Discovering Ontological Knowledge in Unstructured Recipes of a Portuguese Monk from the 16th Century

Orlando Belo[1][a], Bruno Silva[1][b] and Anabela Barros[2][c]
[1]ALGORITMI Research Centre/LASI, University of Minho, 4710-059 Braga, Portugal
[2]CEHUM, Centre for Humanistic Studies, University of Minho, 4710-059 Braga, Portugal

Keywords: Ontology Learning, Knowledge Extraction, Unstructured Textual Data, Natural Language Processing, Linguistic Ontologies, Graph Databases.

Abstract: Ontology learning is often applied to textual data sources with the aim of identifying, extracting and representing their various data elements, as well as their semantic relationships. Ontologies are excellent instruments for the representation of knowledge about one or more domains of knowledge, which enable us to study in detail the knowledge of the domain they host. With this in mind, we devised and developed a semi-automatic ontology learning system. It was specifically oriented for discovering the knowledge contained in a set of ancient texts of culinary recipes of a monk of the 16th century. Using the system, we produced an ontology incorporating a large diversity of culinary elements and their relationships, which offer a very rich field of research of the culinary of the 16th century in Portugal – the ontology was exposed and explored using the native mechanisms of a graph database management system.

1 INTRODUCTION

Over the past few years, ontologies have probably been one of the most widely used models of knowledge representation in real-world applications, particularly in organizations that need to acquire and manage knowledge about one or more application domains (van Eijnatten, 2004). Since the emergence of the Semantic Web (Ding et al., 2007), ontological models have been catapulted into the spotlight, coming to be widely used in various computational systems (Davies et al., 2003). Since then, several studies have been carried out in the area of ontologies, due to the determining role they play in data organization and in the representation of the relationships established with it, in any area or domain of application.

An ontology (Guarino et al., 2009) (Keet, 2018) defines a set of concepts related to each other, with the purpose of creating interoperability between them and forming a semantic link to the domain in question. The semantics illustrated by an ontology consist of concepts, relationships, and properties. Concepts relate to each other through the establishment of semantic relationships that characterize the way they can be conjugated. In this way, it is possible to work and reason about the context and application field of concepts that an ontology welcomes, creating a real perception of what the domain presents. Ontologies have a great application level, being transversal to any existing problem in the context of the real world. Application areas can range from health sciences to e-commerce, or any other area that needs to explore shareable and reusable information within a particular knowledge domain. In addition to being a scalable and sustainable way of storing information, an ontology aims at better management of the knowledge extracted from a knowledge domain. An ontology allows for faster analysis and under-standing of what is exposed in the source of knowledge. The process of conceiving an ontology often referred as ontology learning (Cimiano et al., 2009) (Asim et al., 2018), encompasses all the tasks we use to perform to extract the terms, concepts, relationships, rules or axioms from textual sources to get the structure of an

[a] https://orcid.org/0000-0003-2157-8891
[b] https://orcid.org/0009-0000-5760-3193
[c] https://orcid.org/0000-0002-2959-9200

Belo, O., Silva, B. and Barros, A.
Discovering Ontological Knowledge in Unstructured Recipes of a Portuguese Monk from the 16th Century.
DOI: 10.5220/0012189500003598
In *Proceedings of the 15th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2023) - Volume 2: KEOD*, pages 200-207
ISBN: 978-989-758-671-2; ISSN: 2184-3228

ontology. Among all the proposed methods for ontology learning on texts, the proposal of Brewster (2006) stands out significantly, given the simplicity of its task structure and application model. However, the nature of an ontology learning process is quite complex, usually requiring the application of very sophisticated processing techniques, which involve frequently natural language processing techniques (Sharma, 2021) and machine learning algorithms (Zhang et al., 2018).

In this paper, we present and discuss a semi-automatic ontology learning system. It was conceived specifically for extracting a culinary ontology from a textual source of the 16th century, which were collected by the monk José Joaquim de Santa Teresa, recorded in the library of the Monastery of S. Martinho de Tibães, Braga, Portugal, and published by Barros in "As Receitas de um Frade Portuguese no séc. XVI" (“The Recipes of a Portuguese Monk in the 16th century”) (Barros, 2013). The referred data source hosts about three hundred recipes, in disparate formats, their ingredients and preparation processes, which are very important elements for the study of Portuguese cuisine during this time. The ontology produced incorporates a large diversity of culinary elements and their relationships, offering a very rich field of research for students, professors, researchers and individual users and providing a very interesting view of the culinary practices of 16th in Portugal. Additionally, we will present the process of discovering and extraction the ontology we carried out, as well as describe all its main tasks and techniques used, from the preparation of the texts of the recipes to their processing and generation of the ontology. Next, section 2 presents and discusses several aspects related to ontology learning aspects, section 3 introduces the book of recipes, presenting and describing our application case, and exposes how we extracted the ontology of recipes, and, finally, section 4, presents some conclusions and future work.

