PONTIFICAL CATHOLIC UNIVERSITY OF RIO GRANDE DO SUL

FACULTY OF INFORMATICS

GRADUATE PROGRAM IN COMPUTER SCIENCE

AN ADAPTIVE APPROACH FOR ONTOLOGY ALIGNMENT VISUALIZATION

BERNARDO SEVERO DE SOUZA

Dissertation submitted in partial fulfillment of the requirements for the degree of Ph. D. in Computer Science at the Pontifical Catholic University of Rio Grande do Sul.

Advisor: Prof. Renata Vieira

Co-advisor: Prof. Cássia Trojahn dos Santos

Porto Alegre

2017

Ficha Catalográfica

S729a Souza, Bernardo Severo de

An Adaptive Approach for Ontology Alignment Visualization / Bernardo Severo de Souza . – 2017.

111 f.

Tese (Doutorado) – Programa de Pós-Graduação em Ciência da Computação, PUCRS.

Orientadora: Profa. Dra. Renata Vieira.

Co-orientadora: Profa. Dra. Cássia Trojahn dos Santos.

1. alinhamento de ontologias. 2. visualização de alinhamentos de ontologias. 3. avaliação de alinhamentos de ontologias. I. Vieira, Renata. II. Santos, Cássia Trojahn dos. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS com os dados fornecidos pelo(a) autor(a).

Pontifícia Universidade Católica do Rio Grande do Sul
FACULDADE DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

TERMO DE APRESENTAÇÃO DE TESE DE DOUTORADO

Tese intitulada "An Adaptive Approach For Ontology Alignment Visualization" apresentada por Bernardo Severo de Souza como parte dos requisitos para obtenção do grau de Doutor em Ciência da Computação, aprovada em 20 de fevereiro de 2017 pela Comissão Examinadora:

Profa. Dra. Renata Vieira (PPGCC/PUCRS – Orientadora)

Profa. Dra. Cássia Trojahn dos Santos (Université Toulouse II & IRIT - Coorientadora)

Profa. Dra. Milene Selbach Silveira (PPGCC/PUCRS)

Profa. Dra. Carla Maria Dal Sasso Freitas (UFRGS)

Prof. Dr. Sandro José Rigo (UNISINOS)

Homologada em......./......./........., conforme Ata No. .... pela Comissão Coordenadora.

Prof. Dr. Luiz Gustavo Leão Fernandes
Coordenador.

UMA ABORDAGEM ADAPTATIVA PARA VISUALIZAÇÃO DE ALINHAMENTOS DE ONTOLOGIAS

RESUMO

O aumento do volume de dados não estruturados na Web nas últimas décadas tem sido impulsionado pelo surgimento de novos meios de comunicação, dispositivos e tecnologias. Neste contexto se desenvolve a Web Semântica, cujo objetivo é o de atribuir uma camada de representação de conhecimento a esses dados, facilitando o tratamento por processos automatizados. Ontologias são elementos chave da Web Semântica, oferecendo uma descrição dos conceitos e dos relacionamentos entre os mesmos para um domínio específico. Entretanto, ontologias de um mesmo domínio podem divergir em sua estrutura, granularidade ou terminologia, necessitando que um processo de mapeamento entre as mesmas seja realizado, produzindo um conjunto de correspondências entre entidades semanticamente relacionadas (alinhamento). Um número crescente de abordagens de mapeamento tem surgido na literatura e a necessidade de avaliar e comparar qualitativamente os alinhamentos produzidos se faz presente. Tarefas que fazem uso de alinhamentos passaram a demandar melhores representações gráficas dos mesmos. Neste contexto, foi realizada uma pesquisa com especialistas em alinhamentos para identificar os aspectos mais importantes em uma visualização de alinhamentos. Este trabalho apresenta então uma abordagem adaptativa de visualização para alinhamentos, que permite ao usuário escolher como e o que visualizar, de acordo com preferências próprias ou para uma atividade sendo realizada no momento (criação, manipulação, avaliação, etc.). Por fim, um protótipo foi construído com o intuito de validar a solução. Os resultados obtidos da avaliação dos usuários com o protótipo mostram que a abordagem lida com os problemas que se propõe a resolver, com uma margem para trabalhos futuros em formas de visualização de alinhamentos.

Palavras chave: alinhamento de ontologias, visualização de alinhamentos de ontologias, avaliação de alinhamentos de ontologias

AN ADAPTIVE APPROACH FOR ONTOLOGY ALIGNMENT VISUALIZATION

ABSTRACT

The increase in the volume of unstructured web data in recent decades has been driven by the arising of new media, devices and technologies. In this context, the Semantic Web was developed, whose objective is to provide a layer of knowledge representation to that data, facilitating the treatment by automated processes. Ontologies are key elements of the Semantic Web, providing a description of the concepts and relationships between them, for a specific domain. However, ontologies of the same domain may differ in structure, granularity or terminology, requiring a process of matching between them to be performed, producing a set of correspondences between semantically related entities (alignment). A growing number of matching approaches have emerged in the literature, and the need to evaluate and qualitatively compare the produced alignments is presented. Tasks that make use of alignments started to demand better graphical representations for it. In this context, a survey was conducted with alignment specialists to identify the most important aspects in an alignment visualization. This work presents an adaptative approach for alignment visualization, that allows users to choose how and what to visualize, according to their own preferences or the task being performed at that moment (creation, manipulation, evaluation, etc.). Finally, a prototype was built with the purpose of validating the solution. The results obtained from the prototype validation with users show that the approach handles the problems it proposes to solve, with a margin for future work on alignment visualization.

Keywords: ontology matching, ontology alignment visualization**,** ontology alignment evaluation

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AI - Artificial Intelligence

API - Application Programming Interface

GUI - Graphical User Interface

HCI - Human–computer Interaction

IR - Information Retrieval

ISWC - International Semantic Web Conference

NLP - Natural Language Processing

OAEI - Ontology Alignment Evaluation Initiative

OM - Ontology Matching

OWL - Web Ontology Language

XML - eXtensible Markup Language

CONTENTS

1. INTRODUCTION

1.1. Background and problem motivation

Over the last decades, the volume of digital information available has increased, mainly driven by the development of new means of communication, e-commerce, social networks, mobile devices and other new technologies. Most of these data is actually presented in an unstructured form, such as in natural language. This means that most data has no structure or pattern and no associated semantic information, which makes it difficult to interpret the data by automated processes. The projections made for the coming years estimate a continuous growth of this type of data, proportionally larger than structured data [26].

Besides the lack of structure, there is usually heterogeneity in the information presented in multiple sources, a characteristic which makes the analysis of data even more complex. Often important and related data are dispersed, without a semantic organization, making it difficult to access, compare and extract knowledge from it. This heterogeneity and dominance of unstructured data in information sources are a challenge for scientists, and are constantly discussed in recent research topics, such as Big Data, as well as Information Retrieval (IR), Natural Language Processing (NLP) and Semantic Web. In particular, the Semantic Web aims to provide a semantic layer to web content, thus aiming to facilitate the interaction between resources and the automated extraction of knowledge.

In this context, ontologies play a key role. Ontologies represent shared vocabularies that describe domain concepts and relationships between them [20]. Therefore, they help in the formal definition of web content. Using ontologies in the Semantic Web makes it possible to enrich the data by adding meaning, which allows more people (and machines) to reuse them and do more with the data [61]. From this association, systems begin to be able to interact, finding common and related entities, making decisions based on the interpretation of the content, etc.

However, despite being considered as a possible solution to the problem of heterogeneity in Semantic Web, ontologies do not solve it entirely, because ontologies of the same domain still can diverge in structure, content and granularity, depending on the view of the engineer who built it and for what purpose. Some ontologies share the same entities but with different terminology or granularity (being more or less expressive). For the heterogeneity problem between ontologies, the area of ontology matching has been developed as a solution [52]. Ontology matching is a process that aims to establish a set of correspondences between semantically related entities of different ontologies. A correspondence relates two entities of different ontologies, allowing the determination of the type of relationship (equivalence, inclusion, etc.) and its confidence level. The result of a matching process is an alignment, which is a set of correspondences between entities belonging to different ontologies [12].

With the dissemination of alignments, comes the need to evaluate the quality of such alignments. Due to alignments also being automatically generated by tools, many wrong relationships (false positive) are eventually produced in the process [8]. Campaigns with the purpose of evaluating tools that generate alignments were created, among which, one can mention the Ontology Alignment Evaluation Initiative (OAEI) [11], which aims to systematically evaluate alternative techniques and tools, using different datasets.

Another problem found is the lack of approachs/tools to visualize and work with multiple alignments, like those outputted by campaigns such as OAEI. Being analysed separately and without an approach that consider what information and representation of the data is relevant for the use case.

According to the research of Shvaiko and Euzenat [51], that still represents the current challenges in the area, one of the main challenges has been to design ways to actively involve users in the process of creation and visualization of alignments. This interaction would then empower users to define improvements and corrections in the alignments, influence the definition of weights for matching algorithms and critique the quality of intermediate and final results. They point out that among the dozens of applications analyzed and their experiences, only few of these have a graphical interface, essential for end users to interact more intuitively during the visualization and manipulation of alignments. So far, the attention directed to design solutions that help end users in these activities has been insufficient.

1.2. Goals

Given the importance of ontology alignments, the complexity of the task and the lack of current solutions towards user empowerment for this task, the problem to be solved concerns how ontology alignments can be visualized. Considering the research on the state of the art, this works aims to handle the gaps in current approaches with a new approach.

In order to better understand the user needs in this context, a survey with specialists needs to be conducted, complementing and reinforcing what is observed in the current state of the art. Based on the results of this survey, this work intends to propose a new approach to visualize and manipulate multiple alignments, thus enabling better experience towards quantitative and qualitative analysis of ontology alignments. In particular, the approach should address current limitations on alignments visualization when applied to diversified scenarios (where the tasks or the entities on focus change). To achieve this goal the approach needs to be visually adaptive, in which the user opts for the configurations that best fit his/her needs, so that the chosen visualisation is adapted for that user and the task. The approach should allow a graphical way to create, manipulate and evaluate alignments.

The overall aim is to propose a visual approach to the main tasks in the field of study (creation, manipulation, evaluation) and incorporate features to visually handle alignments in its different working scenarios. Therefore, it must present an adaptive/modularized approach, taking into account user preferences and combining different views and sets of information.

Finally, a prototype allowing the validation of the objectives will be developed.

1.2.1. Objectives

Propose an approach that:

- presents a graphical way to work with alignments.
- allows to manipulate and evaluate alignments and their various elements.
- allows visualization of multiple alignments.
- considers user preferences and/or type of task to be performed.

1.3. Outline

The remainder of this work is organized as follows. Chapter 2 provides background on ontologies, ontology matching and ontology matching evaluation. Chapter 3 introduces and compares the most important related work. Chapter 4 details the survey conducted to elicit the requirements for this work. Chapter 5 presents the proposed approach and its implementation. Chapter 6 analyzes the results of the evaluation conducted on the prototype. Chapter 7 presents the final considerations and future work.

2. BACKGROUND

2.1. Ontology

An ontology in computer science, according Gruber [20], is a formal and explicit specification of a shared knowledge. Formal refers to the fact that the ontology should be machine-readable and expressed using a suitable formalism. Explicit means that the concepts used and restrictions on their use are explicitly defined (declared). Shared reflects the notion that an ontology captures a consensus on a knowledge domain (is an idea accepted by a group, not only for an individual). Knowledge refers to a common model of a phenomenon in the world, through the identification of relevant entities for that.

Ontologies are widely applied in the areas of Software Engineering [23] and Artificial Intelligence [65]. They are also used in related areas such as knowledge management, e-commerce, NLP, information integration, IR, project/integration of databases, bioinformatics, education, etc. [15].

Although there are different ways to organize concepts and relationships in a ontology domain, Studer, Benjamins and Fensel [57] defines the following as the minimum set of common components to all:

- **Classes (or concepts):** express domain concepts, representing entities, sets, collections, objects, etc. Classes in an ontology are usually organized into hierarchical taxonomies in which inheritance mechanisms can be applied.
- **Relations (or object properties):** represent some kind of association between the domain concepts. Ontologies usually contain binary relations where the first argument is the domain of the relation, and the second is the range.
- **Attributes (or data properties):** differ from regular relations due to its range being a data type (string, numeric, etc.).
- **Instances:** represent individuals or members of a class in the ontology, they are the "concrete" form of these elements.

OWL (Web Ontology Language) is the standard language for describing expressive ontologies, adopted by the W3C.

Figure 1 shows a fragment of an ontology definition in OWL. In the example a hierarchy is presented, where “Monograph” and “Proceedings” are subclasses of “Book”, that along with “Journal” and “Article” are root classes. Journal has a relation “articles” with the class “Article”, and a data property “periodicity” as string. Finally, the instance “Journal_XYX” presents a concrete relation of type “articles” with an instance “Article_ABC”, and the value “monthly” as its periodicity.

```
...
<!-- Ontology -->
<owl:Ontology rdf:about="http://example.org/onto1.owl"/>

<!-- Classes -->
<owl:Class rdf:about="Book"/>
<owl:Class rdf:about="Monograph">
 <rdfs:subClassOf rdf:resource="Book"/>
</owl:Class>
<owl:Class rdf:about="Proceedings">
 <rdfs:subClassOf rdf:resource="Book"/>
</owl:Class>
<owl:Class rdf:about="Journal"/>
<owl:Class rdf:about="Article"/>

<!-- Relations -->
<owl:ObjectProperty rdf:about="articles">
 <rdfs:domain rdf:resource="Journal">
 <rdfs:range rdf:resource="Article">
</owl:ObjectProperty>

<!-- Attributes -->
<owl:DatatypeProperty rdf:about="periodicity">
 <rdfs:domain rdf:resource="Journal">
 <rdfs:range rdf:resource="string">
</owl:DatatypeProperty>

<!-- Instances -->
<owl:NamedIndividual rdf:about="Journal_XYX"
 <rdf:type rdf:resource="Journal"/>
 <articles rdf:resource="Article_ABC"/>
 <periodicity>monthly</ periodicity>
</owl:NamedIndividual>
...
```

Figure 1 - Fragment of an ontology in OWL.

2.2. Ontology Alignment

Ontologies are often subject to the perspective adopted at the time of its creation, as different engineers may have different views of the domain. In distributed and open systems using ontologies, heterogeneity in its many forms cannot be avoided [12]. The ontology matching is the process of defining correspondences between entities (classes, properties, instances) of different ontologies. As a result of this process, an alignment is produced. These alignments can be used for several tasks, such as ontologies merging, query rewriting, automatic data translation, browsing the Semantic Web and Linked Open Data, etc.

For this work, the definition of ontology matching proposed by Euzenat and Shvaiko [12] will be adopted, which defines a matching process as a function: $f(O_1, O_2, P, R, A)$. This function takes as input two ontologies O_1 and O_2 for matching, and optionally a set of parameters P and a set of resources R. As input, it is also possible to provide an alignment A, which can be analyzed and later complemented during the matching process. The resulting alignment of an ontology matching process consists of a set of correspondences between the involved ontologies.

According to Euzenat and Shvaiko [12], a correspondence can be defined as $<E_1, E_2, R, N>$, such that: E_1 and E_2 are entities of O_1 and O_2, respectively. R is the type of relation held between the aligned entities. Although the matching algorithms typically use the equivalence relation type, other relations such as more general, less general, disjoint, can also be expressed. In addition, N is a confidence measure number in the [0; 1] range, which expresses how much the author or algorithm believes in the fact that the relation exists.

Different approaches to ontology matching have been proposed in the literature [12]. The difference between each of them is primarily the type of knowledge encoded in each ontology, and the way this knowledge is used in the identification of correspondences between elements of ontologies. Terminological approaches perform a lexical comparison on strings (tokens or n-grams) used when naming entities (or labels and comments); while semantic approaches use the semantics of the theoretical model to determine whether there is a match between two entities. Approaches may consider internal ontological structure, such as the range of properties (attributes and relations), cardinality, transitivity or the symmetry of its properties. Alternatively, the external ontology structure can be used, such as the position of the entities within the ontology hierarchy. Finally, data instances can also be analysed, called the extensional level (while intensional level comprises the classes, relations and attributes). Many ontology alignment systems do not rely on a single approach and combine multiple approaches.

An alignment, as well as an ontology, are structured documents that can be represented in a textual form, for example as an XML, like the model defined and adopted by the Alignment API [9], as illustrated in Figure 2.

```
...
<Alignment>

 <onto1>http://example.org/onto1.owl</onto1>
 <onto2>http://example.org/onto2.owl</onto2>

 <map>
   <Cell>
     <entity1 rdf:resource='http://example.org/onto1.owl#Article/>
     <entity2 rdf:resource='http://example.org/onto2.owl#Paper'/>
     <measure
rdf:datatype='http://www.w3.org/2001/XMLSchema#float'>1.0</measure>
     <relation>=</relation>
   </Cell>
 </map>
...
```

Figure 2 - Fragment of an alignment in the Alignment API format.

In this format, there is the "Alignment" element, which encompasses the alignment, and nested in it the following:

- **onto1** and **onto2:** URIs that identify the ontologies being analyzed.
- **map** / **Cell:** a list of elements "**map**" storing correspondences found between two entities of ontologies. In each map, one "**Cell**" element will contain the attributes "**entity1**" and "**entity2**”, that stores the URIs of the entities involved in the match. Whereas the element of "**measure**" contains a value between 0 and 1, reflecting confidence in the mapped interface, which is defined by the "**relation**" tag.

2.3. Ontology Alignment Evaluation

Alignments are often created through automatic alignment generation tools (matchers), and their algorithms are subject to create erroneous correspondences [17]. In this context, the need to evaluate these algorithms and tools through the analysis of generated alignments has emerged.

2.3.1. OAEI

The International Semantic Web Conference (ISWC) [66] is an international forum for communities working with the Semantic Web and more recently the Linked Open Data initiative. In the last years, co-located with the ISWC occurs the International Workshop on Ontology Matching (OM) [67], a workshop focused on ontology matching research. Associated with the OM workshop is the Ontology Alignment Evaluation Initiative (OAEI) [11]. Since 2004, OAEI organizes campaigns to evaluate ontology matching technologies. This need originated from the growing number of methods available to perform integrations between ontologies and schema matching [11].

The main goal is to compare systems and algorithms openly and on the same basis, in order to allow anyone to draw conclusions about the best matching strategies. From such evaluations, tool developers can then improve their systems [6]. The ontologies available are described in OWL and alignments submitted by the participants defined in a custom format (defined by the Alignment API [9]). The data sets used in the evaluations involve different domains, languages and sizes, emphasizing the different characteristics that ontologies and alignments can express. There are alignments data covering topics such as anatomy, organization of conferences, biological / medical classifications, social sciences, etc. The results are automatically evaluated using a reference alignment.

2.3.2. Evaluation Metrics

In alignments evaluation such as those performed in OAEI, quantitative measures are typically analyzed (resulting from the comparison of the alignments), such as precision, recall and F-measure. To calculate these metrics, the analysis of the following terms is necessary:

- **True positive:** correspondences found that are correct.
- **False positive:** correspondences found that are incorrect.
- **True negative:** correspondences determined as not existing correctly (not declared).
- **False negative:** correspondences determined as not existing incorrectly (missing).

To evaluate alignments, known measures of IR were adapted to the context. These measures are calculated by comparing an alignment with a reference alignment (or gold standard), typically built and revised manually.

Precision is the set of retrieved data that is relevant to the query, calculated using the following formula:

$$\textbf{precision} = \frac{\text{true positive}}{\text{true positive + false positive}}$$

OR

$$\textbf{precision} = \frac{\{\text{correct correspondences}\} \cap \{\text{produced correspondences}\}}{\{\text{produced correspondences}\}}$$

The precision then represents the percentage of correspondences correctly classified as positive, of all those that were classified as positive. Therefore, it can be said that the precision is a measure of quality over the correspondences suggested.

While the recall is the set of data relevant to the query that could be recovered, calculated using the following formula:

$$\textbf{recall} = \frac{\text{true positive}}{\text{true positive + false negative}}$$

OR

$$\textbf{recall} = \frac{\{\text{correct correspondences}\} \cap \{\text{produced correspondences}\}}{\{\text{correct correspondences}\}}$$

The recall in this context is the percentage of correct correspondences discovered among the total number of existing correspondences to be discovered. It can be classified as a measure of quantity over the correspondences suggested.

Finally, the f-measure (or F1 score) is a measure of the correctness (accuracy) of a test, taking into account both precision and recall. This measure can be interpreted as the harmonic mean of precision and recall. Its calculation is given by the following formula:

$$\textbf{f-measure} = 2 \text{ x } \frac{\text{precision x recall}}{\text{precision + recall}}$$

2.4. Ontology and Alignment Visualization

As presented in this chapter, ontologies and alignments have natively a structured format, represented by standards in a textual form. They are important to enable machine interpretation, but such representations are often difficult to read and manage by users, especially when dealing with large volume of data. Therefore, a visualization on these structures and their entities contributes on the user experience, facilitating its interaction. Graphical representations adapt the information available to present it, having different advantages and disadvantages, depending of several factors. Many researches, viewed on Chapter 3, were conducted to discuss and evaluate the best approaches for ontology and alignment visualization.

3. RELATED WORK

3.1. Ontology Visualization

Several forms of graphical visualization of an ontology structure and its various elements have been proposed in the literature. Each one typically focuses or facilitates the display of a set of ontology elements (classes, instances, relations, etc.) according to the type of activity to be performed, the domain, or even the user profile for which it was constructed [28]. In this field, the work of Silva [53] presents multiple, coordinated views for exploring the intensional and extensional levels of an ontology. What is new here is the use of degree of interest notion in order to reduce the complexity of the representation, drawing the user attention to the main concepts for a given task.

A survey about ontologies visualization methods was conducted by Katifori et al. [28], which produced a classification of visualization types, where each category refers to a set of particular characteristics to visually represent the structure of an ontology. An analysis was made for each category, considering the characteristics of ontologies presented (instances, relations, multiple inheritance, etc.), their advantages and disadvantages. The main proposed categories are:

- **Indented list:** format commonly found in manipulation tools and ontology visualization. This format presents the taxonomy of the ontology (determined by subclass relation "is a") in the form of a tree, where the nodes are the classes. Classes with more than one parent class (multiple inheritance) are represented by repeating the class below each parent. According to Akrivi et al. [29], this approach is familiar and intuitive for all types of users, due to its similarity to a program to explore files and directories. However, a limitation of this technique is that because it represents a tree (and not a graph, which supports the representation of more connections), usually the relation type displayed is the "is a", leaving out other relation types (sometimes displaying them separately, with no connection to the tree). Katifori et. al also point out that this type of visualization is not useful in tasks such as identifying the depth of the hierarchy, searching nodes with many children, identification of sibling nodes, etc.

A widely used tool to manipulate ontologies is Protégé [38] (several other visualization tools are built in the form of plugins for it), which uses as the basis of its visualization and navigation the indented list format, as shown in Figure 3. Some other visualization tools that also uses this format are: OntoEdit [58], OntoRama [10] and Orient [63].

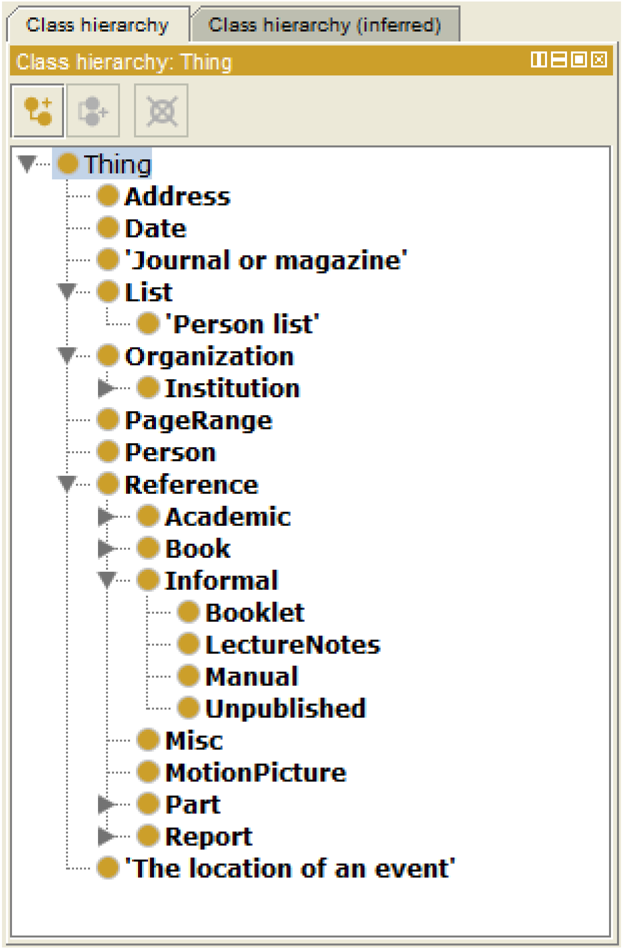


Figure 3 - Example of indented list visualization in the Protégé software [38].

- **Node-link and tree:** represent the ontology through interconnected nodes forming a tree, in formats that can be oriented from right to left or top to bottom. With an intuitive view, this type of visualization makes it easy to analyze the overall structure of the ontology, its depth and length. A limitation of this approach is the view for larger ontologies, given the necessary space on the screen to view each class. Usually, navigation becomes necessary for ontologies with a few tens or hundreds entities. Plaisant et al. [42] also observed that the side where the tree root is located have an inefficient use of screen space (only the root node is displayed, centered), while the other side ends up visually loaded (leaf nodes occupy the entire space).

