOTIC CELL ǀ HOST CELL ǀ MOTHER CELL ǀ TARGET CELL
No. of *part of plant* concepts	1
***Part of plant* concepts**	PLANT CELL
No. of *particle* concepts	5
***Particle* concepts**	COARSE SAND ǀ FINE SAND ǀ FINE SEDIMENT ǀ INTRACLAST ǀ SAND
No. of *plant* concepts	11
***Plant* concepts**	BT MAIZE ǀ COMMERCIAL MAIZE ǀ FERTILE TRANSGENIC MAIZE ǀ HIGH OIL MAIZE ǀ HYBRID MAIZE ǀ INBRED MAIZE ǀ MAIZE ǀ OPAQUE-2 MAIZE ǀ QUALITY PROTEIN MAIZE ǀ TRANSGENIC MAIZE ǀ TROPICAL MAIZE
No. of *rock* concepts	5
***Rock* concepts**	CLAY ǀ CONSOLIDATED SEDIMENT ǀ DIAMICTITE ǀ PELITE ǀ SILT
No. of *soil* concepts	34
***Soil* concepts**	AGRICULTURAL SOIL ǀ BACKFILL SOIL ǀ BULK SOIL ǀ CALCAREOUS SOIL ǀ CAMBISOL ǀ CLAY SOIL ǀ COHESIVE SOIL ǀ CONTAMINATED SOIL ǀ CONTRACTIVE SOIL ǀ EXPANSIVE SOIL ǀ FERTILE SOIL ǀ FOREST SOIL ǀ GRANULAR SOIL ǀ LEPTOSOL ǀ LITHOSOL ǀ LOAM ǀ METASTABLE SOIL ǀ MINERAL SOIL ǀ MOIST SOIL ǀ MOLLISOL ǀ ORGANIC SOIL ǀ REGOSOL ǀ RESIDUAL SOIL ǀ SALINE SOIL ǀ SANDY SOIL ǀ SATURATED SOIL ǀ SHALLOW SOIL ǀ SOFT SOIL ǀ SOIL ǀ SUBSOIL ǀ SURFACE SOIL ǀ TROPICAL SOIL ǀ URBAN SOIL ǀ WET SOIL
No. of *solid matter* concepts	20
***Solid matter* concepts**	ACIDIFYING POLLUTANT ǀ ANTHROPOGENIC POLLUTANT ǀ API SEPARATOR SLUDGE ǀ BULKING SLUDGE ǀ DEWATERED SLUDGE ǀ INORGANIC POLLUTANT ǀ LEAD ǀ METAL POLLUTANT ǀ OILY SLUDGE ǀ ORGANIC POLLUTANT ǀ PERSISTENT ORGANIC POLLUTANT ǀ PESTICIDE ǀ POLLUTANT ǀ PRIORITY POLLUTANT ǀ RAW SLUDGE ǀ SEWAGE SLUDGE ǀ SLUDGE ǀ SOIL POLLUTANT ǀ UNDIGESTED SLUDGE ǀ WATER POLLUTANT

No. of *structure* concepts	1
***Structure* concepts**	ARTIFICIAL REEF
No. of *water* concepts	25
***Water* concepts**	ABATTOIR WASTEWATER ǀ BLACKWATER ǀ DAIRY WASTEWATER ǀ EFFLUENT ǀ FOOD PROCESSING WASTEWATER ǀ GREYWATER ǀ HYDRAULIC FRACTURING WASTEWATER ǀ INDUSTRIAL WASTEWATER ǀ INFLUENT ǀ LEACHATE ǀ MOLASSES WASTEWATER ǀ OLIVE MILL WASTEWATER ǀ PETROCHEMICAL WASTEWATER ǀ PTA WASTEWATER ǀ RAW WASTEWATER ǀ SEPTIC WATERS ǀ SETTLED WASTEWATER ǀ SEWAGE ǀ STORMWATER ǀ SUGAR BEET WASTEWATER ǀ SURFACE RUNOFF ǀ TEXTILE WASTEWATER ǀ TREATED WASTEWATER ǀ URBAN WASTEWATER ǀ WASTEWATER

Hyponymy subtypes	
No. of hyponymy subtypes	19
Hyponymy subtypes	*ability-based* ǀ *color-based* ǀ *composition-based* ǀ *degree-based* ǀ *denomination-based* ǀ *effect-based* ǀ *function-based* ǀ *height-based* ǀ *location-based* ǀ *method-based* ǀ *moisture-based* ǀ *movement-based* ǀ *origin-based* ǀ *relation-based* ǀ *shape-based* ǀ *size-based* ǀ *state-based* ǀ *status-based* ǀ *technology-based*
No. of *ability-based* hyponyms	13
***Ability-based* hyponyms**	AEROBIC BACTERIUM ǀ ANAEROBIC BACTERIUM ǀ AUTOTROPHIC BACTERIUM ǀ CHROMATOPHORE ǀ FACULTATIVE AEROBIC BACTERIUM ǀ FERTILE SOIL ǀ FERTILE TRANSGENIC MAIZE ǀ HETEROTROPHIC BACTERIUM ǀ HOST CELL ǀ METAL POLLUTANT ǀ METHANOGENIC BACTERIUM ǀ PHOTOSYNTHETIC BACTERIUM ǀ TARGET CELL
No. of *color-based* hyponyms	5
***Color-based* hyponyms**	BLACKWATER ǀ CYANOBACTERIUM ǀ GREYWATER ǀ RED BLOOD CELL ǀ WHITE BLOOD CELL
No. of *composition-based* hyponyms	52
***Composition-based* hyponyms**	ACTIVATED SLUDGE ǀ AEOLIAN DEPOSIT ǀ AMMONIACAL NITROGEN AROMATIC POLLUTANT ǀ BENZENE ǀ BIOHERM ǀ CALCAREOUS SOIL ǀ CARBON DIOXIDE ǀ CARBONATE SEDIMENT ǀ CHEMICAL SEDIMENT ǀ CLASTIC SEDIMENT ǀ CLAY SOIL ǀ COHESIVE SEDIMENT ǀ COHESIVE SOIL ǀ CORAL REEF ǀ DIAMICTITE ǀ DICHLORODIPHENYLTRICHLOROETHANE ǀ DIOXIN ǀ EUKARYOTIC CELL ǀ FLOATING BREAKWATER ǀ GRANULAR SOIL ǀ GRAVEL ǀ HIGH OIL MAIZE ǀ INORGANIC POLLUTANT ǀ INTRACLAST ǀ LOAM ǀ LOESS ǀ METHANE ǀ METHYL BROMIDE ǀ MINERAL NITROGEN ǀ MINERAL SOIL ǀ NITROGEN DIOXIDE ǀ NITROGEN OXIDE ǀ NONCOHESIVE SEDIMENT ǀ OPAQUE-2 MAIZE ǀ ORGANIC POLLUTANT ǀ ORGANIC SOIL ǀ OXIDIZED NITROGEN ǀ OZONE ǀ PELITE ǀ PEROXYACYL NITRATE ǀ POLYCHLORINATED BIPHENYL ǀ QUALITY PROTEIN MAIZE ǀ REACTIVE NITROGEN ǀ ROCKY REEF ǀ RUBBLE-MOUND

	BREAKWATER ǀ SALINE SOIL ǀ SAND ǀ SANDY SOIL ǀ SOFT SEDIMENT ǀ SULFUR DIOXIDE ǀ TERRIGENOUS SEDIMENT
No. of *degree-based* hyponyms	1
***Degree-based* hyponyms**	AFTERSHOCK
No. of *denomination-based* hyponyms	17
***Denomination-based* hyponyms**	BACILLUS ǀ BACILLUS SUBTILIS ǀ CAMBISOL ǀ CHROMATIUM ǀ ESCHERICHIA COLI ǀ HAEMOPHILUS ǀ LEAD ǀ LEPTOSOL ǀ LITHOSOL ǀ MOLLISOL ǀ MYCOBACTERIUM ǀ PRIORITY POLLUTANT ǀ PSEUDOMONAS ǀ REGOSOL ǀ SALMONELLA ǀ STAPHYLOCOCCUS ǀ STREPTOCOCCUS
No. of *effect-based* hyponyms	18
***Effect-based* hyponyms**	ACIDIFYING POLLUTANT ǀ BT MAIZE ǀ CANCER CELL ǀ CONTRACTIVE SOIL ǀ EXPANSIVE SOIL ǀ FIXED NITROGEN ǀ GREENHOUSE GAS ǀ INORGANIC NITROGEN ǀ LACTIC ACID BACTERIUM ǀ MORAINE ǀ NITRIFYING BACTERIUM ǀ NITROGEN-FIXING BACTERIUM ǀ ORGANIC NITROGEN ǀ PATHOGENIC BACTERIUM ǀ PERSISTENT ORGANIC POLLUTANT ǀ PHOTOCHEMICAL OXIDANT ǀ SULFUR BACTERIUM ǀ TSUNAMIGENIC EARTHQUAKE
No. of *function-based* hyponyms	13
***Function-based* hyponyms**	AGRICULTURAL SOIL ǀ B CELL ǀ BACKFILL SOIL ǀ CHARACTERISTIC EARTHQUAKE ǀ COMMERCIAL MAIZE ǀ DRY-SCREEN SAND ǀ EOSINOPHIL ǀ LYMPHOCYTE ǀ MACROPHAGE ǀ PESTICIDE ǀ SCENARIO EARTHQUAKE ǀ STEM CELL ǀ T CELL
No. of *height-based* hyponyms	11
***Height-based* hyponyms**	DEEP EARTHQUAKE ǀ HIGH-CRESTED BREAKWATER ǀ LOW-CRESTED BREAKWATER ǀ SHALLOW CRUSTAL EARTHQUAKE ǀ SHALLOW EARTHQUAKE ǀ SHALLOW REEF ǀ SHALLOW SOIL ǀ SUBMERGED BREAKWATER ǀ SUBMERGED REEF ǀ SURFACE SOIL ǀ UPLIFTED REEF
No. of *location-based* hyponyms	55
***Location-based* hyponyms**	AIR POLLUTANT ǀ ALLUVIAL SEDIMENT ǀ ATMOSPHERIC NITROGEN BARRIER REEF ǀ BEACH SEDIMENT ǀ BLOOD CELL ǀ BULK SOIL ǀ CENTRAL MORAINE ǀ COASTAL BREAKWATER ǀ COASTAL REEF ǀ COASTAL SEDIMENT ǀ CONTINENTAL SEDIMENT ǀ DEEP-SEA SEDIMENT ǀ DEPOSITED SEDIMENT ǀ ENTERIC BACTERIUM ǀ EPIDERMAL CELL ǀ EPITHELIAL CELL ǀ FLUVIAL SEDIMENT ǀ FOREST SOIL ǀ FRINGING REEF ǀ GRANULOSA CELL ǀ GROUND MORAINE ǀ HEADLAND BREAKWATER ǀ INLAND CRUSTAL EARTHQUAKE ǀ INTRAPLATE EARTHQUAKE ǀ LAKE SEDIMENT ǀ LATERAL MORAINE ǀ LOCAL EARTHQUAKE ǀ MARINE BACTERIUM ǀ MARINE NITROGEN ǀ MARINE SEDIMENT ǀ MUSCLE CELL ǀ NATAL REEF ǀ NERVE CELL ǀ OFFSHORE BREAKWATER ǀ OFFSHORE REEF ǀ ORAL BACTERIUM ǀ OUTER REEF ǀ RIVERINE NITROGEN ǀ SEAWARD REEF ǀ SEWAGE ǀ SHALLOW SEDIMENT ǀ SHORE-PARALLEL BREAKWATER ǀ SOIL BACTERIUM ǀ SOIL POLLUTANT ǀ SUBDUCTION EARTHQUAKE ǀ SUBSOIL ǀ TERMINAL MORAINE ǀ TISSUE NITROGEN ǀ TROPICAL MAIZE ǀ TROPICAL REEF ǀ

	TROPICAL SOIL ǀ TROPOSPHERIC OZONE ǀ URBAN SOIL ǀ WATER POLLUTANT
No. of *method-based* hyponyms	6
***Method-based* hyponyms**	AANP BACTERIUM ǀ CHEMICAL SLUDGE ǀ GRAM-NEGATIVE BACTERIUM ǀ GRAM-POSITIVE BACTERIUM ǀ PRIMARY SLUDGE ǀ SECONDARY SLUDGE
No. of *moisture-based* hyponyms	7
***Moisture-based* hyponyms**	DEWATERED SLUDGE ǀ LIQUID SLUDGE ǀ METASTABLE SOIL ǀ MOIST SOIL ǀ SATURATED SOIL ǀ WET SLUDGE ǀ WET SOIL
No. of *movement-based* hyponyms	4
***Movement-based* hyponyms**	EFFLUENT ǀ INFLUENT ǀ LEACHATE ǀ SHEET FLOW SEDIMENT
No. of *origin-based* hyponyms	45
***Origin-based* hyponyms**	ABATTOIR WASTEWATER ǀ AEROSOL ǀ ANTHROPOGENIC NITROGEN ǀ ANTHROPOGENIC POLLUTANT ǀ API SEPARATOR SLUDGE ǀ ARTIFICIAL REEF ǀ BIOGENIC SEDIMENT ǀ BOTTOM SEDIMENT ǀ CLAY ǀ CRUSTAL EARTHQUAKE ǀ DAIRY WASTEWATER ǀ EXCESS SLUDGE ǀ FOOD PROCESSING WASTEWATER ǀ HYBRID MAIZE ǀ HYDRAULIC FRACTURING WASTEWATER ǀ ICE REEF ǀ INBRED MAIZE ǀ INDUSTRIAL WASTEWATER ǀ MOLASSES WASTEWATER ǀ MUNICIPAL SLUDGE ǀ OILY SLUDGE ǀ OLIVE MILL WASTEWATER ǀ ORIGINAL SAND ǀ PETROCHEMICAL WASTEWATER ǀ PRIMARY POLLUTANT ǀ PTA WASTEWATER ǀ RECESSIONAL MORAINE ǀ REEF PATCH ǀ RESIDUAL SOIL ǀ RETURNED SLUDGE ǀ SECONDARY POLLUTANT ǀ SEWAGE SLUDGE ǀ SILICICLASTIC ǀ SILT ǀ STORMWATER ǀ STREAM SEDIMENT ǀ SUGAR BEET WASTEWATER ǀ SURFACE RUNOFF ǀ TECTONIC EARTHQUAKE ǀ TEXTILE WASTEWATER ǀ TILL ǀ TRANSGENIC MAIZE ǀ URBAN WASTEWATER ǀ VOLCANIC EARTHQUAKE ǀ ZYGOTE
No. of *relation-based* hyponyms	8
***Relation-based* hyponyms**	ANIMAL CELL ǀ COMBINED NITROGEN ǀ DAUGHTER CELL ǀ MOTHER CELL ǀ PARTICULATE NITROGEN ǀ PLANT CELL ǀ SYMBIOTIC BACTERIUM ǀ YEAST CELL
No. of *shape-based* hyponyms	4
***Shape-based* hyponyms**	ATOLL ǀ COLIFORM BACTERIUM ǀ COMPLETELY LAND RINGED ATOLL ǀ FILAMENTOUS BACTERIUM
No. of *size-based* hyponyms	4
***Size-based* hyponyms**	COARSE SAND ǀ COARSE SEDIMENT ǀ FINE SAND ǀ FINE SEDIMENT
No. of *state-based* hyponyms	3
***State-based* hyponyms**	GASEOUS NITROGEN ǀ GASEOUS POLLUTANT ǀ LIQUID NITROGEN
No. of *status-based* hyponyms	18
***Status-based* hyponyms**	BULKING SLUDGE ǀ CONSOLIDATED SEDIMENT ǀ CONTAMINATED SOIL ǀ DIGESTED SLUDGE ǀ DISSOLVED INORGANIC NITROGEN ǀ DISSOLVED ORGANIC NITROGEN ǀ POROUS BREAKWATER ǀ RAW SLUDGE ǀ RAW WASTEWATER ǀ SEPTIC WATERS ǀ SETTLED WASTEWATER ǀ SOFT SOIL ǀ STABILIZED SLUDGE ǀ SUSPENDED SEDIMENT

	ǀ TREATED SLUDGE ǀ TREATED WASTEWATER ǀ UNCONSOLIDATED SEDIMENT ǀ UNDIGESTED SLUDGE
No. of *technology-based* hyponyms	13
***Technology-based* hyponyms**	BERM BREAKWATER ǀ BOX-TYPE BREAKWATER ǀ CAISSON BREAKWATER ǀ COMPOSITE BREAKWATER ǀ DUAL CYLINDRICAL CAISSON BREAKWATER ǀ PERFORATED-WALL CAISSON BREAKWATER ǀ PERMEABLE BREAKWATER ǀ REEF BREAKWATER ǀ SEMICIRCULAR BREAKWATER ǀ SKIRT BREAKWATER ǀ SLOPING BREAKWATER ǀ S-SLOPE BREAKWATER ǀ VERTICAL BREAKWATER

Hyponymy levels	
No. of hyponymy levels	6
No. of first-level hyponyms	131
First-level hyponyms	AEOLIAN DEPOSIT ǀ AEROBIC BACTERIUM ǀ AFTERSHOCK ǀ AIR POLLUTANT ǀ ANAEROBIC BACTERIUM ǀ ANTHROPOGENIC NITROGEN ǀ ANTHROPOGENIC POLLUTANT ǀ ARTIFICIAL REEF ǀ ATMOSPHERIC NITROGEN ǀ AUTOTROPHIC BACTERIUM ǀ BACKFILL SOIL ǀ BIOGENIC SEDIMENT ǀ BIOHERM ǀ BOTTOM SEDIMENT ǀ BULK SOIL ǀ CALCAREOUS SOIL ǀ CAMBISOL ǀ CHEMICAL SEDIMENT ǀ CLASTIC SEDIMENT ǀ CLAY ǀ COARSE SEDIMENT ǀ COASTAL BREAKWATER ǀ COHESIVE SEDIMENT ǀ COHESIVE SOIL ǀ COMBINED NITROGEN ǀ COMMERCIAL MAIZE ǀ CONSOLIDATED SEDIMENT ǀ CONTAMINATED SOIL ǀ CONTINENTAL SEDIMENT ǀ CORAL REEF ǀ DAUGHTER CELL ǀ DEEP EARTHQUAKE ǀ EFFLUENT ǀ EUKARYOTIC CELL ǀ EXPANSIVE SOIL ǀ FERTILE SOIL ǀ FILAMENTOUS BACTERIUM ǀ FINE SEDIMENT ǀ FLOATING BREAKWATER ǀ GASEOUS NITROGEN ǀ GRAM-NEGATIVE BACTERIUM ǀ GRAM-POSITIVE BACTERIUM ǀ GRANULAR SOIL ǀ GRAVEL ǀ HETEROTROPHIC BACTERIUM ǀ HIGH OIL MAIZE ǀ HOST CELL ǀ HYBRID MAIZE ǀ ICE REEF ǀ INBRED MAIZE ǀ INDUSTRIAL WASTEWATER ǀ INFLUENT ǀ INORGANIC NITROGEN ǀ INORGANIC POLLUTANT ǀ INTRACLAST ǀ LEACHATE ǀ LEPTOSOL ǀ LITHOSOL ǀ LOCAL EARTHQUAKE ǀ MARINE BACTERIUM ǀ MARINE NITROGEN ǀ MARINE SEDIMENT ǀ METAL POLLUTANT ǀ METASTABLE SOIL ǀ MINERAL SOIL ǀ MOLLISOL ǀ MORAINE ǀ MOTHER CELL ǀ NATAL REEF ǀ NITROGEN-FIXING BACTERIUM ǀ NONCOHESIVE SEDIMENT ǀ OFFSHORE REEF ǀ OILY SLUDGE ǀ ORAL BACTERIUM ǀ ORGANIC NITROGEN ǀ ORGANIC POLLUTANT ǀ ORGANIC SOIL ǀ PARTICULATE NITROGEN ǀ PATHOGENIC BACTERIUM ǀ PELITE ǀ PESTICIDE ǀ POROUS BREAKWATER ǀ PRIORITY POLLUTANT ǀ QUALITY PROTEIN MAIZE ǀ RAW WASTEWATER ǀ REGOSOL ǀ RESIDUAL SOIL ǀ RIVERINE NITROGEN ǀ ROCKY REEF ǀ RUBBLE-MOUND BREAKWATER ǀ SALINE SOIL ǀ SAND ǀ SANDY SOIL ǀ SATURATED SOIL ǀ SCENARIO EARTHQUAKE ǀ SEAWARD REEF ǀ SETTLED WASTEWATER ǀ SEWAGE SLUDGE ǀ SHALLOW EARTHQUAKE ǀ SHALLOW REEF ǀ SHALLOW SOIL ǀ SHEET