2 RELATED WORK

Ontologies (Guarino et al., 2009) (Keet, 2018) have long proven their usefulness and versatility in knowledge acquisition and management. Currently, its application is very widespread. We can find ontological systems in the most varied fields of application, such as medicine, management, economics, or linguistics. Gruber (1995) proposed the most consensual definition in the field of Computer Science for an ontology, saying that it “is an explicit specification of a conceptualization”. However, many other authors proposed their own definitions of ontologies, such as presented in (Borst, 1997) or (Uschold and Gruninger, 1996). However, all of them based their definition in a specification of a shared conceptualization of the knowledge of one or more specific application domains. They are very important instruments in processes of analysis and application of knowledge. However, contrary to the process of exploring the knowledge of an ontology, all those who are or have been involved in a process of defining and characterizing an ontology know that this is not an easy process. In order for us to be able to successfully learn an ontology, we have to have at our disposal credible data sources, as well as people with the expertise and knowledge in the domain of the knowledge involved and in the process of building the required ontology, whether manual or automatic.

The manual construction of an ontology is often too time-consuming for experts in the field, as it requires a broad analysis of the context concerned and the corresponding identification of all the elements relevant to the ontological constitution. The research of this theme, nowadays, falls mainly on the semi-automatic and automatic construction, in order to overcome the problems identified in the process of manual construction of the ontology. The processes of ontology learning (Mishra and Jain, 2014) aim to represent the semantic relationships between the various elements present in a data set. Today, it is common to choose to carry out this work in textual, unstructured data sources, given the enormous richness of its content (Choudhary and Tomar, 2014) (Belhoucine and Mourchid, 2020). In this domain, one of the great purposes of ontologies is to identify and characterize relevant elements in the contents of texts, as well as to establish the relationships between them. Then, through control and automation mechanisms, ensure the execution of fast and effective analysis processes of the contents addressed and treated in these texts.

In (Brewster, 2006) it was used an image of a “layer cake” for representing the order that an ontology learning process must execute the extraction tasks (Figure 1). The author ordered the extraction tasks based on their difficulty to perform. The Brewster’s proposal puts on the base of the cake the most simply task, the identification of terms (task 1) containing in the texts. In the second and third layer appears, respectively, the tasks of synonyms (task 2) and concept identification (task 3), which in some cases can be performed simultaneously. Next, we need to establish the concept hierarchies (task 4), followed by the identification of relationships (task 5) between concepts. Finally, at the top level, appears

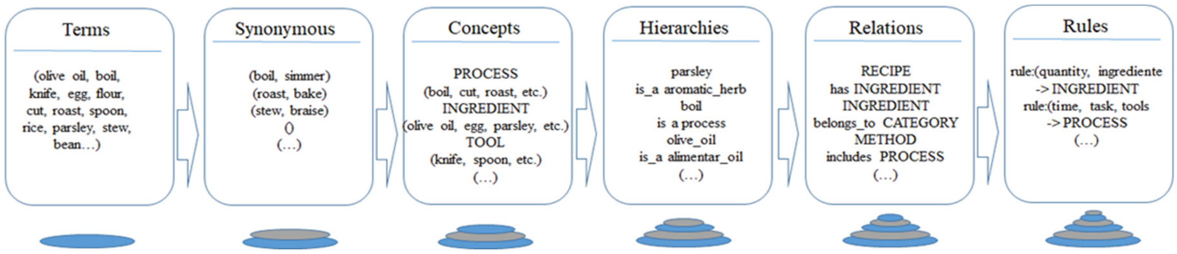

Figure 1: The tasks of an ontology learning process according to (Brewster, 2006).

the task for discovering rules (task 6). As happens in other fields, this is not an exclusive method for organizing an ontology learning process. To note, (Tiwari and Jain, 2014) added some value to the proposal done in (Brewster, 2006) saying that some of the tasks can be executed simultaneously, which speed up the extraction process time. Later, in (Asim et al., 2018) were added some new tasks to the proposal made in (Tiwari and Jain, 2014). It added to the process two very important tasks: data pre-processing and ontology evaluation. They formally included two tasks we really do in practice before starting and after executing a conventional extraction process.

