European Commission

Project Deliverable D4.4 Criteria, protocols and results for assessing the mapping methodologies on the identified gold-standard (Report)

PMKI–Public Multilingual Knowledge Management Infrastructure

Date: 02/10/2018
Doc. Version: 1.0
Template Version: 2.5

PM² PROJECT MANAGEMENT METHODOLOGY

Document Control Information

Settings	Value
Document Title:	WP4 – Implementation and test of the technical infrastructure – Project Deliverable D4.4 Criteria, protocols and results for assessing the mapping methodologies on the identified gold-standard (Report)
Project Title:	PMKI–Public Multilingual Knowledge Management Infrastructure
Document Author:	<Armando Stellato>
Project Owner:	<Peter SCHMITZ>
Project Manager:	<Najeh HAJLAOUI>
Doc. Version:	1.0
Sensitivity:	<PMKI team>
Date:	02/10/2018

Document Approver(s) and Reviewer(s):

NOTE: All Approvers are required. Records of each approver must be maintained. All Reviewers in the list are considered required unless explicitly listed as Optional.

Name	Role	Action	Date

Document history:

The Document Author is authorized to make the following types of changes to the document without requiring that the document be re-approved:

- Editorial, formatting, and spelling
- Clarification

To request a change to this document, contact the Document Author or Owner.

Changes to this document are summarized in the following table in reverse chronological order (latest version first).

Revision	Date	Created by	Short Description of Changes
1.0	02/10/2018	Armando Stellato	Submitted version

Configuration Management: Document Location

The latest version of this controlled document is stored in <link[1]>.

[1] SharePoint accessible from OP environment.

TABLE OF CONTENTS

Introductive summary 4

1. INTRODUCTION 5

2. MOTIVATION FOR THE PRODUCTION OF GOLD-STANDARD MAPPINGS 6

3. STANDARDS FOR THE REPRESENTATION OF THE ALIGNED RESOURCES AND OF THE ONTOLOGY ALIGNMENTS 6

4. ALIGNMENT COMPLETENESS AND MINIMALITY 10

5. THE PROTOCOL FOR CREATING THE DATASETS CORPUS 11

6. THE EXPERIENCE IN ALIGNING EUROVOC AND TESEO 12

6.1 Report on observed phenomena 13

7. THE ALIGNMENT BETWEEN EUROVOC AND THE IATE TERMINOLOGY 14

8. CRITERIA FOR THE EVALUATION OF ALIGNMENTS AGAINST THE PRODUCED GOLD-STANDARD ALIGNMENTS 15

9. CONCLUSIONS 15

Introductive summary

The PMKI project facilitates the public sector and SMEs in the multilingualisation of the Digital Single Market of the EU, by merging the available linguistic resources in order to have cross border accessibility of public administration services and e-commerce solutions. The language technology industry provides support for the EU economy in particular to SMEs to overcome the language barriers. This will help to unlock the e-Commerce potential within the EU, allowing several public services to use the services offered by PMKI.

In this delivery we:

1. report on the experience gathered in producing two gold-standard alignment datasets between the European Union thesaurus EuroVoc and two other notable resources adopted in legal environments: the thesaurus of the Italian Senate TESEO and the IATE European terminological resource.
2. describe the requirements that have been set and the protocols we adopted for producing the above-mentioned resources
3. describe the criteria that we expect to adopt for evaluating automatically produced alignments against the produced gold-standards

1. INTRODUCTION

The Semantic Web [1] has offered a powerful stack of standard languages and protocols for modeling, sharing and reuse of knowledge on the Web. However, the advantages brought by metadata standards and infrastructures for shareability of information cannot avoid (but can support) the reconciliation work needed on the information content. Different, domain-overlapping, redundant to different extents, ontologies, thesauri, vocabularies, datasets etc. are expected to emerge on the Web in order to satisfy specific needs and exigencies and, at different points in time, are expected as well to be – somehow – "reconciled" out of their heterogeneities, for interoperability's sake.

This "reconciliation" takes the form of alignments, that is, sets of correspondences between the different entities that populate lexical and semantic resources on the Web. The expression "ontology alignment" is often used in a broader sense than the one that the first word of the term would suggest. "Ontology" is in this case a synecdoche for ontologies, thesauri, lexicons and any sorts of knowledge resources modeled according to core knowledge modeling languages for the Semantic Web, which shared and made available on the Web itself. The expression ontology alignment thus defines the task of discovering and assessing alignments between ontologies and other data models of the RDF family; alternative expressions are *ontology mapping* or *ontology matching* (as the produced alignments are also referred to as *matches*). In the RDF jargon, and following the terminology adopted in the VoID metadata vocabulary [2], a set of alignments is also called a *Linkset*.

The production of alignments is an intensive and error prone task; for this reason, several approaches for automating the task have been devised [3] since the early years of the Semantic Web. An Ontology Alignment Evaluation Initiative[2] [4] is also held every year since 2004 with the intent of evaluating available tools against benchmarks consisting in well-assessed alignments between notable semantic resources (mostly ontologies, and some thesauri). The task are also divided into T-Box/Schema matching, dealing – as the name suggests – with the alignment of ontology vocabularies, i.e. word models, including class and properties, and instance matching, involving the creation of links between domain objects represented in different datasets.

