eChallenges e-2011 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2011
ISBN: 978-1-905824-27-4

Thesaurus Mapping for Promoting Semantic Interoperability of European Public Services[1]

M.A. BIASIOTTI, S. FARO, E. FRANCESCONI
ITTIG-CNR, via de' Barucci 20, Florence, 50127, Italy
Tel: +39 055 43995, Fax: +39 055 4399605, Email:
{biasiotti,faro,francesconi}@ittig.cnr.it

Abstract: Interoperability of eGovernment information systems is essential to provide advanced services to citizens. This work proposes a framework for implementing interoperability among thesauri for promoting cross-collection and cross-language information retrieval, as well as a specific approach within such framework on a case study aimed at mapping five thesauri of interest for the European Union institutions (EUROVOC, ECLAS, GEMET, UNESCO Thesaurus and ETT) having only schema information available. In particular the standards as well as the approach for the interoperability case study implementation are presented; finally the interoperability assessment criteria and experimental results for the case study are shown.

1. Introduction

Public administrations use ICTs in a systematic way for providing several public services to different stakeholders (from citizens to SMEs). This practice has been adopted in the last 10 years in a massive way evolving from time to time according to the progress of new technologies.

Currently, as underlined by [19], the systematic implementation of ICTs in Public administrations daily activities has, as direct consequence, the need to face some challenges arising from the latest evolution of the Semantic Web approach. More specifically: (i) Knowledge representation (Governments need to capture information about themselves and describe themselves to their citizens and stakeholders); (ii) Integration of information (the information resources in e-government contexts are a wide set, including data, documents (including multi-media), files for download, transactions, links, services, and user-provided or user-related items (such as credit card details); (iii) Search and discovery (data in most government organizations are stored in many databases in different departments and locations, yet to realize the potential gains from these data it is necessary to bring them together across government and elsewhere on the Web to achieve synergy and maximize value).

In order to cope with some aspects of the above mentioned challenges, Governments need to provide themselves with semantic tools such as ontologies, thesauri and other Knowledge Organizational System able to represent contents of public data rendering them more understandable for users and more easily machine-processable. Currently each public body providing public services/data has its own semantic tool, of different level, degree and granularity.

[1] M.A. Biasiotti and S. Faro are authors of Sections 1, 3, 5, 8; E. Francesconi is author of Sections 2, 3.1, 4, 6, 7.

When providing public services different departments need to interact within the same context: therefore enhancing the interoperability among organization and exchanged data becomes an essential factor to maximise the social and economic potential of ICTs.

This is even more important migrating from the national perspective into the transnational one, i.e. cross-border public sector services supplied by public administrations, either to one another or to European businesses and citizens. In this perspective more awareness should be focused on the risk of creating new electronic barriers if opting for solutions that are not interoperable.

In December 2010 the European Commission enacted a communication on interoperability for European public services [20]. It includes a European interoperability strategy and European interoperability framework.

Starting from the consideration that "lack of agreement and guidance on the meaning and format of information to be exchanged between Member States is a stumbling block" and that "semantic interoperability is jeopardised by different interpretations of the information exchanged between people, applications and administrations", both documents underline the relevance of semantic interoperability. It "enables organisations to process information from external sources in a meaningful manner". Moreover it "ensures that the precise meaning of exchanged information is understood and preserved throughout exchanges between parties".

In this vein a solution enhancing interoperability for processing information among organizations each acting according to its own semantic tool might arise from the mapping of these resources. Starting from a case-study about interoperability of thesauri of interest of the European institutions, this paper will propose an approach for estabishing semantic interoperability to be extended as a pilot methodology to all possible applications related to public services.

In particular this work proposes a methodological framework for semantic mapping between thesauri as well as a specific approach within such framework on a case study aimed at mapping five thesauri (EUROVOC, ECLAS, GEMET, UNESCO Thesaurus and ETT) of interest for the European Union institutions having only schema information available.

2. Formal characterization of thesaurus mapping

Thesaurus mapping can be seen as the process of identifying terms, concepts and hierarchical relationships that are approximately equivalent [3]. The problem therefore is to define the meaning of "equivalence" between concepts.

In literature [4] [14] "concept equivalence" is defined in terms of set theory: according to this vision two concepts are deemed to be equivalent if they are associated with, or classify the same set of objects [2] (*Instance-based mapping* [3]).