Among the tools with this visualization, one can mention OntoViz [54] and OWLViz [24] (Figure 4), Protégé plugins that let the user work with a focus on a group of entities for visualization and choose which elements to display, thereby softening the problem of data volume on the screen. Other tools use different approaches to address the disadvantages of this technique (not all being applied for ontologies); SpaceTree [42] replaces part of the tree for an icon representing the size of subtree present there; IsaViz [41] has a "radar", allowing the user who is scrolling to locate itself, like in a map; Cone Tree [47], Tree Viewer [30] and OntoSphere [5] adopt a 3D view to improve navigation and use of the screen; other tools like TreePlus [5], OntoTrack [36], GoSurfer [64], Gobar [35], GOMiner [62].

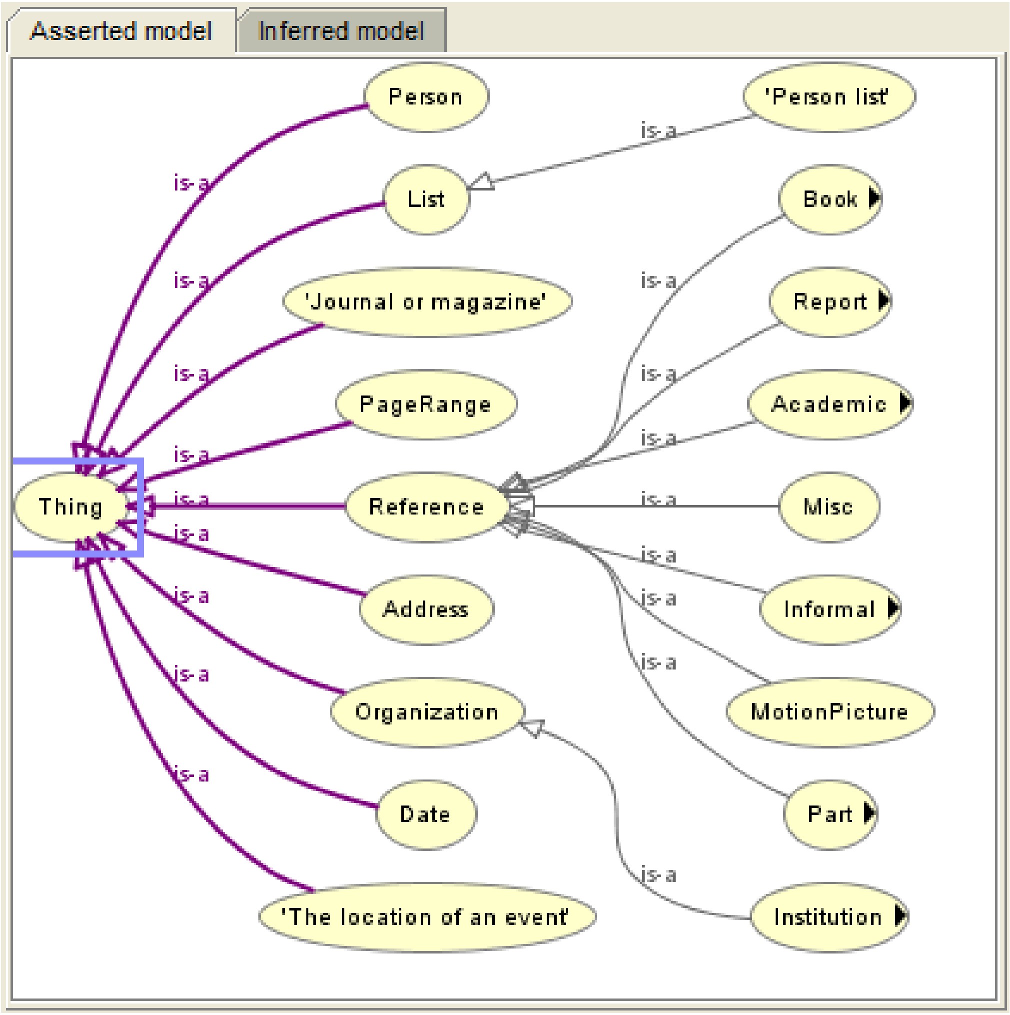


Figure 4 - Example of node-link visualization in the OWLViz software [24].

- **Zoomable visualizations:** display the lower nodes of the hierarchy embedded in their parent entities, allowing users to perform a "zoom" on children nodes, making one of them the current node. This visualization is efficient for a richer view of a certain level (the current one in the zoom). Due to lack of an overview of the ontology, authors like Rivadeneira [45] suggest improvements in the navigation tracking, like a level indicator (for depth analysis) and the information on what nodes were already visited.

  One tool that implement this type of view is Jambalaya [56] (Figure 5), a popular plugin for Protégé with the possibility to switch between different graphical representations of this type. There are also other tools with this representation, like Grokker [45] and CropCircles [40].

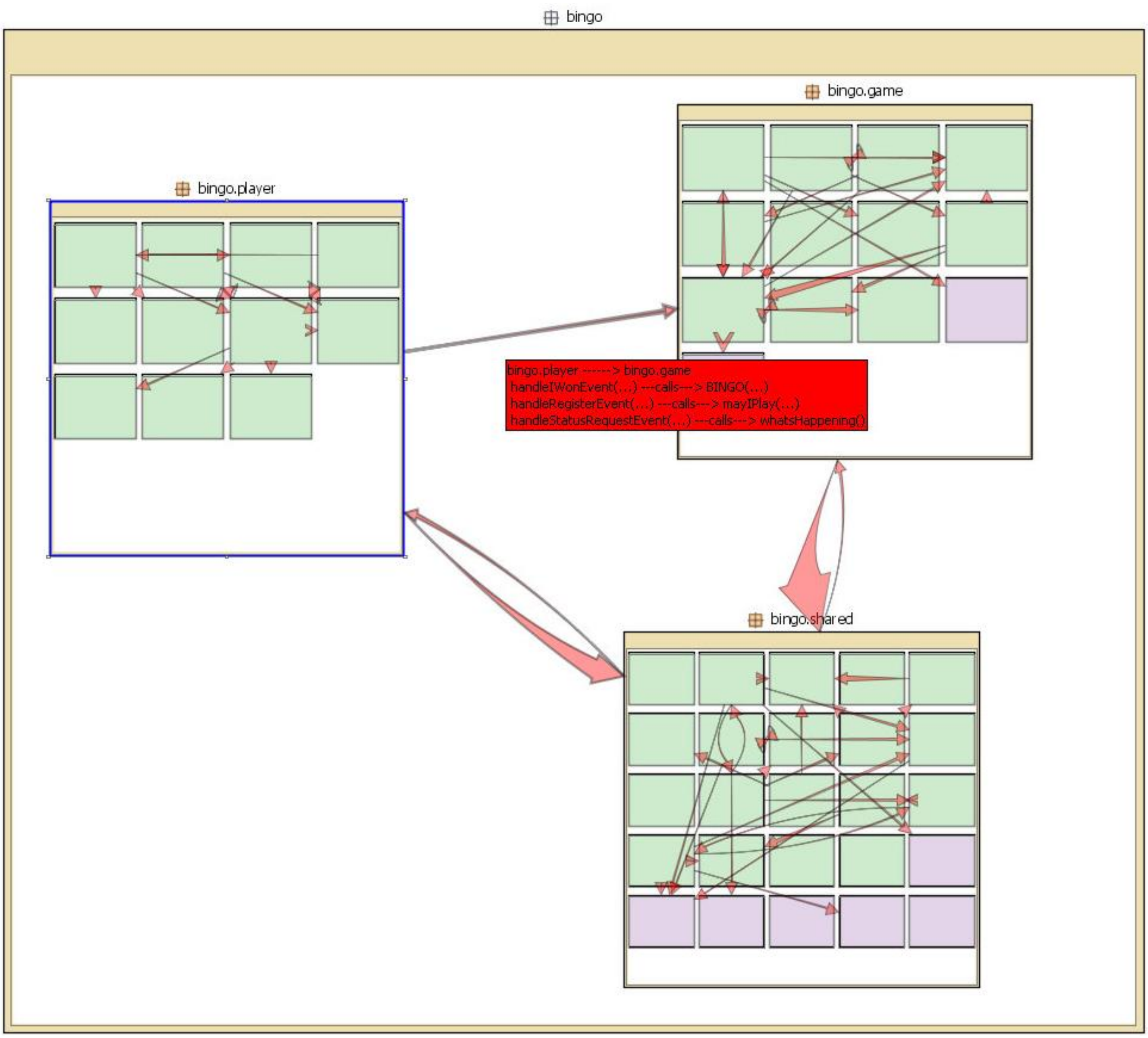


Figure 5 - Example of zoomable visualizations in the Jambalaya tool [56].

- **Space filling:** uses an area of the screen and divide it among the children nodes. Particularly useful when the user needs more focus on a predetermined relation or instance, since they could be highlighted in some visual way (color, size, etc.), facilitating its identification and distinction between different children. A consequence in this view is that it is difficult to identify the topology, especially with many levels. Even though it was classified by Katifori [28] as good representation for ontologies, the tools that implement this technique do not use it for ontologies. Some tools with this visualization are Information Slices [2] (Figure 6), TreeMaps [50], SequoiaView [68] and BeamTrees [21].

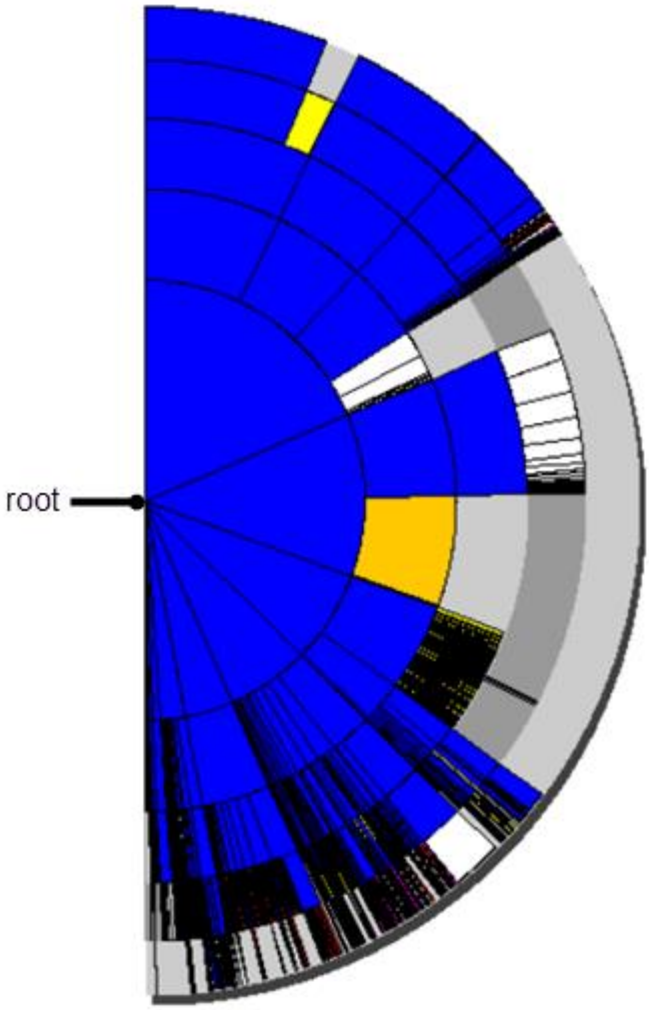


Figure 6 - Example of space filling visualization in the Information Slices software [2].

- **Focus + Context and distortion techniques:** distort the view presented in a part of the screen (usually a graph) to match the context and focus. Typically, an entity is in focus and the others around it are presented in sizes that diminish with distance. The user will then change the focus to navigate. Therefore, it is easy to navigate between related nodes, changing the focus. As in other types of views, the structure of the ontology is not well represented here, especially in a large size (larger than a few hundred) [28].

As implementation for this visualization there is the Protégé plugin TGVizTab [1] (Figure 7), which uses the relations to approximate nodes, making it more grouped and evident semantically similar entities. There are also other implementations such as HyperTree [55], StarTree [33], OntoRama [10] and BiFocal Tree [44].

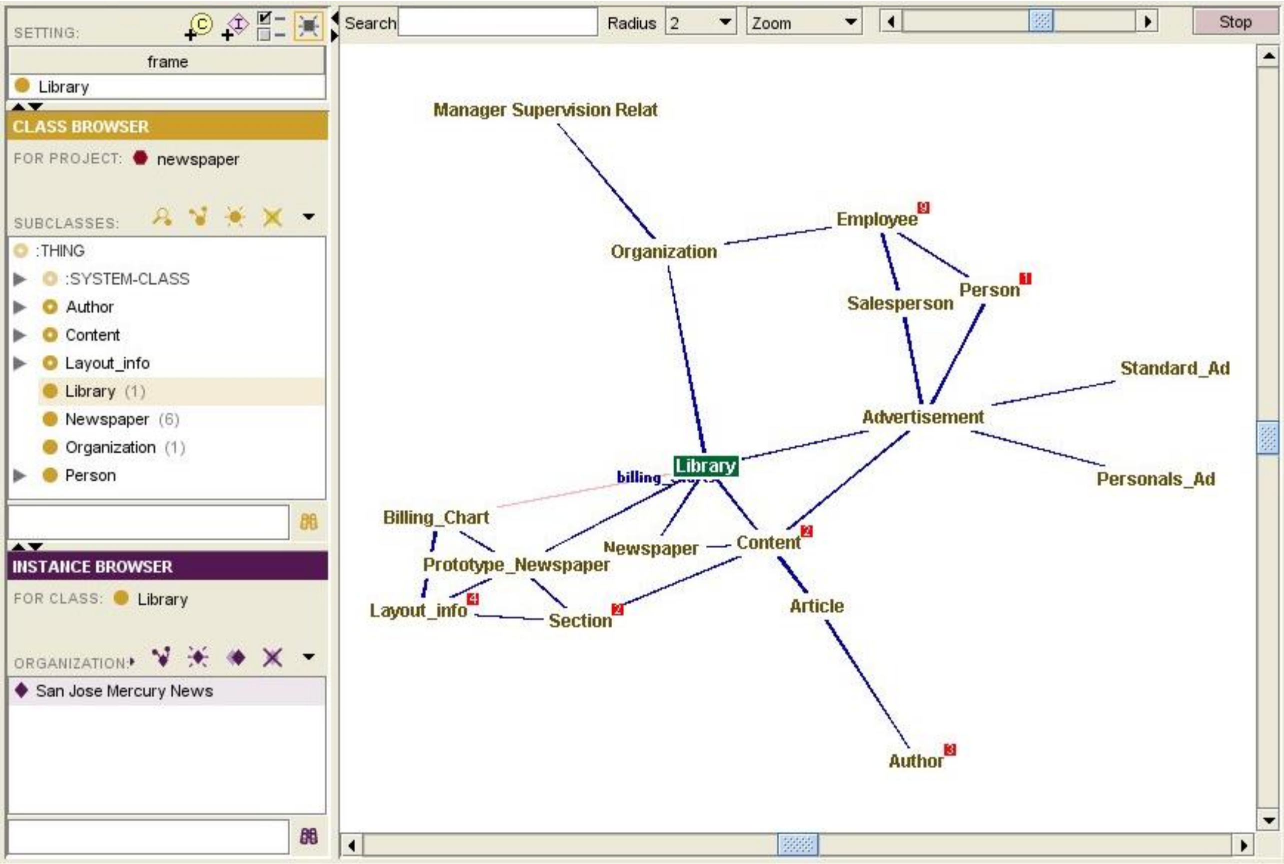


Figure 7 - Example of focus + context visualization on the TGVizTab software [1].

It is also important to note that the conclusion of studies that analyze multiple forms of ontology visualization converges with what was referred to as objective of this thesis: the configuration of the view according to the task performed by the user. The conclusion usually exposed is that the choice of the best way to view an ontology is a big challenge, and that more than one way to view must be disposed, as none can approach all tasks adequately [28] [29] [32].

3.2. Alignment Visualization

In a similar way to ontology visualization, the visualization of alignments should also consider the type of tasks to be performed and the user profile for which it was built for [19]. Several tools allow one to observe the structure of an alignment, typically built to assist the user in the task of matching two ontologies. The most common approach in alignments visualization is the representation of the two ontologies with a technique to display ontologies (such as mentioned in Section 3.1), adding elements that connect the entities of the two structures or enumerate the correspondences in a separate structure (table). There are no much variety of approaches currently offered in the state of the art.

The type of visualization most used to represent alignments is the one that displays ontologies as indented lists, and from each entity (node) lines are drawn connecting those involved in the correspondence. This is consistent with the conclusion of Katifori et al. [29] and Granitzer et al. [19], where the indented lists view was identified as the most familiar and easy for users. But this approach also carries some negative points such as the difficulty to handle a larger number of entities (only part of the structure is displayed at a time) or correspondences (overlapping lines hinders visibility); the difficulty of correctly represent the attributes of a correspondence when displayed along the connection line (Figure 8); and omitting part of the structure of ontologies (usually only classes are displayed, without representation of instances, attributes or relations). Some examples of tools using these forms of representation are COMA++ [3], YAM++ [37], PROMPT [39], COGZ [14] and AgreementMaker [7].

Some of these systems, such as AgreementMaker (Figure 8) add other features besides the representation with indented lists, in order to solve some of the difficulties that this display poses, such as coloring the connection lines to facilitate visualization in cases of a larger quantity of correspondences; displaying attributes of the relation in the connection line to avoid an additional consult for the end user; filters to search every tree, avoiding the manual search; etc.

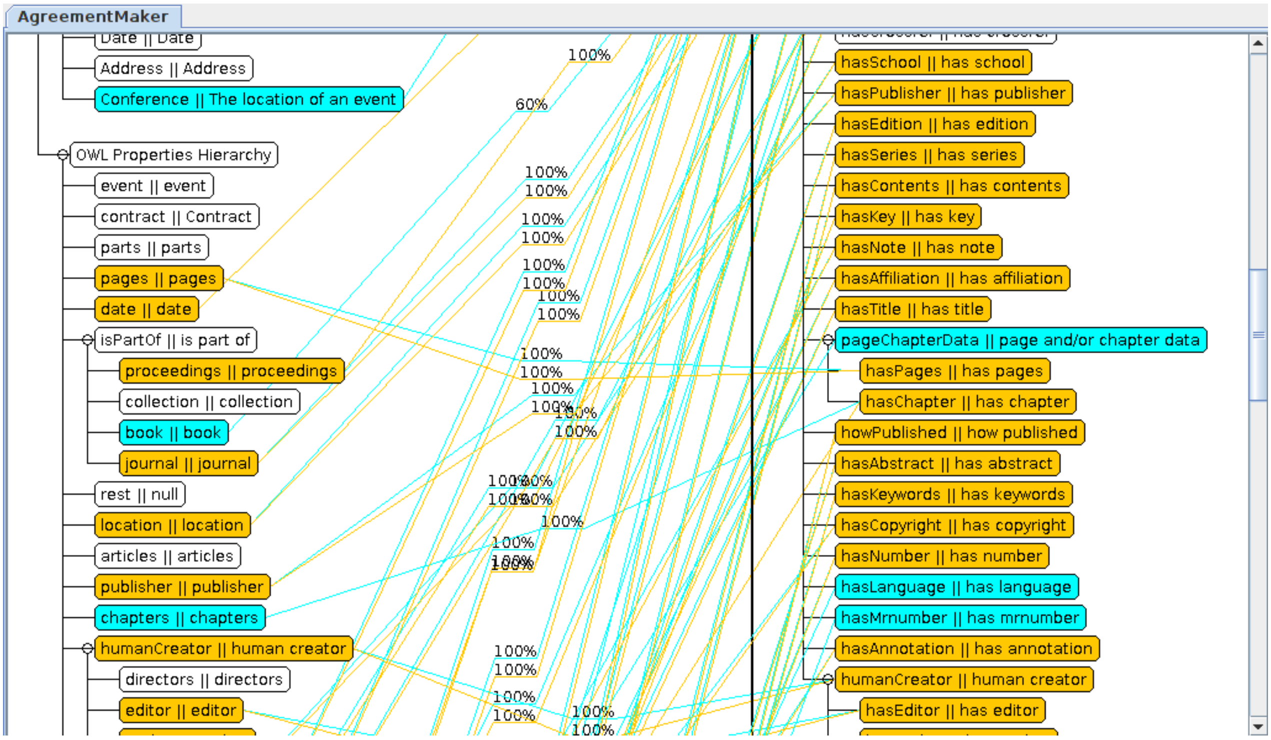


Figure 8 - Example of alignment view on the AgreementMaker tool [7].

Visualization of the category node-link are also employed in the representation of alignments, like in the OPTIMA tool [31] (Figure 9), which displays both ontologies this way, highlighting nodes containing a correspondence with a different color. This form of visualization allows a better overview than the indented lists representation. OLA [13] represents each element with a different geometric shape, but the correspondences have no visual representation since it displays ontologies separately. The difficulty in this category is having to select each node of interest to check the entity relations.

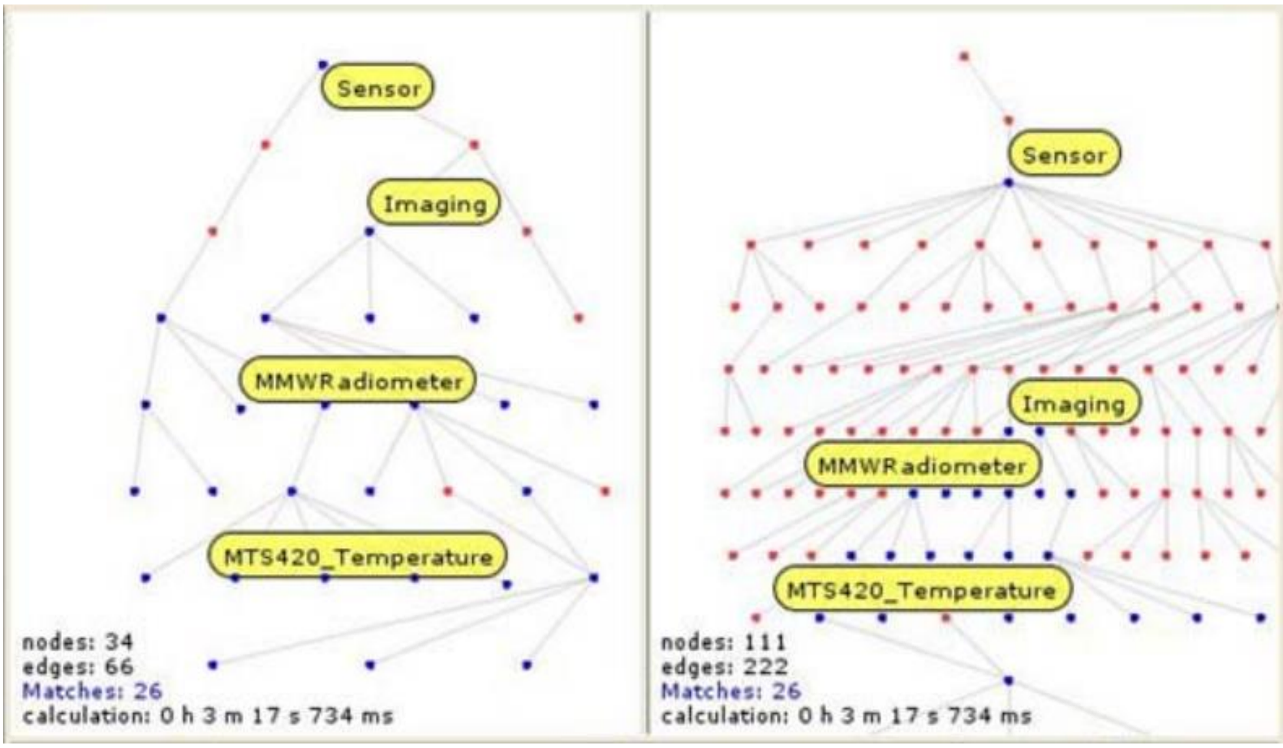


Figure 9 - Example of alignment view with the OPTIMA tool [31].

In the Focus + Context group, there is the Alviz [34] (Figure 10) approach, which represents ontologies by positioning the strongly related node entities in an aggregated manner (considering the entity in focus). Alviz also colors and change the size of the nodes according to the number of related concepts and matches. As with the connected nodes group, the difficulty is to see data connections and attributes of entities. HOMER [60] (Figure 11) tries to approach a solution to the view of correspondences, representing each ontology in the Focus + Context form, and linking the entities with correspondences through lines.