We report here on the experience gathered in producing two gold-standard alignment datasets between EuroVoc[3] – EU's multilingual thesaurus covering the activities of the European Union – and two other notable resources adopted in legal environments: the thesaurus of the Italian Senate TESEO[4] (TEsauro SEnato per l'Organizzazione dei documenti parlamentari) and IATE[5] (InterActive Terminology for Europe), the EU's multilingual terms base.

In this deliverable, the standards used to represent aligned resources and a methodology for the production of gold-standard alignments between the language resources mentioned above are presented.

[2] http://oaei.ontologymatching.org/

[3] http://eurovoc.europa.eu/

[4] http://www.senato.it/3235?testo_generico=745

[5] http://iate.europa.eu

2. MOTIVATION FOR THE PRODUCTION OF GOLD-STANDARD MAPPINGS

As far as producing mappings is an intensive and error prone task, it is also a very difficult one which cannot be easily automated. If alignments could be logically inferred, there would be processes generating them automatically with guaranteed precision. In fact, the necessity for alignments comes into play when there is no a-priori semantic agreement between two resources. The identification of proper alignments is thus a discovery process (some approaches [5] treat it as an Information Retrieval problem indeed) based on the analysis of the content. This content is mostly, on a first step, pure textual information, in order to gather the first anchors between the two resources, as common language expressions are indeed the last stand for creating hypothesis about alignments [6, 7], which is later followed by inference mechanisms based on the created hypotheses. This kind of analysis obviously benefits from background knowledge – owned by the mapping agent – that goes beyond the one available in both datasets and from the capability to process natural language content.

As of the state of the art, the importance and main contribution of automatic alignment consists in providing in a short time a seed that would be hard to produce manually, in order to support the subsequent refinement performed by domain experts. Even though meant to be completed by means of human supervision, the quality of the automatically produced alignments is important: besides the trivial aspect that a good automatic process shortens the following validation, it is also true that, especially in alignments of large corpora, an unprecise seed will in any case provide a bias on the validator: a single “almost perfectly matched” concept would drift the attention of the validator away from some best representative that is still laying somewhere in the target dataset. Even more important, an incomplete mapping between two datasets (i.e. missing some important matches) can be validated in the precision of the produced matches, but missing matches would be no less hard to be discovered than in a point-blank situation. For this reason, even though precision is important, recall (as in every Information Retrieval process) is fundamental, as users may easily discard wrong results but their trust in the system depends on the feeling that no important information is left unretrieved.

A proper evaluation of alignment processes demands for gold-standard alignments, guaranteeing maximum precision (quality-produced matches) and recall (completeness of the alignment). Another possible exploitation of such resources is their use as training sets for mapping algorithms based on learning techniques. Also, to avoid the negative effects of in-vitro experimentation, alignments between real datasets should be produced.

With these requirements in mind and with the legal domain as an important topic for the project that has been the context for this work; we chose to create two alignments for subsets of three important resources of the European Union: EuroVoc, TESEO and IATE. In order to create reliable gold-standard alignments, we selected a range of circa 300 concepts per each resource, so that the task of manual alignment could be performed entirely by humans without any bias nor errors induced by machine processing.

3. STANDARDS FOR THE REPRESENTATION OF THE ALIGNED RESOURCES AND OF THE ONTOLOGY ALIGNMENTS

The PMKI project considers resources that are modeled and published according to Semantic Web standards, thus adopting RDF (and the various vocabularies for creating ontologies, thesauri, etc.) and Linked Open Data best practices for publication. Different actions of the project also aim at providing services (transformation, hosting etc.) for bringing non-RDF resources into the PMKI platform, thus leveling all of them to the desired standards.

Within this context, the kind of resources that have been considered (at least for a first bootstrap of the project) are multilingual language resources, including thesauri and lexicons. W3C offers two core modeling vocabularies for dealing with these kind of resources: for thesauri there is the W3C recommendation SKOS [8] (and its extension SKOS-XL [9], for modeling reified labels) while a community group under the W3C umbrella has recently developed the OntoLex-Lemon (https://www.w3.org/2016/05/ontolex/) suite of vocabularies for modeling lexicons.

SKOS is a vocabulary defined in OWL introducing a set of terms characterizing thesauri: skos:Concepts are described as representing ideas or notions, units of thought. However, as the reference manual clarifies, what constitutes a unit of thought is subjective, and this definition is meant to be suggestive, rather than restrictive. In a few words, while OWL classes provided a formal classification mechanism for describing entities, adopting a first-level logic perspective made of objects and predicated over them, SKOS concepts are meant to provide a simplified view, with shallow semantics, where entities can be generally described with values. Concepts are, in this sense, all objects with respect to an OWL perspective (they are indeed instances of the skos:Concept class), whether they are merely conceptual entities or concrete objects of a domain. A more detailed description of differences between SKOS and OWL (or better, when to model something as a thesaurus and when as an ontology) is provided in [10]. Specific properties are also available for representing the hierarchical organization of concepts and the membership of concepts to schemes (i.e. different views over a thesaurus).