On the other hand concepts may also be associated with semantic features and mappings can be based on equivalences between feature sets (*Schema-based mapping* [3]). This approach is the only possible when only schema information is available. .

Taking into account this classification and elementary techniques described in literature, in this paper a framework is proposed along with a methodology to implement thesaurus mapping for the case study, which can be classified as *Schema-based mapping* problem [3], since documents classified with such thesauri are not available. The problem is to identify the conceptual/semantic similarity between a concept (expressed by simple or complex terms[2]) in the source thesaurus and candidate concepts in a target thesaurus. This can be done providing similarity measures according to specific semantic representations of concepts and related metrics for comparison. The characteristics of this problem allow us to

[2] for example *Parliament* is a simple term, *President of the Republic* is a complex term.

propose a characterization of the schema-based Thesaurus Mapping (*TM*) as a problem of Information Retrieval (*IR*): the aim is to find concepts in target thesaurus, better matching the semantics of a concept in a source thesaurus. The isomorphism between *TM* and *IR* ($TM \equiv IR$) can be established once we consider a source concept as a *query* of the *IR* problem, and a target concept as a *document* of the *IR* problem.

Therefore, the *TM* problem can be viewed and formalized, like the *IR* problem, as a 4-uple $TM = [D, Q, F, R(q, d)]$ [5], where:

1. *D* is the set possible representations (*logical views*) of a concept in a target thesaurus (a document to be retrieved in the *IR* problem);
2. *Q* is the set of the possible representations (*logical views*) of a concept in a source thesaurus (a query in the *IR* problem);
3. *F* is the framework of concepts representation in source and target thesauri (for example an euclidean space)
4. *R(q, d)* is a ranking function, which associates a real number with (*q, d*) where $q \in Q$ (source concept), $d \in D$ (target concept), giving an order of relevance to the concepts in a target thesaurus with respect to a concept of a source thesaurus.

In this framework the implementation of a thesaurus mapping procedure is represented by the instantiation of the previous 4 components. Before going in this direction, a possible approach to represent theasuri and promote interoperability, within such framework, is presented (Section 3).

3. Standard representation of thesauri concepts and relations

ISO has defined two international standards ISO5964/ISO2788[3] which are useful to ensure consistency in the development of mono/multilingual thesauri within or between indexing agencies. Such standards, which are to be replaced by ISO25964, provide guidelines for concepts and relations but do not provide guidelines for adopting specific thesaurus data formats. In order to manage, process and compare different thesauri structures, as well as to share them in a machine readable way, the use of a common standard, able to keep the semantics of their native data formats, is essential. The Semantic Web community has developed the SKOS[4] standard that uses RDF to represent different knowledge organization systems such as thesauri, classification schemes, subject heading systems and taxonomies, as well as to share them in a distributed environment. Following SKOS recommendations[5], a knowledge organization system can be viewed as a *concept scheme* including a set of concepts. The SKOS vocabulary deems a concept (identified by the skos:concept class) as the most elementary unit. A concept can be connected with any number of strings, in any natural language, but with only one preferred label (in every language) while it can have infinite alternative descriptions. By the use of the skos:prefLabel and skos:altLabel properties, the preferred and the alternative descriptions are tied to the concepts. Furthermore, one or more notations (`skos:notation`), including a string of characters in any natural language, can be assigned to a SKOS concept in order to identify it in the application field of another concept scheme.

While SKOS provides a standard way to represent thesauri descriptors and relationships, in literature different approaches to represent thesauri have been proposed [15] [16] [17], but, so far, no standardized architecture for translating thesauri from their proprietary format has emerged. Therefore, in this case study, a knowledge architecture for

[3] ISO5964/ISO2788 *Guidelines for the establishment and development of multilingual/monolingual thesauri*
[4] Simple Knowledge Organization System (http://www.w3.org/2004/02/skos/)
[5] developed from the 2005's to the 2009's versions.

representing thesauri using SKOS is proposed on the basis of similar experiences reported in literature.