In the field of cooking, several proposals for ontologies have been presented, some of them quite specific. For example, we have the works of (Noy and McGuinness, 2001) or (Graça et al., 2005), which presented ontologies for the enology field, or (Markantonatou et al., 2021), which modelled the domains of dishes that figured in a large number of menus quite diverse. In the field of cooking recipes, we find also several interesting examples of ontologies. For example, in (Villarías, 2004), it is presented an ontology for cooking recipes to be used by a Web semantic querying system, or in (Batista et al., 2006) it was devised and implemented an ontology for the cooking domain, involving concepts such as actions, food, recipes, and utensils. As well, in (Monica et al., 2014) is presented an ontology for culinary processes defined in recipes and the authors of (Nanba et al., 2014) created an ontology for the establishment of hyponymy, synonymy, attributes, and meronymy in cuisine recipes. The definition of another ontology for cooking recipes was not the motivation of this work, but rather the extraction process implied from the nature of a set of texts written in classical Portuguese of the 16th century. We wanted to demonstrate the application of various techniques, models and tools used today in ontology learning, and apply them over ancient textual data sources. In particular, on a set of very peculiar texts: cuisine recipes. Then, we want to verify whether the culinary processes and the ingredients used at the time would be similar (or not) to those we use today. In addition, the interest in carrying out this work was reinforced by the utility that an ontology of the 16th century cuisine has for all those who develop their research work in the field of Food History in Portugal. In the next section, we will present and discuss the process of discovering ontological knowledge in unstructured recipes of a Portuguese monk from the 16th century.

3 DISCOVERING THE RECIPE ONTOLOGY

3.1 The Monk's Recipes

The analysis and editing of a vast set of ancient cooking recipes, from the second oldest culinary manuscript known in Portugal (the first notebook of codex 142 of the District Archive of Braga), originated a source of textual data, containing very rich and diverse information about Portuguese cuisine in the 16th century. The manuscript contains the texts of the recipes, referring ingredients and a diverse set of combinations of culinary processes. A monk, José Joaquim de Santa Teresa, collected the recipes. He decided to gather various manuscripts scattered throughout its congregation, and proceeded to register them in the library of the Monastery of S. Martinho de Tibães (Braga, Portugal). In 2013, Anabela Barros, edited and published the recipes in the book "As Receitas de um Frade Português no séc. XVI" ("The Recipes of a Monk Portuguese in the century. XVI") (Barros, 2013). This book provides a rich source for studying the various culinary elements – recipes, ingredients, methods, processes, etc. –, included and dispersed throughout the various unstructured texts of the source. All these elements are very important for studying the 16th century Portuguese cuisine. As already referred, the creation and development of a domain ontology, specifically oriented to the reception of the knowledge associated with these elements would provide a very useful instrument for all who wanted to learn about the cuisine that was taking place in that century. In a first stage of our work, we studied the texts of the recipes, both in terms of contextual and structural level – Table 1 presents an example of a recipe – "Sopas de Vaca Contrafeitas" ("Counterfeit Cow Soup"). The aim was establishing a knowledge base and understanding writing patterns, as well as characterizing the natural language used in the texts.

Table 1: The text of a recipe in classic Portuguese.

Recipe Nr. and Title **(Block 1)**	**42** **Sopas de vaca contrafeitas**
Ingredients **(Block 2)**	2 molhos de cheiros (coentro, endro, segurelha e hortelã) Azeite (muito bom) Água 1 ou 2 olhos de couve Pão 1 ramo de coentros
Method **(Block 3)**	Tomarão um molho de cheiros, coentro, endro, segurelha, hortelã, e deitá-lo-ão dentro de uma panela nova com um pouco de azeite muito bom e sua água, e depois de ter dado uma boa fervura, lhe metem dentro um ou dois olhos de couves segundo a quantidade de sopas que querem fazer, e outro molho de cheiros como o primeiro, e como a couve é cozida fazem as sopas com aquele caldo, pondo-lhes a couve em cima com sua capela de coentros.