Another important aspect of SKOS lies in its terminological properties. Preferred labels (skos:prefLabel) describe the lexical expressions that are most recurring for referring to a certain concept. Alternative labels (skos:altLabel) provide alternative expressions, less common than the preferred ones. Hidden labels (skos:hiddenLabel) cover common misspells and other expressions that should not be explicitly shown while being at the same adopted for technical aspects such as indexing and retrieval. SKOS also provides several properties for expressing notes about the elements of a thesaurus. These notes can be in-domain or extra-domain, representing respectively the description (e.g. skos:definition) of the object of the domain represented by the SKOS concept or notes taken by the editors (e.g. skos:editorialNote) that concern the concept as an element of the thesaurus construct and not as its denoted domain object.

A third important aspect of SKOS lies in its mapping vocabulary: a set of properties describing relationships of similarity or, at least, connection, between concepts from two different thesauri. skos:exactMatch provides a shallow semantics (i.e. no logical entailment) approach to identity, expressing the fact that two SKOS concepts denote the same object of the world. skos:closeMatch implies a strong closeness between two SKOS concepts even though they are not considered to be exactly the same thing. skos:broad/narrowMatch express a "more specific"/"more generic" relationship between two concepts from different thesauri while skos:related connects two concepts denoting two things that, while being clearly different, are somehow related to each other.

In 2012, the OntoLex W3C Community Group was chartered to define an agreed specification informed by the aforementioned models, whose designers are all involved in the community group.

The OntoLex-Lemon model is primarily based on the ideas found in Monnet *lemon*, which was already adopted by a number of lexicons (Eckle-Kohler, McCrae, & Chiarcos, 2015; Borin, Dannélls, Forsberg, & McCrae, 2014; Navigli & Ponzetto, 2012). More specifically, OntoLex-Lemon consists of a number of vocabularies corresponding to different modules: core, synsem, decomp, vartrans, lime.

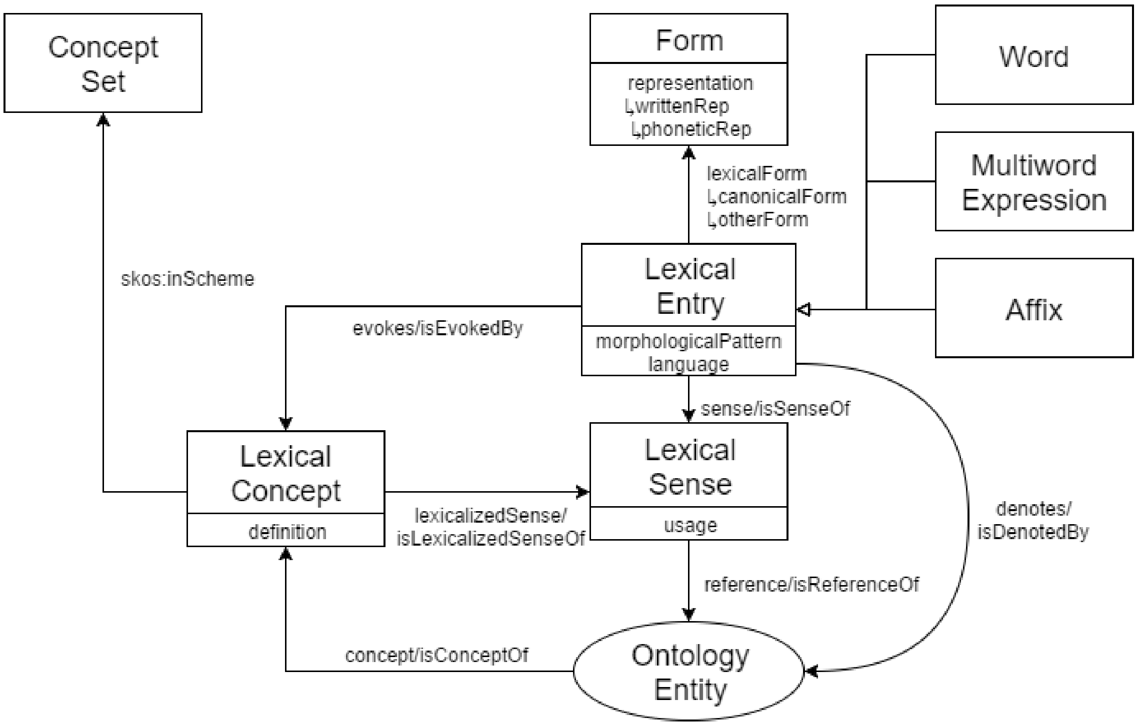

Figure 1 The OntoLex-Lemon core module

The core module (

Figure 1) retains from Monnet *lemon* the separation between the lexical and the ontological layer, where the ontology describes the semantics of the domain and the lexicon describes the morphology, syntax and pragmatics of the words used to express the domain in a language. A lexicon consists of lexical entries with a single syntactic class (part-of-speech) to which a number of forms are attached (e.g. the singular/plural forms of a noun), and each form has a number of representations (string forms), e.g. written or phonetic representation.