	FLOW SEDIMENT ǀ SILICICLASTIC ǀ SILT ǀ SOFT SEDIMENT ǀSUSPENDED SEDIMENT ǀ SOIL BACTERIUM ǀ SOIL POLLUTANT ǀ SUBMERGED BREAKWATER ǀ SUBMERGED REEF ǀ SUBSOIL ǀ SULFUR BACTERIUM ǀ SURFACE RUNOFF ǀ SURFACE SOIL ǀ SYMBIOTIC BACTERIUM ǀ TARGET CELL ǀ TECTONIC EARTHQUAKE ǀ TILL ǀ TISSUE NITROGEN ǀ TRANSGENIC MAIZE ǀ TREATED WASTEWATER ǀ TROPICAL MAIZE ǀ TROPICAL REEF ǀ TROPICAL SOIL ǀ TSUNAMIGENIC EARTHQUAKE ǀ UNCONSOLIDATED SEDIMENT ǀ URBAN SOIL ǀ URBAN WASTEWATER ǀ VOLCANIC EARTHQUAKE ǀ WATER POLLUTANT ǀ WET SLUDGE
No. of second-level hyponyms	101
Second-level hyponyms	ABATTOIR WASTEWATER ǀ ACIDIFYING POLLUTANT ǀ AEROSOL ǀ AGRICULTURAL SOIL ǀ ALLUVIAL SEDIMENT ǀ AMMONIACAL NITROGEN ǀ ANIMAL CELL ǀ API SEPARATOR SLUDGE ǀ AROMATIC POLLUTANT ǀ ATOLL ǀ BACILLUS ǀ BARRIER REEF ǀ BERM BREAKWATER ǀ BLACKWATER ǀ BOX-TYPE BREAKWATER ǀ BT MAIZE ǀ BULKING SLUDGE ǀ CARBONATE SEDIMENT ǀ CENTRAL MORAINE ǀ CHARACTERISTIC EARTHQUAKE ǀ CHROMATIUM ǀ CLAY SOIL ǀ COARSE SAND ǀ COASTAL REEF ǀ COASTAL SEDIMENT ǀ CONTRACTIVE SOIL ǀ CRUSTAL EARTHQUAKE ǀ DEEP-SEA SEDIMENT ǀ DEWATERED SLUDGE ǀ DICHLORODIPHENYLTRICHLOROETHANE ǀ DIGESTED SLUDGE ǀ DISSOLVED INORGANIC NITROGEN ǀ DISSOLVED ORGANIC NITROGEN ǀ DRY-SCREEN SAND ǀ ENTERIC BACTERIUM ǀ FACULTATIVE AEROBIC BACTERIUM ǀ FERTILE TRANSGENIC MAIZE ǀ FINE SAND ǀ FIXED NITROGEN ǀ FOOD PROCESSING WASTEWATER ǀ FOREST SOIL ǀ FRINGING REEF ǀ GASEOUS POLLUTANT ǀ GREYWATER ǀ GROUND MORAINE ǀ HAEMOPHILUS ǀ HEADLAND BREAKWATER ǀ HIGH-CRESTED BREAKWATER ǀ HYDRAULIC FRACTURING WASTEWATER ǀ INTRAPLATE EARTHQUAKE ǀ LACTIC ACID BACTERIUM ǀ LAKE SEDIMENT ǀ LATERAL MORAINE ǀ LEAD ǀ LIQUID NITROGEN ǀ LIQUID SLUDGE ǀ LOAM ǀ LOESS ǀ LOW-CRESTED BREAKWATER ǀ METHANOGENIC BACTERIUM ǀ METHYL BROMIDE ǀ MINERAL NITROGEN ǀ MOIST SOIL ǀ MUNICIPAL SLUDGE ǀ MYCOBACTERIUM ǀ NITRIFYING BACTERIUM ǀ OFFSHORE BREAKWATER ǀ OPAQUE-2 MAIZE ǀ ORIGINAL SAND ǀ OUTER REEF ǀ OXIDIZED NITROGEN ǀ PERMEABLE BREAKWATER ǀ PERSISTENT ORGANIC POLLUTANT ǀ PETROCHEMICAL WASTEWATER ǀ PHOTOSYNTHETIC BACTERIUM ǀ PLANT CELL ǀ PRIMARY POLLUTANT ǀ PSEUDOMONAS ǀ PTA WASTEWATER ǀ RAW SLUDGE ǀ REACTIVE NITROGEN ǀ REEF BREAKWATER ǀ REEF PATCH ǀ SECONDARY POLLUTANT ǀ SEWAGE ǀ SHALLOW SEDIMENT ǀ SLOPING BREAKWATER ǀ S-SLOPE BREAKWATER ǀ STABILIZED SLUDGE ǀ STAPHYLOCOCCUS ǀ STORMWATER ǀ STREPTOCOCCUS ǀ SUBDUCTION EARTHQUAKE ǀ TERMINAL MORAINE ǀ TEXTILE

	WASTEWATER ǀ TREATED SLUDGE ǀ UNDIGESTED SLUDGE ǀ UPLIFTED REEF ǀ VERTICAL BREAKWATER ǀ WET SOIL ǀ YEAST CELL
No. of third-level hyponyms	42
Third-level hyponyms	AANP BACTERIUM ǀ BACILLUS SUBTILIS ǀ BEACH SEDIMENT ǀ BENZENE ǀ BLOOD CELL ǀ CANCER CELL ǀ CHROMATOPHORE ǀ COLIFORM BACTERIUM ǀ COMPLETELY LAND RINGED ATOLL ǀ COMPOSITE BREAKWATER ǀ CYANOBACTERIUM ǀ DAIRY WASTEWATER ǀ DEPOSITED SEDIMENT ǀ DIOXIN ǀ EPIDERMAL CELL ǀ EPITHELIAL CELL ǀ FLUVIAL SEDIMENT ǀ GRANULOSA CELL ǀ GREENHOUSE GAS ǀ INLAND CRUSTAL EARTHQUAKE ǀ MUSCLE CELL ǀ NERVE CELL ǀ NITROGEN DIOXIDE ǀ NITROGEN OXIDE ǀ OLIVE MILL WASTEWATER ǀ PHOTOCHEMICAL OXIDANT ǀ POLYCHLORINATED BIPHENYL ǀ PRIMARY SLUDGE ǀ RECESSIONAL MORAINE ǀ SALMONELLA ǀ SECONDARY SLUDGE ǀ SEPTIC WATER ǀ SHALLOW CRUSTAL EARTHQUAKE ǀ SHORE-PARALLEL BREAKWATER ǀ SKIRT BREAKWATER ǀ SOFT SOIL ǀ STEM CELL ǀ STREAM SEDIMENT ǀ SUGAR BEET WASTEWATER ǀ SULFUR DIOXIDE ǀ TERRIGENOUS SEDIMENT ǀ ZYGOTE
No. of fourth-level hyponyms	13
Fourth-level hyponyms	ACTIVATED SLUDGE ǀ CAISSON BREAKWATER ǀ CARBON DIOXIDE ǀ CHEMICAL SLUDGE ǀ DIAMICTITE ǀ ESCHERICHIA COLI ǀ METHANE ǀ MOLASSES WASTEWATER ǀ OZONE ǀ PEROXYACYL NITRATE ǀ RED BLOOD CELL ǀ SEMICIRCULAR BREAKWATER ǀ WHITE BLOOD CELL
No. of fifth-level hyponyms	8
Fifth-level hyponyms	DUAL CYLINDRICAL CAISSON BREAKWATER ǀ EOSINOPHIL ǀ EXCESS SLUDGE ǀ LYMPHOCYTE ǀ MACROPHAGE ǀ PERFORATED-WALL CAISSON BREAKWATER ǀ RETURNED SLUDGE ǀ TROPOSPHERIC OZONE
No. of sixth-level hyponyms	2
Sixth-level hyponyms	B CELL ǀ T CELL

Hyponymic contexts	
No. of concepts with hyponymic contexts	107
Concepts with hyponymic contexts	AEROBIC BACTERIUM ǀ AFTERSHOCK ǀ AIR POLLUTANT ǀ ALLUVIAL SEDIMENT ǀ AMMONIACAL NITROGEN ǀ ANAEROBIC BACTERIUM ǀ ANIMAL CELL ǀ ANTHROPOGENIC POLLUTANT ǀ ARTIFICIAL REEF ǀ AUTOTROPHIC BACTERIUM ǀ B CELL ǀ BACILLUS ǀ BACILLUS SUBTILIS ǀ BACTERIUM ǀ BENZENE ǀ BIOGENIC SEDIMENT ǀ BLOOD CELL ǀ BREAKWATER ǀ CANCER CELL ǀ CARBON DIOXIDE ǀ CELL ǀ CHARACTERISTIC EARTHQUAKE ǀ CLAY SOIL ǀ COARSE SEDIMENT ǀ COASTAL SEDIMENT ǀCONSOLIDATED SEDIMENT ǀ COLIFORM BACTERIUM ǀ CONTAMINATED SOIL ǀ CORAL REEF ǀ CYANOBACTERIUM ǀ DAIRY WASTEWATER ǀ DEEP-SEA SEDIMENT ǀ DEWATERED SLUDGE ǀ DIAMICTITE ǀ DICHLORODIPHENYLTRICHLOROETHANE ǀ DIOXIN ǀ

	DISSOLVED ORGANIC NITROGEN I EARTHQUAKE I ENTERIC BACTERIUM I EOSINOPHIL I ESCHERICHIA COLI I EUKARYOTIC CELL I FILAMENTOUS BACTERIUM I FINE SEDIMENT I FOOD PROCESSING WASTEWATER I GRAM-NEGATIVE BACTERIUM I GRAVEL I GREENHOUSE GAS I HAEMOPHILUS I HETEROTROPHIC BACTERIUM I HIGH OIL MAIZE I LACTIC ACID BACTERIUM I LAKE SEDIMENT I LEAD I LOCAL EARTHQUAKE I LYMPHOCYTE I MACROPHAGE I MAIZE I METHANE I METHANOGENIC BACTERIUM I METHYL BROMIDE I MOLLISOL I MORAINE I MYCOBACTERIUM I NITROGEN I NITROGEN DIOXIDE I NITROGEN OXIDE I NITROGEN-FIXING BACTERIUM I ORAL BACTERIUM I ORGANIC POLLUTANT I OZONE I PARTICULATE NITROGEN I PATHOGENIC BACTERIUM I PELITE I PEROXYACYL NITRATE I PERSISTENT ORGANIC POLLUTANT I PESTICIDE I PHOTOSYNTHETIC BACTERIUM I POLLUTANT I POLYCHLORINATED BIPHENYL I PRIMARY POLLUTANT I PRIORITY POLLUTANT I PSEUDOMONAS I REEF I RESIDUAL SOIL I SALMONELLA I SAND I SEDIMENT I SEWAGE I SHORE-PARALLEL BREAKWATER I SLUDGE I SOFT SEDIMENT I SOIL I SOIL POLLUTANT I STAPHYLOCOCCUS I STORMWATER I STREAM SEDIMENT I STREPTOCOCCUS I SULFUR BACTERIUM I SULFUR DIOXIDE I SYMBIOTIC BACTERIUM I T CELL I TEXTILE WASTEWATER I TROPOSPHERIC OZONE I UNCONSOLIDATED SEDIMENT I WASTEWATER I WHITE BLOOD CELL

Table 56. Summary of the information in all hyponymy-based terminological entries

5.3.2. SUMMARY OF ALL CONCEPTUAL CATEGORIES AND HYPONYMY SUBTYPES

This section gives the percentages of conceptual categories and hyponymy subtypes in the terminological entries in HypoLexicon. **Figure 100** shows the percentages for the conceptual categories in the entries, in decreasing order: *chemical substance* (13.30%), *deposit* (12.28%), *fluid matter* (9.97%), *microorganism* (8.70%), *soil* (8.70%), *landform* (6.65%), *water* (6.39%), *defense structure* (6.14%), *solid matter* (5.12%), *part of animal* (4.60%), *earth/soil movement* (3.84%), *gas* (3.58%), *plant* (2.81%), *part of lifeform* (1.53%), *layer* (1.28%), *mineral* (1.28%), *particle* (1.28%), *rock* (1.28%), *model* (0.51%), *part of fungus* (0.26%), *part of plant* (0.26%), and *structure* (0.26%).

In this case, there is no need to analyze the numbers per specialized domain for two reasons: (i) in certain entries, such as those of BACTERIUM or BREAKWATER, there is only one conceptual category, so these amounts depend on how populated the entries are; and (ii) concepts can belong to more than one conceptual category.

The final volumes can thus depend on how granulated the concepts are in each entry, which is not representative in general terms.

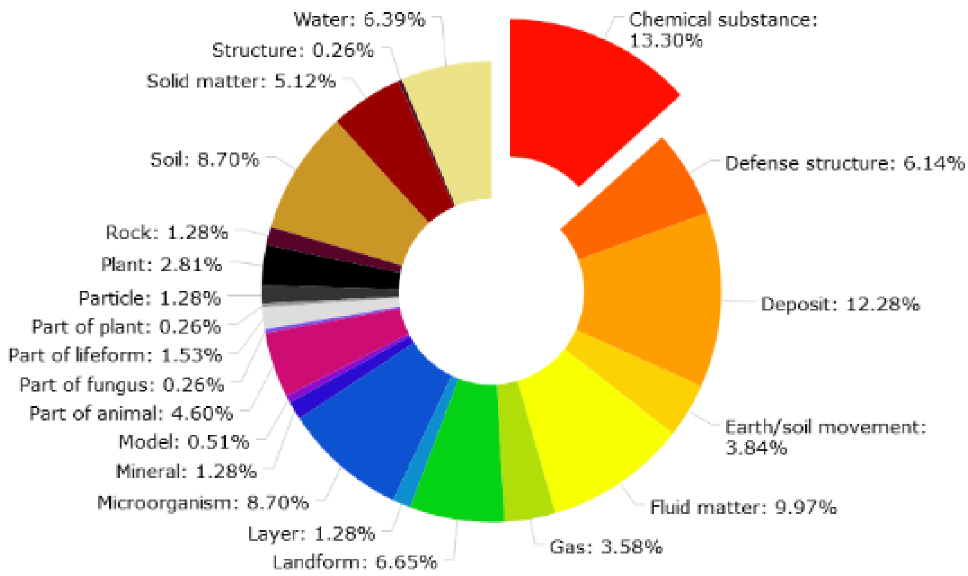


Figure 100. Statistics of the conceptual categories in all terminological entries

In contrast, **Figure 101** shows the percentage of the hyponymy subtypes for all entries. In decreasing order, they are the following: *location-based* (18.52%), *composition-based* (17.51%), *origin-based* (15.15%), *effect-based* (6.06%), *status-based* (6.06%), *denomination-based* (5.72%), *ability-based* (4.38%), *function-based* (4.38%), *technology-based* (4.38%), *height-based* (3.70%), *relation-based* (2.69%), *moisture-based* (2.36%), *method-based* (2.02%), *color-based* (1.68%), *movement-based* (1.35%), *shape-based* (1.35%), *size-based* (1.35%), *state-based* (1.01%), and *degree-based* (0.34%) hyponymy.

As can be observed, the most common hyponymy subtypes in the entries are those referring to the location, composition, and origin of the hyponyms. This information is consistent with previous findings regarding the most common hyponymic nuances in environmental concepts (Gil-Berrozpe *et al*. 2017). In the corpus-based study carried out by Gil-Berrozpe *et al.* (2017), *location-based*, *composition-based*, and *function-based* hyponymy were the most prevalent hyponymy subtypes. The fact that only *function-based* hyponymy clashes with the current data is easily explained, because this type of hyponymy is very common for artificial entities or artifacts, but in this research most of the general hypernyms are natural entities. This also explains why *origin-based* hyponymy was found among the top three subtypes in this research.

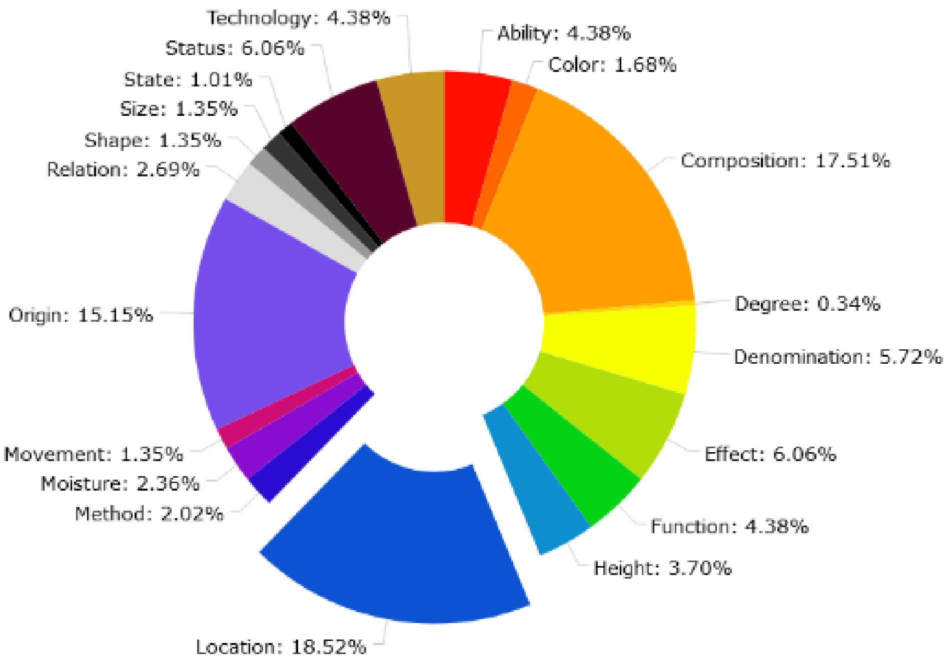


Figure 101. Statistics of the hyponymy subtypes in all terminological entries

Therefore, depending on the specialized knowledge domain and the nature of the concepts (e.g., whether they are natural or artificial, whether they are entities or processes), one type of hyponymy or another is activated (Gil-Berrozpe *et al*. 2017). For this reason, the information in **Figure 101** is divided in the following figures, according to the specialized domains of the terminological entries: BIO, CHEM, CIV, and GEO.

Figure 102 shows the percentages of the hyponymy subtypes in the BIO entries. In decreasing order, they are the following: *location-based* (23.38%), *ability-based* (12.99%), *denomination-based* (12.99%), *effect-based* (7.79%), *function-based* (7.79%), *relation-based* (7.79%), *composition-based* (5.19%), *origin-based* (5.19%), *shape-based* (5.19%), *color-based* (3.90%), *height-based* (3.90%), and *method-based* (3.90%) hyponymy. Therefore, the three most representative hyponymy subtypes of the BIO terminological entries are *location-based*, *ability-based*, and *denomination-based* hyponymy.

As can be observed, *location-based* hyponymy is one of the most frequent ways of expressing generic-specific relations in all three cases (i.e., BACTERIUM, REEF, and CELL) because it indicates the environment in which a concept characteristically takes place. For example, REEF hyponyms specify their location with respect to the shore or the sea, and CELL hyponyms indicate which part of an organism they are belong to. In contrast, *ability-based* hyponymy and *denomination-based* hyponymy are typical of BACTERIUM hyponyms. Their designations thus indicate either their capacity to survive or to produce certain elements, or directly refer to their scientific names in Latin.

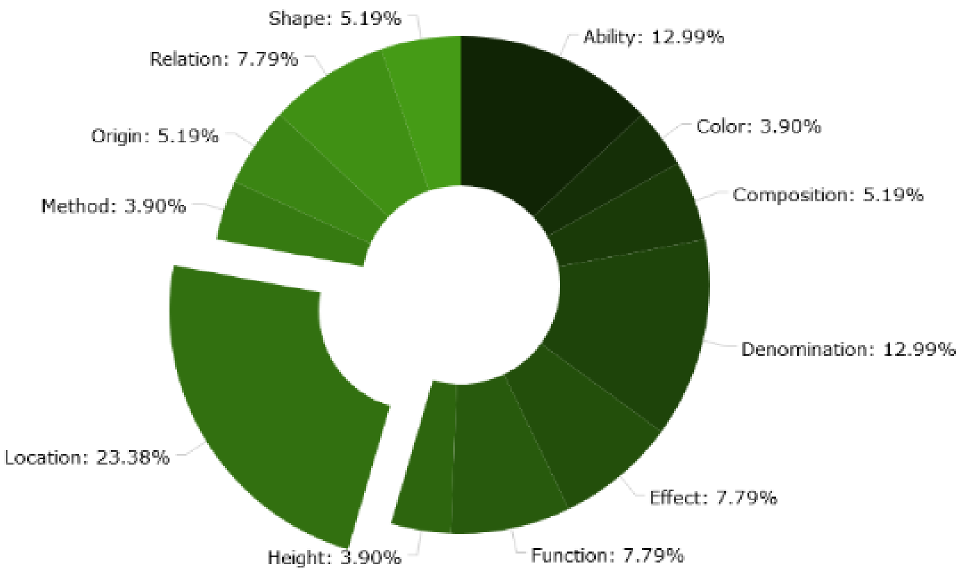


Figure 102. Statistics of the hyponymy subtypes in the BIO terminological entries

Figure 103 shows the percentages of the hyponymy subtypes in the CHEM entries. In decreasing order, they are the following: *origin-based* (21.28%), *composition-based* (17.02%), *status-based* (17.02%), *location-based* (10.64%), *effect-based* (8.51%), *method-based* (6.38%), *moisture-based* (6.38%), *relation-based* (4.26%), *state-based* (4.26%), *ability-based* (2.13%), and *function-based* (2.13%) hyponymy.

As can be observed, the three most representative hyponymy subtypes of the CHEM entries are *origin-based*, *composition-based*, and *status-based* hyponymy. For example, *origin-based* hyponymy is one of the most prevalent subtypes in the SLUDGE and MAIZE hierarchies because it alludes to the waste treatment or farming processes that produce them. Moreover, *composition-based* hyponymy is prevalent in the conceptual hierarchies of NITROGEN and MAIZE, since it is related to their chemical components. Finally, *status-based* hyponymy is frequent in the conceptual hierarchies of SLUDGE and NITROGEN, since it refers to the physical conditions that transform these entities after they have undergone certain processes.

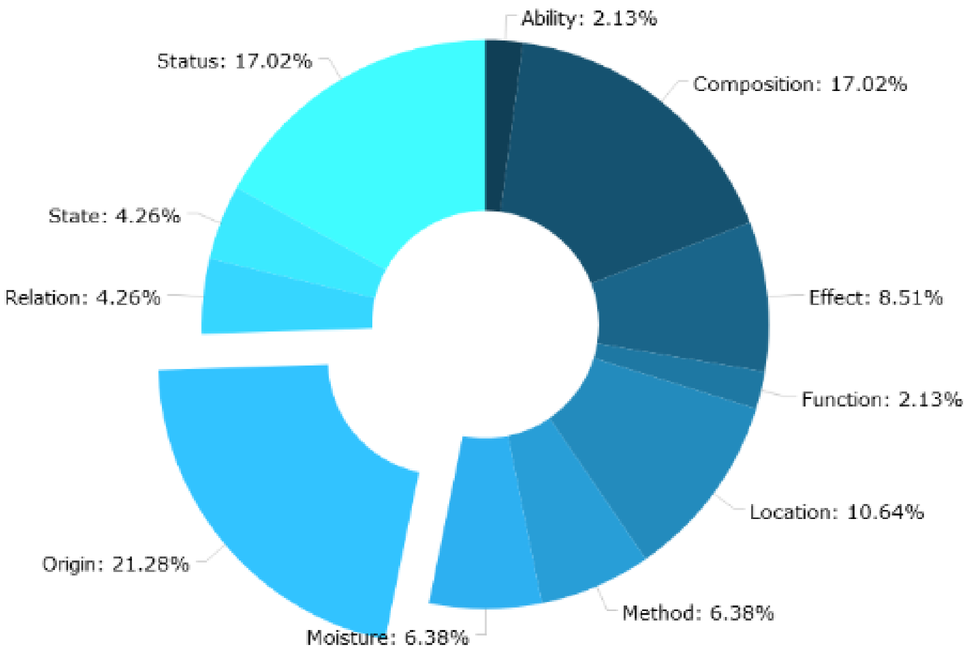


Figure 103. Statistics of the hyponymy subtypes in the CHEM terminological entries

Figure 104 shows the percentages of the hyponymy subtypes in the CIV entries. In decreasing order, they are the following: *origin-based* (22.78%), *composition-based* (21.52%), *technology-based* (16.46%), *location-based* (11.39%), *status-based* (6.33%), *effect-based* (5.06%), *height-based* (3.80%), *movement-based* (3.80%), *color-based* (2.53%), *denomination-based* (2.53%), *ability-based* (1.27%), *function-based* (1.27%), and *state-based* (1.27%) hyponymy.