Next, we defined the process of extraction of the ontology, the way we would carry it out, taking into account the particularities that the texts presented, such as writing patterns, linguistic structure, frequency of terms, and heterogeneity of the texts. Finally, we exported the ontological structures created to a graph database management system.

3.2 Extracting the Ontology

The process of extracting an ontology from a set of unstructured texts is not easy to accomplish (Tiwari and Jain, 2014). Despite the numerous computational tools that exist today, its automation is not yet completely automatic. Each process of development of an ontology has its own specificities, as well as the texts that serve as its source of knowledge. Usually, in practice these processes use to involve a hybrid approach, involving work performed by experts in the domain of ontology (the manual part) and work performed by natural language processing tools, text mining and machine learning (the automatic part). Although the automatic part is very attractive and saves temporal and human resources in cases of large textual data sources, the manual approach is always recommended for ensuring the quality of the ontology, which depends a lot on the complexity and granularity of the data elements included in the texts (Wong et al., 2012).

In the development of the ontology learning process, we followed a supervised approach, particularly in the initial part of the ontology definition, leaving the automatic part of the process to perform syntactic analysis and term extraction tasks. In any case, we follow the approach proposed in (Brewster, 2006), organizing our process in such a way as to perform the tasks of extracting terms, synonyms, concepts, hierarchies, relationships and rules in the way he indicated. Thus, we first extracted all the terms present in the texts of the recipes, with the aim of creating a set of terms to be cataloged. We then grouped each of the identified terms with all their synonyms found in the previous set. This significantly increased the accuracy of the concept identification process and greatly reduced the number of redundant terms. After having carried out the task of grouping terms, we associated them with concepts for establishing a concrete correspondence between the language of the domain of knowledge and the ontological system to be created. Then, we established the hierarchical relationships and defined the relationships between concepts, specifying the domain and the scope of the relationship. However, in order to carry out successfully all the extraction tasks, we had to analyze preliminarily, in detail, the texts of the recipes, in order to organize the extraction process and facilitate the treatment of the data. Therefore, we decided to divide each recipe into three distinct parts (Table 1): block 1, which contains the identifier and designation of the recipe; block 2, which has a description of the ingredients used by the recipe and, sporadically, some additional notes; and block 3, which describes the recipe preparation process. It should be noted that the information in block 2 does not appear explicitly in the manuscript, having been prepared previously by the author of (Barros, 2013) during the edition of the manuscript, who presented them in modern Portuguese.

Table 2: Some concepts identified in the recipes.

	Concept	Description
1	Recipe	*Represents each of the culinary records or texts for the confection of a dish and its general characteristics.*
2	Ingredient	*Identifies an ingredient and some additional notes.*
3	Note	*Represents a culinary note, made by the editor (who removed it or deduced it from the text of the recipe) and inserted in the block of ingredients.*
4	Procedure	*Describes all the steps (processes) of the preparation of the recipe.*
5	Process	*Represents each process identified in the recipe preparation block.*
6	Index	*Allows for the enumerating all the words present in the texts of the recipes.*

The existence of this block greatly facilitated the process of identifying ingredients in the recipes. Originally, these elements figured in very variable positions throughout the manuscript. The definition of the most pertinent classes in culinary ontology depends highly on the representative units that a recipe presents. Initially, we recognized and established six classes (concepts) to incorporate the entities referred in each text. Table 2 presents and describes the concepts we identified – nothing new if we analyze some of the current cuisine recipes ontologies. The formalization of an ontology is often a too complex process to do through logical definitions formally written. To overcome the need to simplify the entire process of creation and evolution of an ontology, several tools were developed having the ability for expressing this formalization in an organized. Next, we will see how the remaining tasks of the ontology extraction process developed, identifying whenever necessary the tools that were being used.

Textual documents have their own characteristics and specificities, which vary according to temporal, external and personal factors. Just as a human being has fingerprints, a person who writes a text somehow incorporates his own identity into the writing of the text. It is something like a signature. To recognize this kind of signature, we need to study how the sentences in the text are articulated and constructed syntactically. To extract the syntactic and lexical elements present in the writing of a given author, we can define a set of specific patterns to identify linguistic marks (terms), that have been used by the author of the texts of the recipes. Subsequently, these terms have to be evaluated to determine whether they are acceptable or of interest in the context of the ontology of cooking recipes. To establish these patterns we relied on the study of the various recipes we had available.

Table 3: Examples of patterns used for identifying terms.