While an entry can be linked directly to an entity in an ontology, usually the binding between them is realized by a lexical sense resource where pragmatic information such as domain or register of the connection may be recorded. Lexical concepts were introduced in the model to represent the "semantic pole of linguistic units, mentally instantiated abstractions which language users derive from conceptions" (Evans, 2006). They are intended to represent abstractions in existing lexical resources such as synsets in wordnets.

The synsem module (left side of Figure 2) allows to associate a lexical entry with a syntactic frame (representing a stereotypical syntactic context for the entry), while an ontology mapping can be used to bind syntactic and semantic arguments together.

The decomp module (right side of Figure 2) is concerned with the decomposition of a lexical entry into its constituents (i.e. tokens). The class decomp:Component models these constituents, which in turn correspond to lexical entries. This indirection allows recording inside a component information such as the fact that the entry “autonomo”@es occurs with feminine gender inside “comunidad autonoma”@es. We can also represent parse trees, by subdividing a component into its constituents.

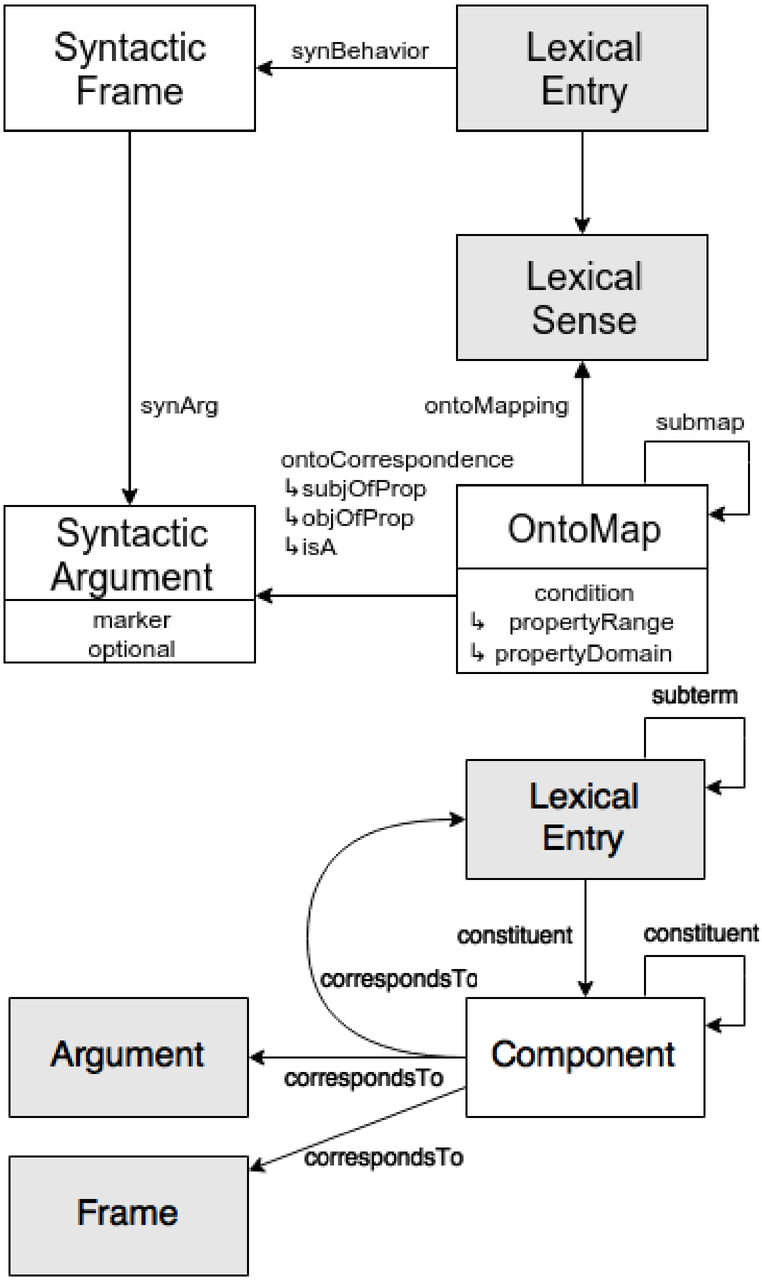

Figure 2 The syntax-semantics module (synsem) on the left and the decomposition module (decomp) on the right

The vocabulary that has been used for the alignments is the one created and adopted for the Alignment API (David, Euzenat, Scharffe, & Trojahn dos Santos, 2011), and it is a base for the more elaborated EDOAL mapping language. In the Alignment API ontology, mappings are reified into resources (see figure 1), in order to be decorated with further metadata: for instance, besides the type of relation (expressed by the property :relation) the vocabulary foresees a property :measure, representing the confidence of the mapping. However, other additional information can then be added as well (e.g. rdfs:comments). The vocabulary is also extensible, as new relations can be added in custom extensions of the vocabulary.

```
<map>
  <Cell>
    <entity1 rdf:resource="http://iasted#Student registration fee"/>
    <entity2 rdf:resource="http://sigkdd#Registration Student"/>
    <measure rdf:datatype="xsd:float">0.75555555555</measure>
    <relation>=</relation>
  </Cell>
</map>
```

Figure 3. an excerpt of the RDF/XML code for an alignment expressed for the Alignment API

For instance, the basic vocabulary only foresees an equivalence relation, probably due to the fact that it is not expected for a machine to be so precise in distinguishing among subtle variations of a match, thus postponing this specification to later refinement by means of human supervision. However, in the context of the maintenance of the EuroVoc thesaurus, the Publications Office of the European Union adopted an extension of the alignment ontology where a property semantically close to skos:closeMatch has been added to the range of available ones, thus making the porting to SKOS mapping relations a simple 1-1 transformation.