3.1 A knowledge architecture to represent thesauri using SKOS

[17] is an interesting work which proposes an architecture to represent thesaural concepts and relations: it describes a structured method to convert thesauri to SKOS, evaluating the applicability of SKOS meta model to represent existing thesauri. Starting from the method given in [17], as well as SKOS specifications, in our case-study a methodology for thesauri conversion to SKOS is proposed. In particular the following criteria have been followed: thesauri descriptor labels are represented by `skos:prefLabel`; *used-for* relations are represented by `skos:altLabel`, and different kind of *notes* are mapped to the correspondent `skos:scopeNote` and `skos:editorialNote` elements. *Broader*, *narrower* and *related* relations are directly mapped to the corresponding SKOS properties. Moreover, the native multilingual tools provided by SKOS made it easy to handle the multilingual labels connected to the concepts.

Structural patterns which do not have a direct counterpart in SKOS are represented providing extensions according to SKOS specifications [18]. For example, usually and in particular in our case-study, thesauri have a conceptual structure, organized in hierarchical levels. Therefore, in order to describe their native semantics, the `skos:Concept` class has been extended into to 3 additional classes: `eu:Descriptor`, `eu:Microthesaurus` and `eu:Domain` where 'eu' is the namespace defined in this work for the Publication Office of the European Union thesauri SKOS extension (Fig. 1).

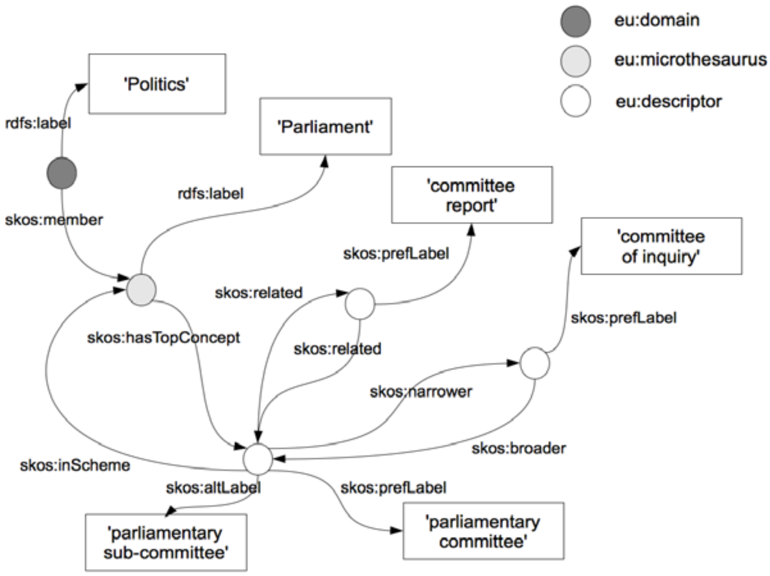

Fig. 1 SKOS representation of an EUROVOC excerpt.

4. Logical views (Q and D) of *descriptors*

Mapping between thesaural concepts is a process that aims at matching concept semantics rather than their lexical equivalences. In traditional thesauri *descriptors* and *non-descriptors* are represented by different terms (`skos:prefLabel` and `skos:altLabel`, according to SKOS) expressing the same meaning. To effectively map thesaural concepts, term (simple or complex) semantics has to be captured and represented. In *IR* a query is usually constructed as a context (set of keywords). Similarly in *TM* the semantics of a thesaural concept is conveyed not only by its terms, but also by the context in which the concept is used as well as by the relations with other concepts. In *TM* problem, *Q*, *D* and *F* are exactly aimed at identifying logical views and related framework for concept representations able to capture the semantics of terms in source and target thesauri, as well as to measure their conceptual similarity.

We propose to represent the semantics of a term in a thesaurus, according to an ascending degree of expressiveness, by: its Lexical Manifestation, its Lexical Context, its Lexical Network.

Lexical Manifestation of a thesaurus concept is its expression as a string of characters, normalized according to some standard pre-processing steps (stemming, stopwords elimination, etc.).

In this work we propose to represent the semantics of a thesaural concept by a vector *d* of binary[6] entries composed by the term itself, relevant terms in its definition, in the alternative labels, as well as terms of directly related thesaural concepts (broader, narrower, related concepts).

Lexical Context of a thesaurus concept is represented by a vector *d* of term binary entries (statistics on terms to obtain weighted entries are not possible since document collections are not available) composed by the term itself, relevant terms in its definition and linked terms.