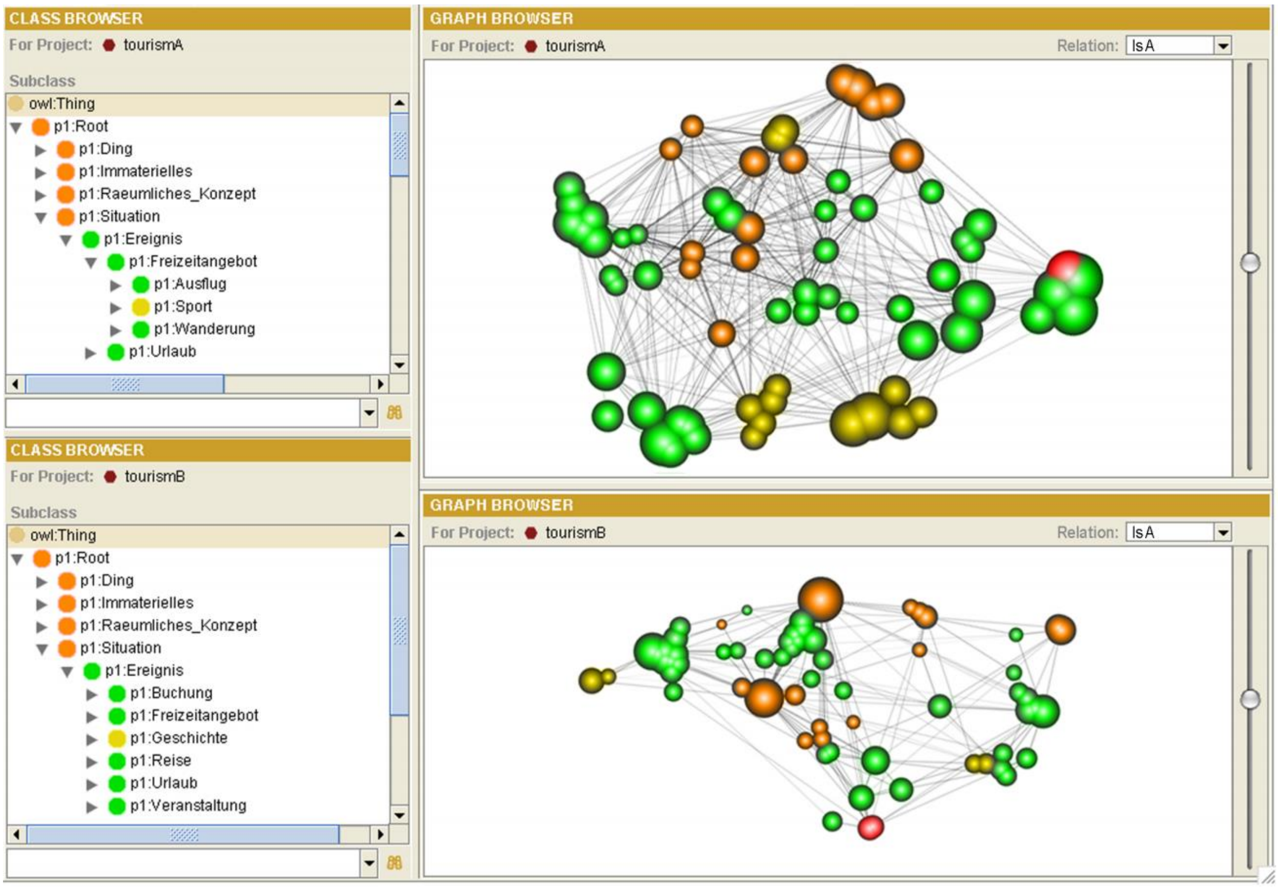


Figure 10 - Example of alignment view in the Alviz tool [34].

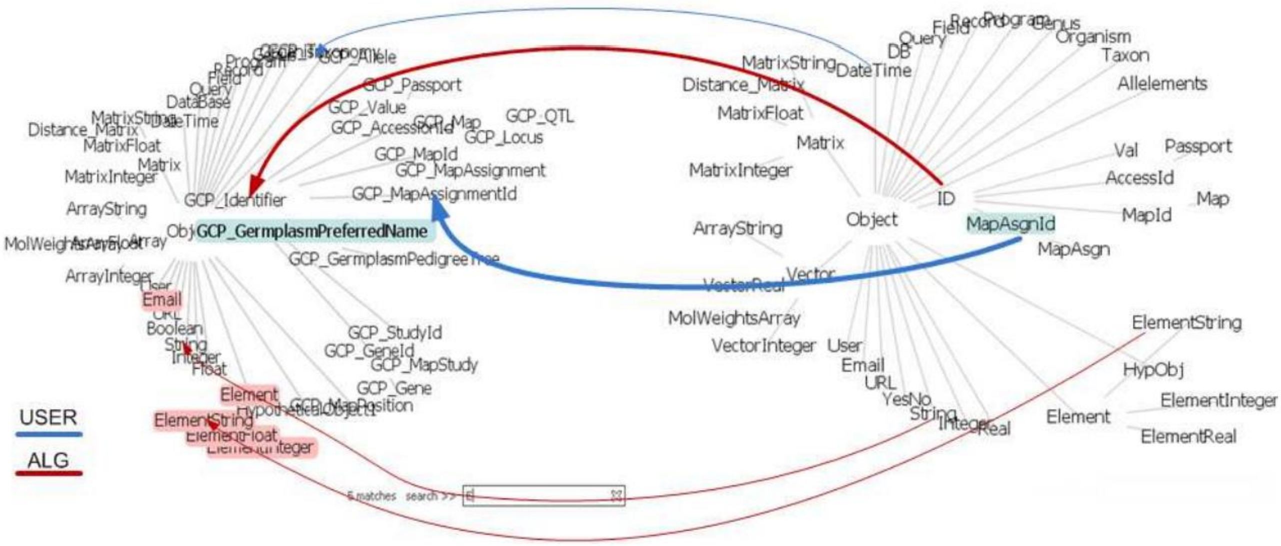


Figure 11 - Example of alignment view in the HOMER tool [60].

Finally, the space filling category also has representation for alignments with the COGZ tool [14] in a form of TreeMap (Figure 12). Due to be limited to represent hierarchical structures only, it uses alternatives to show other entity data (including correspondences), such as colors and internal symbols. COGZ uses colors to indicate areas where candidates for alignment can be found.

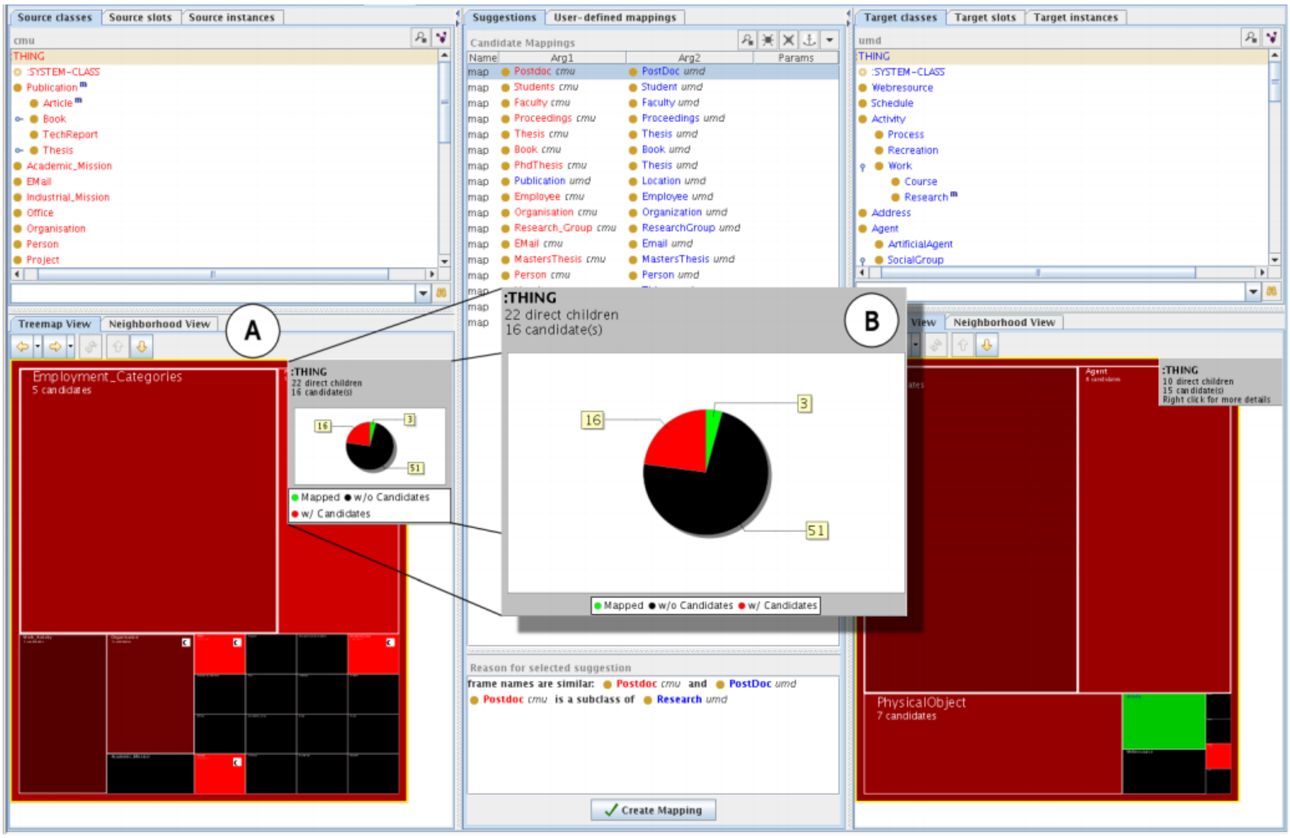


Figure 12 - Example of alignment view in the COGZ tool [14].

Many other tools for aligning ontologies provide an interface only for execution and configuration, without graphical display, such as Taxomap [22], LogMap [27] and Alignment Server [9] (shows the XML only). Therefore, this area still requires significant advances to meet the limitations of existing approaches and to answer the needs in ontology matching.

3.3. Evaluation of Visualization for Ontologies and their Alignments

Besides researches that analyze types of visualization available for ontologies and alignments, there are also works that evaluate their visualization, conducted with users of different profiles and technical levels. As result, they help to emphasize the most significant advantages and disadvantages of different types of visualization, as well as suggestions for improvement in the use of graphical techniques.

3.3.1. Evaluation of Ontology Visualization

Two works stand out here, the first conducted by Katifori et al. [29] which analyzes graphical tools (representing the categories described in Section 3.1); and the second led by Kriglstein [32], which reports the findings of interviews and online evaluations with users for the construction of ontology visualization tools.

The first evaluation [29], focused on four graphical tools representatives of some categories were chosen by the authors, all being graphical plugin implementations in the Protégé tool. This decision was made in order to keep the focus on the characteristics of the visualizations, not the implementation itself. The evaluation was conducted with 13 users of different levels of computer experience, but familiar with the ontology used in the experiment (University of Athens domain). Information retrieval tasks from different levels of complexity were then proposed and the results observed. Some examples of tasks are: to inform the year of birth of a teacher with a particular name; report the number of departments from a given faculty, etc.

The observed results for each of the visualization categories, as well as characteristics for tool development were the following:

- **Indented list:** received positively by most users, given the familiarity with the format of presentation used by file exploring programs from operating systems. Observed difficulties of some users to expand the nodes, where they expected to only click on the entity name would be enough to perform the action. Suggested the addition of buttons to expand and collapse all.

- **Node-link and tree:** received more negatively, users primarily complain about the lack of interaction with the presentation. They reported that accidental clicks on entities were common and caused a loss of focus on the desired item. It was suggested the possibility of ordering the nodes according to a criteria. They emphasized that it is a good representation for small ontologies.

- **Zoomable visualizations:** considered a good visualization by users, they highlighted the search and the transition between nodes as positive, as long as it does not take too long. Pointed out here that just as with the Context + Focus technique, when too many entities are presented, their labels end up overlapping, which is aggravated in this view by the relations that are graphically displayed as well.

- **Context + Focus and distortion techniques:** some users complained about a feeling of "chasing" the concept due to motion effects. Regarding the overview of the ontology, there were divided opinions between being positive and being chaotic. As points to be improved, it was suggested to have a search and a treatment to the overlapping captions.

In the second evaluation, conducted by Kriglstein [32], they seek out to raise the desired requirements for ontologies visualization, independent of the domain ontology to be used. The study was conducted with 16 people with intermediate and advanced levels of familiarity with ontologies. As such, more specific questions of the structure could be made, such as how relevant is the display of relations, annotations, search and filter, etc. Below are some important observations found in the research:

- **Understanding and structure:** stated that good visual representations of data help in understanding the ontology, particularly for non-specialist users. In addition, they hold a vital role in the development of ontologies, because without graphical resources the design of an ontology can be very time consuming and subject to errors.

- **Tools:** problems were found in general when dealing with very large ontologies. Many tools evaluated did not support user interaction for editing and navigation.

- **Expectations:** it was reported the desire to see instances in the overview of the ontology and how many instances exist in total, as well as search and filter elements. Navigations within the ontology were desired, like checking a relation and being able to navigate to the referenced element.

- **Considerations:** one should seek for a balance in the visualization, like having an overview of the ontology as well as a detailed option. A suggestion was to make it possible to visualize relations between subclasses and instances, for instance.

3.3.2. Evaluation of Ontology Alignment Visualization

Few publications evaluate forms of visualization for alignments. In particular, two studies were more relevant: the study led by Bo Fu et al. [16], which addresses two types of ontology visualization, pointing out the advantages and disadvantages of each applied to the task of evaluating alignments; and the research of Granitzer et al. [19], focused on visual forms of alignments to support semi-automatic alignment techniques.

In the usability study conducted by Bo Fu et al. [16], two forms of ontologies visualization were selected (indented list and node-link) to expose 36 users without experience with ontologies to the task of checking correspondence between entities. The support given by each view in understanding the semantics of the ontologies was evaluated (there was no representation of the alignment itself). In the end, they calculated the effectiveness, efficiency, workload and satisfaction, but there were no statistically significant differences between the display forms. This means that both forms served similarly to the needs of the scenario (navigate/explore the ontologies in order to find correspondences), having their differences in user preferences only.

Focusing on visual forms for generating alignments in a semi-automatic way (with user interaction), the work of Granitzer et al. [19] evaluated three kinds of visualizations: indented lists, node-link and space filling (in particular the graphical representation form TreeMap). Each visualization was evaluated against a list of requirements for working in a semi-automatic way, where none was able to meet all requirements. These results reinforce how difficult it is to find a single representation to meet all specific requirements that a user/scenario may require.

3.4. Concluding Remarks

This chapter presented state of the art works related to the field of study of this thesis. The main findings were the most common/important types of visualization for the structures studied and how these visualizations are perceived. After analysis of these data, it was determined the following visualization categories as the most relevant: indented trees, graphs (sometimes presented as the variation node-link) and context + focus. We noticed that the question of multiple alignments, however important, is mostly unattended (with exception of one reviewed tool). The results of this chapter were also important to highlight patterns and variations on preferences, and their conclusions reinforced the idea that several factors can influence on what must be shown and how.

Even though the collected information in these works is relevant and will be considered in the proposed approach, the current researches lack the analysis on how these visualizations are evaluated for different tasks. Therefore, a complementary research must be done to elucidate the best application of visualizations and information per activity performed. This is conducted as a survey with specialists on ontology alignments, considering the findings of these works, such as the main visualizations. The survey is described and analysed on Chapter 4.

4. ALIGNMENT VISUALIZATION SURVEY

The literature review regarding forms of visualizing ontologies and alignments (presented in Chapter 3) led to the need of eliciting with alignment users the most important requirements for related tasks, in order to complement what was presented. To the best of our knowledge, there is no similar research in the state of the art on the needs for the main tasks performed with alignments (creation, manipulation and evaluation). Related works usually evaluate one task directed to a specific domain. Moreover, the current visualization approachs/tools analysed were built for specific activities as well, thus, demanding a new evaluation to enable an analysis and comparison of requirements for each task.

Since most of the target users (alignments specialists) are located in different places, the format of a questionnaire was chosen in order to facilitate the contact. An online survey was then conducted aiming to identify relevant aspects and user preferences for visualization of alignments. The survey was sent by email to known ontology matching researchers and groups of interest on the subject.

4.1. Survey design

The survey was designed as a questionnaire divided in 10 sections with a total of 32 questions. The target of this research were subjects that work with alignments, specialists in the field, so it was built with the premise of previous knowledge on ontologies and alignments. The sections were organized as follows:

1. **Agreement:** terms and conditions for the survey.
2. **Profile and Expertise:** basic user information profiling.
3. **Ontology Alignment Creation (automated or semi-automated task):** enquires about the relevant information for visualization considering the task of creating new alignments with ontology matchers (with or without human intervention).
4. **Ontology Alignment Creation (manual task):** enquires about the relevant information for visualization considering the task of creating new alignments manually.
5. **Ontology Alignment Manipulation:** enquires about the relevant information for visualization considering the task of editing existing alignments.
6. **Ontology Alignment Visualization (approaches):** enquires about graphical representations of alignments (positive/negative aspects).

7. **Ontology Alignment Visualization (task):** enquires about the relevant information for comprehension of the alignments through visualization. Also enquires about existing forms for visualization of multiple alignments together.
8. **Ontology Alignment Evaluation:** enquires about relevant information, metrics and tools for visualization considering the task of evaluating (qualitatively and quantitatively) an alignment.
9. **Ontology Alignment Tools:** enquires about tools used to perform alignment related tasks (creation, manipulation, evaluation).
10. **Conclusion:** open space for final considerations and comments.

The structure of the survey includes a set of questions applied to different tasks, like creation, manipulation and evaluation, allowing later comparison of what information is more important for each one. To avoid misleading answers, most sections presented a question to validate if the user works with that kind of task (e.g. if a user does not perform manual creation of alignments, that section would be skipped), and some questions had a ‘not applicable’ or equivalent option. The proposed survey was first evaluated in a pilot testings with alignment users for minor adjustements. The complete questionnaire is available in the Appendix A.

4.2. Results and analysis

The results of this survey were used to assist and direct the development of an adaptative approach for visualising and manipulating alignments, guiding which aspects should have more attention, and what unforeseen aspects should be addressed to add greater value to the end user experience.

4.2.1. Profiling

During the period of research (50 days), 12 responses were received from 10 different countries, as seen in Table 1. According to the data in Table 2, the majority of respondents work with ontologies as researchers (multiple choices were possible), and evaluate their level of expertise with ontology matching as high (Table 3). This confirms a target audience, which is specialist expert on ontologies and alignments.

Table 1 - Survey repondents countries.

Country	Responses	
Algeria	1	8.33%
Brazil	2	16.67%
China	1	8.33%
Croatia	1	8.33%
Czech Republic	1	8.33%
France	2	16.67%
Italy	1	8.33%
Portugal	1	8.33%
Sweden	1	8.33%
United Kingdom of Great Britain and Northern Ireland	1	8.33%
Total Respondents: 12		

Table 2 - Survey repondents context for working with ontologies.

Activity	Responses	
Researcher	8	66.67%
Professor/Lecturer	5	41.67%
Non-Academic activity	0	0.00%
Other	1	8.33%
Total Respondents: 12		

Table 3 - Survey repondents auto-evaliation on expertise with ontology matching.

Level	Responses	
Low	0	0.00%
Medium	4	33.33%
High	8	66.67%
Total Respondents: 12		

When asked about tasks performed when working with alignments (Table 4), most users answered that they perform activities involving automatic creation of alignments, followed by evaluation. This suggests those as being the tasks that are performed the most in the field, having events like the OAEI dedicated for it.

Table 4 - Survey repondents performed tasks

Level	Responses	
Alignment creation (automated)	10	83.33%
Alignment creation (manual)	6	50.00%
Alignment manipulation	6	50.00%
Alignment evaluation	8	66.67%
Total Respondents: 12		

The size of ontologies that the users usually work with (for each activity performed with alignments) was enquired, using the definition for size according to the number of entities *E*, where “Small” is E ≤1000, “Medium” is 1000 < E < 10000 and “Large” is E ≥ 10000. In Figure 13 it is possible to see a balanced distribution (being “Large” the most common), where the predominant sizes per task were:

- Alignment creation (auto or semi): Large
- Alignment creation (manual): Small
- Alignment manipulation: Small / Medium
- Alignment evaluation: Large
- Alignment visualization: Large

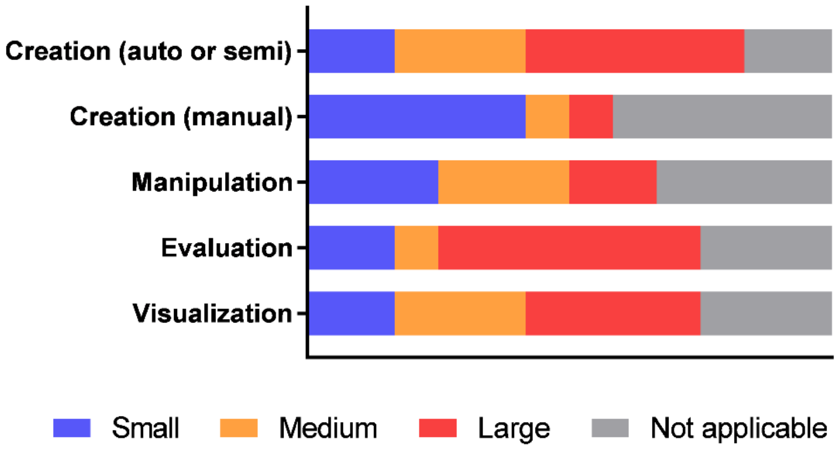


Figure 13 - Survey repondents working ontology size per task.

4.2.2. Importance and relevance analysis

For each of the main alignment tasks contemplated by this survey (creation, manipulation, evaluation and visualization), a comparative evaluation was created to measure how the users rate the relevance of information against the availability of it in existing tools. The ontology information inquired was classes, object properties, data properties and instances; while the alignment information inquired was correspondences, correspondences confidence, correspondences type and number of correspondences (total, per type and per entity). The choice of these elements was based on what was found in related work and state of the art as relevant for visualization.

The analysis of this data (importance and relevance) allows the identification of cases that need improvement, where the relevance for a given information when compared to their availability in tools are pointed by users as having a significant gap. In an ideal scenario, a proportional distribution when comparing the two evaluations for a given information would be expected (if evaluated as important, should be graded as good in tools).

For evaluation of importance and relevance measures in this comparison, Likert scales were used. Likert scaling is a bipolar method (a balanced scale of answer choices on both sides of a neutral option), measuring either positive or negative response to a statement. In order to present the comparison, results were grouped by task in a diverging stacked bar chart, following the recommendation from the work of Robbins and Heiberger [46]. The percentages of respondents who evaluated the statement in a positive way (relevance as important and availability in tools as good) are shown to the right of the zero line; while the percentages who evaluated in a negative way are shown to the left. The percentages for neutral responses are split down the middle.

Evaluation was made for each task considering ontology-centric data (classes, object properties, data properties and instances), alignment-centric data (correspondences, their confidence and relation type) and alignments counts (total of correspondences, and subtotals by entity and type). Besides this information inquired in the relevance analysis, some users also pointed out that would be important to visualize logical conflicts and a history of changes. The charts for comparison by task are presented and analyzed below:

- **Ontology alignment creation (automated or semi-automated task):** the results represented in Figure 14 allow identifying classes and correspondences as the most important information in this task (their relevance is presented predominantly at the right side of the chart). What also draws attention is that all correspondences count (total, by entity or by type) have a significant percentage of evaluations as "moderately important" (neutral), indicating that they are probably not very relevant for the task.

  When comparing availability against the relevance of information, the most significant difference (highly evaluated as important, but very poorly evaluate for its availability in tools) was found for "instances", which means that this information is the one that deserves more attention. Other ontology related information were also evaluated as having significant poor representations, suggesting that this must be a point for improvement.

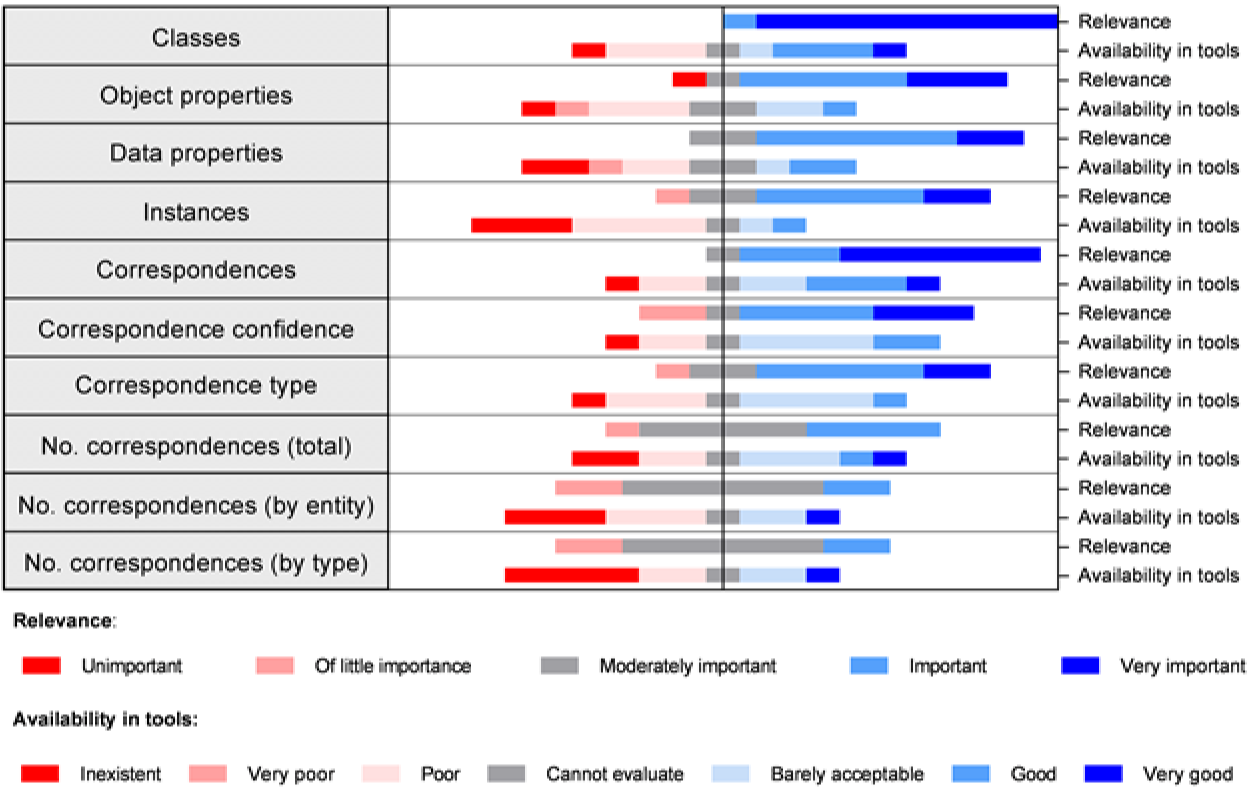


Figure 14 - Evaluation on information for automated alignment creation

- **Ontology alignment creation (manual task):** the results represented in Figure 15 show a significant drop in the quality of how most information related to manual ontology creation is shown on tools when compared to the evaluation for automated creation.