4. ALIGNMENT COMPLETENESS AND MINIMALITY

As outlined in section 2, an important aspect when considering an alignment effort is, besides the (trivially clear in its essence, yet not easy to achieve) precision of the outcome, its completeness. As most alignment vocabularies consider at least three kind of relations: equivalence and the two inverse more specific/more general relations, plus usually a relation to tell that two elements are not matchable, in an alignment effort between two datasets D_A and D_B it would be theoretically possible to explicitly state a relationship between each element of D_A and D_B, thus filling the cartesian product of $|D_A| \times |D_B|$ possible alignments. However, there are some aspects that should be taken into consideration:

- *Direction of the alignment*. An alignment can be given a direction, from one source dataset to a "target" dataset, so ensuring that at least all entities of the "source" dataset are aligned.
- *"Nobleness" of the relationship*. If a concept $c_A \in D_A$ is aligned with a relation of a certain "nobleness" to a concept $c_B \in D_B$, then there is no need to establish other "*less noble*" relationships between c_A and any other concept in D_B. "nobleness" is a total order over the available relations in a given mapping vocabulary. E.g. in the SKOS mapping properties, the following order would hold: exactMatch, closeMatch, narrowMatch, broadMatch, related. We consider then two sets relevant sets of relations: equivalence relations and tolerance relations. *Equivalence* relations are reflexive, symmetric and transitive. Tolerance relations are reflexive, symmetric but not necessarily transitive. skos:exactMatch match is an example of a equivalence relation, while skos:closeMatch is a tolerance relation. Notably, in SKOS neither :exactMatch nor :closeMatch have been defined as reflexive. The absence of reflexivity is due to historical reasons (the SKOS model is older than OWL2, where reflexivity has been introduced). From the definition it follows that an *Equivalence* relation is also a *Tolerance* relation
- Note that more relationships of the same type (if allowed by the relation itself) would be admitted. For instance, a concept $c_A \in D_A$ could be declared to be in a skos:narrowMatch relation with two concepts $c_{1B}, c_{2B} \in D_B$. The rationale for this

double assignment is that being no better (more noble) match for c_A in D_B, then as many concepts from D_B, which (together) provide a best representation for c_A, are linked to it.

- *Linkability and Reachability*. If a concept $c_A \in D_A$ is aligned with a *tolerance relation* to a concept $c_B \in D_B$, then $\forall\, c_i \in D_A : c_i < c_A$ it holds that $c_i \in reach(c_B)$.
  In other words, in absence of an explicit relationship for c_i, c_i can still be considered *linked* to c_B (thus being able to "reach" c_B*)* through a traversal of the hierarchy in D_A until concept c_A that is perfectly matching c_B is met.

We can thus provide the following definitions:

Definition 1 (linked concept). *Given two datasets D_A and D_B linked through linkset L, a concept $c_A \in D_A$* is *said to be linked to D_B iff at least one of the two following statements is true:*

- *1.* $\exists\, l(c_A, c_B) : c_A \in D_A, c_B \in D_B$
- *2.* $\exists\, c_B \in D_B : c_A \in reach(c_B)$

In *particular:*

- *c_A is said to be perfectly linked to D_B iff c_A is linked and statement 1 is false only in case there is no possible alignment for a match on c_A based on a tolerance relation.*
- *c_A is said to be redundantly linked to D_B whenever both statements 1 and 2 are true and the asserted alignment is not a match based on a tolerance relation*

Definition 2 (properly unlinked concepts). *Given two datasets D_A and D_B linked through linkset L, a concept $c_A \in D_A$* is *said to be properly unlinked from D_B if there is no possible assignment of alignments ∈ L so that c_A can be linked to D_B.*

So, the notion of *perfectly linked* concept guarantees that no concept is simply linked as the consequence of a tree-traversal when it could have been better linked through an equivalence link. This is necessary for defining the completeness of an alignment. Conversely, the notion of redundantly linked concept supports the definition of minimal and complete alignment.

Definition 3 (Mapping completeness). *A linkset L between two datasets D_A* and *D_B is said to be complete with respect to dataset D_A iff each concept in D_A is perfectly linked to or is properly unlinked from D_B.*

Definition 3 (Minimal complete mapping). *A linkset L between two datasets D_A* and *D_B is said to be complete and minimal with respect to dataset D_A iff each concept in D_A is either perfectly and non-redundantly linked to D_B or is properly unlinked from D_B.*

5. THE PROTOCOL FOR CREATING THE DATASETS CORPUS

As the objective for the golden-standard corpus is to reach a number of 300 concepts per dataset, a protocol for selecting the 300 concepts without de facto performing a mapping (which we wanted to perform on a second phase, after the cut for the corpus had been done) had to be established.