As can be observed, the three most representative hyponymy subtypes of the CIV terminological entries are *origin-based*, *composition-based*, and *technology-based* hyponymy. For example, *origin-based* hyponymy predominates in the WASTEWATER hierarchy, since most of its hyponyms are conceptualized depending on the agent or process that generates or causes them. However, this subtype is also prevalent in the POLLUTANT hierarchy, since many of its hyponyms are classified in terms of who or what causes the pollution.

Moreover, *composition-based* hyponymy is also characteristic of the POLLUTANT hierarchy since its hyponyms are often polluting agents characterized by their chemical compounds. Lastly, *technology-based* hyponymy is undoubtedly the main way of conceptualizing hyponymy in the BREAKWATER hierarchy. Since a BREAKWATER is a defense structure, its more specific hyponyms are distinguished by the type of technology used to make them more effective.

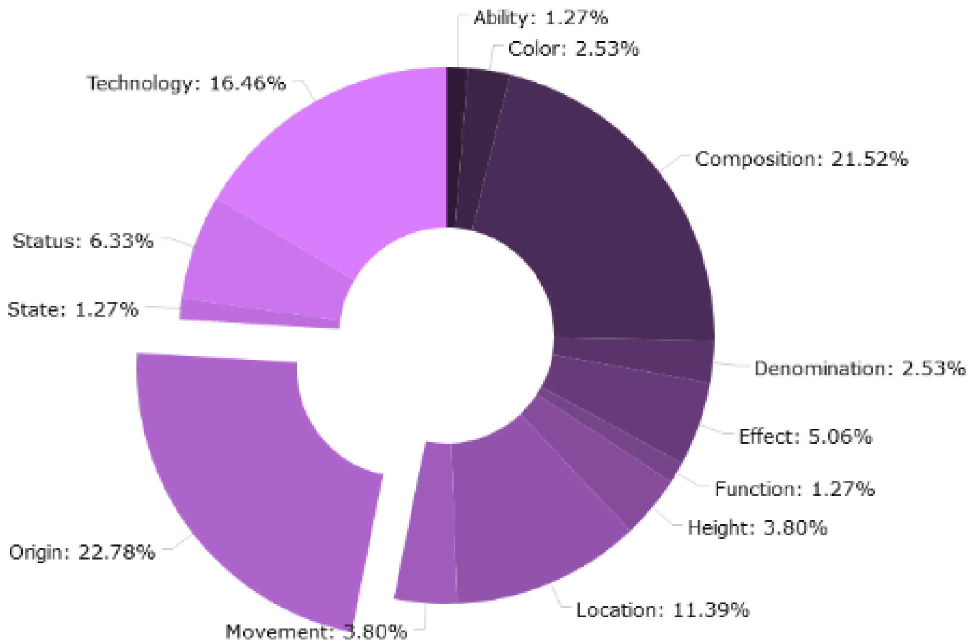


Figure 104. Statistics of the hyponymy subtypes in the CIV terminological entries

Figure 105 shows the percentages of the hyponymy subtypes in the GEO entries. In decreasing order, they are the following: *composition-based* (24.47%), *location-based* (24.47%), *origin-based* (13.83%), *denomination-based* (5.32%), *function-based* (5.32%), *height-based* (5.32%), *status-based* (5.32%), *effect-based* (4.26%), *moisture-based* (4.26%), *size-based* (4.26%), *ability-based* (1.06%), *degree-based* (1.06%), and *movement-based* (1.06%) hyponymy.

As can be observed, the three most representative hyponymy subtypes of the BIO terminological entries are *composition-based*, *location-based*, and *origin-based* hyponymy. Both *composition-based* and *location-based* hyponymy are, at the same level, equally relevant especially in conceptualizing the hyponyms of SEDIMENT and SOIL. In fact, most of these hyponyms characterize new realities thanks to the components that integrate them or to the place where they are found. In the case of EARTHQUAKE, however, composition is not a characteristic nuance because it is a process and not an entity. Instead, what is relevant is the place where an earthquake occurs. This means that *location-based* hyponymy is also important in its hierarchy. *Origin-based* hyponymy is present in the hyponyms of both EARTHQUAKE and SEDIMENT. It identifies the causes of certain types of earthquakes, and the processes that produce certain types of sediment.

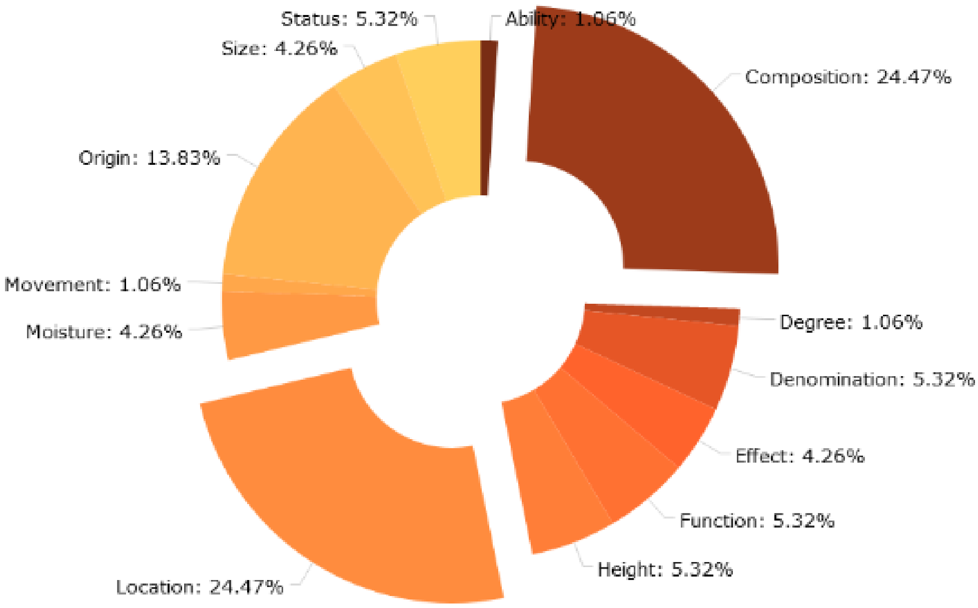


Figure 105. Statistics of the hyponymy subtypes in the GEO terminological entries

5.3.3. Comparison with EcoLexicon

Since HypoLexicon and the hyponymy-based terminological entries improve the description, categorization, and representation of hyponymy in EcoLexicon, this section compares the original information in EcoLexicon and the results obtained in this research. In fact, once the twelve general hypernyms were identified, their hierarchies in EcoLexicon were taken as a starting point.

However, it is evident that the treatment of generic-specific relations in conceptual hierarchies is much more exhaustive in the new terminological entries than in the original resource, since this research exclusively focused on that aspect. It is also true that EcoLexicon has a different purpose, in that it collects a much larger number of concepts and provides a greater variety of other linguistic information, including multilingual terms and phraseological entries.

For the sake of comparison, **Table 57** shows the statistical data for each hierarchy of the twelve hypernyms in HypoLexicon and EcoLexicon, respectively. This comparison thus focuses on terminological information (i.e., number of concepts, terms, definitions, conceptual categories, hyponymy subtypes, hyponymy levels, and hyponymic contexts). At the end, there is a comparison of the total statistics.

The first column in **Table 57** refers to HypoLexicon statistics; the second to EcoLexicon statistics; and the third to the difference between the two resources. The MAIZE hierarchy is only in HypoLexicon because it does not currently exist in

EcoLexicon. Regarding content, both hyponymy subtypes and hyponymic contexts are also innovations that are not present in EcoLexicon. Our research substantially increased the quantity of conceptual, linguistic, and hyponymic information for these twelve hierarchies: +201 concepts, +292 terms, +214 definitions, +5 conceptual categories, +19 hyponymy subtypes, +3 hyponymy levels, and +107 hyponymic contexts. As a result, the way in which hyponymy-based terminological entries in HypoLexicon deal with hyponymy is both a quantitative and qualitative improvement to terminological resources.

	HYPOLEXICON	ECOLEXICON	DIFFERENCE
BACTERIUM hierarchy			
No. of concepts	34	2	+ 32
No. of terms	49	2	+ 47
No. of definitions	34	1	+ 33
No. of conceptual categories	1	1	0
No. of hyponymy subtypes	8	0	+ 8
No. of hyponymy levels	4	1	+ 3
No. of hyponymic contexts	27	0	+ 27
REEF hierarchy			
No. of concepts	20	12	+ 8
No. of terms	23	17	+ 6
No. of definitions	20	11	+ 9
No. of conceptual categories	2	2	0
No. of hyponymy subtypes	5	0	+ 5
No. of hyponymy levels	3	3	0
No. of hyponymic contexts	3	0	+ 3
CELL hierarchy			
No. of concepts	26	4	+ 22
No. of terms	39	4	+ 35
No. of definitions	26	0	+ 26
No. of conceptual categories	4	2	+ 2
No. of hyponymy subtypes	8	0	+ 8
No. of hyponymy levels	6	1	+ 5
No. of hyponymic contexts	11	0	+ 11
SLUDGE hierarchy			
No. of concepts	20	5	+ 15
No. of terms	46	7	+ 39
No. of definitions	20	5	+ 15
No. of conceptual categories	2	2	0
No. of hyponymy subtypes	5	0	+ 5
No. of hyponymy levels	5	1	+ 4
No. of hyponymic contexts	2	0	+ 2
NITROGEN hierarchy			
No. of concepts	19	1	+ 18
No. of terms	32	2	+ 30
No. of definitions	19	1	+ 18
No. of conceptual categories	3	1	+ 2
No. of hyponymy subtypes	7	0	+ 7

No. of hyponymy levels	2	0	+ 2
No. of hyponymic contexts	4	0	+ 4
MAIZE hierarchy			
No. of concepts	11	0	+ 11
No. of terms	16	0	+ 16
No. of definitions	11	0	+ 11
No. of conceptual categories	1	0	+ 1
No. of hyponymy subtypes	6	0	+ 6
No. of hyponymy levels	2	0	+ 2
No. of hyponymic contexts	2	0	+ 2
WASTEWATER hierarchy			
No. of concepts	25	11	+ 14
No. of terms	44	28	+ 16
No. of definitions	25	11	+ 14
No. of conceptual categories	1	1	0
No. of hyponymy subtypes	5	0	+ 5
No. of hyponymy levels	4	3	+ 1
No. of hyponymic contexts	6	0	+ 6
BREAKWATER hierarchy			
No. of concepts	24	9	+ 15
No. of terms	35	13	+ 22
No. of definitions	24	9	+ 15
No. of conceptual categories	1	1	0
No. of hyponymy subtypes	5	0	+ 5
No. of hyponymy levels	5	1	+ 4
No. of hyponymic contexts	2	0	+ 2
POLLUTANT hierarchy			
No. of concepts	33	22	+ 11
No. of terms	55	43	+ 12
No. of definitions	33	15	+ 18
No. of conceptual categories	4	3	+ 1
No. of hyponymy subtypes	8	0	+ 8
No. of hyponymy levels	5	3	+ 2
No. of hyponymic contexts	24	0	+ 24
EARTHQUAKE hierarchy			
No. of concepts	15	5	+ 10
No. of terms	18	7	+ 11
No. of definitions	15	5	+ 10
No. of conceptual categories	2	1	+ 1
No. of hyponymy subtypes	6	0	+ 6
No. of hyponymy levels	3	1	+ 2
No. of hyponymic contexts	4	0	+ 4
SEDIMENT hierarchy			
No. of concepts	48	25	+ 23
No. of terms	71	36	+ 35
No. of definitions	48	25	+ 23
No. of conceptual categories	5	5	0
No. of hyponymy subtypes	8	0	+ 8
No. of hyponymy levels	4	3	+ 1
No. of hyponymic contexts	17	0	+ 17

SOIL hierarchy				
No. of concepts	34	12		+ 22
No. of terms	37	14		+ 23
No. of definitions	34	12		+ 22
No. of conceptual categories	2	2		0
No. of hyponymy subtypes	10	0		+ 10
No. of hyponymy levels	3	1		+ 2
No. of hyponymic contexts	5	0		+ 5

TOTAL				
No. of concepts	**309**	**108**		**+ 201**
No. of terms	**465**	**173**		**+ 292**
No. of definitions	**309**	**95**		**+ 214**
No. of conceptual categories	**22**	**17**		**+ 5**
No. of hyponymy subtypes	**19**	**0**		**+ 19**
No. of hyponymy levels	**6**	**3**		**+ 3**
No. of hyponymic contexts	**107**	**0**		**+ 107**

Table 57. Comparison of the statistics in HypoLexicon vs. EcoLexicon

6. CONCLUSIONS

This research focused on the description, categorization, and representation of hyponymy in Terminology. It started with a review of the general theoretical framework, composed of the disciplines related to the treatment of hyponymy from a conceptual perspective, namely, Cognitive Linguistics and Frame-based Terminology. Since Cognitive Linguistics is the branch of Linguistics that incorporates premises of Psychology and Neuroscience, it was the most suitable approach to hyponymy from a conceptual or cognitive perspective.

Because of their influence on Terminology theory, special importance was given to the following: (i) conceptualization, understood as the process by which certain entities are organized based on their salient characteristics, and which includes topics such as conceptual systems, cognition, and frames; (ii) categorization, understood as the mental process that enables humans to classify elements of the world by perceiving similarities and differences between them while mentally storing concepts in long-term memory; and (iii) conceptual relations, understood as the links between concepts, which can be hierarchical or non-hierarchical.

Frame-based Terminology, on which this research was based, is a cognitive approach to Terminology that focuses on semantic aspects of specialized knowledge representation. This theory was first placed in the context of other theoretical approaches, such as the General Terminology Theory, Socioterminology, the Communicative Terminology Theory, and the Sociocognitive Terminology Theory.

This was followed by a description of the theoretical and practical foundations of Frame-based Terminology, with special attention to the following elements: (i) conceptual organization, based on frames to structure specialized domains and create non-language-specific representations, and on events to describe the processes within a domain of expertise; (ii) multidimensionality, understood as the conceptual classification that arises when concepts are perceived in more than one way within a conceptual system based on different characteristics; (iii) frame-based definitional templates, used for describing real-world entities through a generic concept and differentiating features; and (iv) knowledge extraction, based on the premises of Corpus Linguistics, particularly on corpus analysis strategies and techniques, so as to obtain terminological information from large specialized corpora.

A description was given of EcoLexicon, the environmental terminological knowledge base that is the practical application of Frame-based Terminology. Its key features were analyzed in terms of the representation of conceptual, linguistic, multimodal, ontological, and phraseological information. It was also necessary to place EcoLexicon within the context of other terminological resources.

This section on the general theoretical framework highlighted the following:

- The aspects of Cognitive Linguistics related to conceptualization, categorization, and representation of conceptual relations are essential to understand all dimensions of hyponymy.
- The chronological review of Terminology theories provided a better understanding of the contributions made by Frame-based Terminology.
- Frame-based Terminology is the most suitable theory for the treatment of conceptual relations, specifically hyponymy, thanks to the importance given to conceptual organization and multidimensionality.
- The methodology of Frame-based Terminology provides a sound basis for the creation of definitions and knowledge extraction through corpus analysis.

This was followed by the description, categorization, and representation of hyponymy. As is well-known, hyponymy is the conceptual relation between a hypernym and a hyponym. All conceptual hierarchies are based on hyponymy and it plays the most important role in our thinking about word meaning, hence its importance from both a cognitive and terminological point of view.

Firstly, the description of hyponymy was analyzed from the perspectives of Terminology and Ontology. Hyponymy has been widely studied in different linguistic and terminological theories because of its inference-invoking nature, its importance in definitions, and its relevance to selectional restrictions in grammar. The description of hyponymy also included four related phenomena: (i) taxonomy, a variant of hyponymy which describes a classification relation between the hypernym and the hyponym; (ii) incompatibility, an exclusion relation present in certain conceptual hierarchies; (iii) troponymy, a temporally inclusive relation used to describe verb hyponymy; and (iv) autohyponymy, which occurs when a term has both a default general sense and a contextually restricted sense.

Since hyponymy is also important in Computer Science and Ontology Engineering because of its influence in network models, the generic-specific relation was also analyzed by focusing on the following terminology-related aspects: (i) ontologies, understood as databases describing the concepts of a knowledge field, their properties, and relations; (ii) termontography, a multidisciplinary approach combining theories and methods of the Sociocognitive Terminology Theory and Ontology Engineering; (iii) ontoterminology, a model that describes a terminology whose conceptual system is a formal ontology relying on epistemological principles; and (iv) the ontological knowledge enhancement in EcoLexicon in which all concepts were classified in a hierarchy of conceptual categories.

Secondly, we analyzed the categorization of hyponymy, which has different types with specific nuances, such as the following: (i) taxonomic and functional hyponymy, where hyponyms express what they are or what they are used for; and (ii) direct and indirect hyponymy, based on the conceptual proximity or distance from hypernyms to hyponyms. We also presented the case study of hyponymy refinement in EcoLexicon, which consisted of the specification of hyponymy subtypes typical of environmental terminology and the identification of hyponymic knowledge patterns.

Thirdly, the representation of hyponymy was analyzed by reviewing its visualization and expression in various terminological resources. As a conceptual relation of great importance in knowledge organization, hyponymy is difficult to represent in more traditional resources because of their format. It was thus analyzed in dictionaries, encyclopedias, term banks (i.e., IATE and TERMIUM Plus), and terminological knowledge bases (i.e., WIPO Pearl and EcoLexicon). The criteria for a more comprehensive representation of hyponymy were discussed with a view to proposing a new way of representing hyponymy based on terminological entries with a structure and content focused on the description and categorization of generic-specific relations. These criteria were the following:

- Terminological entries should be hierarchically structured in different levels of hyponymy so as to better represent transitivity from hypernyms to direct hyponyms rather than to indirect hyponyms.
- Terminological entries should contain intensional or terminological definitions of concepts, based on *genus* and *differentiae*, in order to display property inheritance from hypernyms to hyponyms and to clearly distinguish between co-hyponyms.
- Terminological entries should describe concepts according to conceptual categories so as to highlight the ontological changes from certain hypernyms to certain hyponyms.

- Hyponymy should be categorized in subtypes by identifying the nuance that differentiates a hyponym from its direct hypernym, and co-hyponyms should be classified according to the hyponymy subtype involved.
- Terminological entries should reflect hyponymic contexts to show how hyponymy is encoded in KPs and expressed in specialized language.

As a result, analyzing hyponymy from an approach combining Terminology and Ontology highlights the fact that hyponymy is the most important relation for conceptualization. In fact, since hyponymy can be classified according to different dimensions or microsenses, these nuances should be reflected in knowledge resources that represent generic-specific relations. However, as yet, there is no resource able to provide a satisfactory description, categorization, and representation of hyponymy.

The materials and methods of the research were four specialized subcorpora (BIO, CHEM, CIV, GEO) derived from the EcoLexicon English Corpus. However, a selection of specialized terminological resources, namely dictionaries and encyclopedias, was also necessary to help build the terminological definitions and improve the conceptual hierarchies. The software tools used were the following: (i) the EcoLexicon internal application, for compiling the environmental subcorpora; (ii) Sketch Engine, for processing the subcorpora and extracting all the hyponymic information; and (iii) Lexonomy, for designing the hyponymy-based terminological template.

Our research was performed in four stages: (i) extraction and compilation of the four subcorpora; (ii) analysis of the subcorpora to extract, identify, and select the hypernyms and hyponyms; (iii) creation of conceptual hierarchies (including terminological definitions, conceptual categories, hyponymy subtypes, and hyponymic contexts); and (iv) design of the terminological template using the XML-based structure. Our research methodology thus combined corpus analysis for the extraction of information through word sketch and CQL-based queries, and terminography to design of the terminological template for the hyponymy-based terminological entries.

During this research, it was necessary to have validated materials that met professional standards in order to perform high-quality terminology work. Moreover, corpus analysis techniques and strategies greatly facilitated the extraction, identification, and selection of terminological information from large corpora.

Needless to say, the combination of terminological resources and software tools of various types greatly enhanced our results. In this sense, the criteria for the elaboration of definitions and conceptual hierarchies of Frame-based Terminology were particularly useful for the representation of hyponymic information. Only in this way can conceptual hierarchies account for terminological definitions, conceptual categories, hyponymy subtypes, and hyponymic contexts.

In regard to our results, our main contribution was the creation of twelve hyponymy-based terminological entries that provided an accurate description, categorization, and representation of environmental terminology. This was combined with the design and publication of an electronic hyponymy-based terminological resource containing these entries in an accessible interface.

The structure and format of these entries facilitate the representation of their hyponymic content. This was achieved by emphasizing conceptual hierarchies with hyponymy subtypes at different levels and populating the information for each concept with synonyms, conceptual categories, terminological definitions, and hyponymic contexts.

After presenting the twelve entries, the content of each was discussed and summarized. Hyponymic nuances were found that affected the verticality and horizontality of the hierarchies. There were also dimensions or microsenses of co-hyponyms expressed in hyponymy subtypes. Also striking were the changes in characteristics or traits, produced by the addition of conceptual categories at more specific hyponymy levels.

The entries were published in HypoLexicon, the practical application of this research. This terminological resource designed as part of this thesis focuses on the description, categorization, and representation of hyponymy in environmental concepts. It includes definitional, relational, ontological, and contextual information about the specialized hypernyms and hyponyms of environmental terminology.

As such, HypoLexicon is the convergence point of four resources: (i) EcoLexicon, for the basic structure and information of the terminological entries; (ii) the EcoLexicon English Corpus and the four specialized subcorpora, for the population and enhancement of the terminological entries; (iii) Sketch Engine, for the extraction of hyponymic and contextual information through corpus analysis; and (iv) Lexonomy, for the design of the terminological template and for the implementation of all data in the form of an actual terminological resource.