  Besides classes (once again), correspondences, their confidence and type are the most relevant information for this task, which is consistent with the purpose of creating new correspondences from scratch.

  What stands out is that although they are very important information to perform this task, correspondences and its characteristics have poor availability in the current state of the art.

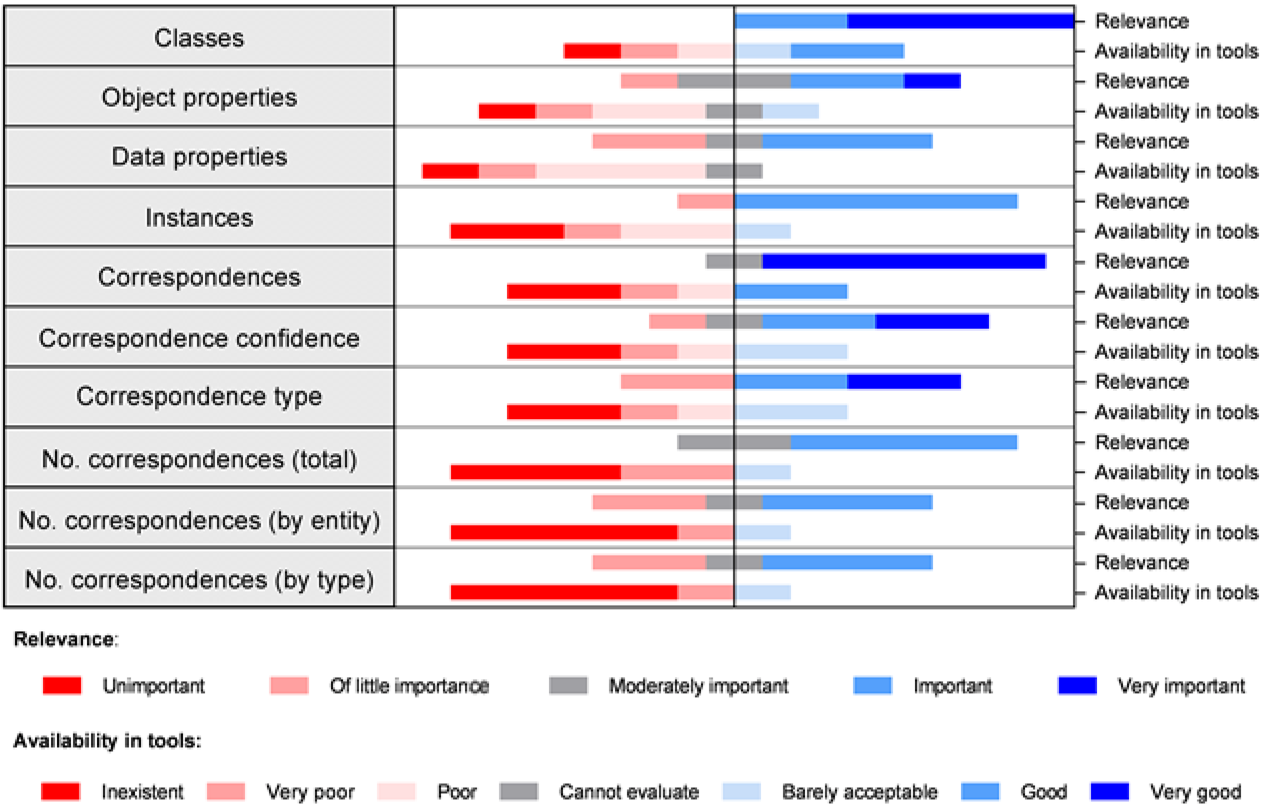


Figure 15 - Evaluation on information for manual alignment creation

- **Ontology alignment manipulation:** the results represented in Figure 16 show that all ontology information were evaluated as important, along with correspondences, what seems consistent, since it is what is actually being manipulated in this task.

  Besides correspondences counts (that in this task also maintain their distributed rate), the only other information that had any evaluation as not important was "correspondence type”. This represents part of what was observed in the current tools in the state of the art (in works from Chapter 3), where the relation of equality is often assumed as the default or only option for type.

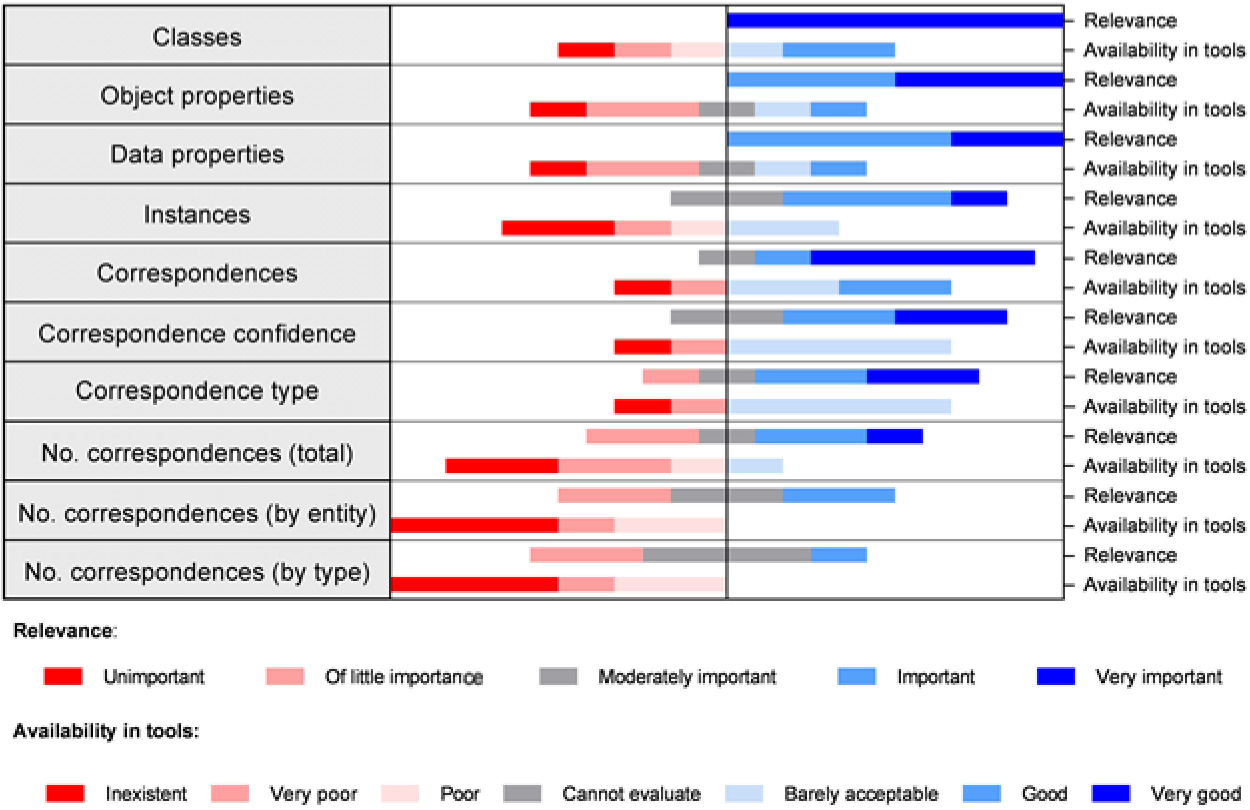


Figure 16 - Evaluation on information for alignment manipulation

- **Ontology alignment visualization:** the results represented in Figure 17 show that ontology information and correspondences are important for visualization, in a similar way that was observed for manipulation.

  Since this task sometimes tries to provide an overview comprehension for the alignment, correspondences counts should have a better availability in tools (helping on understanding size, distribution, etc.), but instead they have the worst rating for information in this task.

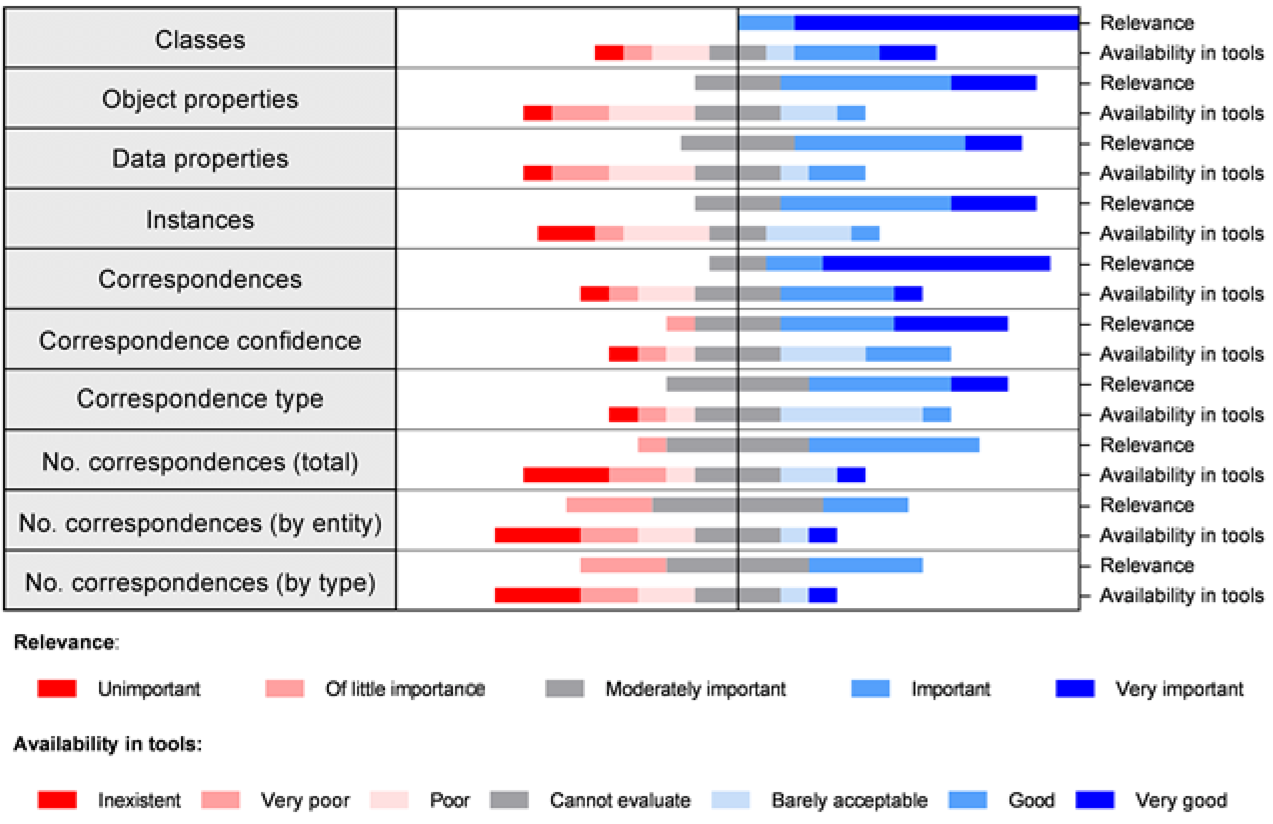


Figure 17 - Evaluation on information for alignment visualization

- **Ontology alignment evaluation:** the results represented in Figure 18 show that classes and instances are vital information for this task, both being evaluated entirely as important (not even neutral evaluations). Since this task may take into consideration all data available, the majority of other information was evaluated as important on some level also, being the exception here once more the correspondences counts.

  What stands out in the current tools is that, except for classes, other ontology information seems to lack better representation. Considering the importance rate for instances, this seems to be the information that lack more improvements for this task.

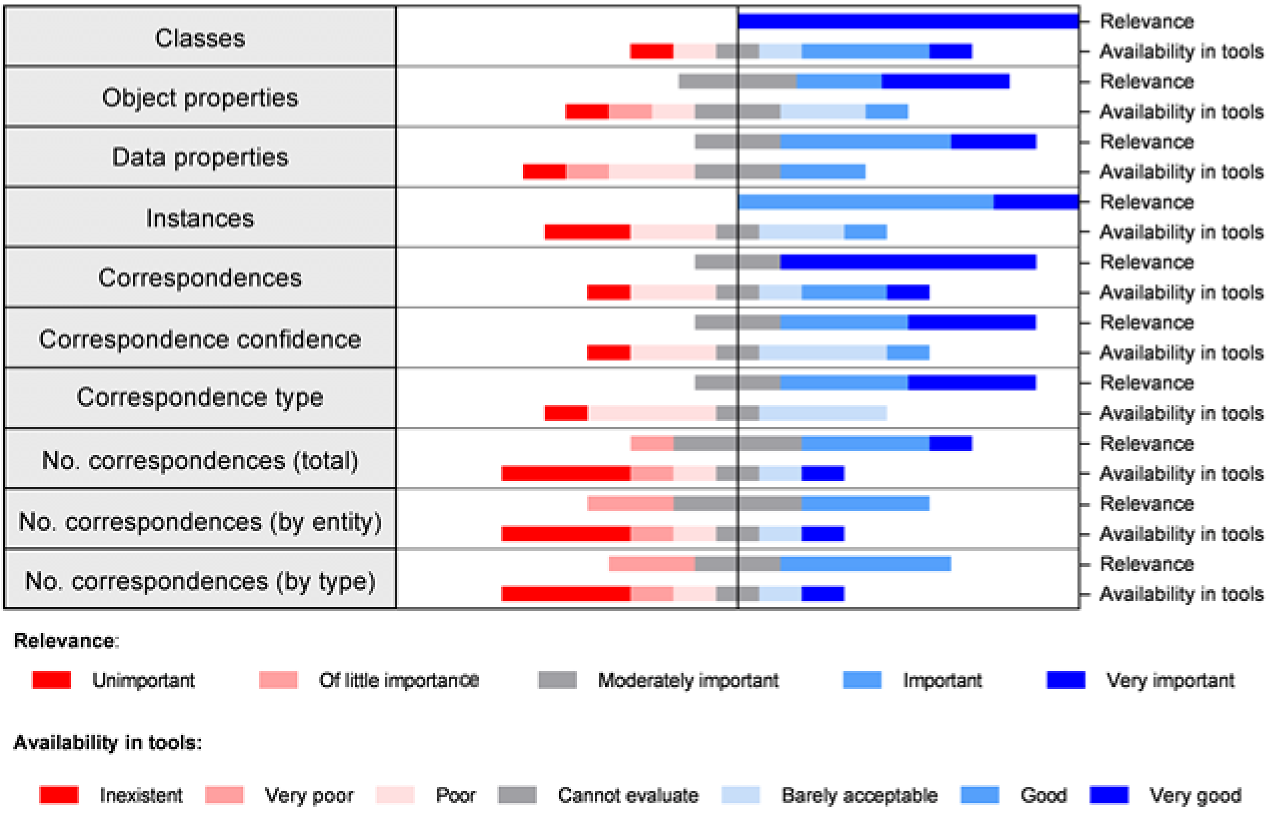


Figure 18 - Evaluation on information for alignment evaluation

4.2.3. Visualization approaches analysis

The most common approaches for visualizing ontology alignments, according to the literature review presented in Chapter 3, were submitted to the users for evaluation of positive and negative aspects. The main findings and defining concepts were identified, and are presented below per approach:

- **Indented trees:** the users perceive this approach as a good visualization for all the tasks that involve alignments, mainly for creation of new ones.
  - Concept for positive aspects: Understanding
    Several concepts related to the understanding of ontologies were emphasized, such as how “easy" the interpretation of this representation is, as well as usefulness for an overview of the entire ontology, with the possibility to navigate through it easily.
  - Concept for negative aspects: Ontology Size
    This makes sense because of the difficulty one has when very large ontologies are visualized in this representation. Some users in this sense mentioned that as a consequence, it became difficult to see information beyond the nodes of the hierarchy (usually only classes), as well as identifying multiple inheritance.

- **Graphs:** less "popular" than the indented trees, but still of interest for all activities, mainly for evaluation and creation. Commented by users as being the least useful of the views proposed in practice, yet interesting for exploration.
  - Concept for positive aspects: Visualization level
    It was reported that graphs make it possible to have a complete and general vision. It is also pointed out how in this view the higher level has a better idea of relationships, including multiple inheritance.
  - Concept for negative aspects: Understanding
    Unlike the visualization of indented trees, this view is often confusing and looks like a "mess". One factor that influences this is the size of the ontologies and alignments, generating large graphs sometimes.

- **Focus + Context:** evaluated as most appropriate primarily to the evaluation activity.
    - Concept for positive aspects: Details
      As the technique proposes, it is easier to see the details of each correspondence or focused node. Makes it possible to deal with alignments and ontologies of a larger size.
    - Concept for negative aspects: Information Suppressed
      It is reported that while enables better view of the entity on focus, it becomes more difficult to have an overview, including the relations between entities.

- **Others:** the users made a few suggestions for visualization improvements, like combining Indented Trees and Focus + Context (common approach in tools); present a count for relations (or represent with edge thickness); cluster displaying (macro nodes that expand); and variations of indented trees.

4.2.4. Evaluation methods analysis

In this survey, most participants reported that they perform alignment evaluation activities. Moreover, among these users, the majority uses techniques that evaluate compliance by metrics (Table 5). Besides that, one user mentioned that in addition to those listed, query-based metrics were also used.

Table 5 - Metrics used for evaluation

Metric	Responses	
Compliance with gold standard (precision, recall, f-measure and weighted variants)	7	87.50%
Logical reasoning based evaluation	4	50.00%
Other	1	12.50%
Total Respondents: 8		

Regarding multiple alignments evaluation (for instance against a gold standard), due to a difficulty in finding tools with the resources to perform this evaluation, users typically perform the activity manually. The minority of users responded that they use some kind of visual resource, and these commented that it is typically a comparative table.

4.2.5. State of the art tools analysis

Since this work aims to handle multiple alignments, users were asked if there was support for this on tools. It was pointed out that most of their currently used tools do not have capabilities to display multiple alignments.

In order to help to develop the proposed approach with the strengths of the current tools and trying to solve possible flaws, users were asked to point positive and negative aspects by task. The main findings and categories generated from grounded theory for each task are presented below:

- **Ontology alignment creation (automated or semi-automated task):** this task is related to the use of matching systems, so users were able to evaluate based on their experience with the task.
    - Concept for positive aspects: Automated techniques
      It was emphasized by users the value that using matchers (and their techniques/algorithms) provide, like dealing with large ontologies, logical repair, better performance (compared to the work that would be spent manually), etc.
    - Concept for negative aspects: Display format
      Reported by users as something typically textual, incomplete and more developer-oriented. Reinforcing the need for a visual approach.

- **Ontology alignment creation (manual task):** users described their experience as a “long process”, sometimes using tools for ontology creation in order to assist it.
    - Concept for positive aspects: Graphical interface
      Here, the few users who have some tool to help in the process see their graphical interface as the most important factor, facilitating the creation and edition (compared to the textual file manipulation).
    - Concept for negative aspects: Visualization for the task
      Tools used here are not designed for this particular task, and therefore it is difficult to find the data and easy lose track of changes. It is further proposed that a tool could suggest correspondences.

- **Ontology alignment manipulation:** users assimilated this task as something very similar to manual creation of alignments, since both make changes in the structure of the alignment.
    - Concept for positive aspects: Graphical interface
      In this evaluation, the users repeated the opinion from manual creation, on how important an interface is to detect inconsistencies and for fast operations.
    - Concept for negative aspects: Visualization for the task
      It lacks, for example, a parallel comparison of the alignments involved. Respondents mentioned again that the tools either are not task-oriented or are not user-friendly to the end user.

- **Ontology alignment visualization:** it was highlighted by the users the value a representation for the alignment has. Some users also pointed the gains that performing manipulation directly on the view would bring.
    - Concept for positive aspects: Visualization of large data
      In addition to the benefits a visualization provides, users also pointed out that it is particularly useful for interpreting large alignments / ontologies, when using appropriate resources (e.g. subgraphs).
    - Concept for negative aspects: Information overload
      It was pointed out that when many data are displayed, it is difficult to find or view a particular information.

- **Ontology alignment evaluation:** users appreciated the existence of any kind of interface to manage results of an evaluation.
    - Concept for positive aspects: Interface benefits
      More than simply having an interface, the features it provides, such as highlighting information with colors, inconsistencies, etc. are the greatest benefits to the tool.
    - Concept for negative aspects: Working Format
      They mentioned that the typically used format is adequate only for developers, making the visualization and interpretation difficult to end users.

4.3. Concluding Remarks

This chapter presented the main findings on the survey submitted to alignment users in order to identify the most relevant aspects on visualization for alignment tasks. In summary, we conclude from this survey some aspects that are important to consider for the visualization of ontology alignments:

- classes are very important, for all tasks;
- instances need to have a better visualization;
- the display of ontology/alignment elements is more relevant than correspondences counts;
- different tasks have a different rate for the importance of a given information, demanding different configurations on what must be shown;
- visualizations have better/worse use depending of the scenario presented (ontology size, task, etc.);
- multiple alignments are very important and barely supported.

These findings, along with the significant amount of poor evaluations on the visual availability of information in tools, supports the proposal of this thesis: the need to have an adaptive approach for graphical representation, that presents what the user needs to work with depending on the use case.

5. AN ADAPTIVE APPROACH FOR ONTOLOGY ALIGNMENT VISUALIZATION

This chapter presents the proposed approach for the visualization of ontology alignments, based on what was observed on the state of the art and on the evaluation performed with specialists.

One of the objectives of this work is to propose an adaptative approach, which that takes into consideration the different needs of users and different ontologies characteristics in order to present a suitable visualization of alignments for each situation. Following the analysis of related work, the state of the art and mainly the results of the conducted survey, it was concluded that a single visualization solution would not be adequate for dealing with the variety of use cases.

We found evidence of similar conclusions throughout the related works from ontology and alignment visualization studies, like the one conducted by Bo Fu et al. [16] which concludes that “their applications should thus be determined upon specific ontology characteristics, visualization needs and user goals”; the one led by Katifori et al. [28] that summarizes the analysis saying that “there is not one specific method that seems to be the most appropriate for all applications and, consequently, a viable solution would be to provide the user with several visualizations”; the one proposed by Silva [53], that states “techniques for visualizing ontologies should be based on effective graphical representations and interaction techniques that support users tasks related to different entities and aspects”; and the evaluation of Granitzer [19] that affirms “no single visual representation is capable of fulfilling all requirements”. Also in open-ended questions of the survey, statements like “a tool has to provide different visualizations to balance for their advantages/disadvantages and since they may be suitable for different tasks, user experience and preferences” once more reaffirms the importance of an adaptive approach.

This chapter will detail this adaptive approach on how to handle many possible variations when trying to display an alignment using the developed prototype (VOAR system).

5.1. Adaptive Approach for Ontology Alignment Visualization

The survey, presented in Chapter 4, provided orientation on what should be the focus of each task and on what improvements should be made on current visualizations. Based on that, an approach composed of a combination of pre-dermined visualizations for optimal use in different pre-determined scenarios was planned. Considering what information is desired per task, the usual ontology size and other information gathered, suggested visualizations would be provided. However, even though the use of these configurations would probably work for most users, under certain circumstances (user preferences, domain-specific characteristics, etc.) these proposed approachs might not be ideal.

Therefore, the proposal was directed to build an approach with a higher level of customization on the interface layout and information displayed. This approach proposes the organization of the visualization in two windows, similar to the focus + context approach, allowing the user to determine how the screen will be split to show the windows: in a horizontal or vertical way (Figure 19). Each of these windows can hold a particular visualization, thus enabling the combination of any available visualization to match the requirements for a specific scenario or preferences/needs. For example, a user can have an overview and a more detailed representation or a space to perform operations on data.