	Pattern	Sentence	Process
1	[^\s]+-[^\s\.]*	"Deitar-se-á em um tachinho..."	deitar
2	[^\s,\.]+ [^\s,\.]+	"...salsa e cebola picados,..."	picar
3	[^\s,\.]+	"...e clarificado com uma clara de ovo..."	clarificar
4	[^\s,\.]+ [^\s,\.]+|se [^\s,\.]+ [^\s,\.]+|lhe [^\s,\.]+ [^\s,\.]+	"...a qual se lhe lançará muito bem..."	lançar
5	[^\s,\.]+ com [^\s,\.]+	"...e nela se frigirão com azeite..."	frigir
6	[^\s,\.]+ em [^\s,\.]+	"...e afogados em azeite se lançará..."	afogar
7	[^\s,\.]+| lhe [^\s,\.]+	"...um fogo brando se derreterá nela..."	derreter
8	[^\s,\.]+ [^\s,\.]+	"...havendo primeiro sido passadas..."	passar

In this context, we identified several patterns, some more generic than others. We defined patterns using specific regular expressions. For example, the pattern '[^\s]+-[^\s\.] *' allows for identifying all verbs with associated pronouns in a sentence. During the process of syntactic analysis of the procedure of a recipe, we found that, generally, a verb followed by a post verbal pronoun is a culinary process. These cases arise essentially at the beginning of sentences. However, this particularity was not considered in the expression, so that it is in some way "permissive". The sentence "Deitar-se-á em um tachinho..." ("It will lie down in a pot...") was one of the phrases that may be instantiated with the referred pattern, in which we can easily identify the culinary process "deitar" (to put). Another pattern we defined was '[^\s,\.] + [^\s,\.] +', which allows the extraction of two words, one relating to ingredients and the other relating to the state of the ingredients. This pattern defines a state that is interpreted as a verb in participle form. Through this pattern, the phrase "...salsa e cebola picados..." ("...chopped parsley and onion...") allowed to identify the "picar" (to chop) process. Many other processes, such as "clarificar" (to clarify), "lançar" (to cast), or "frigir" (to fry), have been identified by other patterns in sentences such as "...e clarificado com uma clara de ovo..." ("... and clarified with an egg white..."), "...a qual se lhe lançará muito bem..." ("... which will be thrown at you very well...", or "...e nela se frigirão com azeite..." ("...and in it they will be fried with olive oil..."), respectively. Other examples of patterns can be found in Table 3.

After extracting the terms with the patterns, it was necessary to classify syntactically the content of the preparation (method) of a recipe. We made the tokenization of the text of the culinary procedure using Spacy (Honnibal and Montani, 2017). Once all the sentences were classified, it was possible to extract the syntactic function assigned to each word, the tokens. This assignment is critical to shaping the rules for acceptance of the terms, in conjunction with the established standards. Each word was instantiated in a given object, whose properties can be its literal meaning, the reduction of the word to its root or its grammatical class, that is, the attribute that allows for identifying the verbs. After segmenting the phrases of the texts in tokens, we associated with each identified process the various ways in which it appeared written in the recipes. In practice, we mapped all the variants of a culinary process, taking into account their verbal form, gender and number. In this way, it was easier to identify the occurrences of a process in a given recipe.

Next, we extracted concepts and aggregated synonyms. The attribution of concepts (or classes) (Table 2) to the extracted terms was, to some extent, quite evident, since their location in the text of the recipe revealed their identity in most cases. The elements contained in block 1 of the recipe (Table 1) are integrated into the concept “Receita” (“Recipe”). This block always contains the same elements: number and name of the recipe. However, the same is no longer the case in the other blocks. At least so directly, since the extraction of the concepts associated here depended directly on the algorithms used for extracting terms, relationships and associated rules.

At this point, we start the automatic part of the ontology learning process. As we process the recipe, we created its identity. Using the contents of block 2 of a recipe (Table 1), we can find out which entities belong to the concept “Ingredient”. In some cases, in block 2, may appear references to entities that belong to the concept “Note”, since they are do not belong to the recipe or to the list of ingredients. However, they indirectly help the preparation of the recipe. Then, in addition to the ingredients, extracted manually by the text editor and placed in block 2, using contents of block 3, we can infer the identification the most representative elements included in the “Method” concept description, although may arise other processes derived from the processing of the content of this class. Finally, we identified the representative entities of the “Process” concept.