Finally, a summary was given of the statistical results for the information in all hyponymy-based terminological entries regarding conceptual categories and hyponymy subtypes. Our data was also compared with the original terminological data in EcoLexicon so as to reveal the improvements implemented in HypoLexicon. The results showed that there was a significant increase in the conceptual, linguistic, and hyponymy-related content of the conceptual hierarchies related to the twelve hyponymy-based terminological entries.

Accordingly, the following conclusions can be derived from this research:

- The hyponymy-based terminological entries proved to be a successful approach to the description, categorization, and representation of hyponymy because of their hierarchical structure and graphical classification of information based on definitional and corpus analysis.
- The visualization of hyponymic information in the hyponymy-based terminological entries permitted the identification of dynamic phenomena regarding generic-specific relations (e.g., hyponymic nuances in the verticality and horizontality of the conceptual hierarchies, different dimensions or microsenses of co-hyponyms, changes in characteristics of concepts through the addition of conceptual categories at more specific hyponymy levels, etc.).
- HypoLexicon, the hyponymy-based terminological resource, has an excellent structure and interface to showcase all the information of the hyponymy-based terminological entries.
- HypoLexicon, which is freely available online, facilitates acquiring and sharing specialized knowledge among all types of users (e.g., terminologists, linguists, translators, technical experts, lay users).

This research has led to a new type of description, categorization, and representation of hyponymy. This methodology is also applicable to any other specialized domain, and may even provide an accessible way of dealing with hyponymy in general language resources as well. Basically, the objective of the methodology and resource proposed is to facilitate knowledge acquisition at all level.

Future work in this research line will take two paths. On the one hand, HypoLexicon can continue to grow and be nourished with more content by creating additional terminological entries with all kinds of hyponymic information extracted from corpus techniques. These new entries, moreover, could belong to the same environmental subdomains or to new ones so as to extend the range of conceptual categories and hyponymy subtypes.

However, perhaps the most innovative idea would be to seek the total integration of HypoLexicon in EcoLexicon. In this way, it would cease to be a stand-alone module or a by-product, and would become an integral part of the original resource. This thesis has thus laid a foundation that can be continued in further research. It has not only opened the door to future work, but also fosters interoperability with other information and with other ontological, linguistic, and terminological tools and resources.

REFERENCES

Allaby, M. (ed.) (2013). *A Dictionary of Geology and Earth Sciences*. 4th ed. Oxford: Oxford University Press.

Andreasen, T., Jensen, P. A., Nilsson, J. F., Paggio, P., Pedersen, B. S., & Thomsen, H. E. (2004). Content-Based Text Querying with Ontological Descriptors. *Data & Knowledge Engineering*, 48, 199–219.

Antia, B. E. (2000). *Terminology and Language Planning: An Alternative Framework of Practice and Discourse*. Amsterdam/Philadelphia: John Benjamins.

Atkins, S., Clear, J., & Ostler, N. (1991). Corpus Design Criteria. *Literary & Linguistic Computing*, 7(1), 1–16.

Atkins, T., & Escudier, M. (2019). *A Dictionary of Mechanical Engineering*. 2nd ed. Oxford: Oxford University Press.

Auger, A., & Barrière, C. (2008). Pattern-based Approaches to Semantic Relation Extraction: A State-of-the-art. *Terminology*, 14(1), 1–19. Amsterdam/Philadelphia: John Benjamins.

Babaie, H. A. (2011). Ontological relations and spatial reasoning in earth science ontologies. In A. K. Sinha, D. Arctur, I. Jackson, & L. Gundersen (eds.), *Societal Challenges and Geoinformatics: Geological Society of America*, Special Paper 482, 13–27.

Baisa, V., & Suchomel, V. (2015). Corpus Based Extraction of Hypernyms in Terminological Thesaurus for Land Surveying Domain. In *9th Workshop on Recent Advances in Slavonic Natural Language Processing*, 69–74. Brno: Tribun EU.

Barbero, C. (2022). CQL Grammars for Lexical and Semantic Information Extraction for Portuguese and Italian. In V. Pinheiro, P. Gamallo, R. Amaro, C. Scarton, F. Batista, D. Silva, C. Magro, & H. Pinto (eds.), *Computational Processing of the Portuguese Language, Proceedings of the 15th International Conference PROPOR 2022*, 376–386. Fortaleza, Brazil.

Barrière, C. (2004a). Building a Conceptual Hierarchy from Corpus Analysis. *Terminology*, 10(2), 241–263. Amsterdam/Philadelphia: John Benjamins.

Barrière, C. (2004b). Knowledge-Rich Contexts Discovery. In *17th Canadian Conference on Artificial Intelligence (AI'2004)*, 3060, 187–201.

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, 18(5–6), 513–562.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B*, 1281–1289.

Barsalou, L. W., Yeh, W., Luka, B. J., & Olseth, K. L. (1993). Concepts and Meaning. In *Chicago Linguistics Society 9: Papers from the Parasessions on Conceptual Representations*, 2, 23–61. Chicago: Chicago Linguistic Society.

Bergenholtz, H., & Tarp. S. (2010). LSP Lexicography or Terminography? The Lexicographer's Point of View. In P. A. Fuertes-Olivera (ed.), *Specialized Dictionaries for Learners*, 27–37. Berlin/New York: Mouton de Gruyter.

Berland, M., & Charniak, E. (1999). Finding Parts in Very Large Corpora. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, 57–64. Morristown (NJ): Association for Computational Linguistics.

Berlin, B., & Kay, P. (1969). *Basic Color Terms: their Universality and Evolution*. Berkeley: University of California at Berkeley.

Bernier-Colborne, G., Barrière, C., & Ménard, P. A. (2017). Fine-grained domain classification of text using TERMIUM Plus. In *Proceedings of Language, Ontology, Terminology and Knowledge Structures Workshop (LOTKS 2017)*, Montpellier, France: Association for Computational Linguistics.

Bielinskiene, A., Boizou, L., Kovalevskaite, J., & Utka, A. (2012). Towards the Automatic Extraction of Term-defining Contexts in Lithuanian. In A. Tavast,

K. Muischnek, & M. Koit (eds.), *Human Language Technologies: The Baltic Perspective*, 18–26. Amsterdam/Berlin/Tokyo/Washington DC: IOS Press.

Bobrowsky, P. T., & Marker, B. (eds.) (2018). *Encyclopedia of Engineering Geology*. New York: Springer.

Boulanger, J. C. (1991). Une lecture socioculturelle de la terminologie. *Cahiers de Linguistique Sociale*, 18, 13–30.

Boulanger, J. C. (1995). Présentation: images et parcours de la socioterminologie. *Meta*, 40(2), 194–205. Montreal: Les Presses de l'Université de Montréal.

Bourigault, D., & Slodzian, M. (1999). Pour une terminologie textuelle. *Terminologies Nouvelles*, 19, 29–32.

Bowker, L. (1997). Multidimensional classification of concepts and terms. In S. E. Wright & G. Budin (eds.), *Handbook of Terminology Management*, 131–143. Amsterdam/Philadelphia: John Benjamins.

Bowker, L. (2003). Lexical Knowledge Patterns, Semantic Relations, and Language Varieties: Exploring the Possibilities for Refining Information Retrieval in an International Context. *Cataloging & Classification Quarterly*, 37(1–2), 153–171.

Bowker, L. (2022). Multidimensionality. In P. Faber & M. C. L'Homme (eds.), *Theoretical Perspectives on Terminology: Explaining terms, concepts and specialized knowledge*, *Terminology and Lexicography Research and Practice*, 23, 127–147. Amsterdam/Philadelphia: John Benjamins.

Bowker, L., & Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. London: Routledge.

Bradshaw, R. A., & Stahl, P. D. (eds.) (2016). *Encyclopedia of Cell Biology*. Amsterdam: Elsevier.

Buttigieg, P. L., Morrison, N., Smith, B., Mungall, C. J., & Lewis, S. E. (2013). The Environment Ontology: Contextualising Biological and Biomedical Entities. *Journal of Biomedical Semantics*, 4(43).

Buttigieg, P. L., Pafilis, E., Lewis, S. E., Schildhauer, M. P., Walls, R. L., & Mungall, C. J. (2016). The Environment Ontology in 2016: Bridging Domains with Increased Scope, Semantic Density, and Interoperation. *Journal of Biomedical Semantics*, 7(57).

Cabezas-García, M. (2019). *Los compuestos nominales en terminología: formación, traducción y representación*. Doctoral Thesis. Granada: Universidad de Granada.

Cabezas-García, M. (2020). *Los términos compuestos desde la Terminología y la Traducción*. Berlin: Peter Lang.

Cabré, M. T. (1993). *La terminología: Teoría, metodología, aplicaciones*. Barcelona: Antártida/Empúries.

Cabré, M. T. (1999). *La terminología: Representación y comunicación*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.

Cabré, M. T. (2000). Elements for a Theory of Terminology: Towards an Alternative Paradigm. *Terminology*, 6(1), 35–57. Amsterdam/Philadelphia: John Benjamins.

Cabré, M. T. (2001). Terminología y lenguas minoritarias. In M. X. Bugarín López, X. Cajide Val, A. Dosil Maceira, G. Ferreiro Fente, M. González González, & M. A. Santos Rego (eds.), *Actas da VIII Conferencia Internacional de Linguas Minoritarias*, 89–102. Santiago de Compostela, Spain.

Cabré, M. T. (2003). Theories of terminology: Their description, prescription and explanation. *Terminology*, 9(2), 163–199. Amsterdam/Philadelphia: John Benjamins.

Cabré, M. T., & Feliu, J. (2001). *La terminología científico-técnica: reconocimiento, análisis y extracción de información formal y semántica*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.

Cabré, M. T., Estopà, R., & Lorente Casafont, M. (1996). Terminología y Fraseología. In *Actas del V Simposio de Terminología Iberoamericana*, 67–81. Mexico City: Colegio de México.

Caffrey, C., & Valentini, C. (2019). Applications of technology in the Patent Cooperation Treaty (PCT) Translation Division of the World Intellectual Property Organization (WIPO). In M. O'Hagan (ed.), *The Routledge Handbook of Translation and Technology*, 127–147. London: Routledge.

Condamines, A. (2002). Corpus Analysis and Conceptual Relation Patterns. *Terminology*, 8(1), 141–162. Amsterdam/Philadelphia: John Benjamins.

Creighton, T. E. (ed.) (1999). *Encyclopedia of Molecular Biology*. New York: Wiley.

Croft, W., & Cruse, D. A. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Cruse, D. A. (1995). Polysemy and Related Phenomena from a Cognitive Linguistic Viewpoint. In P. Saint-Dizier & E. Viegas (eds.), *Computational Lexical Semantics*. Cambridge: Cambridge University Press.

Cruse, D. A. (2000). *Meaning in Language*. 1st ed. Oxford: Oxford University Press.

Cruse, D. A. (2002). Hyponymy and Its Varieties. In R. Green, C. A. Bean, & S. H. Myaeng (eds.), *The Semantics of Relationships: An Interdisciplinary Perspective*, 3–22. Dordrecht/Boston/London: Kluwer Academic Publishers.

Cruse, D. A. (2006). *A Glossary of Semantics and Pragmatics*. Edinburgh: Edinburgh University Press.

Daintith, J. (ed.) (2008). *A Dictionary of Chemistry*. 6th ed. Oxford: Oxford University Press.

Damasio, A., & Damasio, H. (1994). Cortical systems for retrieval of concrete knowledge: the convergence zone framework. In C. Koch & J. L. Davis (eds.), *Largescale Neuronal Theories of the Brain*, 61–70. Cambridge (MA)/London: MIT Press.

Dubuc, R., & Lauriston, A. (1997). Terms and Contexts. In S. E. Wright & G. Budin (eds.), *Handbook of Terminology Management*, 1, 80–88. Amsterdam/ Philadelphia: John Benjamins.

Durán-Muñoz, I. (2016). Producing frame-based definitions: A case study. *Terminology*, 22(2), 223–249. Amsterdam/Philadelphia: John Benjamins.

Eck, K., & Meyer, I. (1995). Bringing Aristotle into the 20th Century: Computer-Aided Definition Construction in a Terminological Knowledge Base. In S. E. Wright & R. A. Strehlow (eds.), *Standardizing and Harmonizing Terminology: Theory and Practice*, 83–100. West Conshohocken (PA): ASTM International.

Eckard, E., Barque, L., Nasr, A., & Sagot, B. (2012). Dictionary-Ontology Cross-Enrichment Using TLFi and WOLF to enrich one another. In *Proceedings of the 3rd Workshop on Cognitive Aspects of the Lexicon (CogALex-III), COLING 2012*, 81–94. Mumbai, India.

Edo Marzá, N. (2012). Lexicografía especializada y lenguajes de especialidad: fundamentos teóricos y metodológicos para la elaboración de diccionarios especializados / Specialised Lexicography and Specialised Languages: Theoretical and Methodological Foundations for the Elaboration of Specialised Dictionaries. *Lingüística*, 27, 98–114.

Evans, V., & Green. M. (2006). *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Evans, V., Bergen, B. K., & Zinken, J. (2007). The cognitive linguistics enterprise: an overview. In V. Evans *et al*. (eds.), *The Cognitive Linguistics Reader*, 2–36. London: Equinox.

Faber, P. (2009). The Cognitive Shift in Terminology and Specialized Translation. *MonTI (Monografías de Traducción e Interpretación)*, 1, 107–134. Valencia: Universitat de València.

Faber, P. (2011). The Dynamics of Specialized Knowledge Representation: Simulational Reconstruction or the Perception–Action Interface. *Terminology*, 17(1), 9–29. Amsterdam/Philadelphia: John Benjamins.

Faber, P. (2015). Frames as a Framework for Terminology. In H. J. Kockaert & F. Steurs (eds.), *Handbook of Terminology*, 1, 14–33. Amsterdam/Philadelphia: John Benjamins.

Faber, P. (2022). Frame-based Terminology. In P. Faber & M. C. L'Homme (eds.), *Theoretical Perspectives on Terminology: Explaining terms, concepts and specialized knowledge*, *Terminology and Lexicography Research and Practice*, 23, 353–376. Amsterdam/Philadelphia: John Benjamins.

Faber, P. (ed.) (2012). *A Cognitive Linguistics View of Terminology and Specialized Language*. Berlin/Boston: De Gruyter Mouton.

Faber, P., & León-Araúz, P. (2010). Dinamismo conceptual en las bases de conocimiento terminológico: el caso de EcoLexicon. *Íkala, Revista de lenguaje y cultura*, 15(25), 75–100.

Faber, P., & León-Araúz, P. (2014). Specialized knowledge dynamics: From cognition to culture-bound terminology. In R. Temmerman & M. Van Campenhoudt (eds.), *Dynamics and Terminology: An Interdisciplinary Perspective on Monolingual and Multilingual Culture-bound Communication*, *Terminology and Lexicography Research and Practice,* 16, 135–158. Amsterdam/Philadelphia: John Benjamins.

Faber, P., & León-Araúz, P. (2021). Designing Terminology Resources for Environmental Translation. In C. Meng Ji & S. Laviosa (eds.), *The Oxford Handbook of Translation and Social Practices*, 587–616. Oxford: Oxford University Press.

Faber, P., & López Rodríguez, C. I. (2012). Terminology and specialized language. In P. Faber (ed.), *A Cognitive Linguistics View of Terminology and Specialized Language*, 9–32. Berlin/Boston: De Gruyter Mouton.

Faber, P., León-Araúz, P., & Prieto-Velasco, J. A. (2009). Semantic Relations, Dynamicity, and Terminological Knowledge Bases. *Current Issues in Language Studies*, 1(1), 1–23.

Faber, P., León-Araúz, P., & Reimerink, A. (2014a). Representing Environmental Knowledge in EcoLexicon. In E. Bárcena, T. Read, & J. Arús (eds.), *Languages for Specific Purposes in the Digital Era*, 19, 267–301. Berlin/Heidelberg: Springer.

Faber, P., León-Araúz, P., & Reimerink, A. (2016). EcoLexicon: New Features and Challenges. In I. Kernerman, I. Kosem Trojina, S. Krek, & L. Trap-Jensen (eds.), *GLOBALEX 2016: Lexicographic Resources for Human Language Technology in Conjunction with the 10th Edition of the Language Resources and Evaluation Conference*, 73–80. Portorož, Slovenia.

Faber, P., López Rodríguez, C. I., & Tercedor Sánchez, M. (2001). La utilización de técnicas de corpus en la representación del conocimiento médico. *Terminology*, 7(2), 167–197. Amsterdam/Philadelphia: John Benjamins.

Faber, P., Márquez Linares, C., & Vega Expósito, M. (2005). Framing Terminology: A process-oriented approach. *Meta*, 50(4). Montreal: Les Presses de l'Université de Montréal.

Faber, P., Montero Martínez, S., Castro Prieto, M. R., Senso Ruiz, J., Prieto Velasco, J. A., León-Araúz, P., Márquez Linares, C., & Vega Expósito, M. (2006). Process-oriented terminology management in the domain of Coastal Engineering. *Terminology*, 12(2), 189–213. Amsterdam/Philadelphia: John Benjamins.

Faber, P., Verdejo-Román, J., León-Araúz, P., Reimerink, A., & Guzmán Pérez-Carrillo, G. (2014b). Neural Substrates of Specialized Knowledge Representation: An fMRI study. *Revue française de linguistique appliquée*, 19(1), 15–32.

Faber, P., Verdejo-Román, J., León-Araúz, P., Reimerink, A., & Guzmán Pérez-Carrillo, G. (2017). Specialized knowledge processing in the brain: an fMRI study. In P. Faini (ed.), *Terminological Approaches in the European Context*, 168–182. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Faralli, S., Lefever, E., & Ponzetto, S. P. (2018). MISA: Multilingual "ISA" Extraction from Corpora. In N. Calzolari *et al*. (eds.), *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, 2040–2044. Miyazaki, Japan: European Language Resources Association (ELRA).

Felber, H. (1984). *Terminology Manual*. Paris: UNESCO (Infoterm).

Fellbaum, C. (1998). *WordNet: An electronic lexical database*. Cambridge (MA)/London: MIT Press.

Fellbaum, C. (2002). On the semantics of troponymy. In R. Green, C. A. Bean, & S. H. Myaeng (eds.), *The Semantics of Relationships: An Interdisciplinary Perspective: Information Science and Knowledge Management Series*, 22–34. Dordrecht: Kluwer Academic Publishers.

Fernández, M., Gómez-Pérez, A., & Jurista, N. (1997). METHONTOLOGY: From ontological art towards ontological engineering. In *Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering*, 33–40. Stanford (CA): Stanford University.

Fillmore, C. J. (1968). The case for case. In E. Bach & R. T. Harms (eds.), *Universals in Linguistic Theory*, 1–88. New York: Holt Rinehart and Winston.

Fillmore, C. J. (1975). An alternative to checklist theories of meaning. In *Papers from the First Annual Meeting of the Berkeley Linguistics Society*, 123–132. Berkeley (CA): University of California at Berkeley.

Fillmore, C. J. (1977). Scenes-and-frames semantics. *Linguistic Structures Processing*, 5, 55–81.

Fillmore, C. J. (1982). Frame Semantics. In Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*, 111–137. Seoul: Hanshin.

Fillmore, C. J. (1985). Frames and the semantics of understanding. *Quaderni di Semantica*, 6(2), 222–254.

Fillmore, C. J., & Atkins, S. (1992). Towards a frame-based organization of the lexicon: The semantics of risk and its neighbors. In A. Lehrer & E. F. Kittay (eds.), *Frames, Fields, and Contrast: New Essays in Semantics and Lexical Organization*, 75–102. Hillsdale (NJ): Lawrence Erlbaum.

Fillmore, C. J., & Atkins, S. (1998). FrameNet and lexicographic relevance. In A. Rubio, N. Gallardo, R. Castro, & A. Tejada (eds.), *Proceedings of the ELRA Conference on Linguistic Resources*, 417–423. Granada: Universidad de Granada.

Fillmore, C. J., & Kay, P. (1987). The goals of Construction Grammar. *Berkeley Cognitive Science Report*, 50. Berkeley: University of California at Berkeley.

Fillmore, C. J., Petruck, M., Ruppenhofer, J., & Wright, A. (2003). FrameNet in action: the case of attaching. *International Journal of Lexicography*, 16(3), 298–332.

Gambier, Y. (1987). Problèmes terminologiques des pluies acides: pour une socio-terminologie. *Meta*, 32(2), 314–320. Montreal: Les Presses de l'Université de Montréal.

Gambier, Y. (1991). Présupposés de la terminologie: vers une remise en cause. *Cahiers de linguistique sociale*, 18, 31–58.

García Aragón, A., Buendía Castro, M., & López Rodríguez, C. I. (2014). Evaluación de una base de conocimiento terminológica sobre el medio ambiente en el aula de la traducción especializada. In C. Vargas Sierra (ed.), *TIC, trabajo colaborativo e interacción en Terminología y Traducción*, 477–487. Granada: Comares.

Gaudin, F. (1993). *Pour une socioterminologie: des problèmes sémantiques aux pratiques institutionnelles*. Rouen: Publications de l'Université de Rouen.

Gaudin, F. (2003). *Socioterminologie: une approche sociolinguistique de la terminologie*. Brussels: Duculot.

Gaudin, F. (2005). Socioterminologie. *Langages*, 157, 80–92.

Giacomini, L. (2014). Testing User Interaction with LSP e-Lexicographic Tools: A Case Study on Active Translation of Environmental Terms. In *Proceedings of the 12th edition of the Konvens Conference*, 77–85. Hildesheim, Germany.

Giacomini, L. (2015). Management and exploitation of conceptual data and information in technical termbases: the electrotechnical vocabulary. In I. Kosem, M. Jakubiček, J. Kallas, & S. Krek (eds.), *Proceedings of the eLex 2015 Conference – Electronic Lexicography in the 21st Century: Linking Lexical Data in the Digital Age*, 186–197. Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd.

Gibbs, R. W. (1996). What's cognitive about cognitive linguistics. In E. H. Casad (ed.), *Cognitive Linguistics in the Redwoods*, 27–53. Berlin/New York: Mouton de Gruyter.

Gibbs, R. W. (2003). Embodied experience and linguistic meaning. *Brain and Language*, 84(1), 1–15.

Gil-Berrozpe, J. C. (2016). *Extending the Conceptual Systems in EcoLexicon to Enhance Multidimensionality*. BA Thesis. Granada: Universidad de Granada.