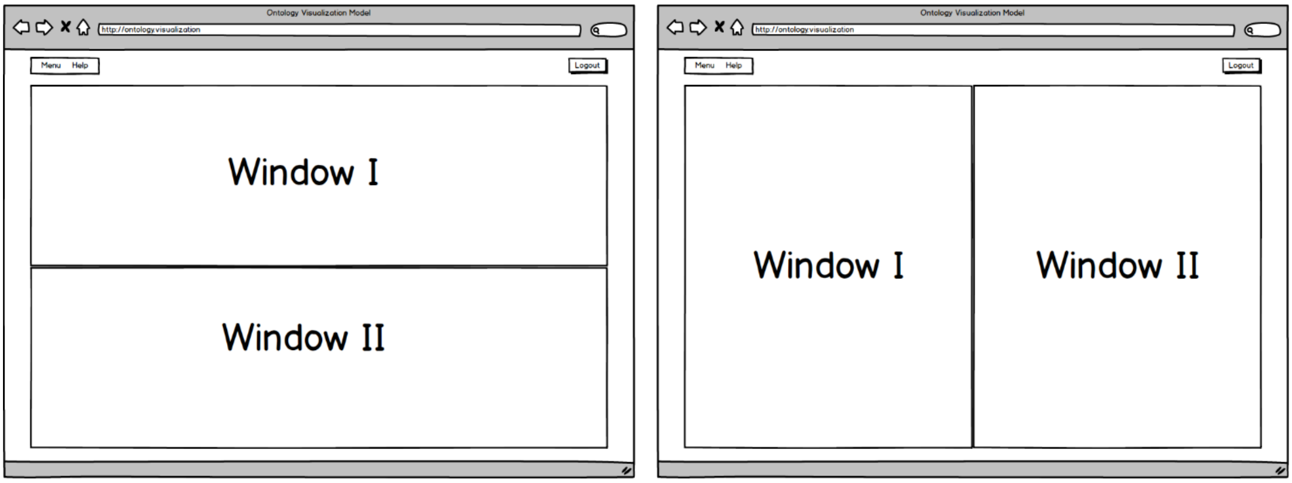


Figure 19 - Horizontal (left) and vertical (right) splits on the proposed approach

The configurable split of the screen also allows a choice according to preferences or particularities of a given visualization (e.g. indented trees would probably fit better in a vertical split). With this split, the approach enables to complement weak points of one visualization with another parallel representation of the data, allowing the user to focus on the best aspects of each.

The concept of two visualizations for the alignments, each one in a window, will contemplate some representation for their ontologies and correspondences. Allowing the users to choose the most adequate visualization for his/her working alignments, the ontologies involved (considering the size, number of relations, etc.) or even individual preferences (a preferred type of visualization).

Given the importance of multiple alignments for alignment related tasks and the results observed from the survey (almost no current approach / tool supports, being done manually in many cases), all visualizations were planned to support multiple alignments. The proposal for working with multiple alignments considers several alignments (A_1, A_2, A_3 ... A_n) for the same pair of ontologies source O_1 and target O_2. It should be up to the user to choose at any time which alignments to plot on the visualization, adding and removing them in time of visualization. This should facilitate the analysis and comparison of the impact that an alignment causes to a set.

It is believed that with multiple alignments, this work contributes to tasks where it is necessary to work with parallel alignments and make comparisons. Such as the alignment creation activity, where an engineer may be interested in observing which correspondences are currently an agreement in other alignments to build a reference one; or in the alignment evaluation activity, where this comparison assists in the evaluation of the quality of the correspondences.

According to the survey results, the following visualizations were proposed for the initial configuration: indented trees, graphs, node-link, correspondences list, quantitative evaluation and qualitative evaluation. Besides visualization, manipulation of alignments is also possible is some cases.

Below these initial visualizations are detailed:

- **Indented trees visualization:** perceived in the conducted survey as good visualization for all tasks, and evaluated as familiar and easy to understand, the visualization of an indented tree (Figure 20) shows the ontologies as hierachical trees. Relations and instances will be presented at a level below in the class hierarchy, identied by icons provided for each entity type (as described in Table 6).

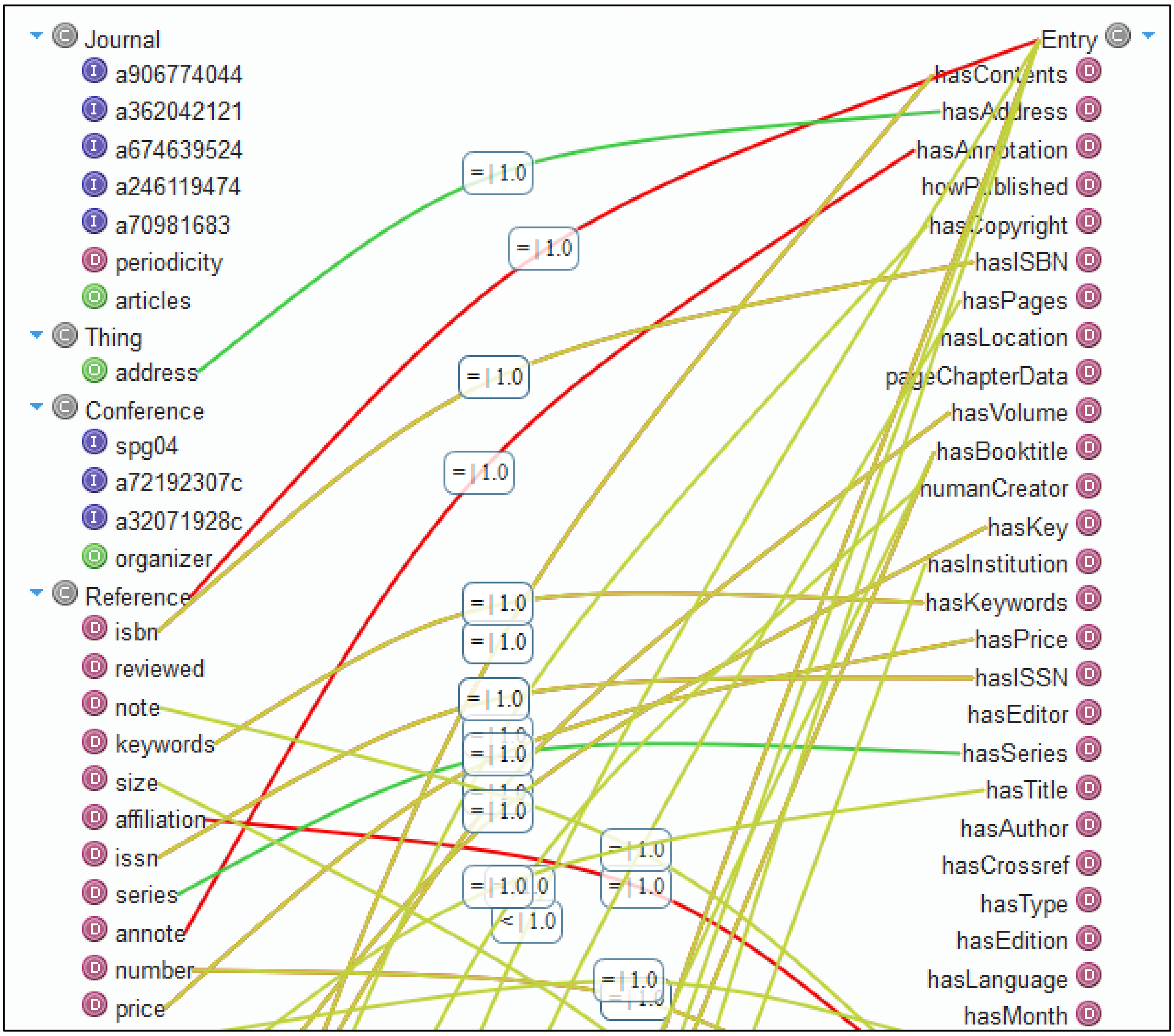


Figure 20 - Example of multiple alignments in the indented trees visualization of VOAR

Table 6 - Visual representation of ontology entities on the VOAR system

Icon	Entity
	Class
	Data property
	Object property
	Instance

Correspondences are shown as lines connecting two entities, and correspondences information, such as confidence and type, are displayed above this link. To identify the alignment that contains a correspondence when facing multiple alignments in the representation, a visual feature is used: the color assigned to each alignment in the library is applied to the visual representation of the correspondences.

- **Graph visualization:** a good representation for a complete overview of the alignments, is pointed in the survey as a “more realistic ontology visualization”, since it can connect all entities, and represent better multiple inheritance (Figure 21).

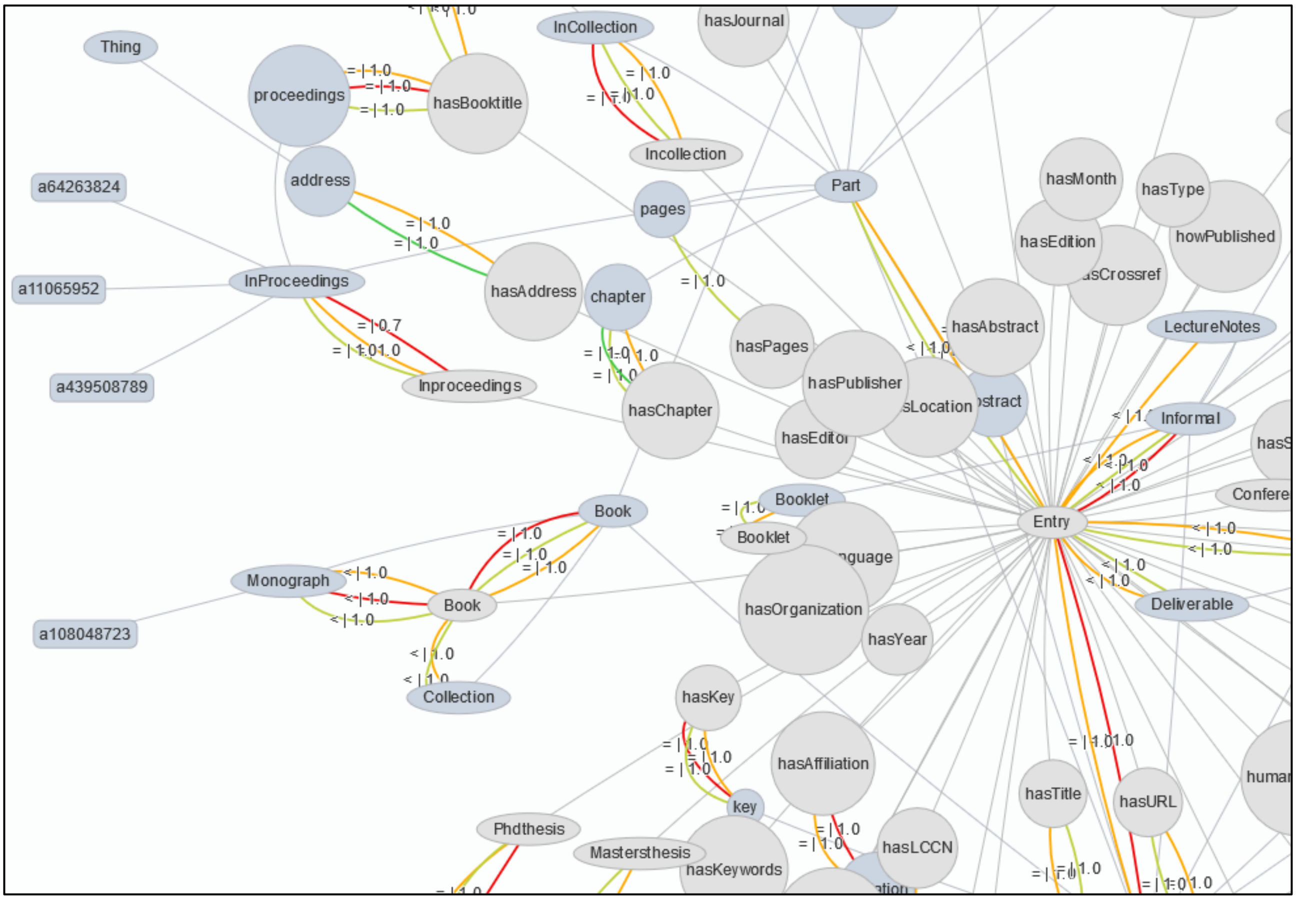


Figure 21 - Example of multiple alignments in the graph visualization of VOAR

In this visualization, each ontology is given a different neutral color, and each entity type is represented in a different shape (ellipses for classes, circles for properties and squares for instances) facilitating the identification by the user (since it lacks a hierarchical structure and everything is a node). To discern affiliation between entities (e.g. classes to instances) from correspondences, the same logic of color per alignment from indented trees is used to draw links. This way, correspondences are represented as a link between nodes, using the alignment color from the library.

The user can also change the focus of the view and zoom in/out, focusing on a particular part or in the overview.

- **Node-link visualization:** similar to the graph visualization, it provides a more oriented view to display the nodes in the graph (Figure 22). The representations of entities by shape, correspondences by alignment color and correspondences data on over the link are maintained. Although it helps in the display of the ontologies structure, the connections that represent correspondences tend to make the representation overloaded in certain points.

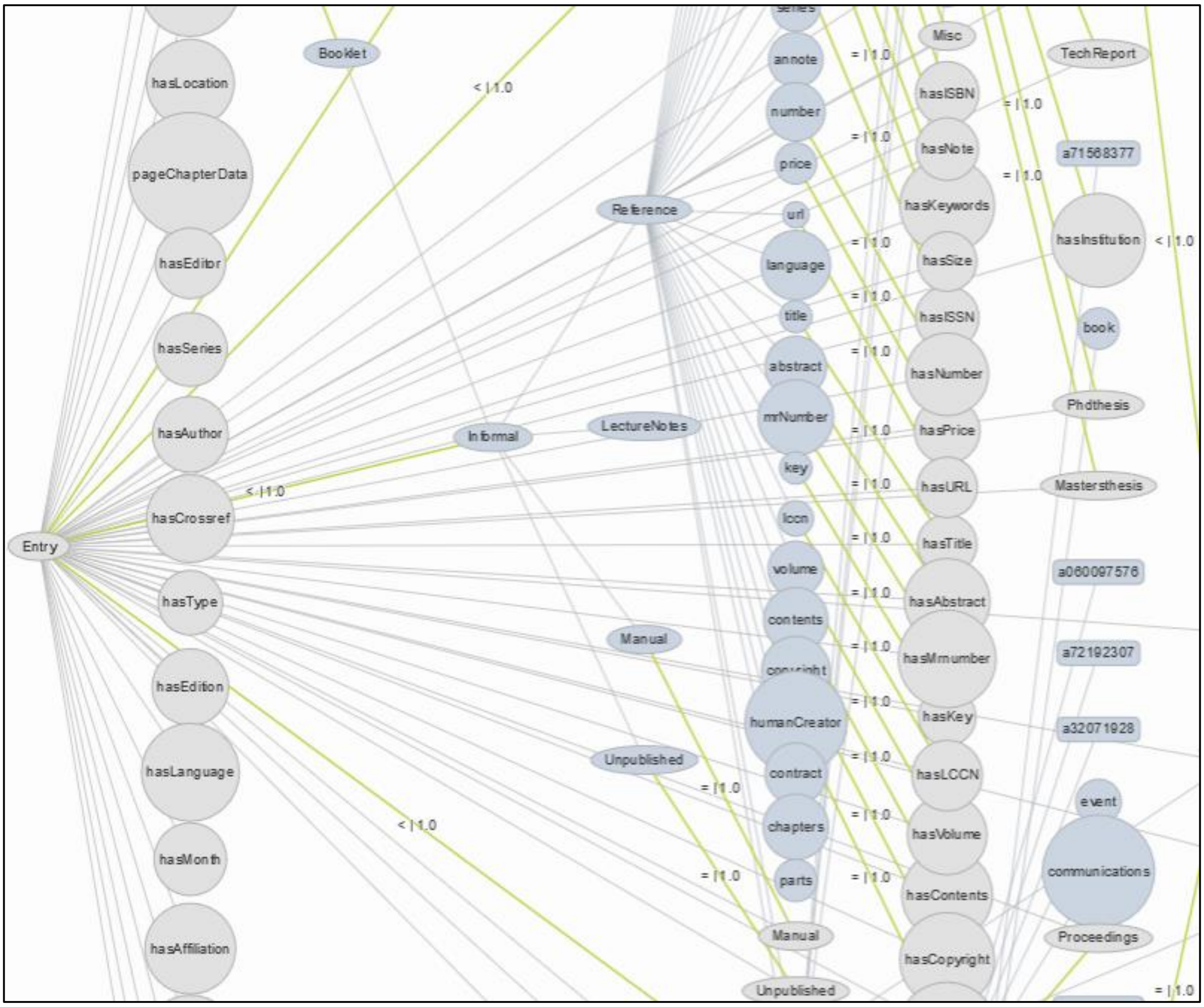
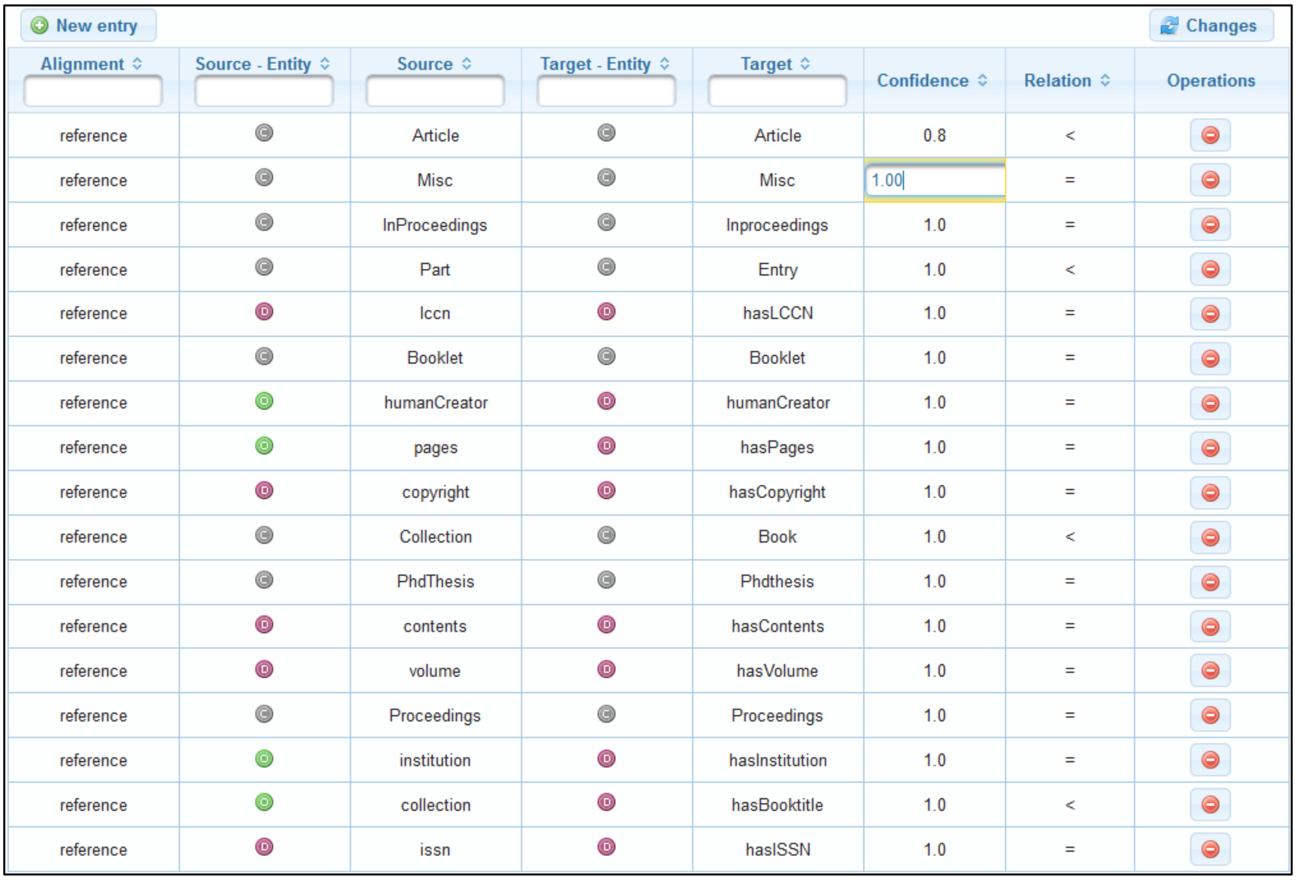


Figure 22 - Example of multiple alignments in the node-link visualization of VOAR

- **Correspondences list visualization:** instead of graphically displaying the alignment and ontologies, this visualization lists its correspondences in a table (Figure 23). This type of visualization is used in many state of the art tools that have some kind of representation for alignments. Users in the visualization survey also mentioned it as familiar.



New entry — Changes

Alignment	Source - Entity	Source	Target - Entity	Target	Confidence	Relation	Operations
reference	C	Article	C	Article	0.8	<	
reference	C	Misc	C	Misc	1.00	=	
reference	C	InProceedings	C	Inproceedings	1.0	=	
reference	C	Part	C	Entry	1.0	<	
reference	D	lccn	D	hasLCCN	1.0	=	
reference	C	Booklet	C	Booklet	1.0	=	
reference	O	humanCreator	D	humanCreator	1.0	=	
reference	O	pages	D	hasPages	1.0	=	
reference	D	copyright	D	hasCopyright	1.0	=	
reference	C	Collection	C	Book	1.0	<	
reference	C	PhdThesis	C	Phdthesis	1.0	=	
reference	D	contents	D	hasContents	1.0	=	
reference	D	volume	D	hasVolume	1.0	=	
reference	C	Proceedings	C	Proceedings	1.0	=	
reference	O	institution	D	hasInstitution	1.0	=	
reference	O	collection	D	hasBooktitle	1.0	<	
reference	D	issn	D	hasISSN	1.0	=	

Figure 23 - Example of alignment in the correspondences list visualization of VOAR

As mentioned before, sometimes, editing is allowed. Here the user is able to modify correspondences in any alignment, adding new correspondences, removing existing ones and editing confidence/type. It also facilitates the location of information, allowing search and ordering on its columns.

- **Quantitative Evaluation Visualization:** indicated by users on the survey as the main metrics used to perform evaluation (also perceived on the state of the art and related conferences), this visualization enables the user to generate metrics based on a compliance with a gold standard (Figure 24). Since the approach allows the user to work with multiple alignments, one should be selected as the reference alignment for an evaluation.

Reference alignment: reference

Alignment	Precision	Recall	F-measure
matcher B results	0.8928571428571429	0.9433962264150944	0.9174311926605505
matcher C results	0.5	0.03773584905660377	0.07017543859649122
matcher A results	0.8928571428571429	0.4716981132075472	0.617283950617284

Figure 24 - Example of multiple alignments in the quantitative evaluation visualization of VOAR

The results can be sorted for better analysis. Ideally, this visualization should be paired with one with a graphical representation for analysis and comprehension of the results.

- **Qualitative Evaluation Visualization:** the focus of most existing options ends up being mainly for quantitative evaluation. However, to help understand the differences between alignments, it is more efficient to use graphical features. With this representation, there is the possibility of qualitatively evaluating alignments, associating a “status” to each entity according to the correspondences from the reference alignment (Figure 25).

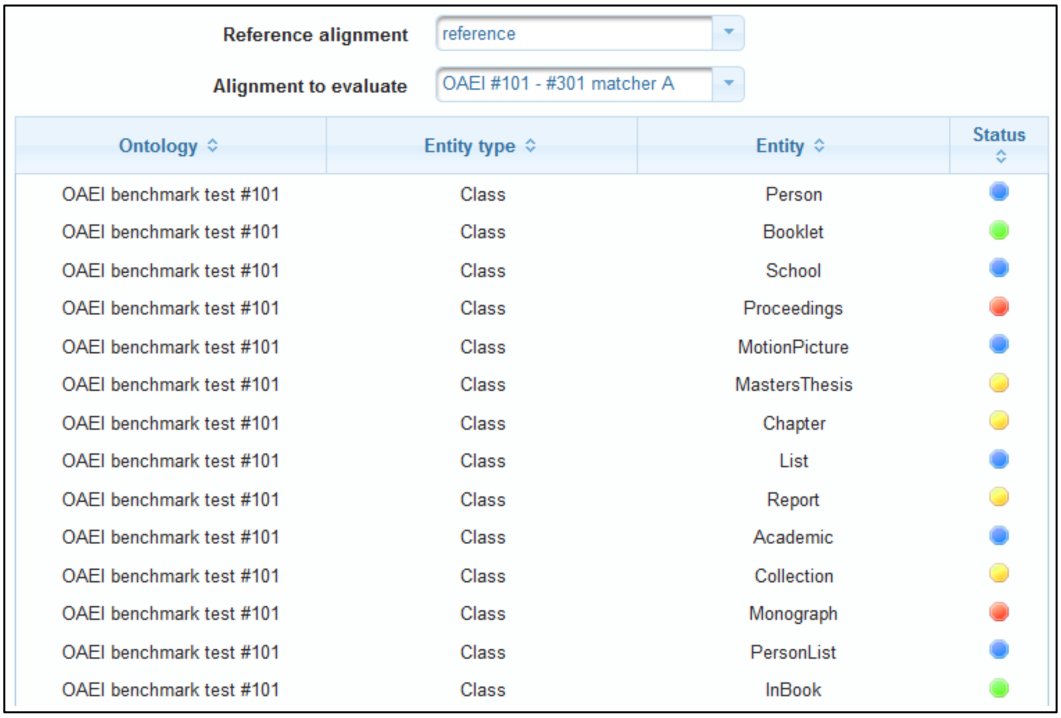


Figure 25 - Example of alignment in the qualitative evaluation visualization visualization of VOAR

According to the status of the entity (described on Table 7), one of the four predetermined icons should be associated, being used to visually represent the result of the qualitative evaluation. Such representation can be later associated with many graphical approaches (in this approach it is listed on a table). This allows the evaluator to have a notion of the individual result, something not found in the state of the art.