If no process was identified, we assumed that we are in the presence of a poor or incomplete procedural description of a recipe, an incomplete method. Any process identified by the system is an abstraction of a more specific definition, such as an action performed by a person preparing a recipe or, simply, by a procedure merely exclusive to the domain of cooking. For example, it is correct to say that the “deitar” (“throwing”) process allows for inferring an action of the user and that the “ferver” (“boiling") process represents an action relative to an ingredient. However, this differentiation was not explicitly integrated into the ontology, since the defined patterns were not developed to make this hierarchical distinction.

To finish the process of extracting the ontology, we needed to confirm the identity of each term, as well as the role it will play in the semantic area. Relationships allow for establishing chains of knowledge and affinity between two entities, linking them through one or more characteristics. Like the properties of a concept, the relationship is an integral part of the process of determining the role that a given element plays in the ontology. When it is not possible to ascertain the properties of a particular element, it is possible to make the recognition of the entity concerned if we identified a relationship. In the recipes we analyzed, all the relationships between concepts were created by the terms found, and all the terms found relate to the concept “Receita” (“Recipe”). This logic is justified by the fact that all the elements found are a part of a recipe and by the direct relationship between a given concept and all the recipes introduced into the system. Each concept found is about a recipe, in which the relationship assumes a degree of possession and integral part. When the domain is "Recipe", the relation is expressed as a possessive bond, and when this is characterized as the initial set, the relationship is identified as a part of the recipe. Figure 2 shows a conceptual definition of the ontology we extracted. The ontology’s structure is simple and very similar to the other recipe ontologies we studied before.

```
Concepts ::                Relationships ::
RECIPE (Id,                RECIPE has INGREDIENT;
Description);              INGREDIENT of_a RECIPE;
INGREDIENT                 RECIPE has NOTE;
(Description);             NOTE of_a RECIPE;
NOTE (Description,         RECIPE has PROCEDURE;
Contents);                 PROCEDURE of_a RECIPE;
PROCEDURE                  RECIPE has PROCESS;
(Description);             PROCESS of_a RECIPE;
PROCESS (Name,             PROCEDURE has PROCESS;
Variations);               PROCESS of_a PROCEDURE;
INDEX (Name);              RECIPE has INDEX;
                           INDEX of_a RECIPE;
```

Figure 2: A conceptual definition of the recipe ontology.

3.3 Exploring the Ontology

After we finished the extraction process, we tried to preserve the ontological structures created. The preservation of these structures was carried out in a graph database management system, the Neo4J (Neo4J, 2023). Through it, we can perform all the preservation operations we need, as well as explore the stored ontology through its interrogation language: Cypher. The registration of the extracted entities can be done through a combination of pre-existing data with the knowledge to be imported. If there are no references yet, all the new information in the ontology is inserted.

The improvement of the ontology was made after the import of the ontological structures into the Neo4J system, and was carried out in two different ways: through the database system or through an OWL file. The difference between these two ways of improving the ontology is only in terms of semantic inference. In Neo4j it is done using queries and in OWL using HermiT (Glimm et al., 2014). Hermit allows completing any information that may not have been associated with a given concept. However, it is extremely time-consuming, especially if the determination of the new relationships is made after the insertion of several records.

To allow the dissemination of results and the experimentation of the ontological system created, we designed and developed a specific Web platform. The development of this platform was done through the Flask micro framework, written in Python and architected based on an MVC (Model-View-Controller) pattern, not only for the organization of the application between the logical layer and the graphic layer that this pattern establishes, but also for the ease it gives us to add or modify system functionalities. The Web platform developed allows for exposing in a simplified way all the information of the ontology extracted. In a real application scenario, this platform substantially facilitates the process of study and research, both of a recipe and its components, as well as of a component and all the recipes linked to it. In Figure 3, we can see a small view of the ontology in Neo4J environment.

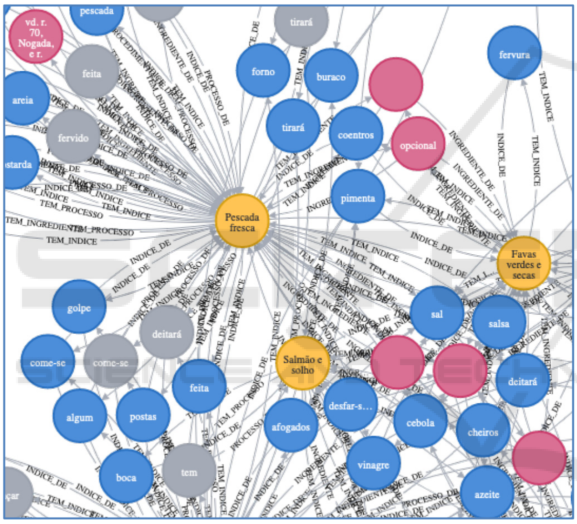

Figure 3: A view of the ontology provided by the Neo4J system.