1. Having the two concept trees at hand or simply, by knowing the domains covered by the two thesauri, a domain is chosen so that the selected concepts will be generally expected to be connected with it.
2. In order to generate cuts of the thesauri that can be still considered as real resources, macro portions of the hierarchy should be selected, without performing too much pruning. The user performing the cut should thus select a concept relevant for the domain (and general enough to subsume other relevant concepts) and perform a bird's

eye inspection (again, the real alignment has to be performed later) to verify that roughly a non-trivial number of concepts is relevant to the domain (and thus possibly linkable to concepts in the cut of the other thesaurus). If the number of concepts in the retrieved branch is too high, some cuts will be performed on it. Conversely, if it is too low, a new branch can be selected – with the same approach – from the original thesaurus in order to increment the number of chosen units.

3. Once the two cuts have been produced, a further bird's eye over their content should confirm if there is an acceptable overlap between the two selected portions of thesauri.

An important aspect to clarify, when working with alignment corpora based on real thesauri cuts, is that a perfectly sound linkset which is also minimal and complete with respect to the aligned excerpts, cannot be considered as a subset of a linkset (with same properties of soundness and completeness) between the two original thesauri. The act of cutting both source and target thesauri implies that the search for alignments will incur in local minimums, dictated by the restricted search space.

6. THE EXPERIENCE IN ALIGNING EUROVOC AND TESEO

The alignment work has been organized as according to the following protocol:

- The pivot language for the alignment has been the Italian language, as TESEO is only expressed in this idiom.
- A domain expert has manually developed a first version of the linkset by analyzing all concept combinations from the two thesauri. The search space is not trivial (300x300=90000 combinations). For this reason, an a-priori identification of clearly exact-matched concepts has been performed in order to reduce its dimension.
- A Semantic Web expert has revised the work performed by the domain expert, by both validating/rejecting/modifying the alignments produced in the first step and by performing a further search (notably reduced in size thanks to the first step and the following validation), discovering new alignments.

The following table provides some figures for the outcome of the alignment process:

Table 1. Results of the alignment

alignments produced by the domain expert	74
alignments discarded by the SW expert	18
alignments modified (type of relation) by the SW expert	3
alignments added by the SW expert	45
total amount of alignments at the end of SW expert's review	101

As it is possible to observe, the Semantic Web expert discarded a non-trivial fraction (~25%) of the original alignments, while being able to discover many more that had been overlooked by the domain expert (though he later re-confirmed these additions).

Had we more resources, a further improvement we would have added to the procedure would have consisted in doubling the first step with two domain experts. This would have allowed us to report on standard measures, such as inter-annotator agreement [12], which are a valid instrument for evaluating the easiness of the task (e.g. if humans agree on a given percentage of the results after carefully performing the job, we cannot expect a machine to perform better than them). In retrospect, considering the numerous changes and improvements that the Semantic Web expert brought during the second step, the inter-annotator agreement would have been extremely low.

This probably reflects the inherent difficulty for humans in performing the alignment task (especially if their familiarity with the domain is not matched by an equal proficiency with thesauri and computer systems) more than the specific case study. In this case, the inter-annotator agreement could probably have not been considered a reliable upper bound for machine performance.

6.1 Report on observed phenomena

Besides the direct outcome of producing the linkset resource, aligning two thesauri is also an occasion to delve into the details of the examined resources and to discover many inconsistencies and issues (because of the deep attention in identifying the nature and identity of concepts) that would have been overlooked otherwise.

The experience in linking the two (cut) thesauri EuroVoc and TESEO has presented a series of results that are indeed not surprising, if some background knowledge about the origins, invested human resources and profile of the two datasets is known a priori. We list here a few considerations emerged during the alignment work:

Rigor. Compared to TESEO, the EuroVoc thesaurus is characterized by a stricter rigor and better precision in the representation and organization of concepts. Vice versa, despite multilinguality is one of its stronger characteristics, EuroVoc loses in precision when we consider labels for languages that are not English or French (respectively, the de facto lingua franca spoken in EU and the first language used in the offices of the European Commission in Luxembourg). This is probably due to the natural interpretation gap generated by different users dealing with the conceptual content and with its translation in different languages (whereas, in a monolingual thesaurus such as TESEO, the two roles, at least for the preferred labels selected at concept creation, converge into the same user). For instance, the domain expert initially expressed a match between “aviazione militare” from EuroVoc and “aeronautica militare” from TESEO. However, despite the two terms show an apparently almost-synonymic expression in Italian, “aviazione militare” in EuroVoc revealed to be a wrong translation for “military aircraft”, whereas the best representative in EuroVoc should have been “forze aree” (“air force”).