Firstly a vocabulary of normalized terms from target thesaurus is constructed, where 'normalization' in this context means string pre-processing, in particular word stemming and stopwords eliminations. Being *T* the dimension of such vocabulary, both source and target concepts *d* are represented in a vector space of *T*-dimension ($d = [x_1, x_2, \ldots, x_T]$); the entry x_i gives information on the presence/absence of the corresponding ith vocabulary term among the terms characterizing the concept *d*. In Fig. 2 a binary vector representation of a EUROVOC concept is sketched. In such representation the framework *F* is composed of *T*-dimensional vectorial space and linear algebra operations on vectors.

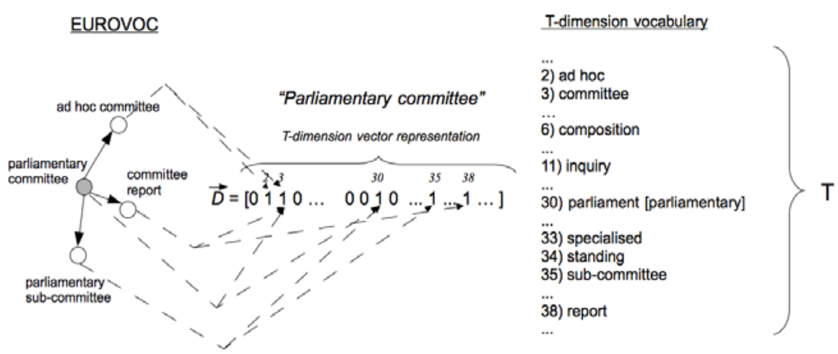

Fig. 2 T-dimension vectorial representation of a thesaural descriptor.

Lexical Network of a thesaurus concept is a direct graph where nodes are the related thesaurus descriptors as well as the descriptors and non-descritors in relation with it, and the labeled edges are semantically characterized relations between concepts.

For Lexical Contexts and Lexical Networks, the concepts connection degree is based on a strict adjacency relations with a descriptor.

4.1 The proposed Framework (F)

Having identified thesaurus concepts logical views, the frameworks in which the *TM* problem can be modelled are also identified.

---

[6] Statistics on terms to obtain weighted entries are not possible since document collections are not available (*schema-based thesaurus mapping*)

For Lexical Manifestations, the framework is composed of strings representing terms and the standard operations on strings. For Lexical Contexts, the framework is composed of *T*-dimensional vectorial space and linear algebra operations on vectors. For Lexical Networks, the framework is composed by graphs (described by nodes, edges and related labels) and the algebra operations on graphs.

The frameworks identified can also provide the intuition for constructing a ranking function *R*, which will be linked to the chosen representation of the space elements (descriptors).

4.2 The proposed Ranking Functions (R)

The ranking function *R* will be able to provide a similarity measure between a concept in a source thesaurus and an associated one in a target thesaurus; when extended to a set of target concepts such a function may provide a matching order among such concepts. With respect to the logical views on concepts identified in Section 4, possible ranking functions to measure the degree of mapping between thesaurus concepts are proposed.

Ranking function for Lexical Manifestations. In this case string-based techniques are used, in particular the Edit distance/similarity (or Levenshtein distance/similarity) applied on pre-processed strings through language-based techniques (as Tokenization, Lemmatization (Stemming) and Elimination (Stopword elimination)) normalized with respect to the longest string (therefore this measure varies in the interval [0,1]).

Ranking function for Lexical Contexts. Having represented the semantics of thesaural concepts as a binary vector, their similarity can be measured as the related binary vectors correlation, quantified, for instance, as the cosine of the angle between them

$$sim(q, d) = q \text{ x } d \ / \ |q| * |d|$$

where $|q|$ and $|d|$ are the norms of the vectors representing concepts in source and target thesauri, respectively.

Ranking function for Lexical Networks. Having represented the semantics of a thesaurus concepts as a Lexical Network, basically a direct graph, a vast literature exists in graph theory [7] to measure the distance between two graphs [8] [9]. In literature the more frequently addressed graph similarity measure is the Graph Edit Distance, namely the minimum number of nodes and edges deletions/insertions/substitutions to transform a graph g_1 into a graph g_2 [10]. Because of computational complexity we have considered three variants of the Graph Edit Distance: Conceptual similarity, Relational similarity and Graph similarity.