4 CONCLUSIONS AND FUTURE WORK

In this paper, we presented a semi-automatic ontology learning system conceived for extracting a culinary ontology from a large set of recipes of the 16th century published in (Barros, 2013). The implementation of the system covered all the stages of an ontology extraction process, from text preprocessing to the materialization of the structures of ontology. In addition, it provides a Web platform for consulting and analyzing the knowledge incorporated in the ontological system. This platform is a very useful tool for all those who wish to study the Portuguese cuisine of the 16th century, through the knowledge extracted from the cooking recipes collected. It facilitates a lot the process of research and analysis, whether of a recipe and its components, or of a component and all the recipes related to it. In semantic terms, the ontology allows for establishing a bridge between the Portuguese culinary processes of the 16th century and the processes of the modern era that, curiously, in one way or another, are similar or, sometimes, the same. An identical reasoning can be applied to the ingredients that were used then in recipes, but that today are already replaced by others. This ontology, validated by the author of (Barros, 2013), also provides means to identify culinary patterns of the 16th century and compare them with current culinary practices, which we believe that is valuable for researchers in the area of Food History, which are dedicated to investigating and comparing them over time.

The version of the ontology we have today is not complete. However, it reflects adequately the knowledge contained in the recipes gathered by the Portuguese monk José Joaquim de Santa Teresa, in the 16th century. One of the improvements we plan to make in a near future involves the development of extraction mechanisms that will allow us to distinguish what are cooking processes and user actions. This is because we found that some of the verbs used in the recipes identify processes inherent in the ingredients and others correspond to steps that must be followed by the person who is making the dish. This distinction will impose a more precise and detailed syntactic analysis, since we will need to analyze all the elements that precede or succeed a verb.

ACKNOWLEDGEMENTS

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020, and the PhD grant: 2022.12728.BD.

REFERENCES

Asim, M., Wasim, M., Khan, M., Mahmood, W., Abbasi, H. (2018). A survey of ontology learning techniques and applications. Database, Volume 2018. DOI: https://doi.org/10.1093/database/bay101.

Barros, A. L. (with Seiça, M., Veloso, J. & Aguiar, M.) (2013). As receitas de cozinha de um frade português do século XVI. Coimbra: Imprensa da Universidade de Coimbra.

Batista, F, Pardal, J., Mamede, N., Vaz, P., Ribeiro, R. (2006). Ontology construction: cooking domain. Technical Report, INESC-ID, February.

Belhoucine, K., Mourchid, M. (2020). A Survey on Methods of Ontology Learning from Text. pages 113–123.

Borst, W. (1997). Construction of engineering ontologies for knowledge sharing and reuse. Twente, sep.

Brewster, C. (2006). Ontology Learning from Text: Methods, Evaluation and Applications Paul Buitelaar, Philipp Cimiano, and Bernado Magnini (editors) (DFKI Saarbrücken, University of Karlsruhe, and ITC-irst), Amsterdam: IOS Press (Frontiers in Artificial Intelligence and appl. Computational Linguistics, 32(4):569–572. DOI: https://doi.org/ 10.1162/coli.2006.32.4.569

Choudhary, J., Tomar, D. (2014). Semi-Automated Ontology building through Natural Lan-guage Processing. International Journal of Computers & Technology, 13(8), 4738–4746. DOI: https://doi.org/10.24297/ijct.v13i8.7072.

Cimiano, P., Mädche, A., Staab, S., Völker, J. (2009). Ontology Learning. In: Staab, S., Studer, R. (eds) Handbook on Ontologies. International Handbooks on Information Systems. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-92673-3_11.

Davies, J., Fensel, D., van Harmelen, F. (2003). Conclusions: Ontology-driven Knowledge Management – Towards the semantic web?, Towards the Semantic Web, pp. 265–266, John Wiley & Sons, Ltd.