Gil-Berrozpe, J. C. (2017). *Corpus-based Identification of Hyponymy Subtypes and Knowledge Patterns in the Environmental Domain*. MA Thesis. Granada: Universidad de Granada.

Gil-Berrozpe, J. C. (2020). Attribute-based Approach to Hyponymic Behavior in Botanical Terminology. In C. Roche (ed.), *TOTh 2019 – Terminology & Ontology: Theories and Applications*, *Terminologica*, 93–108. Chambéry: Éditions de l'Université de Savoie Mont Blanc.

Gil-Berrozpe, J. C., & Faber, P. (2016). Refining Hyponymy in a Terminological Knowledge Base. In F. Khan, S. Vintar, P. León-Araúz, P. Faber, F. Frontini, A. Pavizi, L. Grčić-Simeunović, & C. Unger (eds.), *Proceedings of the 2nd Joint Workshop on Language and Ontology (LangOnto2) & Terminology and Knowledge*

Structures (TermiKS) at the 10th Edition of the Language Resources and Evaluation Conference (LREC 2016), 8–15. Portorož, Slovenia.

Gil-Berrozpe, J. C., & Faber, P. (2017). The Role of Terminological Knowledge Bases in Specialized Translation: The Use of Umbrella Concepts. In M. A. Candel-Mora & C. Vargas-Sierra (eds.), *Temas actuales en terminología y estudios sobre el léxico*, *Interlingua*, 172, 1–25. Granada: Comares.

Gil-Berrozpe, J. C., León-Araúz, P., & Faber, P. (2017). Specifying Hyponymy Subtypes and Knowledge Patterns: A Corpus-based Study. In I. Kosem, C. Tiberius, M. Jakubíček, J. Kallas, S. Krek, & V. Baisa (eds.), *Proceedings of the eLex 2017 Conference – 5th International Conference on Electronic Lexicography in the 21st Century*, 63–92. Brno: Lexical Computing CZ s.r.o.

Gil-Berrozpe, J. C., León-Araúz, P., & Faber, P. (2018). Subtypes of Hyponymy in the Environmental Domain: Entities and Processes. In C. Roche (ed.), *TOTh 2016 – Terminology & Ontology: Theories and Applications*, *Terminologica*, 39–54. Chambéry: Éditions de l'Université de Savoie Mont Blanc.

Gil-Berrozpe, J. C., León-Araúz, P., & Faber, P. (2019). Ontological Knowledge Enhancement in EcoLexicon. In I. Kosem, T. Zingano Kuhn, M. Correia, J.P. Ferreria, M. Jansen, I. Pereira, J. Kallas, M. Jakubíček, S. Krek, & C. Tiberius (eds.), *Proceedings of the eLex 2019 conference – 6th International Conference on Electronic Lexicography in the 21st century*, 177–197. Brno: Lexical Computing CZ, s.r.o.

Girju, R., Badulescu, A., & Moldovan, D. (2003). Learning Semantic Constraints for the Automatic Discovery of Part-Whole Relations. In *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 1–8. Edmonton, Canada.

Goddard, C., & Schalley, A. C. (2010). Semantic Analysis. In N. Indurkhya & F. J. Damerau (eds.), *Handbook of Natural Language Processing*, 2nd ed., 93–120. Boca Raton/London/New York: CRC Press.

Gödert, W., Hubrich, J., & Nagelschmidt, M. (2014). *Semantic Knowledge Representation for Information Retrieval*. Berlin/Boston: De Gruyter Saur.

Goldberg, A. (1995). *A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago.

Goldberg, A. (2005). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2003). *Ontological Engineering: with Examples from the Areas of Knowledge Management, e-Commerce and the Semantic Web*. London/Berlin/Heidelberg: Springer-Verlag.

Gorse, C., Johnston, D., & Pritchard, M. (2012). *A Dictionary of Construction, Surveying and Civil Engineering*. Oxford: Oxford University Press.

Grinev, S., & Klepalchenko, I. (1999). Terminological approach to knowledge representation. In P. Sandrini (ed.), *Proceedings of the 5th International Conference on Terminology and Knowledge Engineering (TKE '99)*, 147–151. Vienna: TermNet.

Guarino, N. (1998). Formal Ontology and Information Systems. *Formal Ontology in Information Systems (FOIS'98)*, 3–15. Trento: IOS Press.

Hahn, U., & Chater, N. (1997). Concepts and Similarity. In K. Lamberts & D. Shanks (eds.), *Knowledge, Concepts, and Categories*, 43–92. Cambridge (MA)/London: MIT Press.

Hampton, J. (1991). The combination of prototype concept. In P. Schwanenflugel (ed.), *The Psychology of Word Meanings*, 91–116. Hillsdale (NJ): Lawrence Erlbaum.

Hearst, M. A. (1992). Automatic Acquisition of Hyponyms from Large Text Corpora. In *Proceedings of COLING-92*, 539–545. Morristown (NJ): International Committee on Computational Linguistics.

Hearst, M. A. (1998). Automated Discovery of Word-Net Relations. In C. Fellbaum (ed.), *WordNet: An Electronic Lexical Database*, 132–152. Cambridge (MA)/London: MIT Press.

Hine, R. (ed.) (2019). *A Dictionary of Biology*. 8th ed. Oxford: Oxford University Press.

Horn, L. R. (1984). Ambiguity, Negation, and the London School of Parsimony. *North East Linguistics Society*, 14(1).

ISO (2000). ISO Standard 704:2000. *Terminology Work: Principles and Methods*. International Organization for Standardization. Geneva, Switzerland.

ISO (2009). ISO Standard 704:2009. *Terminology Work: Principles and Methods*. International Organization for Standardization. Geneva, Switzerland.

ISO (2019). ISO Standard 1087:2019. *Terminology Work and Terminology Science*. International Organization for Standardization. Geneva, Switzerland.

Jackendoff, R. (1983). *Semantics and Cognition*. Cambridge (MA)/London: MIT Press.

Jackendoff, R. (1990). *Semantic Structures*. Cambridge (MA)/London: MIT Press.

Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge (MA)/London: MIT Press.

Johnson, I., & MacPhail, A. (2000). IATE – Inter-Agency Terminology Exchange: Development of a Single Central Terminology Database for the Institutions and Agencies of the European Union. In *Proceedings of the Workshop on Terminology Resources and Computation (LREC 2000 Conference)*. Athens, Greece.

Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.

Jouis, C. (2006). Hierarchical Relationships "is-a": Distinguishing Belonging, Inclusion and Part/of Relationships. In N. Calzolari *et al*. (eds.), *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC*

2006), 571–574. Genoa, Italy: European Language Resources Association (ELRA).

Kageura, K. (1997). Multifaceted/Multidimensional concept systems. In S.E. Wright & G. Budin (eds.), *Handbook of Terminology Management: Basic Aspects of Terminology Management*, 119–132. Amsterdam/Philadelphia: John Benjamins.

Kageura, K. (2002). *The Dynamics of Terminology*. Amsterdam/Philadelphia: John Benjamins.

Kay, P., & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: The *what's X doing Y?* construction. *Language*, 75(1), 1–33.

Keizer, E. (2007). *The English Noun Phrase: The Nature of Linguistic Categorization*. Cambridge: Cambridge University Press.

Kerremans, K., Temmerman, R., & Tummers, J. (2004). Discussion on the requirements for a workbench supporting Termontography. In G. Williams & S. Vessier (eds.), *Proceedings of the 11th EURALEX International Congress,* 559–570. Lorient: Lorient Université de Bretagne-Sud.

Kerremans, K., Temmerman, R., & Zhao, G. (2005). Terminology and knowledge engineering in fraud detection. In B. N. Madsen & H. E. Thomsen (eds.), *Terminology and Content Development*, 101–112. Copenhagen: Litera.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1, 7–36.

Kilgarriff, A., Rychlý, P., Smrz, P., & Tugwell, D. (2004). The Sketch Engine. In G. Williams & S. Vessier (eds.), *Proceedings of the 11th EURALEX International Congress*, 105–116. Lorient: Lorient Université de Bretagne-Sud.

Kittay, E. F., & Lehrer, A. (1992). Introduction. In A. Lehrer & E. F. Kittay (eds.), *Frames, Fields, and Contrasts: New Essays in Semantics and Lexical Organization*, 1–18. Hillsdale (NJ): Lawrence Erlbaum.

Kleineberg, M. (2017). Integrative Levels. In B. Hjørland (ed.), *Reviews of Concepts in Knowledge Organization*, 44(5), 349–379.

Krishna, A., & Schwarz, N. (2014). Sensory Marketing, Embodiment, and Grounded Cognition: A Review and Introduction. *Journal of Consumer Psychology*, 24(2), 159–168.

Kurtz, J. P. (2004). *Dictionary of Civil Engineering*. New York: Kluwer Academic Publishers.

L'Homme, M. C. (2020). *Lexical Semantics for Terminology*. Amsterdam/Philadelphia: John Benjamins.

L'Homme, M. C., & Marshman, E. (2006). Terminological Relationships and Corpus-based Methods for Discovering them: An Assessment for Terminographers. In L. Bowker (ed.), *Lexicography, Terminology and Translation – Text-based Studies in Honour of Ingrid Meyer*, 67–80. Ottawa: University of Ottawa Press.

L'Homme, M. C., Heid, U., & Sager, J. C. (2003). Terminology during the past decade (1994–2004). *Terminology*, 9(2), 151–161. Amsterdam/Philadelphia: John Benjamins.

Labov, W. (1973). The boundaries of words and their meanings. In C. J. N. Bayley & R. W. Shuy (eds.), *New Ways of Analysing Variation in English*, 340–373. Washington (DC): Georgetown University Press.

Lackie, J. M. (ed.) (2013). *The Dictionary of Cell and Molecular Biology*. 5th ed. Amsterdam: Elsevier.

Lafourcade, M., & Ramadier, L. (2016). Semantic Relation Extraction with Semantic Patterns Experiment on Radiology Reports. In N. Calzolari *et al*. (eds.), *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2016)*, 4578–4582. Portorož, Slovenia: European Language Resources Association (ELRA).

Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.

Lakoff, G. (1990). The Invariance Hypothesis: is abstract reason based on image-schemas? *Cognitive Linguistics*, 1(1), 39–74.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (ed.), *Metaphor and Thought*, 202–251. Cambridge: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

LaMoreaux, J. W. (ed.) (2019). *Environmental Geology*. 2nd ed. New York: Springer.

Landry, A. (1988). The Termium termbank: today and tomorrow. In C. Picken (ed.), *Translating and the Computer*, 9, 130–144.

Langacker, R. (1987). *Foundations of Cognitive Grammar, vol. I*. Stanford: Stanford University Press.

Langacker, R. (1991). *Foundations of Cognitive Grammar, vol. II*. Stanford: Stanford University Press.

Lawrence, E. (ed.) (2008). *Henderson's Dictionary of Biology*. 14th ed. Harlow: Pearson.

Leake, D., Maguitman, A., & Reichherzer, T. (2004). Understanding knowledge models: modeling assessment of concept importance in concept maps. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society (CogSci)*, 795–800. Mahwah (NJ): Lawrence Erlbaum.

Lee, D. (2001). *Cognitive Linguistics: An Introduction*. Oxford: Oxford University Press.

Lefeuvre, L., Coustot, K., Condamines, A., & Rebeyrolle, J. (2017). *MAR-REL: Liste de candidats-marqueurs français pour les relations d'hypernymie, de méronymie et de cause*. Toulouse: Laboratoire Cognition, Langues, Langage, Ergonomie (CLLE).

Lefever, E., Van de Kauter, M., & Hoste, V. (2014). HypoTerm: Detection of hypernym relations between domain-specific terms in Dutch and English. *Terminology*, 20(2), 250–278. Amsterdam/Philadelphia: John Benjamins.

Lennarz, W. J., & Lane, M. D. (2013). *Encyclopedia of Biological Chemistry*. 2nd ed. Amsterdam: Elsevier.

León-Araúz, P. (2009). *Representación multidimensional del conocimiento especializado: el uso de marcos desde la macroestructura hasta la microestructura*. Doctoral Thesis. Granada: Universidad de Granada.

León-Araúz, P., & Faber, P. (2010). Natural and Contextual Constraints for Domain-Specific Relations. In V. B. Mititelu, V. Pekar, & E. Barbu (eds.), *Proceedings of the Workshop Semantic Relations, Theory and Applications*, 12–17. Valletta, Malta.

León-Araúz, P., & Reimerink, A. (2010). Knowledge extraction and multidimensionality in the environmental domain. In Ú. Bhreathnach & F. de Barra Cusack (eds.), *TKE 2010 Proceedings - Presenting Terminology and Knowledge Engineering Resources Online: Models and Challenges*. Dublin: Dublin City University.

León-Araúz, P., & Reimerink, A. (2018). Evaluating EcoLexiCAT: A terminology-enhanced CAT tool. In N. Calzolari *et al*. (eds.), *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, 2374–2381. Miyazaki: European Language Resources Association (ELRA).

León-Araúz, P., & San Martín, A. (2018). The EcoLexicon Semantic Sketch Grammar: from Knowledge Patterns to Word Sketches. In I. Kerneman & S. Krek (eds.), *Proceedings of the LREC 2018 Workshop Globalex 2018 – Lexicography & WordNets*, 94–99. Miyazaki: Globalex.

León-Araúz, P., Faber, P., & Montero Martínez, S. (2012). Specialized language semantics. In P. Faber (ed.), *A Cognitive Linguistics View of Terminology and Specialized Language*, 95–175. Berlin/Boston: De Gruyter Mouton.

León-Araúz, P., Reimerink, A., & Faber, P. (2013a). Multidimensional and Multimodal Information in EcoLexicon. In A. Przepiórkowski, M. Piasecki,

K. Jassem, & P. Fuglewicz (eds.), *Computational Linguistics*, 143–161. Berlin: Springer.

León-Araúz, P., Reimerink, A., & Faber, P. (2017). EcoLexiCAT: a Terminology-enhanced Translation Tool for Texts on the Environment. In I. Kosem, C. Tiberius, M. Jakubíček, J. Kallas, S. Krek, & V. Baisa (eds.), *Proceedings of eLex 2017 – 5th International Conference on Electronic Lexicography in the 21st Century*, 321–341. Brno: Lexical Computing CZ s.r.o.

León-Araúz, P., Reimerink, A., & Faber, P. (2019). EcoLexicon and by-products: integrating and reusing terminological resources. In A. Alcina, R. Costa, & C. Roche (eds.), *Special issue of Terminology and e-dictionaries*, *Terminology*, 25(2), 222–258. Amsterdam/Philadelphia: John Benjamins.

León-Araúz, P., Reimerink, A., & García Aragón, A. (2013b). Dynamism and Context in Specialized Knowledge. *Terminology*, 19(1), 31–61. Amsterdam/ Philadelphia: John Benjamins.

León-Araúz, P., San Martín, A., & Faber, P. (2016). Pattern-based Word Sketches for the Extraction of Semantic Relations. In *Proceedings of the 5th International Workshop on Computational Terminology*, 73–82. Osaka, Japan.

León-Araúz, P., San Martín, A., & Reimerink, A. (2018). The EcoLexicon English Corpus as an open corpus in Sketch Engine. In J. Čibej, V. Gorjanc, I. Kosem, & S. Krek (eds.), *Proceedings of the 18th EURALEX International Congress*, 893–901. Ljubljana: EURALEX.

Leonardi, P. M. (2010). Digital Materiality? How Artifacts Without Matter, Matter. *First Monday*, 15(6).

Lewis, M. (2019). Compositional Hyponymy with Positive Operators. In *Proceedings of Recent Advances in Natural Language Processing*, 638–647. Varna, Bulgaria.

Li, J., Lv, X., & Liu, K. (2014). Extracting Hyponymy of Ontology Concepts from Patent Documents. In *Proceedings of the 10th International Conference on Computational Intelligence and Security*, 283–287. Kunming, Yunnan, China.

Liu, L., Cao, C., & Wang, H. (2006). Extracting Hyponymic Relations from Chinese Free Corpus. In *Proceedings of the 5th WSEAS International Conference on Applied Computer Science*, 963–969. Hangzhou, China.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

Madsen, B. N., & Thomsen, H. E. (2009). Terminological Concept Modelling and Conceptual Data Modelling. *International Journal of Metadata, Semantics and Ontologies (IJMSO)*, 4(4), 239–249.

Madsen, B. N., Pedersen, B. S., & Thomsen, H. E. (2001). Defining semantic relations for ontoquery. In P. A. Jensen & P. Skadhauge (eds.), *Proceedings of the 1st International OntoQuery Workshop: Ontology-based Interpretation of NPs*. Kolding: University of Southern Denmark.

Mahesh, K., & Nirenburg, S. (1995). A Situated Ontology for Practical NLP. In *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, International Joint Conference on Artificial Intelligence (IJCAI-1995)*, 1–10. Montreal, Canada.

Malaisé, V., Zweigenbaum, P., & Bachimont, B. (2005). Mining defining contexts to help structuring differential ontologies. *Terminology*, 11(1), 21–53. Amsterdam/Philadelphia: John Benjamins.

Maroto, N., & Alcina, A. (2009). Formal Description of Conceptual Relationships with a View to Implementing Them in the Ontology Editor Protégé. *Terminology*, 15(2), 232–257. Amsterdam/Philadelphia: John Benjamins.

Marshman, E. (2002). The Cause-Effect Relation in Biopharmaceutical Texts: Some English Knowledge Patterns. In *Proceedings of Terminology and Knowledge Engineering (TKE 2002)*, 89–94. Nancy, France.

Marshman, E., Gariépy, J., & Harms, C. (2012). Helping Language Professionals Relate to Terms: Terminological Relations and Termbases. *The Journal of Specialised Translation*, 18, 30–56.

Marshman, E., Morgan, T., & Meyer, I. (2002). French Patterns for Expressing Concept Relations. *Terminology*, 8(1), 1–29. Amsterdam/Philadelphia: John Benjamins.

Martin, E. A. (ed.) (2010). *A Dictionary of Science*. 6th ed. Oxford: Oxford University Press.

McGraw-Hill (ed.) (2003a). *Dictionary of Chemistry*. 2nd ed. New York: McGraw-Hill.

McGraw-Hill (ed.) (2003b). *Dictionary of Geology & Mineralogy*. 2nd ed. New York: McGraw-Hill.

Měchura, M. B. (2017). Introducing Lexonomy: An Open-Source Dictionary Writing and Publishing System. In I. Kosem, C. Tiberius, M. Jakubíček, J. Kallas, S. Krek, & V. Baisa (eds.), *Proceedings of the eLex 2017 conference – 5th International Conference on Electronic Lexicography in the 21st Century*, 662–679. Brno: Lexical Computing CZ s.r.o.

Medin, D. L. (1998). Concepts and conceptual structure. In P. Thagard (ed.), *Mind Readings: Introductory Selections on Cognitive Science*. Cambridge (MA)/London: MIT Press.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology*, 51, 121–147.

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of Age: A Review of Embodiment and the Neuroscience of Semantics. *Cortex*, 48, 788–804.

Meyer, I. (2001). Extracting Knowledge-rich Contexts for Terminography: A Conceptual and Methodological Framework. In D. Bourigault, C. Jacquemin, & M. C. L'Homme (eds.), *Recent Advances in Computational Terminology*, 279–302. Amsterdam/Philadelphia: John Benjamins.

Meyer, I., Bowker, L., & Eck, K. (1992). COGNITERM: An Experiment in Building a Knowledge-based Term Bank. In *Proceedings of the 5th EURALEX International Congress*, 159–172. Tampere, Finland.

Meyers, R. A. (ed.) (2004). *Encyclopedia of Molecular Cell Biology and Molecular Medicine*. 2nd ed. Weinheim: Wiley-VCH.

Michalski, R. S. (1991). *Concepts as Flexible and Context-Dependent Sets: The Two-Tiered View*. Fairfax (VA): George Mason University.

Miller, G. A. (1986). Dictionaries in the Mind. *Language and Cognitive Processes*, 1, 171–185.

Miller, G. A. (1998). Nouns in WordNet. In C. Fellbaum (ed.), *WordNet: An Electronic Lexical Database*, 23–46. Cambridge (MA): MIT Press.

Mizoguchi, R., Vanwelkenhuysen, J., & Ikeda, M. (1995). Task ontology for reuse of problem solving knowledge. In N. Mars (ed.), *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing (KBKS'95)*, 46–57. Amsterdam: IOS Press.

Montero-Martínez, S., Faber, P., & Buendía, M. (2011). *Terminología para traductores e intérpretes*. Granada: Tragacanto.

Montiel-Ponsoda, E. (2022). Terminology and ontologies. In P. Faber & M. C. L'Homme (eds.), *Theoretical Perspectives on Terminology: Explaining terms, concepts and specialized knowledge*, *Terminology and Lexicography Research and Practice*, 23, 149–173. Amsterdam/Philadelphia: John Benjamins.

Moreno-Ortiz, A., & Pérez-Hernández, C. (2000). Reusing the Mikrokosmos Ontology for Concept-based Multilingual Terminology Databases. In N. Calzolari *et al*. (eds.), *Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC 2000)*, 1061–1067. Athens, Greece.

Moreno-Ortiz, A., & Pérez-Hernández, C. (2002). Ontological Semantics and Lexical Templates: Mowing the Grass from the Other Side of the Fence. In R. Mairal-Usón & M. J. Pérez-Quintero (eds.), *New Perspectives on Argument Structure in Functional Grammar*, 179–212. Berlin/New York: Mouton de Gruyter.

Mortchev-Bouveret, M. (2006). Fonctions lexicales pour le typage de relations syntagmatiques et paradigmatiques: Une approche lexicographique du terme. *Terminology*, 12(2), 235–259. Amsterdam/Philadelphia: John Benjamins.

Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge (MA)/London: MIT Press.

Murphy, G. L., & Lassaline, M. E. (1997). Hierarchical Structure in Concepts and Basic Level of Categorization. In K. Lamberts & D. Shanks (eds.), *Knowledge, Concepts, and Categories*, 93–131. Cambridge (MA)/London: MIT Press.

Murphy, M. L. (2003). *Semantic Relations and the Lexicon: Antonymy, Synonymy and Other Paradigms*. Cambridge: Cambridge University Press.

Murphy, M. L. (2006). Hyponymy and Hyperonymy. In K. Brown (ed.), *Encyclopedia of Language and Linguistics*, 1, 446–448. New York: Elsevier.

Murphy, M. L. (2010). *Lexical Meaning*. Cambridge: Cambridge University Press.