Table 7 - Qualitative evaluation of entities

Icon	Status	Description
	Incorrect Correspondences	The alignment contains incorrect correspondences for this entity when compared to the reference alignment. (False positive)
	Missing Correspondences	Correspondences are missing in the alignment for this entity when compared to the reference alignment. (False negative)
	Correct Correspondences	The alignment contains all correspondences for this entity when compared to the reference alignment. (True positive)
	No Correspondences	Neither the proposed alignment nor the reference alignment contains correspondences for this entity. (True negative)

5.2. Adaptive Profiles

According to the lessons learned in the survey, it is proposed a way for users to decide what information they want to display in a particular visualization. For this purpose, the concept of profile was created in this approach (Figure 26). A profile can be defined as <*l, v, v', i, i'*>, such that: *l* stands for an identifying label; *v* and *v'* are visualizations for window I and window II respectively (as defined in Section 5.1); and *i* and *i'* are a set of what information should be displayed for window I and window II respectively.

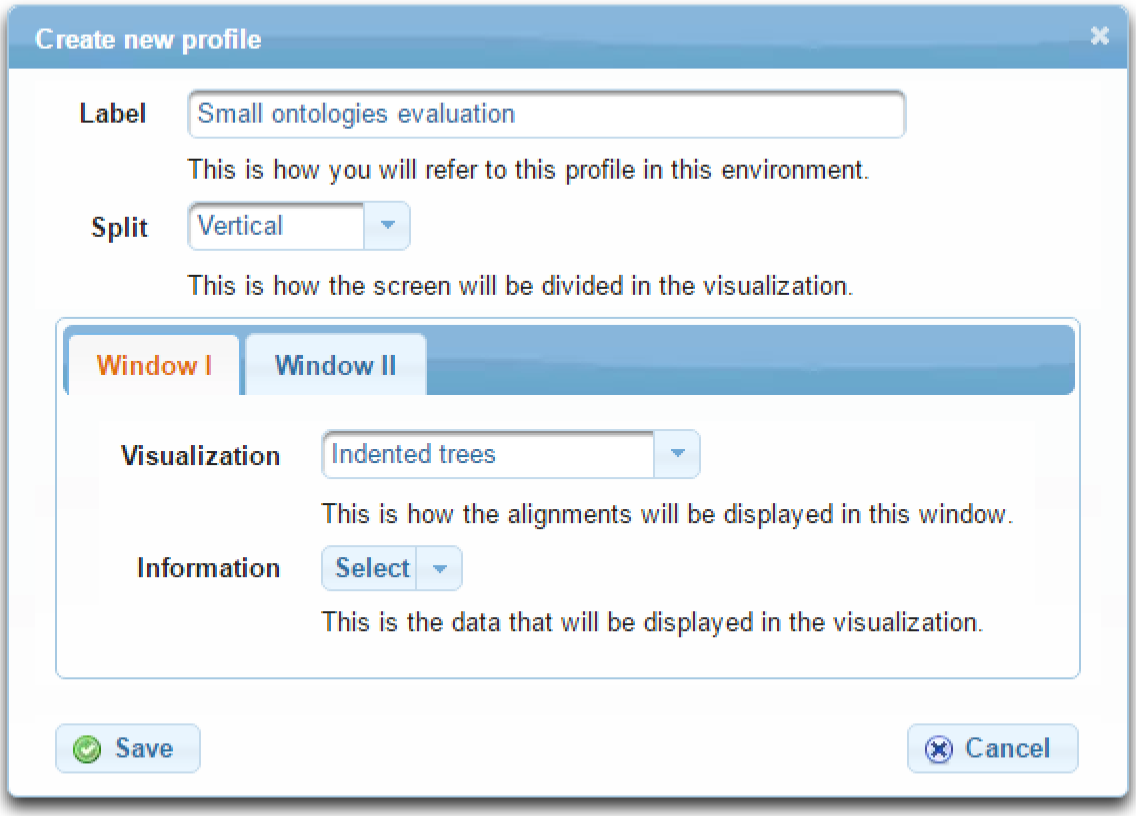
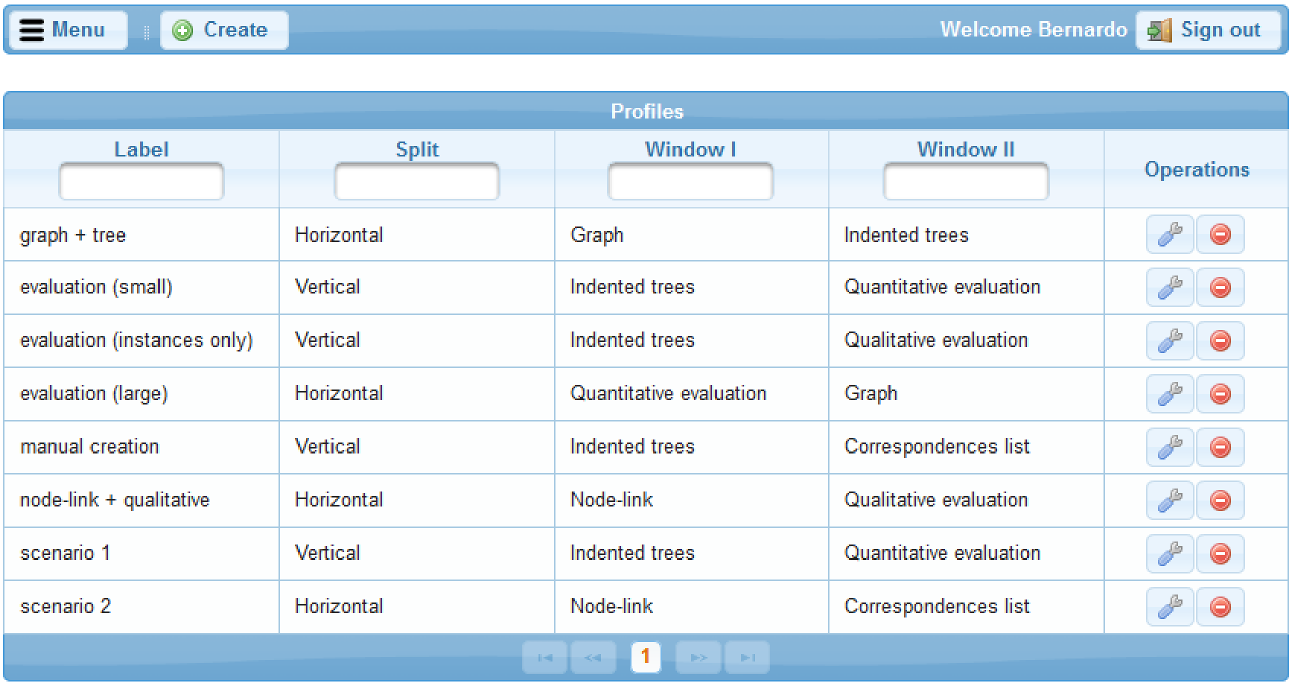


Figure 26 - Profile creation on the VOAR system

A single user can configure multiple profiles, each for a combination of factors that might influence its visualization (ontologies/alignments size, task, etc.) or simple preference. These profiles are maintained a library (Figure 27). Later, during the action of visualization, a user can change the active profile at any moment to adapt to it needs.



Label	Split	Window I	Window II	Operations
graph + tree	Horizontal	Graph	Indented trees	
evaluation (small)	Vertical	Indented trees	Quantitative evaluation	
evaluation (instances only)	Vertical	Indented trees	Qualitative evaluation	
evaluation (large)	Horizontal	Quantitative evaluation	Graph	
manual creation	Vertical	Indented trees	Correspondences list	
node-link + qualitative	Horizontal	Node-link	Qualitative evaluation	
scenario 1	Vertical	Indented trees	Quantitative evaluation	
scenario 2	Horizontal	Node-link	Correspondences list	

Figure 27 - Profiles library on the VOAR system

The available information to display/hide per window is: object properties, data properties, instances, correspondences, correspondences confidence and correspondences relation. The initial list of information proposed was based on those more commonly present in alignment environments considered as state of the art, and submitted for evaluation on the survey. After analyzing the results of the survey, it was clear that classes were a vital information for all tasks, being the only information that was always evaluated at least as "Important". Moreover, since many graphical representations use its hierarchical structure to represent other information, it was not set as optional in profile configurations.

Another information that was suppressed from configuration on profiles was the total counts for correspondences. Usually evaluated in a balanced form (pointed as not important and important in a similar proportion), the information might just overload the screen with more data and for some representations would not fit well. However, since it can be a valuable information for some, it can be displayed in another point of the tool (outside the graphical visualization).

This is one more step in making this approach more adaptive, allowing users to create multiple profiles as necessary or desired. It is believed that these adaptive profiles can solve the problem where even predetermined/optimized visualizations would fail if they show data in an unfavorable way, or show too much information in an already crowded information container.

5.3. Prototype

In order to validate the proposed approach with end users and the viability of its implementation in a tool, a prototype was built. The VOAR (**V**isual **O**ntology **A**lignment Envi**R**onment) system [48] [49] is publicly available[1], offering a GUI to assist users in the tasks of (multiple) alignments visualization, manipulation, and evaluation. VOAR was designed to provide an integrated environment for several operations, aiming to simplify the user activity. The prototype also implements authentication, granting isolation on user data and allowing the storage of the structures described on the remainder of this chapter per user.

In order to promote interoperability and reuse, VOAR is developed on the top of established technologies in the field of ontology matching, such as the Alignment API [9] and OWL API [25]. Therefore, users do not have to adapt its data to work with this system, using standardized formats and being able to export it to continue working on others tools.

During its development, some general design principles and guidelines, along with software design patterns were observed in order to make user experience better. According to Preece et al. [43], design principles are derived from a mix of theory-based knowledge, experience, and common sense. While guidelines are general rules commonly observed in practice [4]. Principles and guidelines act suggesting what to provide and what to avoid, without specifying how to design an actual interface. For the interface itself there are design patterns, that according to Tidwell [59], are best practices within a given design domain, thus, guiding on how to actually build the interface.

5.4. Configurations Scenarios

In order to demonstrate and comprehend possible scenarios where the presented approach can be used, and later even submitted to evaluation with users, the following use cases are presented:

- **Use case 1:** When a given user needs to evaluate multiple alignments from different ontology matchers, with the possibility of analyzing their differences to better comprehend the results. In this case, the alignments are for instances and classes only, and it is known the matchers only generate the equivalence relation (a common case). One possible profile for it could be a configuration like:

[1] http://voar.inf.pucrs.br

- Window I: indented trees visualization (since it was identified by the survey on alignment visualization as familiar for most tasks). On the information selected, only instances, correspondences and correspondences confidence. This will hide the excessive information (like more nodes with object and data properties) and the same relation being repeatedly displayed.
- Window II: quantitative evaluation. Listing the metrics per alignment.

As seen in Figure 28, after uploading the alignment files, this configuration would enable the user to see the calculated metrics for each alignment proposed, and to understand differences by displaying the correspondences plotted next to it. It is splitted on a vertical orientation, taking advantage of the trees layout. Alternatively, the quantitative evaluation could be swapped with the qualitative, depending of the focus of the user during the evaluation.

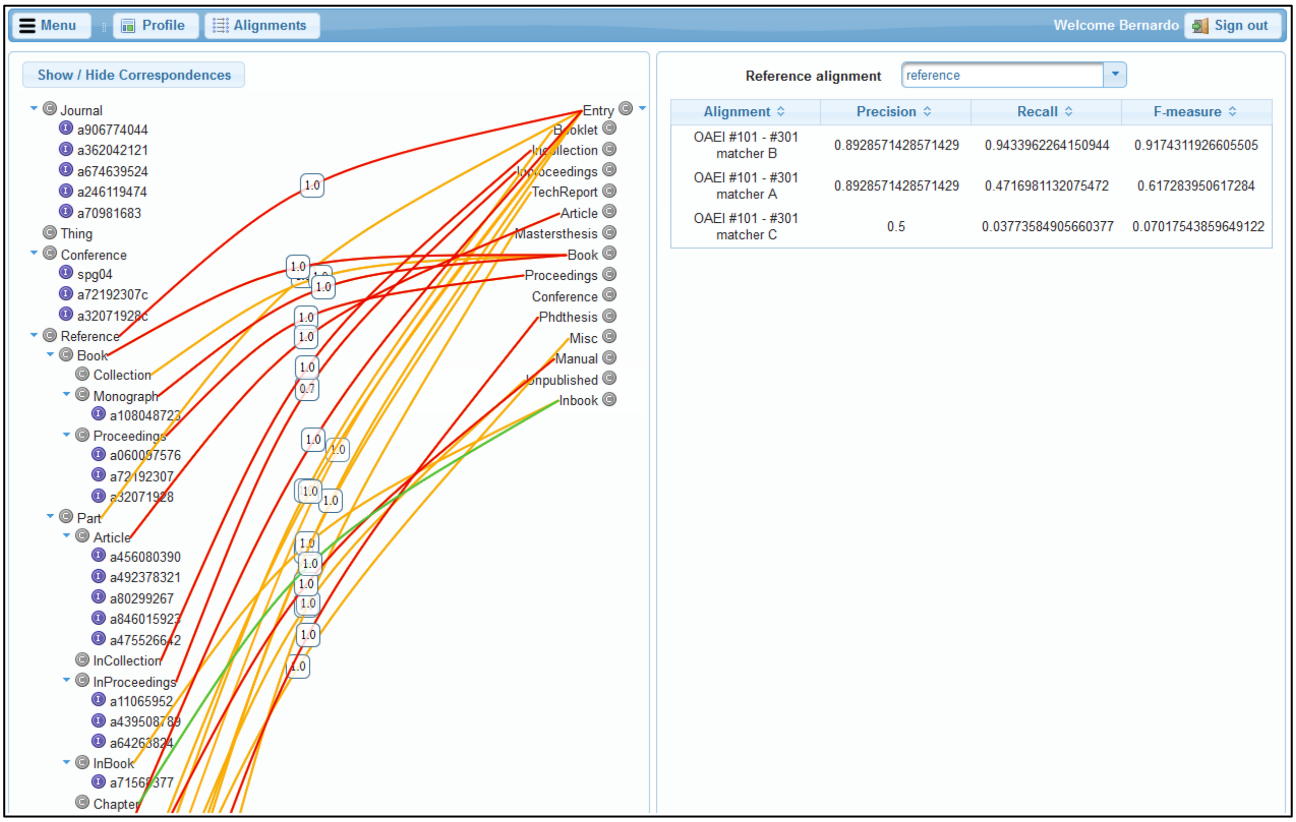


Figure 28 - Multiple alignments evaluation in the VOAR system

- **User case 2:** The user wants to modify an alignment for medium/large ontologies, containing only instances correspondences, which will later be a reference alignment for evaluation. After loading the alignment, a possible configuration would be:
    - Window I: node-link visualization (a type of graph, reported on the survey as a good representation for large data, but with better representation of hierarchy). Showing all ontology information (because the user prefers to see everything to make decisions on a possible correspondence), but none from alignments (the user wants to see only the ontology structure in this view).
    - Window II: correspondences list, displaying only instances, since it is the only entity of interest for ontology data in this scenario; and only correspondences and its relation type for alignment data (the confidence is not relevant, since it is being manually added only when the user is sure of its existence).

As seen on Figure 29, this configuration is splitted in a vertical orientation. It allows the user to focus on all alignment information in the list, while observing/navigating on the graphs to make decision about it, without excessive information in a representation that was reported to easily become confused.

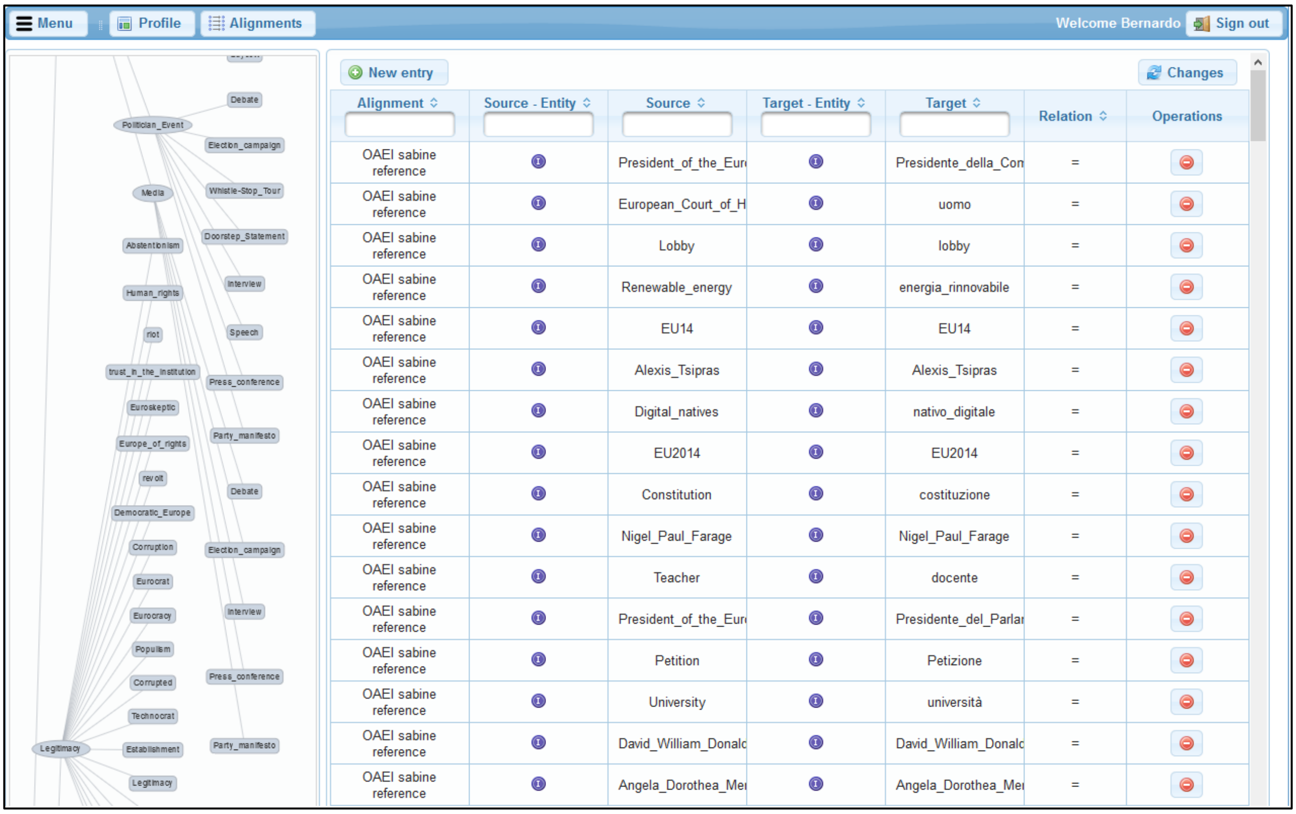


Figure 29 - Alignment manipulation on the VOAR system

Many other scenarios can also be built, and these listed above can be changed according to individual choices, for instance if the users prefer a given visualization over another for interpretation. The point here is to show how adaptive this approach can be to different situations observed during the evaluation made.

5.5. User Libraries

Since this approach is proposed in a way to facilitate user's interaction in a unified place, the concept of library is proposed to enable users to manage the structures that they will be handling on the proposed approach. All working data is uploaded to its respective library, and then operated on the approach. There are two basic libraries in this approach: ontologies library and alignments library.

5.5.1. Ontologies Library

In this library users are capable of manage any ontology that composes their alignments (Figure 30). The ontologies can be searched, ordered and filtered for easy navigation. The possibility of removing and downloading at any moment enables the user to keep a repository of ontologies in one place.

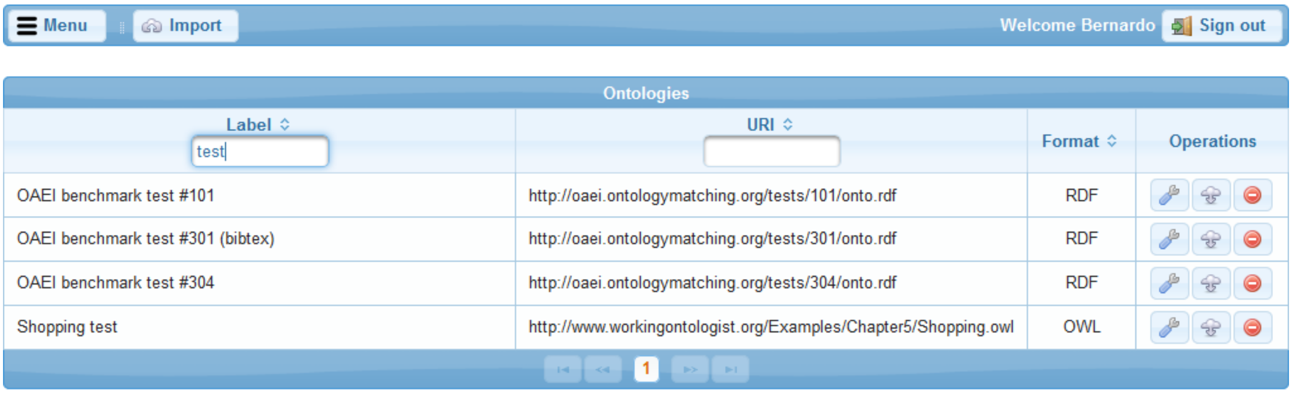


Figure 30 - Ontologies library on the VOAR system

As shown in the VOAR implementation (Figure 31), the process of importing allows the user to select the source of the ontology file (from its URI if available, or through upload of the file), and associate a friendly name to the ontology for later display.

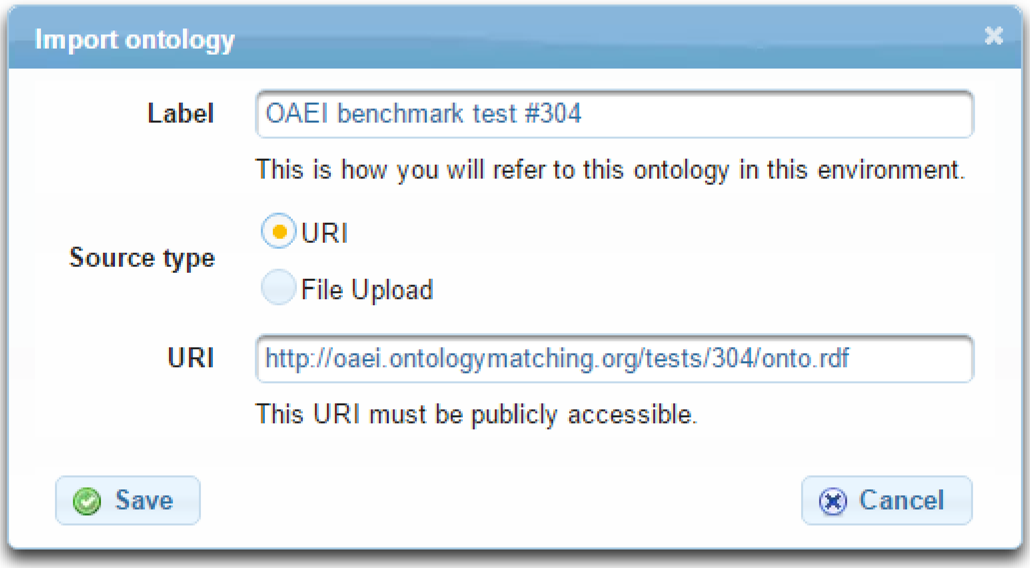


Figure 31 - Ontology import on the VOAR system

5.5.2. Alignments Library

The alignment library allows the management of alignments that a user possesses (Figure 32). Similar to the ontologies library, users are able to iterate over a list of its alignments, doing CRUD operations, filtering and downloading as a file.