Different practices. While SKOS is already a very permissive standard, leaving much liberty in terms of modeling choices (which is also a limit in terms of universal understandability and thus shareability of content), a further degree of freedom is given by different practices at content level that characterize thesauri since the dawn of their time. For instance, in TESEO it is possible to find many cases in which, rather than modeling a domain concept and its more specific interpretations as different formal concepts in SKOS, they are all converged into a single SKOS concept where the preferred label represents the most general interpretation in the domain and alternative labels provide its specializations. In terms of mere representation this approach might be considered a severe mistake. However, if we consider thesauri merely as tools supporting retrieval of annotated resources (usually documents), this conglomeration seems to be justifiable in all cases in which there is no particular interest in exploring the details of the domain concept (thus considering one element as enough), while it is still considered important to broaden the range of its potential lexical expressions. For instance, in TESEO the concept of “conventional weapon” is represented by several labels including "artillery" that is indeed a lexical expressions denoting only one of the possible conventional weapons. In EuroVoc, there seems to be no use (at least, not as broad as in TESEO) of this conceptual collapse, probably due to the larger resources invested in developing the resource, and in the wider scope of its application. TESEO has indeed been explicitly developed for indexing the documents of the Italian Senate, while EuroVoc represents a conceptual hub for all member states, possibly reused in different contexts other than indexing of the documents published by the EU Publications Office.

Differences in the Domain. Aligning two thesauri does not require to deal only with different modeling and lexical choices. The domains of the two resources might not be perfectly overlapping, especially with respect to concrete objects and the classifying concepts that include them. For instance, the domain expert simply mapped two concepts, both labeled as "esercito", as an exact match. However, by looking at the structure of EuroVoc and other translations of the term, it appears that EuroVoc's "esercito" was actually a representation of "armed forces", whilst it is "esercito di terra" that most closely represents the "army", that is what truly the "Esercito" is in Italian. Until now, it could seem another mere translation issue. However, still, if "esercito di terra" has to be interpreted in its broader sense of "ground force" then, strictly speaking, different corps may include ground forces (e.g., US' navy seals belong to the navy while being a ground force, as much as Italian "battaglione San Marco" does). Simply, there is a blurred definition for these elements, as each country has its own military organization and it is not possible to make a 1-1 transposition. For this reason, the new alignment replacing the wrong one has in any case been represented as a close match. Analogous issues happened with all job-related concepts related to workers, rights, separation of legal and physical persons and their abilities. Another troubled area lies in government aspects, as the different perspectives given by the governance of the EU and of a specific government of a member state, brought clearly diverging bias affecting even the most general concepts.

7. THE ALIGNMENT BETWEEN EUROVOC AND THE IATE TERMINOLOGY

The alignment between EuroVoc and IATE followed the same general protocol adopted for the first alignment previously described. We mention here the main differences in terms of knowledge representation models being adopted by this third resource, IATE. IATE is a terminology, giving more emphasis on the lexical aspect rather than the conceptual one. Despite that, lexical entries in IATE are still organized around meanings, which are however not structured into a hierarchy like a traditional thesaurus. IATE is thus being modeled as a OntoLex-lemon lexicon. The lexical entries are described in their details (e.g. term composition) thanks to the above standard, and the meanings are represented as ontolex:LexicalConcepts, a dedicated class of OntoLex (subclassing skos:Concept) expressly thought for representing units of thought in lexicons. The presence of (lexical) concepts allowed us to keep a "concept to concept" approach to alignment representation. Conversely, the lack of a conceptual hierarchical structure forced us to adopt a different strategy for the pre-selection of a set of concepts for the initial development of the corpus to be aligned. Luckily, IATE includes an explicit concept of domain called "subject field", thus we have selected the field "Civil Law" which includes 270 concepts (so we could take them all), and we could proceed with the already experimented process of branch-selection for EuroVoc.

8. CRITERIA FOR THE EVALUATION OF ALIGNMENTS AGAINST THE PRODUCED GOLD-STANDARD ALIGNMENTS

The criteria for the evaluation of alignments follows from the definitions provided in section 4. We define as "alignment oracle" the alignment used as a reference for verifying the quality of a produced linkset.

The following criteria follow:

Given an alignment oracle O that represents a minimal complete mapping between datasets D_A and D_B and a linkset to evaluate L, a single match $l(c_A, c_B) \in L : c_A \in A,\ c_B \in B$
is said to be correct if

$l(c_A, c_B) \in O$

or $l(c_A, c_B)$ is not a tolerance relation and $c_A \in reach(c_B)$

In other words, a computed alignment relationship is correct either if it is explicit in O or if the aligned concept can "reach" the concept in the target dataset and the alignment relation is not a tolerance relation.

9. CONCLUSIONS

In this deliverable, we have described our guidelines for the production of golden-standard alignments with a reasonably-limited amount of resources and reported on our application of the guidelines for the development of such alignments.

CONCLUSIONS

Alexander, K., Cyganiak, R., Hausenblas, M., & Zhao, J. (2011, March 3). *Describing Linked Datasets with the VoID Vocabulary (W3C Interest Group Note).* Retrieved May 16, 2012, from World Wide Web Consortium (W3C): http://www.w3.org/TR/void/

Artstein, R. (2017). Inter-annotator Agreement. In N. Ide, & J. Pustejovsky (Ed.), *Handbook of Linguistic Annotation.* Springer, Dordrecht. doi:https://doi.org/10.1007/978-94-024-0881-2_11

Berners-Lee, T., Hendler, J. A., & Lassila, O. (2001). The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American, 279*(5), 34-43.