Murphy, M. L., & Koskela, A. (2010). *Key Terms in Semantics*. London/New York: Continuum.

Navigli, R., & Ponzetto, S. P. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193, 217–250.

Nazar, R., Vivaldi, J., & Wanner, L. (2012). Automatic Taxonomy Extraction for Specialized Domains Using Distributional Semantics. *Terminology*, 18(2), 188–225. Amsterdam/Philadelphia: John Benjamins.

Nickles, M., Pease, A., Schalley, A. C., & Zaefferer, D. (2007). Ontologies across disciplines. In A. C. Schalley & D. Zaefferer (eds.), *Ontolinguistics – How Ontological Status Shapes the Linguistic Coding of Concepts*, 23–67. Berlin/New York: Mouton de Gruyter.

Nirenburg, S., & Raskin, V. (2004). *Ontological Semantics*. Cambridge (MA)/London: MIT Press.

Nkwenti-Azeh, B. (1998). Information Mediation: The Interface Between Terminology and Translation. In A. Kent (ed.), *Encyclopedia of Library and Information Science*, 62, 157–170. New York/Basel/Hong Kong: Marcel Dekker.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York/Cambridge: Cambridge University Press.

Nuopponen, A. (1994). *Begreppssystem för terminologisk analys*. Doctoral Thesis. Vaasa: University of Vaasa.

Nuopponen, A. (2005). Concept Relations: An Update of a Concept Relation Classification. In B. N. Madsen & H. E. Thomsen (eds.), *7th International Conference on Terminology and Knowledge Engineering (TKE 2005)*, 127–138. Copenhagen, Denmark.

Nuopponen, A. (2014). Tangled Web of Concept Relations: Concept Relations for ISO 1087–1 and ISO 704. In *Ontology, Terminology & Text Mining: TKE 2014, International Conference on Terminology and Knowledge Engineering*. Berlin, Germany.

Nuopponen, A. (2018). Terminological Concept Systems. In J. Humbley, G. Budin, & C. Laurén (eds.), *Languages for Special Purposes*, 453–468. Berlin/Boston: De Gruyter Mouton.

Nuopponen, A. (2022). Conceptual relations: From the General Theory of Terminology to knowledge bases. In P. Faber & M. C. L'Homme (eds.), *Theoretical Perspectives on Terminology: Explaining terms, concepts and specialized knowledge*, *Terminology and Lexicography Research and Practice*, 23, 63–86. Amsterdam/Philadelphia: John Benjamins.

Oxford University Press (2010). *Oxford Dictionary of English*. 3rd ed. Oxford: Oxford University Press.

Oxford University Press (2022). *OED Online*. Oxford: Oxford University Press. Available at: https://www.oed.com/

Pantel, P., & Pennacchiotti, M. (2006). Espresso: Leveraging Generic Patterns for Automatically Harvesting Semantic Relations. In N. Calzolari, C. Cardie, & P. Isabelle (eds.), *Proceedings of COLING/ACL 2006 – 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, 113–120. Sydney, Australia.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where Do You Know What You Know? The Representation of Semantic Knowledge in the Human Brain. *Nature Reviews Neuroscience*, 8, 976–988.

Pavel, S., & Nolet, D. (2001). *Précis de Terminologie / The Handbook of Terminology*. Ottawa: Translation Bureau, Terminologie and Standardization Directorate.

Pearson, J. (1998). *Terms in Context*. Amsterdam/Philadelphia: John Benjamins Publishing.

Pérez, C. (2002). *Explotación de los córpora textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento*. Doctoral Thesis. Malaga: Universidad de Málaga.

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K., & Spivey, M. J. (2013). Computational Grounded Cognition: a new alliance between grounded cognition and computational modeling. *Frontiers in Psychology*, 3, 612.

Pfafflin, J. R., & Ziegler, E. N. (eds.) (2012). *Encyclopedia of Environmental Science and Engineering*. 5th ed. Boca Raton: Taylor & Francis.

Picht, H., & Draskau, J. (1985). *Terminology: An Introduction*. Guildford: University of Surrey.

Pihkala, T. (2001). Socioterminology. Paper presented at the *NordTerm Conference*, 13–16 June. Tuusula: Finnish Terminology Center.

Prieto Velasco, J. A. (2009). To see or not to see: concept visualization in terminological knowledge bases. In *8th International Conference on Terminology and Artificial Intelligence*, 220–231. Toulouse, France.

Quillian, M. (1968). Semantic Memory. In M. Minsky (ed.), *Semantic Information Processing*, 227–270. Cambridge (MA)/London: MIT Press.

Rambousek, A., Jakubíček, M., & Kosem, I. (2021). New developments in Lexonomy. In I. Kosem, M. Cukr, M. Jakubíček, J. Kallas, S. Krek, & C. Tiberius (eds.), *Proceedings of the eLex 2021 Conference – 7th International Conference on Electronic Lexicography in the 21st Century*, 455–462. Brno: Lexical Computing CZ, s.r.o.

Reimerink, A., & Faber, P. (2009). EcoLexicon: A frame-based knowledge base for the environment. In J. Hrebícek *et al*. (eds.), *European Conference of the Czech Presidency of the Council of the EU TOWARDS eENVIRONMENT, Opportunities of SEIS and SISE: Integrating Environmental Knowledge in Europe*, 25–27. Brno: Masaryk University.

Reimerink, A., León-Araúz, P., & Magaña-Redondo, P. J. (2010). EcoLexicon: An Environmental TKB. In N. Calzolari *et al*. (eds.), *Proceedings of the 7th conference on International Language Resources and Evaluation (LREC 2010)*. Valetta: European Language Resources Association (ELRA).

Rittner, D., & Bailey, R. A. (eds.) (2005). *Encyclopedia of Chemistry*. New York: Facts On File.

Rittner, D., & McCabe, T. L. (eds.) (2004). *Encyclopedia of Biology*. New York: Facts On File.

Roche, C. (2007). Le terme et le concept: fondements d'une ontoterminologie. In *Proceedings of the TOTh Conference: Terminology and Ontology (TOTh 2007)*, 1–22. Chambéry: Presses Universitaires Savoie Mont-Blanc.

Roche, C. (2012a). Ontologie: entre terminologie et connaissance de spécialité. In *Colloque GLAT – Terminologies: textes, discours et accès aux savoirs spécialisés*. Genova, Italy.

Roche, C. (2012b). Ontoterminology: How to unify terminology and ontology into a single paradigm. In N. Calzolari *et al*. (eds.), *Proceedings of the 8th International*

Conference on Language Resources and Evaluation (LREC 2012), 2626–2630. Istanbul, Turkey: European Language Resources Association (ELRA).

Roche, C. (2015). Ontological Definition. In H. Kockaert & F. Steurs (eds.), *Handbook of Terminology*, 128–152. Amsterdam/Philadelphia: John Benjamins.

Roche, C., Alcina, A., & Costa, R. (2019a). Terminological resources in the digital age. *Terminology*, 25(2), 139–145. Amsterdam/Philadelphia: John Benjamins.

Roche, C., Calberg-Challot, M., Damas, L., & Rouard, P. (2009). Ontoterminology: A New Paradigm for Terminology. In J. L. G. Dietz (ed.), *Proceedings of the International Conference on Knowledge Engineering and Ontology Development (KEOD 2009)*, 321–326. Madeira, Portugal.

Roche, C., Costa, R., Carvalho, S., & Almeida, B. (2019b). Knowledge-based terminological e-dictionaries: The EndoTerm and al-Andalus Pottery projects. *Terminology*, 25(2), 259–290. Amsterdam/Philadelphia: John Benjamins.

Rogers, M. (2004). Multidimensionality in concepts systems. *Terminology*, 10(2), 215–240. Amsterdam/Philadelphia: John Benjamins.

Rojas-García, J., & Cabezas-García, M. (2019). Use of Knowledge Patterns for the Evaluation of Semi-automatically Induced Semantic Clusters. In I. Simonnæs, Ø. Andersen, & K. Schubert (eds.), *New Challenges for Research on Language for Special Purposes*, *Serie Forum für Fachsprachen-Forschung*, 154, 121–140. Berlin: Frank & Timme.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology*, 104(3), 192–233.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (eds.), *Cognition and Categorization*, 27–48. Hillsdale (NJ): Lawrence Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology*, 7(4), 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3), 382–439.

Rummel, D., & Ball, S. (2001). The IATE Project – Towards a Single Terminology Database for the European Union. *Translating and the Computer*, 23.

Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R., & Scheffdzyk, J. (2010). *FrameNet II: Extended Theory and Practice*. Available at: https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf

Sager, J. C. (1990). *A Practical Course in Terminology Processing*. Amsterdam/ Philadelphia: John Benjamins.

Sager, J. C. (1994). Terminology: Custodian of Knowledge and Means of Knowledge Transfer. *Terminology*, 1(1), 7–15. Amsterdam/Philadelphia: John Benjamins.

Sager, J. C., & Ndi-Kimbi, A. (1995). The conceptual structure of terminological definitions and their linguistic realisations. *Terminology*, 2(1), 61–81. Amsterdam/Philadelphia: John Benjamins.

San Martín, A. (2016). *La representación de la variación contextual mediante definiciones terminológicas flexibles*. Doctoral Thesis. Granada: Universidad de Granada.

San Martín, A., Cabezas-García, M., Buendía-Castro, M., Sánchez-Cárdenas, B., León-Araúz, P., Reimerink, A., & Faber, P. (2020). Presente y futuro de la base de conocimiento terminológica EcoLexicon. *Onomázein*, 49, 174–202. Santiago de Chile: Pontificia Universidad Católica de Chile.

Schmitz, K. D. (2006). Terminology and Terminological Databases. In K. Brown (ed.), *Encyclopedia of Language and Linguistics, vol. 2*. New York: Elsevier.

Selley, R. C., Cocks, L. R. M., & Plimer, I. R. (eds.) (2005). *Encyclopedia of Geology*. Amsterdam: Elsevier.

Seppälä, S. (2015). An ontological framework for modeling the contents of definitions. *Terminology*, 21(1), 23–50. Amsterdam/Philadelphia: John Benjamins.

Shapiro, L. (2004). *The Mind Incarnate*. Cambridge (MA)/London: MIT Press.

Sierra, G., Alarcón, R., Aguilar, C., & Bach, C. (2008). Definitional verbal patterns for semantic relation extraction. *Terminology*, 14(1), 74–98. Amsterdam/Philadelphia: John Benjamins.

Sinclair, J. (1991). *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.

Sinclair, J. (1995). Corpus Typology: A framework for classification. In G. Melchers & B. Warren (eds.), *Studies in Anglistics*, 17–34. Stockholm: Almquist and Wiksell International.

Sinclair, J. (1996a). The empty lexicon. *International Journal of Corpus Linguistics*, 1(1), 99–119.

Sinclair, J. (1996b). The Search for Units of Meaning. *TEXTUS: English Studies in Italy*, 9, 75–106.

Singleton, P., & Sainsbury, D. (2006). *Dictionary of Microbiology and Molecular Biology*. 3rd ed. Chichester: Wiley.

Sloutsky, V. M. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences*, 7(6), 246–251.

Soler, V., & Alcina, A. (2008). Patrones léxicos para la extracción de conceptos vinculados por la relación parte-todo en español. *Terminology*, 14(1), 99–123. Amsterdam/Philadelphia: John Benjamins.

Sowa, J. F. (1984). *Conceptual Structures: Information Processing in Mind and Machine*. Reading (MA): Addison-Wesley.

Stock, W. G. (2010). Concepts and Semantic Relations in Information Science. *Journal of the American Society for Information Science and Technology*, 61(10), 1951–1969.

Stratogiannis, G., Kouris, P., Alexandridis, G., Siolas, G., Stamou, G., & Stafylopatis, A. (2021). Semantic enrichment of documents: a classification perspective for ontology-based imbalanced semantic descriptions. *Knowledge and Information Systems*, 63, 3001–3039. Berlin: Springer.

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge Engineering: Principles and Methods. *IEEE Transactions on Data and Knowledge Engineering*, 25(1–2), 161–197.

Sure, Y., & Studer, R. (2003). A methodology for ontology-based knowledge management. In J. Davies, D. Fensel, & F. Van Harmelen (eds.), *Towards the Semantic Web: Ontology-Driven Knowledge Management*, 33–46. New York: Wiley.

Szostak, R. (2012). Classifying Relationships. *Reviews of Concepts in Knowledge Organization*, 39(3), 165–178.

Talmy, L. (2000). *Toward a Cognitive Semantics*. Cambridge (MA)/London: MIT Press.

Taylor, J. R. (1995). *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Oxford University Press.

Temmerman, R. (2000). *Towards New Ways of Terminology Description: The Sociocognitive Approach*. Amsterdam/Philadelphia: John Benjamins.

Temmerman, R. (2001). Sociocognitive Terminology Theory. In M. T. Cabré & J. Feliu (eds.), *Terminología y cognición*, 75–92. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.

Temmerman, R. (2007). Approaches to terminology: Now that the dust has settled. *Synaps*, 20, 27–36.

Temmerman, R. (2022). Units of understanding in Sociocognitive Terminology. In P. Faber & M. C. L'Homme (eds.), *Theoretical Perspectives on Terminology: Explaining terms, concepts and specialized knowledge, Terminology and Lexicography Research and Practice*, 23, 331–352. Amsterdam/Philadelphia: John Benjamins.

Temmerman, R., & Kerremans, K. (2003). Termontography: Ontology building and the sociocognitive approach to terminology description. In E. Hajicová, A. Kotesovcová, & J. Mírovský (eds.), *Proceedings of the International Congress of Linguists (CIL17)*. Prague: Matfyzpress.

Temmerman, R., Kerremans, K., & Vandervoort, V. (2005). La termontographie en contexte(s): Mots, Termes et Contextes. In D. Blampain, P. Thoiron, & M. Van Campenhoudt (eds.), *Actes des Septièmes Journées Scientifiques du Réseau Lexicologie, Terminologie, Traduction*, 429–439. Paris: Éditions des archives contemporaines.

Tercedor Sánchez, M., Faber, P., & D'Angiulli, A. (2011). The depiction of wheels by blind children: preliminary studies on pictorial metaphors, language and embodied imagery. *Imagination, Cognition and Personality*, 31(1–2), 113–128.

Tercedor Sánchez, M., López Rodríguez, C. I., & Faber, P. (2012). Working with Words: Research Approaches in Translation-oriented Lexicographic Practice. *TTR: Traduction, Terminologie, Rédaction*, 25(1), 181–214. Association canadienne de traductologie.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 830–846.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, 8, 769–800.

Valentini, C. (2013). The PCT Termbase: a scientific and technical resource for the patent community. In C. Roche (ed.), *TOTh 2013 Proceedings - Terminology & Ontology: Theories and applications*, 217–242. Chambéry, France.

Valentini, C., Westgate, G., & Rouquet, P. (2016). The PCT Termbase of the World Intellectual Property Organization: Designing a database for multilingual patent terminology. *Terminology*, 22(2), 171–200. Amsterdam/Philadelphia: John Benjamins.

Van Heijst, G., Schreiber, A. T., & Wielinga, B. J. (1997). Using explicit ontologies in KBS development. *International Journal of Human-Computer Studies*, 45, 183–292.

Veres, C. (2005). Ontology and Taxonomy: why "is-a" still isn't just "is-a". In *Proceedings of the 2005 International Conference on E-Business, Enterprise Information Systems, E-Government, and Outsourcing (EEE 2005)*. Las Vegas, Nevada, USA.

Vezzani, F., & Di Nunzio, G. M. (2020). Methodology for the standardization of terminological resources: Design of TriMED database to support multi-register medical communication. *Terminology*, 26(2), 265–297. Amsterdam/Philadelphia: John Benjamins.

Webster, L. F. (ed.) (1997). *The Wiley Dictionary of Civil Engineering and Construction*. New York: Wiley.

Weigand, H. (1997). Multilingual Ontology-Based Lexicon for News Filtering – The TREVI Project. In K. Mahesh (ed.), *Ontologies and Multilingual NLP: Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI)*, 138–159. Nagoya, Japan.

Wierzbicka, A. (1984). Apples are not a ''kind of fruit'': The semantics of human categorization. *American Ethnologist*, 11, 313–328.

Wiley-VCH (ed.) (2011). *Ullmann's Encyclopedia of Industrial Chemistry*. 7th ed. Weinheim: Wiley-VCH.

Willie, G. (ed.) (2010). *A Comprehensive Dictionary of Chemistry*. Chandigarh: Abhishek.

Winston, M. E., Chaffin, R., & Herrmann, D. (1987). A taxonomy of part-whole relations. *Cognitive Science*, 11, 417–444.

Wittgenstein, L. (1953). *Philosophical investigations*. New York: Macmillan.

Wüster, E. (1968). *The Machine Tool. An Interlingual Dictionary of Basic Concepts*. London: Technical Press.

Wüster, E. (1979). *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*. Vienna: Springer.

Zhan, Q., & Wang, C. (2015). Hyponymy extraction of domain ontology concept based on CCRFs and hierarchy clustering. *International Journal of Web & Semantic Technology (IJWesT)*, 6(3).

Zorrilla-Agut, P., & Fontenelle, T. (2019). IATE 2: Modernising the EU's IATE terminological database to respond to the challenges of today's translation world and beyond. *Terminology*, 25(2), 146–174. Amsterdam/Philadelphia: John Benjamins.

ANNEX I
FULL CONCEPTUAL CATEGORY HIERARCHY IN ECOLEXICON (GIL-BERROZPE *ET AL*. 2019)

A: Attribute

- **A-1:** Ability [e.g., AUTOTROPHIC, PERMEABILITY, TSUNAMIGENIC]
- **A-2:** Direction [e.g., DOWNSTREAM, WINDWARD, ONSHORE]
- **A-3:** Location [e.g., HADOPELAGIC, MESOTIDAL, SUBAQUEOUS]
- **A-4:** Measurement [e.g., QUANTITY, SPECIFIC HEAT CAPACITY, NUTRIENT CONCENTRATION]
  - **A-4.1:** Magnitude [e.g., ALTITUDE, RADICULAR ZONE DEPTH, AMBIENT TEMPERATURE]
    - **A-4.1.1:** Level [e.g., MAXIMUM FLOW, HIGHEST ASTRONOMICAL TIDE, FREEZING POINT]
      - **A-4.1.1.1:** Mean [e.g., MEAN FLOW, MEAN TIDE LEVEL, AVERAGE PRECIPITATION]
- **A-5:** Origin [e.g., ARTIFICIAL, AEOLIAN, LITHOLOGIC]
- **A-6:** Physical attribute [e.g., COLOR, SOIL TEXTURE, XERICITY]
  - **A-6.1:** Composition [e.g., BIOCLASTIC, WOODY, MONOLITHIC]
  - **A-6.2:** Shape [e.g., BACCIFORM, EUHEDRAL, HOOK-SHAPED]
  - **A-6.3:** Size [e.g., BIG, SMALL, GRAIN SIZE]
  - **A-6.4:** State [e.g., CARBONATE EQUILIBRIUM, SLOPE INSTABILITY, UNCONSOLIDATED]
    - **A-6.4.1:** Climate [e.g., BIOCLIMATE, SAVANNA CLIMATE, PERIGLACIALISM]
- **A-7:** Time [e.g., APERIODIC, SEMIDIURNAL, TEMPORARY]

E: Entity

- **E-1:** Creation [e.g., WIND TURBINE GENERATOR SYSTEM, COLLECTOR, SEPTIC SYSTEM]
  - **E-1.1:** Artifact [e.g., CULVERT, DC BUS, STATOSCOPE]
    - **E-1.1.1:** Conduit [e.g., DRAINAGE DITCH, PIPELINE, DUCT]
    - **E-1.1.2:** Container [e.g., CLOUD CHAMBER, SEDIMENTATION TANK, RETENTION BASIN]
    - **E-1.1.3:** Instrument [e.g., CENTRIFUGAL PUMP, FISHING NET, WEATHER SATELLITE]
      - **E-1.1.3.1:** Measuring instrument [e.g., ACCELEROMETER, BAROMETER, SOUNDING MACHINE]
      - **E-1.1.3.2:** Recording instrument [e.g., ALBEDOGRAPH, MARIGRAPH, WATER-LEVEL RECORDER]
      - **E-1.1.3.3:** Sampling instrument [e.g., COLLECTOR, AUTOMATIC SAMPLER, VAN DORN BOTTLE]
      - **E-1.1.3.4:** Transforming instrument [e.g., UPWIND TURBINE, CONVERTER, SOLAR CELL]
    - **E-1.1.4:** Vehicle [e.g., BOAT, DREDGER, ELECTRIC VEHICLE]
  - **E-1.2:** Software [e.g., COMPUTER APPLICATION, CONTOUR GRIDDER, MODFLOW]
  - **E-1.3:** Structure [e.g., SPILLWAY, PIER, ENGINEERING STRUCTURE]
    - **E-1.3.1:** Building [e.g., GEOTHERMAL POWER PLANT, TIDE STATION, OIL REFINERY]
    - **E-1.3.2:** Defense structure [e.g., REEF BREAKWATER, HIGH GROYNE, RETAINING WALL]
- **E-2:** Discipline [e.g., BIOCLIMATOLOGY, HUMAN ECOLOGY, PHYTOPATHOLOGY]
- **E-3:** Force [e.g., TRACTIVE FORCE, TECTONIC FORCE, GRAVITY]

E-3.1: Dynamics [e.g., ATMOSPHERIC DYNAMICS, SLOPE DYNAMICS, COASTAL DYNAMICS]

E-3.2: Energy [e.g., ELECTRICITY, WIND ENERGY, SOLAR ENERGY]

E-3.3: Stress [e.g., FRICTION, DYNAMIC PRESSURE, TENSION]

E-4: Geographic feature [e.g., ENTRY CHANNEL, AQUIFER, BIOME]

E-4.1: Artificial geographic feature [e.g., GROYNE BAY, QUARRY, PORT]

E-4.1.1: Artificial water body [e.g., POOL, POND, RESERVOIR]

E-4.2: Natural geographic feature [e.g., ABYSS, HIGH PLATEAU, BAY]

E-4.2.1: Landform [e.g., FAN DELTA, RIVER GORGE, EMERGENT COAST]

E-4.2.1.1: Natural water body [e.g., SEA CHANNEL, KARST SPRING, LAGOON]

E-4.2.2: Landscape [e.g., TIDAL SHOAL, MONSOON FOREST, MANGROVE SWAMP]

E-5: Human [e.g., PORT AUTHORITY, HUMAN BEING, SOCIAL AGENT]