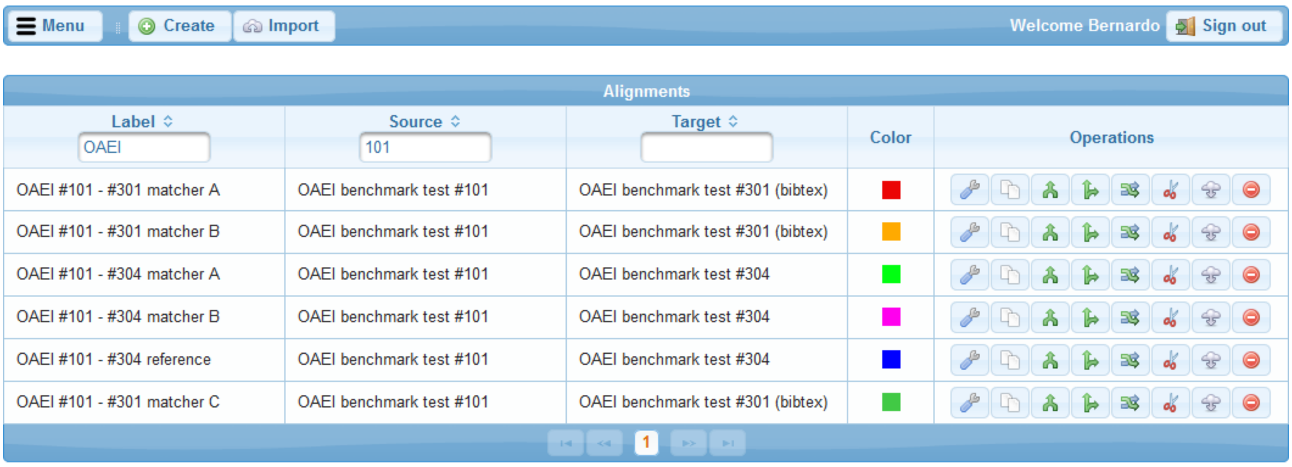
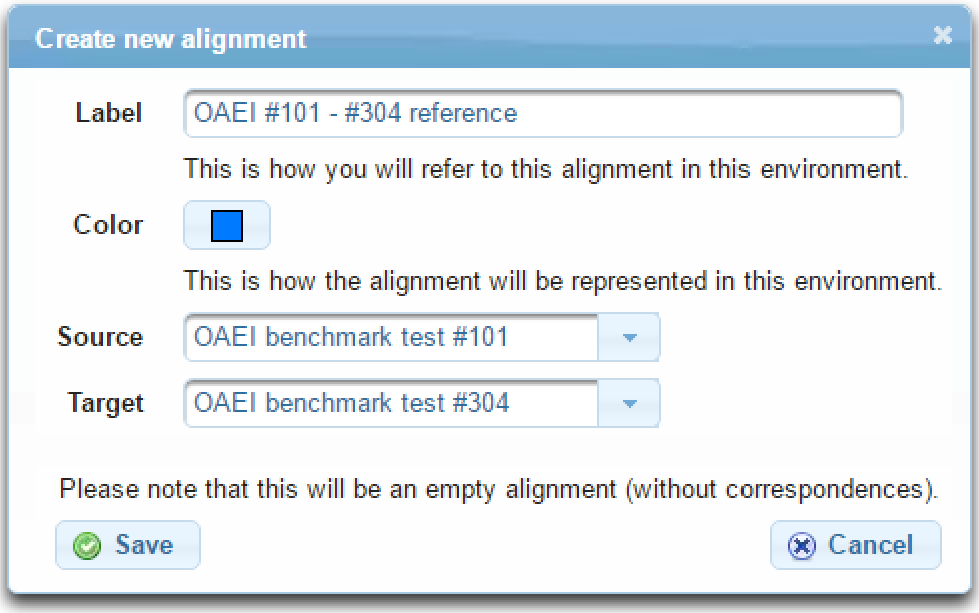


Figure 32 - Alignments library on the VOAR system

Since this library provide a place to manipulate alignments, some basic operations for alignment manipulation can also be applied (displayed on the column “Operations” of Figure 32):

- **Copy:** duplicate the current alignment data, useful to keep the state of the original alignment before the execution of other operations or manipulation.
- **Invert:** the source and target ontologies are swapped, and the correspondences updated accordingly. This facilitates operations where the user has an alignment $A_1(O_1, O_2)$ and finds an alignment $A_2(O_2, O_1)$, allowing them to have the same source as O_1 and target as O_2 in order to interoperate.
- **Trim:** remove all correspondences of the alignment under a given threshold. Useful to remove correspondences generated with low confidence.
- **Union:** by selecting multiple alignments, all correspondences of each are added to the selected one ($A_1 \cup A_2 \cup A_3 \cup ... \cup A_n$).
- **Intersection:** only the correspondences in common between all the alignments selected will be maintained ($A \cap A_2 \cap A_3 \cap ... \cap A_n$).

During the import and creation of alignments (Figure 33), labels and colors are associated to it. These extra attributes are used in the rest of the system to represent a particular alignment. By creating a new alignment no correspondences will be generated (being an empty alignment), being ideal for manual addition of correspondences later (for instance, to make a reference alignment).



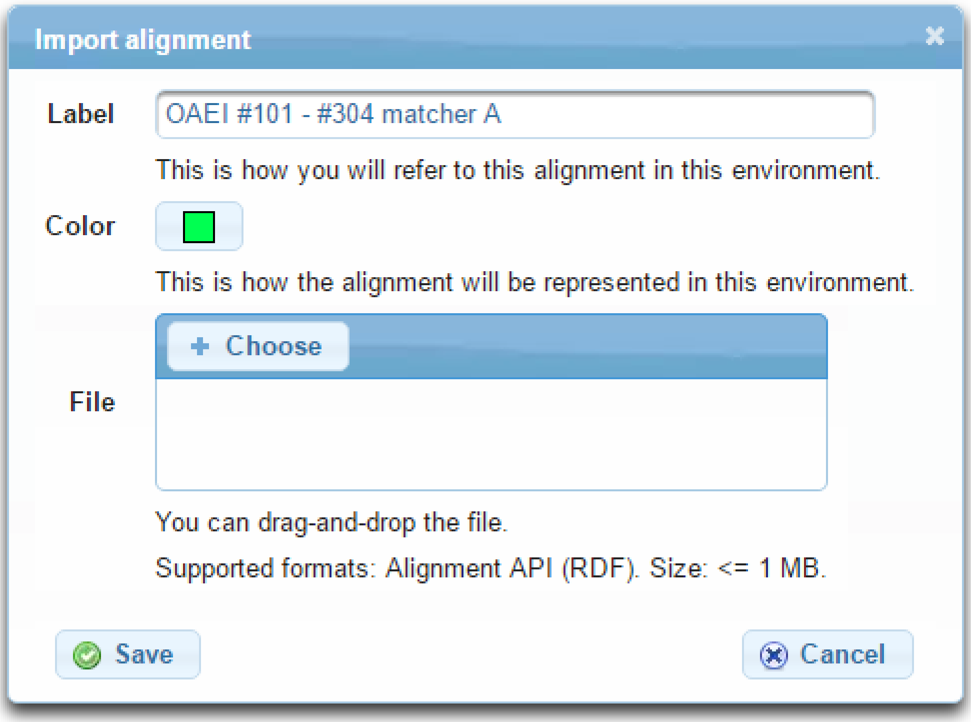


Figure 33 - Alignment creation and import on the VOAR system

6. APPROACH EVALUATION

In order to analyze the value obtained with the work of this thesis, the new adaptive proposal was submitted to evaluation, based on questionnaire with the same target audience of the research on ontology visualization (Section 4).

The online questionnaire was designed with 7 questions divided in two sets, aligned with the objectives of this thesis, aiming to better evaluate the achievements. In the first set, respondents were asked to, considering the approach presented and their experience with ontology alignment tasks, to rate our solution regarding problems they should solve with it. In the second set, it was asked to the users to rate how strongly they agree or disagree with statements, considering the VOAR environment. These questions use Likert scales to measure the answers, and at the end of the questionnaire, an open-ended question allows the user to provide any final considerations. The complete questionnaire is available in the Appendix B.

A video demonstrating the approach on top of the developed prototype (VOAR system) was exposed to the users, as well as providing free exploration of VOAR. The users were asked to consider their experience when testing the environment or the video demo for answering the questionnaire[2].

6.1. Results and analysis

The questionnaire received 10 responses during the period it was open for the users to fulfil it (30 days). Considering the two sets of answers described below, no “negative” evaluations on the many aspects of the approach were made. Therefore, it is possible to consider this work as a positive contribution to the state of the art, according to the opinion of specialists in the field of ontology matching.

The first set of questions asked the user to rate aspects of the proposed approach regarding the problems it is supposed to solve (with options for answers varying from ‘the approach does not handle the issue’ to ‘the approach handles it completely’). The general results presented in Figure 34 suggests that the proposed approach is capable of handling the problems it intended to solve, having mostly positive answers. The analysis of each topic is commented below:

[2] https://www.youtube.com/watch?v=wq-yPBOFN_I

- **Configuration of profiles to handle preferences according to the tasks being performed (creation, manipulation, evaluation, etc.):** the proposal of configuring multiple profiles per task in order to achieve a visualization that fit what is necessary handles the issue very well. Since no other approach in the state of the art is known to allow that per task, this was the response with the highest rate in the research.

- **Configuration of profiles to handle preferences according to ontology/alignment characteristics (size, elements in analysis, etc.):** even being evaluated mainly as handling the issue well, some improvements and extensions might be considered, like more visualizations or changes on the information that is configurable.

- **Configuration of the ontology/alignment elements presented (instances, correspondences confidence, etc.) to adapt the visualization for what is important at the moment:** once again, the approach handles the issue for this scenario, with a margin for improvement. As pointed by a respondent, maybe more ontology information could be added to the visualizations (like restrictions).

- **Use of multiple visualizations (indented tree, graphs, etc.) in order to facilitate the analysis of different scenarios:** like the configuration per task, this seems to be an aspect that users agree on, that handles the problem of sometimes having multiple tools for different use cases. As reported by a respondent, improvements still could be made to make extensions possible as plugins, encouraging the community to create alternatives or their own visualizations.

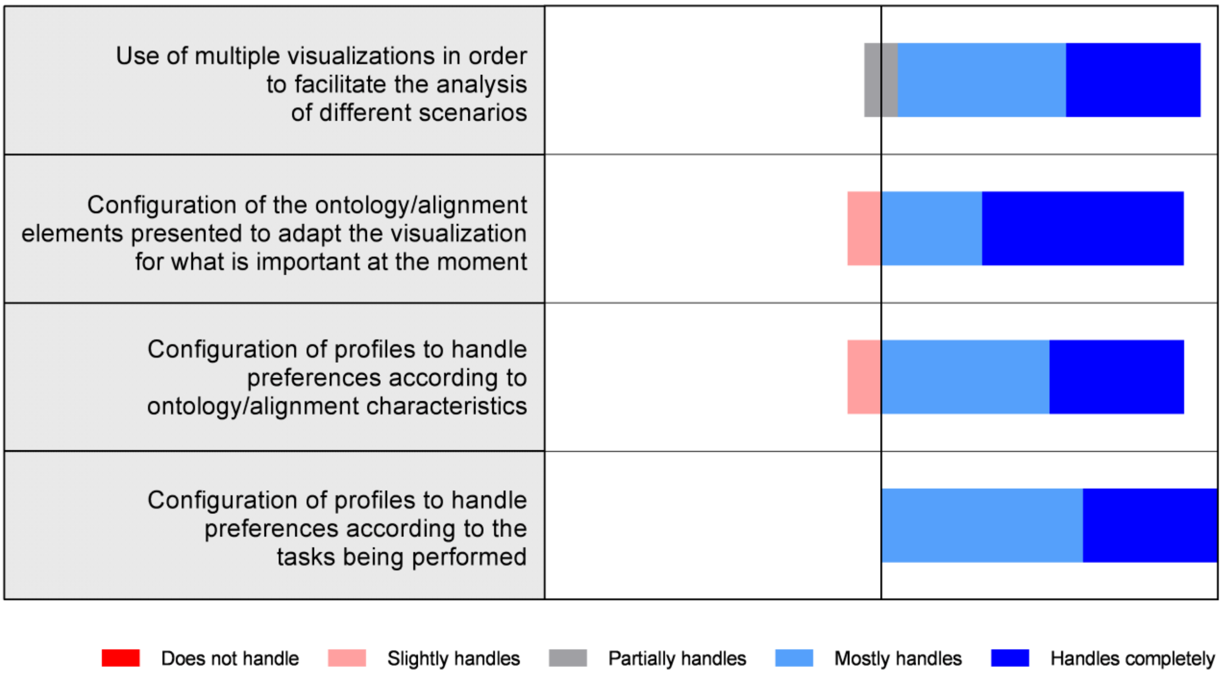


Figure 34 - VOAR evaluation on problem solving

The second set of questions approached more general questions, not related only to profiles. Like the first set, most answers stated that the solution helped in the exposed scenarios (Figure 35). The analysis of each question allows concluding:

- **The approach's visual support helps with alignment evaluation tasks:** as raised on the survey, since many users perform evaluation tasks manually and even without tools, the results being all positives on the support provided for the tasks reaffirm the value of a visual and dynamic way to work with evaluations.

- **The proposed approach helps to better analyze multiple alignments together:** this statement was manly evaluated as positive, meaning that the purpose of enabling users to work with multiple alignments was successfully achieved. Yet, some aspects might need improvement, like the visualizations that sometimes present overlapping correspondences for common matchers among different alignments.

- **The adaptive approach in the approach helps to minimize or solves particular problems with visualizations:** the idea of having two representations for the alignments in order to use the best aspects of each to complement the other seems to work as well. Probably with the addition of more options, it will be even better/easy to find the complement for each visualization in specific scenarios.

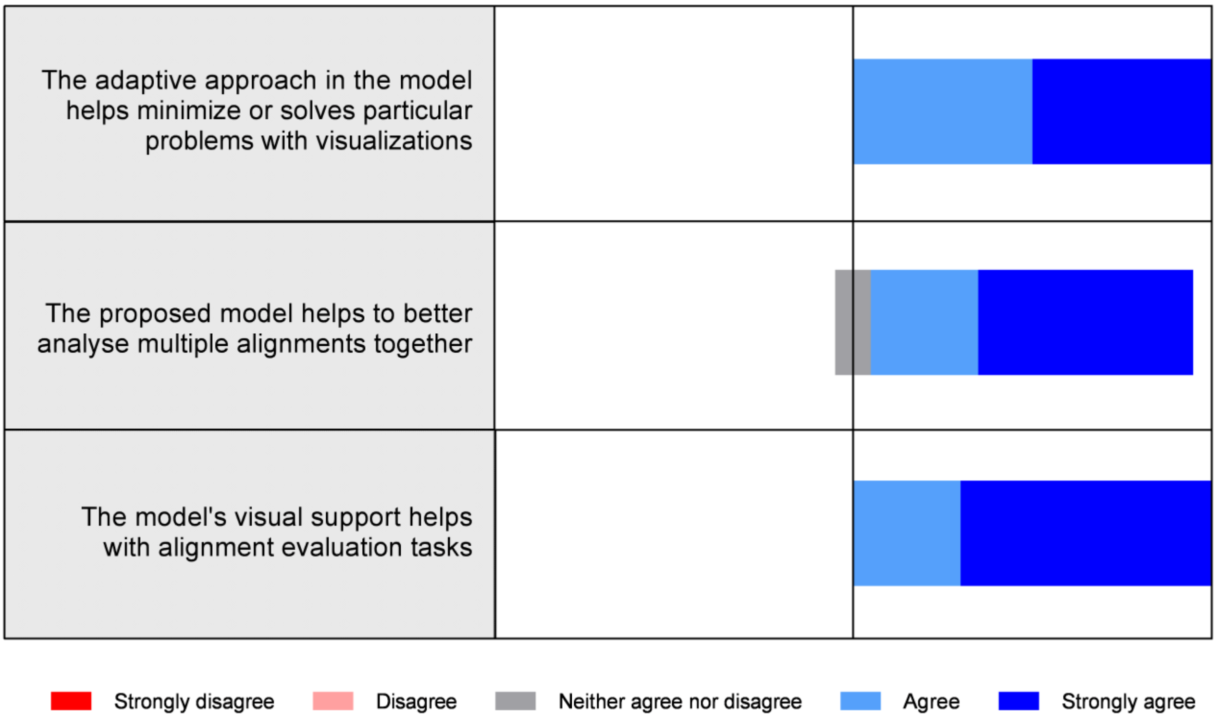


Figure 35 - VOAR evaluation on general topics

6.2. Concluding Remarks

This chapter presented the results on the questionnaire performed with the same target group from the alignment visualization survey, aiming at the validation of the proposed approach. The data collected from the first set of questions allowed to conclude that the problems this thesis proposes to solve in its objectives were handled, with a margin for improvements on the possible configurations (visualizations and information). The results from the second set verified that some achievements, like dealing with multiple alignments, provide visual support for evaluation tasks and the use of multiple visualizations to minimize particular problems were satisfactorily met.

7. CONCLUSION

In this work, a new approach on visualization of ontology alignments was presented. Ontologies and their alignments are valuable structures in several contexts, especially with the growing need to associate meaning to data and enable interoperation on it.

In the literature, current approaches are usually directed towards a specific task (mostly automatic creation, as the support offered by many matching tools). Relevant forms of visualization were presented, but scattered by specific tasks and domains. Reports on these works also pointed to the lack of a combination of multiple approaches to visualize and handle alignments according to the use case.

To better understand user needs and construct an environment for it, a survey was conducted, evaluating requirements for alignment visualization in different scenarios. To the best of our knowledge, this is the first survey regarding the problem of alignment visualization considering specific tasks. The results revealed that different tasks demands different configurations on what must be shown and how. Given visualizations are more prone to attend some tasks (even though this may vary depending on user preferences), like graphs or tables for evaluation, indented trees for creation, etc. Even with information importance tending to change according to tasks, a pattern on classes, as central elements in knowledge representation, being indispensable for all was found. Also, a poor evaluation was observed in the visual availability of several informations in current tools.

The lack of visual resources to work with multiple alignments, and limitation of task-specific existing approachesdemanded a more versatile design. After analysis of the state of the art, in consensus with related work conclusions and opinions from specialists collected in the survey conducted, an adaptive approach to work with alignments in varying conditions and use cases was proposed.

This thesis has contributed to this context by providing a novel approach, through a adaptive approach and a prototype. The approach proposed offers the ability to manipulate, evaluate and visualize multiple alignments in several graphical ways. In order to present the most relevant information using suitable techniques, users can configure “profiles”, structures that determine visualizations and what data must be displayed on it. Alternating personal profiles allows to adequate the presentation to the task and characteristics faced at the moment. The possibility of selecting multiple visualizations and what information to display was not seen in previous approaches.

The proposed approach stands out in the state of the art not only by providing unique features, such as the ability to configure profiles, but also by unifying some features poorly explored, or isolated in some tools. Each tool only disposes one visualization, directed to display one alignment (the only other found that support more alignments is the AgreementMaker [7]). Since multiple alignments are not common, operations over alignments are almost not present also (only available in the Alignment Server [9]).

This approach, available on the web through the prototype, also provides an innovative resource in a unified and public environment (current tools offered are not usually public and easily available). It helps the work of alignment users that deal with several different activities, often requiring several tools or manual handling.

The results obtained in the evaluation over the prototype provided evidence that the approach added significant value to the area of research, by dealing with the problems this thesis proposes to solve. The evaluation on configuration of profiles shown that some improvement could be done to contemplate more scenarios. Aspects of the approach, like the use of one visualization to mitigate the problems of another and the use of multiple alignments to improve the analysis in tasks were solely evaluated by users as being resolved.

The proposal also meets the recommendations of the work of Shvaiko and Euzenat [51] for future challenges in the area, so that users become more active through an interface that adapts to them according to their needs. Yet, it leaves many possible future developments over the proposed approach to continue this work.

7.1. Limitations and Future work

Despite this thesis having presented contributions to the field, some limitations on the current approach, as well as many improvements can be dealt with in the future. As pointed on the questionnaire over the prototype, the configuration of profiles can still be improved, since it might not solve all possible scenarios of use for alignments activities. Like the majority of tools to work with alignments, it also has performance issues dealing with large data.

Below is a list of future work that can be done on top of this thesis. These are not restricted to improvements of the approach, but also technical implementations to the prototype and continuity to the research:

- **More visualization styles:** this work presented just a few possible visualizations for alignments. As saw in Chapter 3, there are many others valuable approaches, such as space filling and zoomable visualizations, which would probably assist in handling even more different scenarios and preferences.

- **Pre-defined / suggested profiles:** although the main idea behind profiles is to enable users to mount views as they desire, some "global" profiles could be created and suggested based on the characteristics of the alignments involved, aiming to save time or working as a base for the users. Other surveys could be applied to identify what information is important in different scenarios, in order to build such pre-defined profiles. Alternatively, the approach could also use some AI techniques on top of the user's usage information to later propose these profiles.

- **Interface evaluation:** the evaluation of HCI is a fundamental activity in any development process that seeks to produce an interactive system with high quality of use. Conducting this evaluation on the VOAR environment would help to validate the proposed interface and identify improvements on usability and user experience.

- **Library sharing:** users could have an option to share an item of the library with another user or group. This would bring many benefits, like reducing the duplication of data and its maintenance, simplify teamwork, etc. It could contribute to collaborative alignment tasks, another challenge in the field.

- **Web services and REST API:** in order to make the data in the environment available to interact with other systems, the designed libraries could be exposed as services for consumption. Modern standards like web services or a REST API could be built and exposed through authentication by user.

- **Visualizations as plugins:** to promote the creation or implementation of more types of visualization by the community, the VOAR system can be adapted to allow new visualizations to be docked in the form of plugins or from a common repository. As mentioned by Katifori et al. [28], some ontology management tools already provide extensions, such as Protégé.

- **Inline search and filter:** the approach could provide in time of visualization of the data a global search and filtering options. For example, search per entity name in order to locate it in the visualization; or filtering correspondences below a given confidence. This would provide means to temporarily restrict the data displayed during an analysis, without the need to create a new alignment to visualize.

- **Inline edition of profile:** a profile could present its configuration in time of visualization, and enable the user to modify it on the fly (and later save the changes over the original or as a new one). Currently it would be necessary to create a new one in the library for each chance desired.

- **Change history:** as suggested by users in the ontology visualization survey, the approach can add control over changes performed. This can be applied to the edition of alignments during visualization (allowing to undo/redo operations) and to changes in the library, thus enabling to see modifications through time.

- **Matchers integration:** in the current state, users have to upload their matchers results to the library in order to visualize and evaluate. A common interface to invoke an ontology matcher, proposed in the context of the SEALS project [18] and in use by some tracks of the OAEI since 2011, could be used for integration. Functional experiments were already made in the past on VOAR for this [48], but due to some stability and security issues, this would need to be redesigned.

- **Cloud storage integration:** popular services nowadays allow users to have a remove drive on the cloud, integrated with many applications. One possible development would be to have the storage of the user library integrated to these services.

- **Dealing with large data:** on the one hand, given some limitations on current technologies used, the performance on large data is not optimal. Some APIs would have to be optimized or replaced/rewritten in order to handle faster large structures. On the other hand, a more compreensive study of approaches for visualizing large ontologies should be conducted. As mentioned in the work of Granitzer et al. [19], handling of very large, complex, evolving ontologies is another challenge, which has appeared on the radar of ontology alignment researchers.

7.2. Publications

During the progress on the thesis, the research and some results were presented and discussed in the following papers:

- **PLATAL - a tool for web hierarchies extraction and alignment. In: International Workshop on Ontology Matching - OM 2013:** presented a tool for extraction of structures from the web and the first idea on visual support to multiple operations/tasks over alignments. Later evolved to the VOAR system.

- **VOAR: A Visual and Integrated Ontology Alignment Environment. In: Language Resources and Evaluation Conference - LREC 2014:** presented the unified web resource to perform visual tasks on alignments. Also sharing the new approach on qualitative evaluation, available on the qualitative visualization of the current version of VOAR.

- **A GUI for Visualising and Manipulating Multiple Ontology Alignments. In: International Semantic Web Conference - ISWC 2015:** discussed the evolution of the first version of the VOAR approach, now enabling the work with multiple alignments.

REFERENCES

[1] Alani, H. “TGVizTab: An Ontology Visualisation Extension for Protégé”. In: Second International Conference on Knowledge Capture, 2003, pp. 3–9.

[2] Andrews, K.; Heidegger, H. “Information Slices: Visualising and Exploring Large Hierarchies using Cascading, Semi-Circular Discs”. In: IEEE Symposium on Information Visualization, 1998, pp. 4–7.

[3] Aumueller, D.; Do, H.; Massmann, S.; Rahm, E. “Schema and ontology matching with COMA++”. In: 2005 ACM SIGMOD International Conference on Management of Data, 2005, p. 906.

[4] Barbosa, S.; Silva, B. “Interação Humano-Computador”. Elsevier Brasil, 2010, 408p.

[5] Bosca, A.; Bonino, D.; Pellegrino, P. “OntoSphere: More than a 3D ontology visualization tool”. In: 2nd Italian Semantic Web Workshop, 2005, p. 15.

[6] Cheatham, M.; Dragisic, Z.; Euzenat, J.; Faria, D.; Ferrara, A.; Flouris, G.; Fundulaki, I.; Granada, R.; Ivanova, V.; Jiménez-Ruiz, E.; Lambrix, P.; Montanelli, S.; Pesquita, C.; Saveta, T.; Shvaiko, P.; Solimando, A.; Trojahn, C.; Zamazal, O. “Results of the Ontology Alignment Evaluation Initiative 2015”. In: 10th International Workshop on Ontology Matching, 2015, p. 56.

[7] Cruz, I. F.; Antonelli, F. P.; Stroe, C. “AgreementMaker: Efficient Matching for Large Real-World Schemas and Ontologies”. VLDB Endowment, vol. 2-2, August 2009, pp. 1586–1589.

[8] Dadzie, A.; Rowe, M. “Approaches to visualising linked data: a survey”. Semantic Web, vol. 2-2, April 2011, pp. 89–124.

[9] David, J.; Euzenat, J.; Scharffe, F.; Dos Santos, C. T. “The alignment API 4.0”. Semantic Web, vol. 2-1, January 2011, pp. 3–10.

[10] Eklund, P.; Roberts, N.; Green, S. “OntoRama: Browsing RDF ontologies using a hyperbolic-style browser”. In: First International Symposium on Cyber Worlds, 2002, pp. 405–411.

[11] Euzenat, J. “Ontology Alignment Evaluation Initiative”. Available on: http://oaei.ontologymatching.org/, 2013.

[12] Euzenat, J.; Shvaiko, P. “Ontology Matching”. Springer, 2007, 333.

[13] Euzenat, J.; Loup, D.; Touzani, M.; Valtchev, P. “Ontology alignment with OLA”. In: 3rd International Semantic Web Conference, 2004, pp. 59–68.

[14] Falconer, S. M.; Storey, M. A. “A cognitive support framework for ontology mapping”. Lecture Notes in Computer Science, vol. 4825 LNCS, 2007, pp. 114–127.

[15] Fensel, D. “Ontologies”. Springer Berlin Heidelberg, 2004, 162p.