Borin, L., Dannélls, D., Forsberg, M., & McCrae, J. P. (2014). Representing Swedish Lexical Resources in RDF with lemon. *Proceedings of the ISWC 2014 Posters & Demonstrations Track a track within the 13th International Semantic Web Conference (ISWC 2014)*, (pp. 329-332). Riva del Garda, Italy. Retrieved from http://ceur-ws.org/Vol-1272/paper_82.pdf

David, J., Euzenat, J., Scharffe, F., & Trojahn dos Santos, C. (2011). The Alignment API 4.0. *Semantic Web Journal, 2*(1), 3-10.

Dragisic, Z., Eckert, K., Euzenat, J., Faria, D., Ferrara, A., Granada, R., . . . Cuenca Grau, B. (2014). Results of the ontology alignment evaluation initiative 2014. *9th International Workshop on Ontology Matching, October 20, 2014. 1317*, p. 61-104. Riva del Garda, Trentino, Italy: CEUR-WS.org.

Eckle-Kohler, J., McCrae, J., & Chiarcos, C. (2015). lemonUby – A large, interlinked, syntactically-rich lexical resource for ontologies. *Semantic Web, VI*(4), 371-378. doi:http://doi.org/10.3233/SW-140159

Enea, R., Pazienza, M. T., & Turbati, A. (2015). GENOMA: GENeric Ontology Matching Architecture. In M. Gavanelli, E. Lamma, & F. Riguzzi (A cura di), *Lecture Notes in Computer Science* (Vol. 9336, p. 303-315). Springer International Publishing. doi:10.1007/978-3-319-24309-2_23

Euzenat, J., & Shvaiko, P. (2013). *Ontology Matching* (2nd ed. ed.). Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-38721-0

Evans, V. (2006). Lexical concepts, cognitive models and meaning-construction. *Cognitive Linguistics*, *XVII*(4), 491-534. doi:https://doi.org/10.1515/COG.2006.016

Fiorelli, M., Pazienza, M. T., & Stellato, A. (2014). A Meta-data Driven Platform for Semi-automatic Configuration of Ontology Mediators. In N. C. Chair), K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, . . . S. Piperidis (Eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), May 2014* (pp. 4178-4183). Reykjavik, Iceland: European Language Resources Association (ELRA). Retrieved from http://www.lrec-conf.org/proceedings/lrec2014/pdf/1059_Paper.pdf

Navigli, R., & Ponzetto, S. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence, CXCIII*, 217-250. doi:https://doi.org/10.1016/j.artint.2012.07.001

Pavel, S., & Euzenat, J. (2013, January). Ontology Matching: State of the Art and Future Challenges. *IEEE Transactions on Knowledge and Data Engineering, 25*(1), 158-176. doi:10.1109/TKDE.2011.253

Pazienza, M. T., Stellato, A., & Turbati, A. (2008). Linguistic Watermark 3.0: an RDF framework and a software library for bridging language and ontologies in the Semantic Web. *5th Workshop on Semantic Web Applications and Perspectives (SWAP2008), Rome, Italy, December 15-17, 2008, CEUR Workshop Proceedings, 426*, p. 11. FAO-UN, Rome, Italy. Retrieved from http://ceur-ws.org/Vol-426/swap2008_submission_39.pdf

Pazienza, M., Sguera, S., & Stellato, A. (2007, December 26). Let's talk about our “being”: A linguistic-based ontology framework for coordinating agents. (R. Ferrario, & L. Prévot, Eds.) *Applied Ontology, special issue on Formal Ontologies for Communicating Agents, 2*(3-4), 305-332.

Schmitz, P., Francesconi, E., Hajlaoui, N., & Batouche, B. S. (2018). Semantic Interoperability of Multilingual Language Resources by Automatic Mapping. In E. Francesconi (A cura di), *Electronic Government and the Information Systems Perspective (Lecture Notes in Computer Science)* (Vol. 11032, p. 153-163). Springer, Cham. doi:https://doi.org/10.1007/978-3-319-98349-3_12

Stellato, A. (2012, May). Dictionary, Thesaurus or Ontology? Disentangling Our Choices in the Semantic Web Jungle. *Journal of Integrative Agriculture, 11*(5), 710-719.

van Ossenbruggen, J., Hildebrand, M., & de Boer, V. (2011). Interactive vocabulary alignment. *Proceedings of the 15th international conference on Theory and practice of digital libraries: research and advanced technology for digital libraries* (pp. 296-307). Berlin, Germany: Springer-Verlag.

World Wide Web Consortium (W3C). (2009, August 18). *SKOS Simple Knowledge Organization System eXtension for Labels (SKOS-XL).* (A. Miles, & S. Bechhofer, Eds.) Retrieved March 22, 2011, from World Wide Web Consortium (W3C): http://www.w3.org/TR/skos-reference/skos-xl.html

World Wide Web Consortium (W3C). (2009, August 18). *SKOS Simple Knowledge Organization System Reference.* (A. Miles, & S. Bechhofer, Eds.) Retrieved March 22, 2011, from World Wide Web Consortium (W3C): http://www.w3.org/TR/skos-reference/