E-5.1: Institution [e.g., METEOROLOGICAL SERVICE, CITY COUNCIL, PUBLIC INSTITUTION]

E-5.2: Specialist [e.g., GEOGRAPHER, GEOLOGIST, OCEANOGRAPHER]

E-6: Information [e.g., PIECE OF DATA, CARTOGRAPHIC INFORMATION, HYDROLOGIC DATA]

E-6.1: Classification [e.g., CLIMATE CLASSIFICATION, CLADE, URBAN HIERARCHY]

E-6.1.1: Scale [e.g., BEAUFORT SCALE, STATE-OF-SEA SCALE, SPECTRUM]

E-6.2: Document [e.g., PLAN, PROTOCOL, TIDE TABLE]

E-6.2.1: Law [e.g., LEGISLATION, WILDLIFE LAW, PRINCIPLE OF ENVIRONMENTAL LAW]

E-6.3: Parameter [e.g., STRUCTURAL CRITERION, QUALITY INDICATOR, K FACTOR]

E-6.4: Record [e.g., BASELINE CARTOGRAPHY, ECHOGRAM, METEOROLOGICAL SERIES]

E-6.5: Representation [e.g., GEODATABASE, AURORAL OVAL, SOIL PROFILE]

E-6.5.1: Graph [e.g., ADIABATIC CHART, STRATIGRAPHIC COLUMN, COMPOUND HYDROGRAPH]

E-6.5.2: Line [e.g., RATING CURVE, ISOHALINE, MERIDIAN]

E-6.5.3: Map [e.g., NAUTICAL CHART, ORIENTATION MAP, ORTHOPHOTOMAP]

E-6.5.4: Mathematical expression [e.g., COEFFICIENT, STANDARD DEVIATION, WAVE EQUATION]

E-6.5.5: Model [e.g., EKMAN SPIRAL, EROSION MODEL, SIMULATION]

E-6.5.6: Picture [e.g., PHOTOMOSAIC, SATELLITE IMAGE, ORTHOPHOTO]

E-6.5.7: Unit [e.g., STERADIAN, FARADAY, MILIMETER]

E-6.6: Theory [e.g., PLATE TECTONICS, EQUILIBRIUM THEORY, STATIONARY WAVE THEORY]

E-7: Lifeform [e.g., DETRITIVORE, NATIVE SPECIES, ORGANISM]

E-7.1: Animal [e.g., AMPHIBIAN, LIVESTOCK, CRUSTACEAN]

E-7.2: Community [e.g., BENTHOS, BIOCENOSIS, BIOLOGICAL COMMUNITY]

E-7.2.1: Animal community [e.g., STYGOFAUNA, COHORT, ZOOPLANKTON]

E-7.2.2: Plant community [e.g., PHYTOBENTOS, FLORA, PHYTOPLANKTON]

E-7.3: Fungus [e.g., BASIDIOMYCOTA, MYCOBIONT, FACULTATIVE PARASITE]

E-7.4: Microorganism [e.g., BACTERIA, FACULTATIVE AEROBE, ENTERIC VIRUS]

E-7.5: Plant [e.g., CHAMAEPHYTE, PHYCOBIONT, MANGROVE]

E-8: Matter [e.g., GREYBODY, ORGANIC MATERIAL, SUBSTANCE]

E-8.1: Chemical substance [e.g., CARBONIC ACID, ARSENIC, NITROGEN DIOXIDE]

E-8.2: Fluid matter [e.g., TAR, LAVA FLOW, MUD]

E-8.2.1: Fluid astronomical body [e.g., HEAVENLY BODY, STAR, SUN]

E-8.2.2: Gas [e.g., POLAR AIR, EXHAUST GAS, SMOG]

E-8.2.3: Water [e.g., RUNOFF WATER, DRINKING WATER, RAINWATER]

E-8.2.3.1: Cloud [e.g., ALTOSTRATUS, STRATOCUMULUS, FRONTAL FOG]

E-8.3: Particle [e.g., VOLCANIC ASH, INTERLEUKIN, ULTRAFINE PARTICLE]

E-8.4: Solid matter [e.g., SOLID FUEL, SOLID WASTE, SOLUTE]

E-8.4.1: Deposit [e.g., ALLUVIUM, SEDIMENT FLOW, AEOLIAN DEPOSIT]

E-8.4.2: Material [e.g., CEMENT, REINFORCED CONCRETE, SEMICONDUCTOR]

E-8.4.2.1: Mineral [e.g., ANTHRACITE, COARSE SAND, ZEOLITE]

E-8.4.2.2: Rock [e.g., LIMESTONE, QUARTZ DIORITE, CLASTIC SEDIMENTARY ROCK]

E-8.4.2.3: Soil [e.g., LEPTOSOL, MOLLISOL, SATURATED SOIL]

E-8.4.3: Snow/ice [e.g., AVALANCHE, SNOWFLAKE, ANCHOR ICE]

E-8.4.4: Solid astronomical body [e.g., ASTEROID, PLANET, SATELLITE]

E-9: Part [e.g., DISCARDS, SECTION, STATOR]

E-9.1: Part of instrument [e.g., ANEMOMETER MAST, WIND TURBINE ROTOR, FLAP]

E-9.2: Part of landform [e.g., BEACH HEAD, BERM CREST, SOIL PROPERTIES]

E-9.3: Part of lifeform [e.g., ALLELE, CELL WALL, TISSUE]

E-9.3.1: Part of animal [e.g., EOSINOPHIL, OTOLITH, VALVE]

E-9.3.2: Part of fungus [e.g., ASCOSPORE, SPOROCARP, PARAPLECTENCHYMA]

E-9.3.3: Part of plant [e.g., BRACTEOLE, CHLOROPLAST, DEHISCENT FRUIT]

E-9.4: Part of structure [e.g., HARBOUR MOUTH, SPILLWAY CREST, GROYNE HEAD]

E-9.5: Part of vehicle [e.g., GUNWALE, HULL, KEEL]

E-9.6: Part of water body [e.g., DOWNSTREAM, APHYTAL ZONE, SEA FLOOR]

E-10: Path [e.g., ROAD, GULLY, VIADUCT]

E-10.1: Imaginary path [e.g., PLANETARY ORBIT, ECLIPTIC PLANE, EARTH'S ELLIPTIC ORBIT]

E-11: Period [e.g., LUNAR DAY, AUTUMN, USEFUL LIFE]

E-11.1: Era [e.g., DEVONIAN, MESOZOIC ERA, PLEISTOCENE EPOCH]

E-12: Space [e.g., CAPILLARY INTERSTICE, MEDIUM, ECOLOGICAL NICHE]

E-12.1: Area [e.g., SEDIMENTARY ENVIRONMENT, PROTECTED AREA, ECOREGION]

E-12.1.1: Administrative area [e.g., CITY, MUNICIPAL BOUNDARY, THE UNITED STATES OF AMERICA]

E-12.1.2: Facility [e.g., BIOMASS POWER PLANT, MEASURING STATION, GAUGING SITE]

E-12.1.3: Land [e.g., BASIN SLOPE, MEADOW, AREA OF LAND]

E-12.2: Layer [e.g., ATMOSPHERE, PLANETARY BOUNDARY LAYER, LOWER MANTLE]

E-12.3: Limit [e.g., WAVE CREST, LIMIT OF UPRUSH, AMPHIDROMIC POINT]

E-12.4: Position [e.g., BIFURCATION, DEPOCENTER, PERIGEE]

E-13: System [e.g., DETRITUS FOOD CHAIN, NETWORK, ISOLATED SYSTEM]

P: Process

- **P-1:** Action [e.g., BIOLOGICAL ACTION, SPAWNING, ENVIRONMENTAL CRIME]
  - **P-1.1:** Analysis [e.g., SEDIMENTOLOGICAL ANALYSIS, ENVIRONMENTAL IMPACT ASSESSMENT, WEATHER FORECAST]
  - **P-1.2:** Chemical reaction [e.g., COMBUSTION, ANABOLISM, DEFLAGRATION]
  - **P-1.3:** Collection [e.g., ENERGY STORAGE, SOIL WATER RETENTION, SAND TRAPPING]
  - **P-1.4:** Interaction [e.g., INTERSPECIFIC COMPETITION, AIR-SEA INTERACTION, ENDOGENIC GEOLOGICAL PROCESS]
  - **P-1.5:** Management [e.g., COASTAL MANAGEMENT, SUSTAINABLE WATER USE, WASTE MANAGEMENT]
  - **P-1.6:** Measurement [e.g., STREAM GAUGING, DENSITOMETRY, STOCHASTIC PROCESS]
  - **P-1.7:** Protection [e.g., ABSORB WAVE ENERGY, SOIL CONSERVATION, FLOOD PREVENTION]
- **P-2:** Activity [e.g., SUBSISTENCE AGRICULTURE, SHIFTING CULTIVATION, FACTORY FARMING]
- **P-3:** Addition [e.g., TECTONIC ACCRETION, ARTIFICIAL NOURISHMENT, PHOSPHATE FERTILIZATION]
- **P-4:** Change [e.g., CLIMATE CHANGE, ECOLOGICAL DEGRADATION, ENVIRONMENTAL IMPACT]
  - **P-4.1:** Change in size/intensity [e.g., TIDE ACCELERATION, CYCLOGENESIS, ANTICYCLOLYSIS]
    - **P-4.1.1:** Decrease [e.g., RETARD LITTORAL DRIFT, WAVE SETDOWN, REDUCTION IN LONGSHORE TRANSPORT]
    - **P-4.1.2:** Increase [e.g., SEA LEVEL RISE, ALGAL BLOOM, RISE OF THE WATER TABLE]
  - **P-4.2:** Change of direction [e.g., DEFLECTION, DENSITY STRATIFICATION, SECULAR VARIATION]
  - **P-4.3:** Change of state [e.g., CONDENSATION, SOIL LIQUEFACTION, SOLIDIFICATION]
  - **P-4.4:** Disease [e.g., BRONCHITIS, YELLOW BAND DISEASE, MONILIA DISEASE]
  - **P-4.5:** Division [e.g., CLEAVAGE, DISPERSION, BREAKING DROPS]
  - **P-4.6:** Transformation [e.g., ACIDIFICATION, METAMORPHISM, TERRITORIAL TRANSFORMATION]
    - **P-4.6.1:** Pollution [e.g., ATMOSPHERIC POLLUTION, OZONE POLLUTION, OCEAN DUMPING]
    - **P-4.6.2:** Restoration [e.g., BIOREMEDIATION, ENVIRONMENTAL RECOVERY, REVEGETATION]
- **P-5:** Cycle [e.g., TIDAL CYCLE, CARBON CYCLE, HYDROLOGIC CYCLE]
- **P-6:** Elimination [e.g., DEFORESTATION, MASS EXTINCTION, ELIMINATION OF SOLID WASTE]
- **P-7:** Emission [e.g., PARTICULATE EMISSION, HYDROMAGMATIC ERUPTION, EVAPOTRANSPIRATION]
- **P-8:** Formation [e.g., BRECCIA FORMATION, ATMOSPHERIC IONIZATION, PRIMARY PRODUCTION]
- **P-9:** Loss [e.g., COASTAL DEGRADATION, INTERNAL EROSION, MECHANICAL WEATHERING]
- **P-10:** Method [e.g., AIR LAYERING, HODOGRAPH METHOD, POLYCULTURE]
- **P-11:** Movement [e.g., DRIFT, OSMOSIS, TRAFFIC]
  - **P-11.1:** Earth/soil movement [e.g., CONTINENTAL DRIFT, SLOPE MOVEMENT, TECTONIC EARTHQUAKE]
  - **P-11.2:** Energy movement [e.g., FORCED CONVECTION, ATMOSPHERIC RADIATION, CLOUD ELECTRIFICATION]
  - **P-11.3:** Fluid movement [e.g., CAPILLARITY, LAMINAR FLOW, MAGMA INTRUSION]
    - **P-11.3.1:** Water movement [e.g., COASTAL CIRCULATION, DRIFT CURRENT, GRAVITY FLOW]
  - **P-11.4:** Transport [e.g., TRANSFER, LONGSHORE TRANSPORT, UPWELL]
  - **P-11.5:** Wave [e.g., REGULAR WAVE, ATMOSPHERIC WAVE, PROGRESSIVE WAVE]
  - **P-11.6:** Wind movement [e.g., SEA BREEZE, ANTICYCLONIC CIRCULATION, WARM FRONT]
- **P-12:** Phase [e.g., KARYOKINESIS, CYTOKINESIS, PRELIMINARY TREATMENT]
  - **P-12.1:** Phase of cycle [e.g., TIDAL STAGE, LITHOGENESIS, OROGENY]

P-12.2: Phase of treatment [e.g., PRIMARY SEDIMENTATION, THERMOPHILIC DIGESTION, PREAERATION]

P-13: Phenomenon [e.g., LUNAR ECLIPSE, EXTREME EVENT, ENVIRONMENTAL NOISE]

P-13.1: Atmospheric phenomenon [e.g., SQUALL, ADVECTIVE THUNDERSTORM, TROPICAL CYCLONE]

P-13.1.1: Precipitation [e.g., HYDROMETEOR, FREEZING RAIN, CONVECTIVE PRECIPITATION]

P-13.2: Optical phenomenon [e.g., RAINBOW, AURORAL STORM, LIGHTNING FLASH]

ANNEX II

HYPONYMY SUBTYPES: DESCRIPTION AND EXAMPLES (GIL-BERROZPE 2017; GIL-BERROZPE *ET AL*. 2017)

HYPONYMY SUBTYPE	DESCRIPTION	EXAMPLES
Ability-based	Hyponyms characterized by own abilities or characteristics.	▪ AUTONOMOUS VEHICLE ▪ HABITABLE PLANET ▪ RENEWABLE RESOURCE
Activity-based	Hyponyms characterized by the activity or stability of their composition.	▪ ACTIVE DUNE ▪ ALKALI METAL ▪ RADIOACTIVE SUBSTANCE
Agent-based	Hyponyms characterized by the agent that causes them.	▪ AIR OXIDATION ▪ SPRINKLER IRRIGATION ▪ STORM TIDE
Amount-based	Hyponyms characterized by their amount or quantity.	▪ RARE METAL ▪ SINGLE STORM ▪ TRACE ELEMENT
Color-based	Hyponyms characterized by their color.	▪ COLORLESS SOLID ▪ RED TIDE ▪ YELLOW LIQUID
Composition-based	Hyponyms characterized by their components or by their material.	▪ CARBONATE SAND ▪ METALLIC ELEMENT ▪ PINE FOREST
Degree-based	Hyponyms characterized by their degree of intensity, size or consequences.	▪ CATACLYSMIC ERUPTION ▪ LOW-MAGNITUDE EARTHQUAKE ▪ MEGA-SCALE EXTRACTION
Denomination-based	Hyponyms characterized by having a particular denomination with a proper noun.	▪ NEW YORK CITY ▪ PACIFIC OCEAN ▪ SAHARA DESERT
Density-based	Hyponyms characterized by their density or particle concentration.	▪ DENSE WATER ▪ HEAVY METAL ▪ LIGHT ELEMENT
Domain-based	Hyponyms characterized by the scientific or knowledge field to which they belong.	▪ AGRICULTURAL PRODUCT ▪ CHEMICAL INDUSTRY ▪ MUSICAL INSTRUMENT
Effect-based	Hyponyms characterized by the effects or consequences that they cause.	▪ GREENHOUSE GAS ▪ HAZARDOUS SUBSTANCE ▪ TOXIC LIQUID
Function-based	Hyponyms characterized by their function or purpose.	▪ DRINKING WATER ▪ MANUFACTURING FACILITY ▪ SURVEILLANCE RADAR
Hardness-based	Hyponyms characterized by their hardness level.	▪ HARD ROCK ▪ HARD STRUCTURE ▪ SOFT WOOD
Height-based	Hyponyms characterized by their height or depth level.	▪ DEEP OCEAN ▪ HIGH TIDE ▪ SHALLOW WATER
Location-based	Hyponyms characterized by their spatial location or position.	▪ OCEAN WATER ▪ SURROUNDING AIR ▪ TROPICAL STORM

Method-based	Hyponyms characterized by the method or the process that they involve.	▪ AEROBIC OXIDATION ▪ DIRECT SUBLIMATION ▪ INDUSTRIAL TREATMENT
Moisture-based	Hyponyms characterized by their moisture level.	▪ ARID DESERT ▪ DRY SOLID ▪ SATURATED AIR
Movement-based	hyponyms characterized by their movement or direction	▪ EBB TIDE ▪ OCEAN-GOING DREDGE ▪ OUTGOING RADIATION
Origin-based	Hyponyms characterized by their origin (i.e., the place where they come from or where they were created).	▪ COUNTRY ROCK ▪ NATURAL RESOURCE ▪ PINE WOOD
Patient-based	Hyponyms characterized by the patient that is affected by them.	▪ COAST EROSION ▪ ICE MELTING ▪ WATER TREATMENT
Relation-based	Hyponyms characterized by being related to other concepts.	▪ COVALENT SOLID ▪ FOREIGN SUBSTANCE ▪ PARENT COMPOUND
Result-based	Hyponyms characterized by the result that they cause, or by being the result of a process.	▪ PAPER INDUSTRY ▪ TSUNAMIGENIC EARTHQUAKE ▪ UNIMOLECULAR DECOMPOSITION
Shape-based	Hyponyms characterized by their shape.	▪ AMORPHOUS SOLID ▪ L-SHAPED GROIN ▪ PARABOLIC DUNE
Size-based	Hyponyms characterized by their size.	▪ COMPACT CAR ▪ GIANT PLANET ▪ TINY CRYSTAL
Speed-based	Hyponyms characterized by their speed.	▪ FLASH EVAPORATION ▪ RAPID EROSION ▪ SPONTANEOUS DECOMPOSITION
State-based	Hyponyms characterized by the state of matter.	▪ FLUID ELEMENT ▪ MOLTEN ROCK ▪ SOLID SUBSTANCE
Status-based	Hyponyms characterized by a particular circumstance or situation.	▪ CONTAMINATED SOIL ▪ REGULATED SUBSTANCE ▪ UNTREATED WOOD
Technology-based	Hyponyms characterized by the technology that they use.	▪ DIGITAL BAROMETER ▪ GREEN TECHNOLOGY ▪ MOTOR VEHICLE
Temperature-based	Hyponyms characterized by their temperature.	▪ COLD AIR ▪ HOT GAS ▪ WARM OCEAN
Texture-based	Hyponyms characterized by their texture.	▪ FINE SAND ▪ SOFT ROCK ▪ VISCOUS LIQUID
Time-based	Hyponyms characterized by their duration, by their age, or by happening in a particular moment.	▪ ANNUAL PRECIPITATION ▪ OLD ROCK ▪ WINTER ICE
Weight-based	Hyponyms characterized by their weight.	▪ HEAVY-DUTY TRUCK ▪ LIGHT TRUCK ▪ LIGHT-DUTY VEHICLE

ANNEX III

HYPONYMIC KNOWLEDGE PATTERNS: DESCRIPTION AND PATTERNS (GIL-BERROZPE 2017; GIL-BERROZPE *ET AL*. 2017)

HYPONYMIC KP TYPE	DESCRIPTION	PATTERNS
Classification	They classify or divide the hypernym into hyponyms.	**1.** HYPER *be* classified into HYPO **2.** HYPER *be* divided into HYPO **3.** HYPO *be* classified in HYPER **4.** types of HYPER *be* classified as HYPO **5.** HYPER *be* classified into types, namely HYPO
Definition	They introduce the hyponym with a definition where the hypernym is the *genus*.	**1.** HYPO *be* defined as HYPER **2.** HYPO, defined as HYPER **3.** HYPO: a HYPER **4.** HYPO: a type of HYPER
Denomination	They introduce the hyponyms as particular denominations.	**1.** a HYPER called HYPO **2.** a type of HYPER called HYPO **3.** a type of HYPER known as HYPO **4.** a type of HYPER referred to as HYPO **5.** a type of HYPER termed HYPO **6.** a type of HYPER, called HYPO **7.** a type of HYPER, known as HYPO **8.** a type of HYPER, named a HYPO **9.** a type of HYPER, termed HYPO **10.** HYPO refers to HYPER **11.** types of HYPER *be* called HYPO **12.** types of HYPER *be* known as HYPO **13.** types of HYPER *be* referred to as HYPO
Enumeration	They show an exhaustive and numbered list of hyponyms for the hypernym.	**1.** # HYPER: HYPO **2.** # types of HYPER *be* considered: HYPO **3.** # types of HYPER *be* distinguished: HYPO **4.** # types of HYPER *be* HYPO **5.** # types of HYPER *be* identified: HYPO **6.** # types of HYPER *be* recognized: HYPO **7.** # types of HYPER *be* required: HYPO **8.** # types of HYPER *be*: HYPO **9.** # types of HYPER occur: HYPO **10.** # types of HYPER: HYPO **11.** of HYPER, # types *be* HYPO **12.** a type of HYPER, divided into # types: HYPO **13.** HYPER *be* classified in # types, HYPO **14.** HYPER *be* classified into # types, HYPO **15.** HYPER *be* classified into # types: HYPO **16.** HYPER *be* divided into # types, HYPO **17.** HYPER *be* divided into # types: HYPO **18.** # types of HYPER called HYPO **19.** # examples of HYPER *be* HYPO **20.** # types of HYPER *be* included: HYPO **21.** # types of HYPER, ranging from HYPO to HYPO **22.** # main types of HYPER *be* HYPO **23.** # main types of HYPER, namely HYPO **24.** # main types of HYPER: HYPO **25.** # types of HYPER, namely HYPO **26.** # types of HYPER, typically HYPO