[16] Fu, B.; Noy, N. F.; Storey, M. “Indented Tree or Graph? A Usability Study of Ontology Visualization Techniques in the Context of Class Mapping Evaluation”. In: 12th International Semantic Web Conference, 2013, p. 18.

[17] Giunchiglia, F.; Yatskevich, M.; Avesani, P.; Shivaiko, P. “A large dataset for the evaluation of ontology matching”. The Knowledge Engineering Review, June 2009, vol. 24-2, p. 137.

[18] Gómez-Pérez, A. “SEALS EU infrastructures project”. Available on: http://www.seals-project.eu/, 2011.

[19] Granitzer, M.; Sabol, V.; Onn, K. W.; Lukose, D.; Tochtermann, K. “Ontology Alignment—A Survey with Focus on Visually Supported Semi-Automatic Techniques”. Future Internet, vol. 2-3, August 2010, pp. 238–258.

[20] Gruber, T. R. “A translation approach to portable ontology specifications”. Knowledge Acquisition, vol. 5-2, June 1993, pp. 199–220.

[21] Van Ham, F.; Van Wijk, J. J. “Beamtrees: Compact visualization of large hierarchies”. In: IEEE Symposium on Information Visualization, 2002, pp. 93–100.

[22] Hamdi, F.; Safar, B.; Niraula, N. B.; Reynaud, C. “TaxoMap alignment and refinement modules: results for OAEI 2010”. In: 5th International Conference on Ontology Matching, 2010, pp. 212–219.

[23] Hesse, W. “Ontologies in the Software Engineering process”. In: Workshop on Enterprise Application Integration, 2005, p. 13.

[24] Horridge, M. “OWLViz”. Available on: http://protegewiki.stanford.edu/wiki/OWLViz, 2010.

[25] Horridge, M.; Bechhofer, S. “The OWL API: A Java API for OWL ontologies”. Semantic Web, vol. 2-1, January 2011, pp. 11–21.

[26] Intel IT center. “Big Data 101 : Unstructured Data Analytics”, Technical Report, 2012, 4p.

[27] Jiménez-Ruiz, E.; Cuenca Grau, B. “LogMap: Logic-Based and Scalable Ontology Matching”. In: 10th International Semantic Web Conference, 2011, pp. 273–288.

[28] Katifori, A.; Halatsis, C.; Lepouras, G.; Vassilakis, C.; Giannopoulou, E. “Ontology visualization methods---a survey”. ACM Computing Surveys, vol. 39-4, November 2007, p. 10.

[29] Katifori, A.; Torou, E.; Halatsis, C.; Lepouras, G.; Vassilakis, C. “A Comparative Study of Four Ontology Visualization Techniques in Protege: Experiment Setup and Preliminary Results”. In: Tenth International Conference on Information Visualisation, 2006, pp. 417–423.

[30] Kleiberg, E.; van de Wetering, H.; Van Wijk, J. J. “Botanical Visualization of Huge Hierarchies”. In: IEEE Symposium on Information Visualization, 2001, p. 8.

[31] Kolli, R.; Doshi, P. “OPTIMA: tool for ontology alignment with application to semantic reconciliation of sensor metadata for publication in SensorMap”. In: IEEE International Conference on Semantic Computing, 2008, pp. 484–485.

[32] Kriglstein, S. “User requirements analysis on ontology visualization”. In: International Conference on Complex, Intelligent and Software Intensive Systems, 2009, pp. 694–699.

[33] Lamping, J.; Rao, R.; Pirolli, P. “A focus+context technique based on hyperbolic geometry for visualizing large hierarchies”. In: SIGCHI Conference on Human factors in computing systems, 1995, pp. 401–408.

[34] Lanzenberger, M.; Sampson, J. “AlViz - A Tool for Visual Ontology Alignment”. In: Tenth International Conference on Information Visualisation, 2006, pp. 430–440.

[35] Lee, J. S. M.; Katari, G.; Sachidanandam, R. “GObar: a gene ontology based analysis and visualization tool for gene sets”. BMC bioinformatics, vol. 6, July 2005, p. 189.

[36] Liebig, T.; Noppens, O. “OntoTrack: Combining Browsing and Editing with Reasoning and Explaining for OWL Lite Ontologies”. In: 3rd International Conference on Semantic Web, 2004, pp. 244–258.

[37] Ngo, D. H.; Bellahsene, Z. “YAM++ : (not) Yet Another Matcher for Ontology Alignment Task”. Bases de Données, vol. 41, December 2016, pp. 30-49.

[38] Noy, N. F.; Fergerson, R. W.; Musen, M. A. “The Knowledge Model of Protégé-2000: Combining Interoperability and Flexibility”. In: International Conference on Knowledge Engineering and Knowledge Management, 2000, pp. 17–32.

[39] Noy, N. F.; Musen, M. A. “PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment”. In: 17th National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence, 2000, pp. 450-455.

[40] Parsia, B.; Wang, T.; Golbeck, J. “Visualizing Web ontologies with CropCircles”. In: 4th International Semantic Web Conference, 2005, p. 8.

[41] Pietriga, E. “IsaViz: A Visual Authoring Tool for RDF”. Available on: https://www.w3.org/2001/11/IsaViz/, 2007.

[42] Plaisant, C.; Grosjean, J.; Bederson, B. B. “SpaceTree: supporting exploration in large node link tree, design evolution and empirical evaluation”. In: IEEE Symposium on Information Visualization, 2002, pp. 57–64.

[43] Preece, J.; Rogers, Y.; Sharp, H. “Interaction Design: Beyond Human-Computer Interaction”. Wiley, 2002, 584p.

[44] Ricardo, A. Cava; Luzzardi, P. R. G.; Freitas, C. M. D. S. “The Bifocal Tree: A technique for the visualization of hierarchical information structures”. In: Workshop on Human Factors in Computer Systems, 2002, p. 11.

[45] Rivadeneira, W. “A Study of Search Result Clustering Interfaces: Comparing Textual and Zoomable User Interfaces”, Technical Report, Maryland (US), 2003, 8p.

[46] Robbins, N. B.; Heiberger, R. M.; Court, C.; Hall, S. “Plotting Likert and Other Rating Scales”. In: Joint Statistical Meetings, 2011, pp. 1058–1066.

[47] Robertson, G. G.; Mackinlay, J. D.; Card, S. K. “Cone Trees”. In: SIGCHI Conference on Human Factors in Computing Systems Reaching Through Technology, 1991, pp. 189–194.

[48] Severo, B.; Trojahn, C.; Vieira, R. “VOAR: A Visual and Integrated Ontology Alignment Environment”. In: Language Resources and Evaluation Conference, 2014, p. 6.

[49] Severo, B.; Trojahn, C.; Vieira, R. “A GUI for visualising and manipulating multiple ontology alignments”. In: International Semantic Web Conference, 2015, p. 4.

[50] Shneiderman, B. “Tree visualization with tree-maps: 2-d space-filling approach”. ACM Transactions on Graphics, vol. 11-1, January 1992, pp. 92–99.

[51] Shvaiko, P.; Euzenat, J. “Ontology Matching: State of the Art and Future Challenges”. In: IEEE Transactions on Knowledge and Data Engineering, vol. 25- 1, January 2013, pp. 158–176.

[52] Shvaiko, P.; Euzenat, J. “Ten Challenges for Ontology Matching”: In: On the Move to Meaningful Internet Systems, 2008, pp. 1164–1182.

[53] Silva, I. C. S. “Visualization of intensional and extensional levels of ontologies”, Master Thesis, Universidade Federal do Rio Grande do Sul, 2014, 101p.

[54] Sintek, M. “OntoViz”. Available on: http://protegewiki.stanford.edu/wiki/OntoViz, 2007.

[55] de Souza, K. X. S.; dos Santos, A. D.; Evangelista, S. R. M. “Visualization of ontologies through hypertrees”. In: Latin American Conference on Human-computer Interaction, 2003, pp. 251–255.

[56] Storey, M.; Musen, M.; Silva, J.; Best, C. “Jambalaya: Interactive visualization to enhance ontology authoring and knowledge acquisition in Protege”. In: International Workshop on Interactive Tools for Knowledge Capture, 2001, pp. 1–9.

[57] Studer, R.; Benjamins, V. R.; Fensel, D. “Knowledge engineering: Principles and methods”. Data & Knowledge Engineering, vol. 25-1, March 1998, pp. 161–197.

[58] Sure, Y.; Angele, J.; Staab, S. “OntoEdit: Guiding Ontology Development by Methodology and Inferencing”. In: On the Move to Meaningful Internet Systems, 2002, pp. 1205–1222.

[59] Tidwell, J. “Designing Interfaces: Patterns for Effective Interaction Design”. OReilly Media, 2005, 352p.

[60] Udrea, O.; Getoor, L.; Miller, R. J. “HOMER: Ontology Alignment Visualization and Analysis”. In: 6th International Semantic Web Conference, 2007, p. 3.

[61] W3C Brasil. “Web Semântica” Available on: http://www.w3c.br/Padroes/WebSemantica, 2011.

[62] Zeeberg, B. R.; Feng, W.; Wang, G.; Wang, M. D.; Fojo, A. T.; Sunshine, M.; Narasimhan, S.; Kane, D. W.; Reinhold, W. C.; Lababidi, S.; Bussey, K. J.; Riss, J.; Barrett, J. C.; Weinstein, J. N. “GoMiner: a resource for biological interpretation of genomic and proteomic data”. Genome Biology, vol. 4-4, March 2003, p. 28.

[63] Zhang, L.; Yu, Y.; Lu, J.; Lin, C.; Tu, K.; Guo, M.; Zhang, Z.; Xie, G.; Su, Z.; Pan, Y. “ORIENT: Integrate Ontology Engineering into Industry Tooling Environment”. In: 3rd International Conference on Semantic Web Conference, 2004, pp. 823–837.

[64] Zhong, S.; Storch, K. F.; Lipan, O.; Kao, M. C.; Weitz, C. J.; Wong, W. H. “GoSurfer: a graphical interactive tool for comparative analysis of large gene sets in Gene Ontology space”. Appl Bioinformatics, vol. 3-4, February 2004, pp. 261–264.

[65] Zicalero, C.; Ruiz, F.; Piattini, M. “Ontologies for Software Engineering and Software Technology”. Springer Berlin Heidelberg, 2006, 173p.

[66] “The Semantic Web – ISWC 2016”. Available on: http://iswc2016.semanticweb.org/, 2016.

[67] “OM-2016: The Eleventh International Workshop on Ontology Matching”. Available on: http://om2016.ontologymatching.org/, 2016.

[68] “SequoiaView”. Available on: http://www.win.tue.nl/sequoiaview/, 2002.

APPENDIX A – Ontology Visualization Survey

Ontology Alignment Visualization

1. Ontology Alignment Visualization Research

Informed Consent

This academic research aims at investigating different aspects and tasks regarding ontology alignment visualization. The goal of this questionnaire is to investigate issues related to this research topic, not the participant's expertise.

1. Anonymity will be preserved in any document published in scientific forums (such as conferences, journals, books and similar) or educational (such as handouts courses, slide shows, and similar).

2. The team is entitled to use this survey data, maintaining the above conditions for any academic purposes, teaching and/or development.

If you have any questions regarding this project, please, contact the authors:
Bernardo Severo - bernardo.severo@acad.pucrs.br
Cassia Trojahn - cassia.trojahn@irit.fr
Renata Vieira - renata.vieira@pucrs.br

Faculty of Informatics - PUCRS
www.pucrs.br/inf/
+55 (51) 3320-3558

Thanks for your collaboration.

1. To continue this questionnaire please confirm it below:

☐ I agree with the terms presented above.

Ontology Alignment Visualization

2. Profile and Expertise

2. What country do you live in?

3. Occupation: regarding the ontology matching field, in which context do you work with it?

☐ Researcher

☐ Professor/Lecturer

☐ Non-Academic activity

☐ Other (please specify)

4. Please indicate how you would rate your level of expertise with ontology matching:

- ◯ Low
- ◯ Medium
- ◯ High

Ontology Alignment Visualization

3. Ontology Alignment Creation: automated or semi-automated task

For the purpose of this research, please considerontology alignment creation (automated or semi-automated) the task of creating new alignments using ontology matchers, with or without human intervention (please note that the manual creation will be covered later).

5. Do you perform tasks related to ontology alignment creation (automated or semi-automated)?

- ◯ Yes
- ◯ No

Ontology Alignment Visualization

4. Ontology Alignment Creation: automated or semi-automated task

Please consider for this page your experience in performingautomated (or semi-automated) ontology alignment creation tasks.

6. For this task, according to your experience and needs, how important is the visualization of the elements and information listed in the following?
In the first column, rate how important is the visualization of each element/information.
In the second column, rate the availability of visual resources provided by the tool(s) you use, for visualizing each of them.

	Importance	Tool(s) visual resources
Classes		
Object properties		
Datatype properties		
Instances		
Correspondences		
Confidence level of each correspondence		
Type of each correspondence (relation)		
Number of correspondences (total)		
Number of correspondences by entity (classes, properties and instances)		
Number of correspondences by type (equivalence, etc.)		

Ontology Alignment Visualization

5. Ontology Alignment Creation: manual task

For the purpose of this research, please consider ontology alignment creation (manual) the task of creating new alignments manually, without the support of a system for automatic or semi-automatic ontology generation.

7. Do you perform tasks related to ontology alignment creation (manual)?

- ○ Yes
- ○ No

Ontology Alignment Visualization

6. Ontology Alignment Creation: manual task

Please consider for this page your experience in performing themanual ontology alignment creation task.

8. For this task, according to your experience and needs, how important is thevisualization of the elements and information listed in the following?
In the first column, rate how important is the visualization of each element/information.
In the second column, rate the availability of visual resources provided by the tool(s) you use, for visualizing each of them.

	Importance	Tool(s) visual resources
Classes		
Object properties		
Datatype properties		
Instances		
Correspondences		
Confidence level of each correspondence		
Type of each correspondence (relation)		
Number of correspondences (total)		
Number of correspondences by entity (classes, properties and instances)		
Number of correspondences by type (equivalence, etc.)		

Ontology Alignment Visualization

7. Ontology Alignment Manipulation

For the purpose of this research, please consider ontology alignment manipulation the task of editing existing alignments. For example, filtering out an alignment according to a confidence level or merging two alignments.

9. Do you perform tasks related to ontology alignment manipulation?

- ◯ Yes
- ◯ No

Ontology Alignment Visualization

8. Ontology Alignment Manipulation

Please consider for this page your experience in performing ontology alignment manipulation tasks.

10. For this task, according to your experience and needs, how important is the visualization of the elements and information listed in the following?
In the first column, rate how important is the visualization of each element/information.
In the second column, rate the availability of visual resources provided by the tool(s) you use, for visualizing each of them.

	Importance	Tool(s) visual resources
Classes		
Object properties		
Datatype properties		
Instances		
Correspondences		
Confidence level of each correspondence		
Type of each correspondence (relation)		
Number of correspondences (total)		
Number of correspondences by entity (classes, properties and instances)		
Number of correspondences by type (equivalence, etc.)		

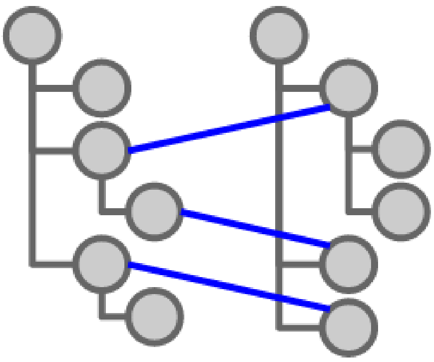
Ontology Alignment Visualization

9. Ontology Alignment Visualization: approaches

For the purpose of this research, please considerontology alignment visualization the task of exploring a graphical version for representation and comprehension of the alignments.

In this page, different visualization methods will be presented, please evaluate each one of them.

Indented Trees: Indentation is used to illustrate super/sub-class relationships, while lines connect the aligned entities between trees.



11. Regarding the Indented Trees visualization style, please indicate in what tasks it would be useful for you:

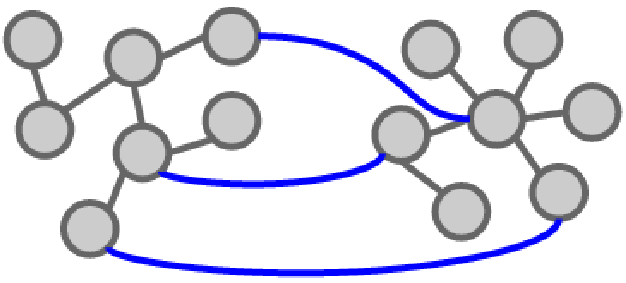
- [ ] Alignment creation
- [ ] Alignment manipulation
- [ ] Alignment evaluation
- [ ] Other (please specify)

12. Regarding the Indented Trees visualization style, please indicate the positive/negative aspects according to your opinion:

Positive

Negative

Graphs: Present nodes with connecting edges that illustrate ontological entities and the relationships that exist among them.



13. Regarding the Graphs visualization style, please indicate in what tasks it would be useful for you:

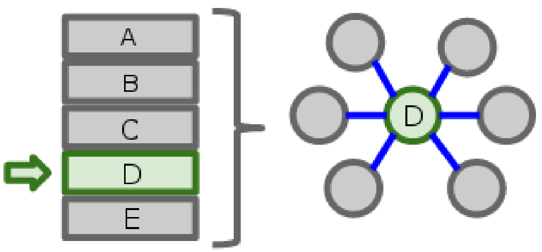
- [ ] Alignment creation
- [ ] Alignment manipulation
- [ ] Alignment evaluation
- [ ] Other (please specify)

14. Regarding the Graphs visualization style, please indicate the positive/negative aspects according to your opinion:

Positive

Negative

Focus + Context: Enable viewers to see the entity of primary interest presented in full detail while at the same time getting an overview–impression of all the surrounding information — or context— available.



15. Regarding the Focus + Context visualization style, please indicate in what tasks it would be useful for you:

- [ ] Alignment creation
- [ ] Alignment manipulation
- [ ] Alignment evaluation
- [ ] Other (please specify)

16. Regarding the Focus + Context visualization style, please indicate the positive/negative aspects according to your opinion:

Positive

Negative

17. What other styles for visualization do you use and/or consider important? Could you provide an exemple of tool with such representation?

Ontology Alignment Visualization

10. Ontology Alignment Visualization: task

Please consider for this page your experience in performing ontology alignment visualization tasks.

18. For this task, according to your experience and needs, how important is the visualization of the elements and information listed in the following?
In the first column, rate how important is the visualization of each element/information.
In the second column, rate the availability of visual resources provided by the tool(s) you use, for visualizing each of them.

	Importance	Tool(s) visual resources
Classes		
Object properties		
Datatype properties		
Instances		
Correspondences		
Confidence level of each correspondence		
Type of each correspondence (relation)		
Number of correspondences (total)		
Number of correspondences by entity (classes, properties and instances)		
Number of correspondences by type (equivalence, etc.)		

19. Are there other ontology or alignment information that you think are important to visualize?

- Yes (specify you answer in the field below)
- No

If "Yes", please specify your answer and how important is this.

20. The tool(s) you use offer resources to visualize multiple alignments together?

- Yes (specify you answer in the field below)
- No

If "Yes", which kind of method is used (table listing the correspondences, ontologies as hierarchical trees and lines between them representing correspondences, graphs, etc.)?

Ontology Alignment Visualization

11. Ontology Alignment Evaluation

For the purpose of this research, please considerontology alignment evaluation the task of evaluating (qualitatively and quantitatively) an alignment. For example, comparing generated alignments with a gold standard.

21. Do you perform tasks related to ontology alignment evaluation?

- ◯ Yes
- ◯ No

Ontology Alignment Visualization

12. Ontology Alignment Evaluation

Please consider for this page your experience in performingontology alignment evaluation tasks.

22. For this task, according to your experience and needs, how important is the visualization of the elements and information listed in the following?
In the first column, rate how important is the visualization of each element/information.
In the second column, rate the availability of visual resources provided by the tool(s) you use, for visualizing each of them.

	Importance	Tool(s) visual resources
Classes		
Object properties		
Datatype properties		
Instances		
Correspondences		
Confidence level of each correspondence		
Type of each correspondence (relation)		
Number of correspondences (total)		
Number of correspondences by entity (classes, properties and instances)		
Number of correspondences by type (equivalence, etc.)		

23. Which metrics do you use to perform this task?

- [ ] Compliance with gold standard (precision, recall, f-measure and weighted variants)
- [ ] Logical reasoning based evaluation
- [ ] Other (please specify):

24. The tool(s) you use offer a visual resource to perform the evaluation of multiple alignments together?

- ( ) Yes (specify your answer in the field bellow)
- ( ) No

If "Yes", which kind of resource (table of results, visual comparison of sets of correspondences, etc.)?

25. Which approach do you use to evaluate multiple alignments together? For example, when comparing a reference alignment against matchers results.

Ontology Alignment Visualization

13. Ontology Alignment Tasks and Tools

26. Apart from the tasks cited before (creation, manipulation, evaluation), do you perform other activities that require visualization of ontology alignments?

- ◯ Yes (specify your answer in the field bellow)
- ◯ No

If "Yes", please describe the task(s).

27. What is the average size of the ontologies you are used to work with for the following tasks?

(E => number of entities)

	Small (E ≤1,000)	Medium (1,000 < E < 10,000)	Large (E ≥ 10,000)	Not Applicable
Alignment Creation (automated or semi-automated)	◯	◯	◯	◯
Alignment Creation (manual)	◯	◯	◯	◯
Alignment Manipulation	◯	◯	◯	◯
Alignment Evaluation	◯	◯	◯	◯
Alignment Visualization	◯	◯	◯	◯

28. What tool(s) do you use for the following tasks? If you need more than one, how do you combine them?

Alignment Creation (automated or semi-automated)

Alignment Creation (manual)

Alignment Manipulation

Alignment Evaluation

Alignment Visualization

29. Please indicate the positive aspects of the tool(s) you use for the following tasks:

Alignment Creation (automated or semi-automated)

Alignment Creation (manual)

Alignment Manipulation

Alignment Evaluation

Alignment Visualization

30. Please indicate the negative aspects and/or possible improvements of the tool(s) you use for the following tasks:

Alignment Creation (automated or semi-automated)

Alignment Creation (manual)

Alignment Manipulation

Alignment Evaluation

Alignment Visualization

Ontology Alignment Visualization

14. Conclusion

31. This is the place where you can make any additional comments. You can request the analyzed results of the research after processing, or make suggestions / questions that you could not place before.

32. If you want a reply to your comments or requests above, please leave your e-mail (optional):

Thank you again for your time.
Please press "Finish" below to submit all answers.

APPENDIX B – Visual Ontology Alignment Environment Survey

Visual Ontology Alignment Environment - Evaluation

Dear researcher,

Earlier this year we performed a survey on ontology alignment visualizations. After analyzing the results, we proposed a visual model to support different user preferences in tasks related to ontology alignments.

The model consists in a library of resources (ontologies and alignments) in the cloud. On top of that, we designed "profiles" for visualization and interaction. Each profile allows you to configure how to split your screen (horizontally or vertically), combining two visualizations options and parameterizing the information on each visualization chosen (to see or hide instances, correspondences, etc.). Visualizations in this context can contain graphical representations of multiple alignments (indented trees, graphs, etc.), evaluation metrics, means for correspondences manipulation, etc. The idea is that you can build multiple profiles, according to your preferences, tasks or the best fit for the data in analysis (large data, instances alignment only, etc.).

The goal of this survey is to evaluate the proposed model based on the environment developed for it, available at http://voar.inf.pucrs.br (in BETA state). If you prefer, you can see a short video demo, where we show how you can upload your files, create/configure profiles for tasks such as manual creation of correspondences and evaluation.
Video demo: https://www.youtube.com/watch?v=wq-yPBOFN_I.

Please consider your experience testing the environment or the video demo for this questionnaire.

Thank you for your collaboration.

Bernardo Severo - bernardo.severo@acad.pucrs.br

Faculty of Informatics - PUCRS - Brazil
www.pucrs.br/inf/

* Required

1. **Considering the model we presented and your experience with ontology alignment tasks, please rate our solution regarding:** *
Mark only one oval per row.

	The model does not handle this issue	The model slightly handles this issue	The model partially handles this issue	The model mostly handles this issue	The model handles this issue completely
Configuration of profiles to handle preferences according to the tasks being performed (creation, manipulation, evaluation, etc.).	◯	◯	◯	◯	◯
Configuration of profiles to handle preferences according to ontology/alignment characteristics (size, elements in analysis, etc.).	◯	◯	◯	◯	◯
Configuration of the ontology/alignment elements presented (instances, correspondences confidence, etc.) to adapt the visualization for what is important at the moment.	◯	◯	◯	◯	◯
Use of multiple visualizations (indented tree, graphs, etc.) in order to facilitate the analysis of different scenarios.	◯	◯	◯	◯	◯

2. **Considering the VOAR environment, please rate how strongly you agree or disagree with each of the following statements:** *
Mark only one oval per row.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
The model's visual support helps with alignment evaluation tasks.	◯	◯	◯	◯	◯
The proposed model helps to better analyse multiple alignments together.	◯	◯	◯	◯	◯
The adaptive approach in the model helps minimize or solves particular problems with visualizations.	◯	◯	◯	◯	◯

3. **Please leave here any additional comments (suggestions, questions, etc.):**

....

....

....

....

Thank you again for your time.