Ding, L., Kolari, P., Ding, Z., Avancha, S. (2007). Using Ontologies in the Semantic Web: A Survey. In: Sharman, R., Kishore, R., Ramesh, R. (eds) Ontologies. Integrated Series in Information Systems, vol 14. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-37022-4_4

Glimm, B., Horrocks, I., Motik, B., Stoilos, G., Wang, Z. (2014). HermiT: An OWL 2 reasoner. Journal of Automated Reasoning, 53(3), 245–269.

Graça, J., Mourao, A., Duarte, M., Anunciação, O., Monteiro, P., Pinto, S., Loureiro, V. (2005). Ontology building process: The wine domain. 5th Conference of the European Fed-eration for Information Technology in Agriculture, Food and Environment.

Gruber, T. (1995). Toward principles for the design of ontologies used for knowledge sharing? International Journal of Human-Computer Studies, 43(5-6):907–928, nov. DOI: https://doi.org/10.1006/ijhc.1995.1081.

Guarino, N., Oberle, D., Staab, S. (2009). What Is an Ontology?, In Staab, S., Studer, R. (eds) Handbook on Ontologies. International Handbooks on Information Systems. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-92673-3_0 2009

Honnibal, M., Montani, I. (2017). spaCy · Industrial-strength Natural Language Processing in Python. Homepage, https://spacy.io/api, last accessed 2023/06/22.

Keet, M. (2018). An Introduction to Ontology Engineering, University of Cape Town. Homepage, https://people.cs.uct.ac.za/~mkeet/OEbook/, last accessed 2023/05/05.

Markantonatou, S., Toraki, K., Minos, P., Vacalopoulou, A., Stamou, V., Pavlidis, G. (2021). "ΑΜΑΛΘΕΙΑ: A Dish-Driven Ontology in the Food Domain" Data 6, no. 4: 41. DOI: https://doi.org/10.3390/data6040041

Mishra, T., Jain, S. (2014). Automatic Ontology Acquisition and Learning. International Jour-nal of Research in Engineering and Technology. 03. 38-43.

Monica, S., Krisnadhi, A., Cong, W., Gallagher, J., Hitzler, P. (2014). An Ontology Design Pat-tern for Cooking Recipes. In Proceedings of the 5th Workshop on Ontology and Se-mantic Web Patterns (WOP2014) co-located with the 13th International Semantic Web Conference (ISWC 2014), Riva del Garda, Italy, October 19.

Nanba, H., Doi, Y., Tsujita, M., Takezawa, T., Sumiya, K. (2014). Construction of a cooking ontology from cooking recipes and patents. In Proceedings of the 2014 ACM Interna-tional Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct). Association for Computing Machinery, New York, NY, USA, 507–516. DOI: https://doi.org/10.1145/2638728.2641328.

Neo4J Graph Data Plataform, Homepage (2023). https://neo4j.com/, last accessed 2023/06/22.

Noy, N., Mcguinness, D. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. Knowledge Systems Laboratory. 32.

Sharma, A. (2021). Natural Language Processing And Sentiment Analysis, In International Re-search Journal of Computer Science, 8(10):237. DOI: https:/doi.org/ 10.26562/irjcs.2021.v0810.001

Tiwari, S., Jain, S. (2014). Automatic Ontology Acquisition and Learning. International Journal of Research in Engineering and Technology, 03(26):38–43, Nov. DOI: https://doi.org/10.15623/ijret.2014.0326008

Uschold, M., Gruninger, M. (1996). Ontologies: principles, methods and applications. The Knowledge Engineering Review, 11(2):93–136, jun. DOI: https://doi.org/10.1017/S0269888900007797.

van Eijnatten, F., Putnik, G. (2004). Chaos, Complexity, Learning, and the Learning Organiza-tion, The Learning Organization, 11(6), pp. 418-429.

Villarías, L. (0420). Ontology-based semantic querying of the web with respect to food reci-pes. Master's thesis, Technical University of Denmark.

Wong, W., Liu, W., Bennamoun, M. (2012). Ontology learning from text. ACM Computing Surveys, 44(4):1–36, Aug. DOI: https://10.1145/2333112.2333115.

Zhang, L. (2018). Wang, S., Liu, B., Deep Learning for Sentiment Analysis: A Survey. WIREs Data Mining Knowledge Discovery, Volume 8, Issue 4, July/August. DOI: https://doi.org/10.1002/widm.1253