Exemplification	They present the hyponyms as examples, types or kinds of the hypernym.	**1.** examples of HYPER *be* HYPO **2.** HYPER (e.g. HYPO) **3.** HYPER (HYPO) **4.** HYPER (such as HYPO) **5.** HYPER like HYPO **6.** HYPER species (such as HYPO) **7.** HYPER species such as HYPO **8.** HYPER species, such as HYPO **9.** HYPER such as HYPO **10.** HYPER types (such as HYPO) **11.** HYPER types such as HYPO **12.** HYPER types, such as HYPO **13.** HYPER, such as HYPO **14.** HYPO *be* an example of a HYPER
Identification	They directly link the hyponym to the hypernym with a copulative verb.	**1.** a HYPER *be* HYPO **2.** a type of HYPER *be* a HYPO **3.** a type of HYPER, a HYPO **4.** HYPO *be* a HYPER **5.** HYPO *be* a type of HYPER **6.** HYPO, a type of HYPER **7.** other HYPER *be* HYPO **8.** types of HYPER *be* HYPO
Inclusion	They present the hyponyms as concepts included in the notion of the hypernym.	**1.** a type of HYPER that includes HYPO **2.** among HYPER *be* HYPO **3.** HYPER include HYPO **4.** HYPER including HYPO **5.** HYPER species including HYPO **6.** HYPER species, including HYPO **7.** HYPER type, which includes HYPO **8.** HYPER types include HYPO **9.** HYPER types including HYPO **10.** HYPER types, including HYPO **11.** HYPER, including HYPO **12.** included in this type of HYPER *be* HYPO
Itemization	They introduce a non-exhaustive list of hyponyms for the hypernym.	**1.** HYPO and other HYPER **2.** HYPO and other HYPER species **3.** HYPO and other HYPER types **4.** HYPO, and other HYPER **5.** HYPO, and other HYPER types **6.** HYPO. These types of HYPER… **7.** types of HYPER: HYPO **8.** types of HYPER *be*: HYPO **9.** HYPO and other HYPER classified as HYPO **10.** HYPO and other HYPER (such as HYPO) **11.** HYPO and other HYPER such as HYPO **12.** HYPO and other HYPER, such as HYPO **13.** HYPO, and other HYPER such as HYPO **14.** HYPER include: HYPO **15.** HYPER types include: HYPO **16.** HYPER types including: HYPO **17.** HYPO and other HYPER (including HYPO) **18.** HYPO and other HYPER including HYPO **19.** HYPO and other HYPER, including HYPO

Range	They establish a span where several hyponyms can be found for the same hypernym.	**1.** HYPER ranging from HYPO to HYPO **2.** HYPER types ranging from HYPO to HYPO **3.** HYPER types, ranging from HYPO to HYPO **4.** HYPER, ranging from HYPO to HYPO
Selection	They highlight main or preferred hyponyms for the hypernym.	**1.** HYPER (especially HYPO) **2.** HYPER primarily HYPO **3.** HYPER species, especially HYPO **4.** HYPER species, mainly HYPO **5.** HYPER, characteristically HYPO **6.** HYPER, especially HYPO **7.** HYPER, generally HYPO **8.** HYPER, mainly HYPO **9.** HYPER, namely HYPO **10.** HYPER, primarily HYPO **11.** HYPER, typically HYPO **12.** HYPER, usually HYPO **13.** HYPO *be* the main types of HYPER **14.** the main types of HYPER *be* HYPO **15.** types of HYPER, mainly HYPO **16.** types of HYPER, primarily HYPO **17.** types of HYPER, typically HYPO **18.** types of HYPER, usually HYPO **19.** the main example of HYPER *be* HYPO **20.** the main types of HYPER include HYPO

ANNEX IV
RESUMEN EXTENSO Y CONCLUSIONES DE LA TESIS EN ESPAÑOL

La terminología es el estudio del lenguaje especializado, es decir, de los términos y de las frases utilizadas en dominios científico-técnicos. Aunque se puede interpretar de muchas formas distintas, la terminología es una materia transdisciplinar que no solamente incluye aspectos lingüísticos, sino también extralingüísticos, tales como elementos relacionados con la percepción humana y procesos computacionales. La terminología, concretamente, nació de la necesidad de unificar conceptos y términos en dominios especializados para poder facilitar la comunicación profesional y la transferencia de conocimiento.

La mayoría de teorías de terminología tienen aplicaciones prácticas, como las enciclopedias, los diccionarios especializados, las bases de conocimiento y otros recursos terminológicos o traductológicos, que son precisamente los puntos clave de los diferentes enfoques. Estos recursos, idóneamente, muestran la información de manera que pueda ser fácilmente adquirida y usada por distintos perfiles de usuario. Esta faceta más práctica de la terminología, la que está enfocada en describir y representar sistemáticamente la información terminológica previamente recabada, es lo que habitualmente se conoce como terminografía.

La terminología ha evolucionado desde teorías prescriptivas y estáticas hasta enfoques cognitivos y dinámicos. Gracias a estos puntos de vista modernos, se han realizado grandes avances en el diseño y en la elaboración de recursos terminológicos. Con el paso de los años, los formatos tradicionales en papel de glosarios y diccionarios se han ido gradualmente reemplazando por versiones electrónicas o digitales, que también pueden actualizarse y modificarse con mayor facilidad. En los últimos años, las bases de conocimiento terminológico se han convertido en uno de los recursos lingüísticos más importantes de todos, ya que muestran una gran variedad de información lingüística y no lingüística a través de interfaces intuitivas.

Un ejemplo representativo de una base de conocimiento terminológico moderna es EcoLexicon. Se trata de un recurso multidimensional y dinámico sobre ciencias medioambientales que proporciona información conceptual, lingüística, fraseológica y multimodal en cada una de sus entradas. EcoLexicon, además de por su enfoque ontológico, también se caracteriza por su representación de las redes conceptuales, que muestran cómo están los conceptos relacionados entre sí mediante diferentes relaciones semánticas o conceptuales: genérico-específicas, partitivas y no jerárquicas. De todas estas relaciones, las genérico-específicas o hiponímicas son especialmente relevantes para la terminología porque están directamente implicadas en la categorización conceptual y en las jerarquías de términos. Por este motivo, se puede mejorar la descripción de los conceptos y términos haciendo énfasis en su información hiponímica.

La hiponimia describe una relación jerárquica entre un concepto genérico o padre (representado por un hiperónimo) y un concepto específico o hijo (representado por un hipónimo). Esta relación conceptual es relevante para la terminología y para el desarrollo de recursos lingüísticos, sobre todo, porque es la piedra angular de todas las relaciones conceptuales. Además, juega un rol fundamental en nuestro pensamiento consciente sobre el significado de las palabras.

Las relaciones hiponímicas son complejas y, por lo tanto, las parejas de hiperónimos e hipónimos se pueden estudiar teniendo en cuenta diferentes criterios. Esta investigación sigue las premisas de la terminología basada en marcos para estudiar la descripción, categorización y representación de la hiponimia en la terminología, centrándose concretamente en la terminología medioambiental. La descripción de la hiponimia se basa en un análisis de su relevancia desde el punto de vista de la terminología y de las ontologías. Su categorización se basa en su clasificación de acuerdo con múltiples perspectivas y diferentes enfoques hacia el refinamiento de la hiponimia. Finalmente, su representación se basa en analizar cómo se suele ilustrar la hiponimia tanto en recursos terminológicos tradicionales como contemporáneos.

Además, como contribución e innovación al estudio de la hiponimia, esta tesis doctoral también contribuye a la descripción, categorización y representación de la hiponimia mediante la presentación de una nueva mejora hiponímica para EcoLexicon. Se conoce como HypoLexicon y se trata de un módulo independiente bajo la forma de un recurso terminológico diseñado para facilitar la representación y la adquisición de información hiponímica.

Por lo tanto, esta tesis doctoral parte de un repaso al marco teórico general, compuesto de las disciplinas relacionadas con el tratamiento de la hiponimia desde una perspectiva conceptual, concretamente, la lingüística cognitiva y la terminología basada en marcos. Dado que la lingüística cognitiva es la rama de la lingüística que incorpora los principios de la psicología y de la neurociencia, está considerada en esta investigación como el mejor enfoque hacia la hiponimia desde una perspectiva conceptual o cognitiva.

Debido a su influencia en la teoría sobre la terminología, se le presta especial atención a los siguientes elementos: (i) la conceptualización, entendida como el proceso por el que ciertas entidades se organizan de acuerdo con sus características más destacadas, y que incluye materias como los sistemas conceptuales, la cognición y los marcos; (ii) la categorización, entendida como el proceso mental que permite que los humanos clasifiquemos los elementos del mundo al percibir similitudes y diferencias entre ellos, al mismo tiempo que se almacenan los conceptos mentalmente en la memoria a largo plazo; y (iii) las relaciones conceptuales, entendidas como los enlaces entre conceptos, que pueden ser tanto jerárquicas como no jerárquicas.

La terminología basada en marcos, la teoría en la que se basa esta investigación, es un enfoque cognitivo hacia la terminología que se centra en aspectos semánticos de la representación de conocimiento especializado. Esta teoría es primero ubicada en el contexto de otros enfoques teóricos, tales como la teoría general de la terminología, la socioterminología, la teoría comunicativa de la terminología y la teoría sociocognitiva de la terminología.

A continuación, se realiza una descripción de los fundamentos teórico-prácticos de la terminología basada en marcos, haciendo énfasis en los siguientes aspectos: (i) la organización conceptual, basada en los marcos que estructuran los dominios especializados y que crean representaciones independientes del lenguaje, así como en los eventos que describen los procesos dentro de un dominio de especialidad; (ii) la multidimensionalidad, entendida como la clasificación conceptual que surge al percibir los conceptos de más de una sola forma dentro de un sistema conceptual en función de sus diferentes características; (iii) las plantillas definicionales basadas en marcos, usadas para describir entidades del mundo real mediante un concepto genérico y características diferenciadoras; y (iv) la extracción de conocimiento, basada en las premisas de la lingüística de corpus, concretamente en las estrategias y técnicas de análisis de corpus, para así obtener fácilmente la información terminológica relevante a partir de grandes corpus especializados.

Tras esto, se proporciona una descripción de EcoLexicon, la base de conocimiento terminológico sobre el medio ambiente que es la aplicación práctica de la terminología basada en marcos. Se analizan sus elementos clave en cuanto a la representación de información conceptual, lingüística, multimodal, ontológica y fraseológica. Para ello, se ubica EcoLexicon en el contexto de otros recursos terminológicos afines.

Esta sección sobre el marco teórico general subraya las siguientes conclusiones:

- Los aspectos de la lingüística cognitiva relacionados con la conceptualización, la categorización y la representación de relaciones conceptuales son esenciales para entender todas las dimensiones de la hiponimia.
- El análisis cronológico de las teorías sobre la terminología proporciona una mejor comprensión de las contribuciones realizadas por la terminología basada en marcos.
- La terminología basada en marcos es la mejor teoría para el tratamiento de las relaciones conceptuales, específicamente de la hiponimia, gracias a la importancia que se le presta a la organización conceptual y a la multidimensionalidad.
- La metodología de la terminología basada en marcos ofrece una base fundamentada para la creación de definiciones y la extracción de conocimiento mediante análisis de corpus.

A este apartado le sigue la descripción, categorización y representación de la hiponimia. Como es bien sabido, la hiponimia es la relación conceptual entre un hiperónimo y un hipónimo. Todas las jerarquías conceptuales se basan en la hiponimia y juega el papel más esencial en nuestro pensamiento sobre el sentido de las palabras; de aquí deriva su importancia tanto desde un punto de vista cognitivo como terminológico.

En primer lugar, la descripción de la hiponimia se analiza desde las perspectivas de la terminología y de la ontología. La hiponimia ha sido ampliamente estudiada en distintas teorías lingüísticas y terminológicas por su naturaleza causante de inferencias, su importancia en las definiciones y su relevancia en las restricciones selectivas en la gramática. La descripción de la hiponimia también incluye cuatro fenómenos relacionados entre sí: (i) la taxonomía, una variante de la hiponimia que describe una relación de clasificación entre el hiperónimo y el hipónimo; (ii) la incompatibilidad, una relación de exclusión presente en ciertas jerarquías conceptuales; (iii) la troponimia, una relación temporalmente inclusiva empleada para describir la hiponimia verbal; y (iv) la autohiponimia, que se manifiesta cuando un término tiene tanto un sentido general por defecto como un sentido restringido contextualmente.

Dado que la hiponimia también es importante para la ciencia computacional y la ingeniería ontológica por su influencia en los modelos de redes, la relación genérico-específica también se analiza centrándose en los siguientes aspectos relacionados con la terminología: (i) las ontologías, entendidas como bases de datos que describen los conceptos de un campo de conocimiento, sus propiedades y sus relaciones; (ii) la termontografía, un enfoque multidisciplinar que combina teorías y metodologías de la teoría sociocognitiva de la terminología y de la ingeniería ontológica; (iii) la ontoterminología, un modelo que describe una terminología cuyo sistema conceptual es una ontología formal basada en principios epistemológicos; y (iv) la mejora de conocimiento ontológico en EcoLexicon, a partir de la cual se clasificaron todos los conceptos en una jerarquía de categorías conceptuales.

En segundo lugar, se analiza la categorización de la hiponimia, que presenta diferentes tipologías con matices específicos, tales como los siguientes: (i) la hiponimia taxonómica y la funcional, según las cuales los hipónimos expresan lo que son o para lo que se usan; y (ii) la hiponimia directa y la indirecta, basadas en la proximidad conceptual o en la distancia entre hiperónimos e hipónimos. También se explica el estudio de caso del refinamiento de la hiponimia en EcoLexicon, que consistió en la especificación de subtipos de hiponimia típicos de la terminología medioambiental y en la identificación de patrones de conocimiento hiponímicos.

En tercer lugar, la representación de la hiponimia se estudia a través de su visualización y manifestación en varios recursos terminológicos. Al tratarse de una relación conceptual de gran importancia para la organización del conocimiento, la hiponimia es difícil de representar en los enfoques más tradicional debido a sus formatos. Por lo tanto, se analiza tanto en diccionarios, enciclopedias y bancos de términos (por ejemplo, IATE y TERMIUM Plus), como en bases de conocimiento terminológico (por ejemplo, WIPO Pearl y EcoLexicon). A continuación, se comenta una serie de criterios para una mejor representación de la hiponimia en vista a proponer una nueva forma de manifestar la hiponimia basada en entradas terminológicas con una estructura y un contenido centrados en la descripción y categorización de relaciones genérico-específicas. Estos son dichos criterios:

- Las entradas terminológicas deberían estar estructuradas jerárquicamente en distintos niveles hiponímicos para representar mejor la transitividad de hiperónimos a hipónimos directos en lugar de a hipónimos indirectos.

- Las entradas terminológicas deberían contener definiciones intensionales o terminológicas de los conceptos, basadas en *genus* y *differentiae*, para así ilustrar la herencia de propiedades de hiperónimos a hipónimos y para distinguir claramente entre cohipónimos.
- Las entradas terminológicas deberían describir los conceptos de acuerdo con categorías conceptuales para así resaltar los cambios ontológicos sucedidos desde ciertos hiperónimos hasta ciertos hipónimos.
- La hiponimia debería estar categorizada en subtipos que identifiquen el matiz que diferencia a un hipónimo de su hiperónimo directo, y los cohipónimos deberían estar clasificados en función del subtipo de hiponimia implicado.
- Las entradas terminológicas deberían reflejar contextos hiponímicos para revelar cómo se codifica la hiponimia en patrones de conocimiento y cómo se expresa en el lenguaje especializado.

Como resultado, analizar la hiponimia con un enfoque que combine la terminología y la ontología resalta el hecho de que la hiponimia es la relación más importante para la conceptualización. De hecho, como la hiponimia puede clasificarse de acuerdo con diferentes dimensiones o microsentidos, estos matices deberían estar reflejados en recursos de conocimiento que representen relaciones genérico-específicas. Sin embargo, hasta ahora no ha existido un recurso capaz de proporcionar una descripción, categorización y representación satisfactorias de la hiponimia.

Los materiales y métodos de esta investigación fueron cuatro subcorpus especializados (BIO, CHEM, CIV, GEO) derivados del corpus en inglés de EcoLexicon. En cambio, también se necesitó utilizar una selección de recursos terminológicos especializados, particularmente diccionarios y enciclopedias, para ayudar en la elaboración de las definiciones terminológicas y para mejorar las jerarquías conceptuales. Las aplicaciones informáticas utilizadas fueron las siguientes: (i) la aplicación interna de EcoLexicon, para compilar los subcorpus medioambientales; (ii) Sketch Engine, para procesar los subcorpus y extraer toda la información hiponímica; y (iii) Lexonomy, para diseñar la plantilla terminológica hiponímica.

Esta investigación se realizó en cuatro fases: (i) la extracción y compilación de los cuatro subcorpus; (ii) el análisis de los subcorpus para extraer, identificar y seleccionar los hiperónimos e hipónimos; (iii) la creación de las jerarquías conceptuales (incluyendo las definiciones terminológicas, las categorías conceptuales, los subtipos de hiponimia y los contextos hiponímicos); y (iv) el diseño de la plantilla terminológica utilizando una estructura basada en XML. Esta metodología, por lo tanto, combinó el uso del análisis de corpus para la extracción de la información a través de búsquedas basadas en *word sketches* y en CQL, y las bases de la terminografía para el diseño de la plantilla terminológica para las entradas terminológicas hiponímicas.

A lo largo de esta investigación, se necesitó validar los materiales para asegurar que cumplían con estándares profesionales y así realizar un trabajo terminológico de alta calidad. Además, las técnicas y estrategias de análisis de corpus facilitaron en gran medida la extracción, identificación y selección de la información terminológica a partir de grandes corpus. No en vano, la combinación de recursos terminológicos y de aplicaciones informáticas de diferentes tipos mejoró en gran medida los resultados obtenidos. En este sentido, los criterios para la elaboración de definiciones y jerarquías conceptuales propios de la terminología basada en marcos fueron especialmente útiles de cara a la representación de la información hiponímica. De hecho, solo de esta forma se consigue que las jerarquías conceptuales representen eficazmente las definiciones terminológicas, las categorías conceptuales, los subtipos de hiponimia y los contextos hiponímicos.

En cuanto a los resultados, la principal aportación fue la creación de doce entradas terminológicas hiponímicas que proporcionan una descripción, categorización y representación adecuadas de la terminología medioambiental. Esto se sumó al diseño y a la publicación de un recurso terminológico hiponímico electrónico que contiene y muestra dichas entradas a través de una interfaz sencilla y accesible.

La estructura y el formato de estas entradas facilitan la representación de su contenido hiponímico. Esto se consiguió haciendo hincapié en jerarquías conceptuales con subtipos de hiponimia a distintos niveles y poblando la información de cada concepto con sinónimos, categorías conceptuales, definiciones terminológicas y contextos hiponímicos.

Tras presentar las doce entradas, se analiza y resume el contenido presente en cada una de ellas. Se identificaron matices hiponímicos que afectan a la verticalidad y a la horizontalidad de las jerarquías. También se hallaron dimensiones o microsentidos expresados a través de los subtipos de hiponimia. Además, merece también prestar atención a las alteraciones en características o rasgos producidas por la adición de categorías conceptuales en niveles hiponímicos más específicos.

Las entradas fueron publicadas en HypoLexicon, la aplicación práctica de esta investigación. Este recurso terminológico, diseñado como parte de esta tesis doctoral, se centra específicamente en la descripción, categorización y representación de la hiponimia en conceptos medioambientales. Incluye información definicional, relacional, ontológica y contextual sobre hiperónimos e hipónimos especializados propios de la terminología medioambiental.

Como tal, HypoLexicon es el punto de convergencia de cuatro recursos: (i) EcoLexicon, por la estructura y la información básicas de las entradas terminológicas; (ii) el corpus en inglés de EcoLexicon y los cuatro subcorpus especializados, por la población y la mejora de las entradas terminológicas; (iii) Sketch Engine, por la extracción de información hiponímica y contextual a través del análisis de corpus; y (iv) Lexonomy, por el diseño de la plantilla terminológica y por la implementación de toda la información bajo la forma de un recurso terminológico de verdad.

Por último, se ofrece un resumen de los resultados estadísticos sobre la información presente en todas las entradas terminológicas hiponímicas relacionado con las categorías conceptuales y con los subtipos de hiponimia. Estos datos se comparan también con la información terminológica original de EcoLexicon para así revelar las mejoras implementadas en HypoLexicon. Los resultados revelaron que hubo un incremento significativo en cuanto al contenido conceptual, lingüístico y relativo a la hiponimia de las jerarquías conceptuales relacionadas con las doce entradas terminológicas hiponímicas.

Por consiguiente, se pueden extraer las siguientes conclusiones de esta investigación:

- Las entradas terminológicas hiponímicas demostraron ser un enfoque exitoso hacia la descripción, categorización y representación de la hiponimia gracias a su estructura jerárquica y a la clasificación gráfica de la información basada en definiciones y en el análisis de corpus.
- La visualización de la información hiponímica en las entradas terminológicas hiponímicas permitió la identificación de fenómenos dinámicos sobre las relaciones genérico-específicas (por ejemplo, matices hiponímicos en la verticalidad y en la horizontalidad de las jerarquías conceptuales, diferentes dimensiones o microsentidos de los cohipónimos, alteraciones en las características de los conceptos mediante la adición de categorías conceptuales a niveles hiponímicos más específicos, etc.).

- HypoLexicon, el recurso terminológico hiponímico, tiene una estructura y una interfaz excelentes para revelar toda la información de las entradas terminológicas hiponímicas.
- HypoLexicon, que está disponible en línea de forma gratuita, facilita la adquisición y la transferencia de conocimiento especializado entre distintos tipos de usuarios (por ejemplo, terminólogos, lingüistas, traductores, expertos técnicos y usuarios legos).

Por lo tanto, esta investigación conduce a un nuevo tipo de descripción, categorización y representación de la hiponimia. Esta metodología también se puede aplicar a cualquier otro tipo de dominio especializado, e incluso podría servir como una manera accesible de tratar con la hiponimia en recursos de lenguaje general igualmente. En resumen, el objetivo de la metodología y del recurso propuestos es facilitar la adquisición a todos los niveles.

Las futuras líneas de trabajo de esta investigación tomarán dos caminos. Por un lado, HypoLexicon continuará creciendo y siendo nutrido con más contenido al crear entradas terminológicas adicionales con todo tipo de información hiponímica extraída con técnicas de corpus. Estas nuevas entradas, además, podrían pertenecer a los mismos subdominios ambientales o incluso a nuevos subdominios para así extender el repertorio de categorías conceptuales y de subtipos de hiponimia.

Sin embargo, tal vez la idea más innovadora sería más bien buscar la integración total de HypoLexicon en EcoLexicon. De esta manera, dejaría de ser un módulo o un producto independiente y pasaría a convertirse en una parte integral del recurso original en el que se basa. Por este motivo, esta tesis doctoral sienta un fundamento que podría continuarse en futuras investigaciones. No solamente deja la puerta abierta a estas nuevas líneas de trabajo, sino que también pretende fomentar la interoperabilidad con otras fuentes de información y con otras herramientas y recursos ontológicos, lingüísticos y terminológicos.