The Conceptual similarity sim_c expresses how many concepts two graphs g_1 and g_2 have in common by an expression analogous to the Dice coefficient [11]:

$$sim_c = 2n(g_c) / n(g_1) + n(g_2)$$

where g_c is the maximum common subgraph of g_1 and g_2 (it defines the parts of both graphs which are identical to one another, and, intuitively, it describes the intersection between the two graphs) and $n(g)$ is the number of concept nodes of a graph g. This expression varies from 0 when the two graphs have no concepts in common to 1 when the two graphs consist of the same set of concepts.

The Relational similarity sim_r indicates how similar the relations between the same concepts in both graphs are, that is, how similar the information communicated in both texts about these concepts is. In a way, it shows how similar the contexts of the common concepts in both graphs are. Using a modified formula for the Dice coefficient, sim_r can be obtained as:

$$sim_r = 2m(g_c) / m_{gc}(g_1) + m_{gc}(g_2)$$

where $m(g_c)$ is the number of the arcs (the relation nodes in the case of conceptual graphs) in the graph g_c, and $m_{gc}(g_i)$ is the number of the arcs in the immediate neighborhood of the graph g_c in the graph g_i. The immediate neighborhood of g_c □ g_i in g_i consists of the arcs of g_i with at least one end belonging to g_c.

Considering a graph g to be matched with a graph g_T as reference, a possible similarity measure able to sum-up the previous two is the Graph similarity [12]:

$$sim_g = Nc\ (g, g_T) + Ec\ (g, g_T) / N\ (g_T) + E(g_T)$$

where $Nc\ (g, g_T)$ is the number of nodes shared by graph g and g_T; $Ec\ (g, g_T)$ is the number of edges common to g and g_T; $N\ (g_T)$ is the number of nodes in graph g_T; $E(g_T)$ is the number of edges in g_T.

4.3 Ranking among candidate terms and mapping implementation

Concepts of the target thesaurus, represented according to one of the described models, are matched with the chosen concepts in a source thesaurus, represented with the same model, using a proper similarity measure. Candidate concepts of the target thesaurus are ranked according to the similarity measure values sim □ $[0, 1]$ and a semantics to the mapping relation is assigned using proper heuristic threshold values ($T1$, $T2$ □ $[0, 1]$)

if $sim < T1$	exactMatch
if $T1 < sim < T2$	partial match (broadMatch or narrowMatch)
if $T2 < sim$	No Match

Then the representation of the established relations between thesaurus concepts is expressed using RDF SKOS mapping properties.

5. Interoperability assessment through a "gold standard"

Interoperability between thesauri is specifically assessed on a data sample selected from the thesauri of interest. In order to evaluate the performance of automatic mapping algorithms an intellectual activity is needed to create a "gold standard", namely a groundtruth file of thesauri concept mapping examples (one for each couple of thesauri having EUROVOC as pivot) which represents the ideal set of expected correct mappings. It is aimed at 1) tuning system heuristics (similarity measure threshold values are tuned to obtain the best results with respect to the gold standard (performance convergence)), 2) evaluating the performances of automatic mapping algorithms, comparing the ideal set of mapping examples with system predictions.

For the purpose of comparison between the "gold standard" and the algorithms prediction, mapping relations of the "gold standard" is described using SKOS mapping properties, limited to the exactMatch, broadMatch and narrowMatch relations. In particular when a concept in EUROVOC corresponds exactly to one or more concepts in a target thesaurus according to the expert judgment, the relation is an exact match. A broad equivalence is one where the source concepts is more specific in some sense than the target one. Similarly a narrow equivalence is one where the source concept is more general in some sense than the target concept or expression [13]. Following [1], when it is possible, mappings that are at least complete, and ideally optimal, have to be established.

A complete mapping [13] is one where a source concept, having no exact equivalence in the target thesaurus, is matched to at least one target concept that is semantically broader and at least one semantically narrower.

An optimal mapping [13] is one where the aforementioned broader target concept is the nearest broader concept to the source term, and where the aforementioned narrower target concept is the nearest narrower concept to the source concept.

In our case study the "gold standard" construction intellectual activity has been carried out by experts of the specific chosen domains, using tools to make their work easier. Some solutions have been evaluated to implement the "gold standard", in particular Protégé

PROMPT merging tool[7], Vocabulary Integration Environment (VINE)[8] and AIDA. All these software have been tested. For different reasons they have been considered too problematic and not enough user-friendly. Therefore we have developed a specific application (THesauri ALigning ENvironment (THALEN)) able to provide a user-friendly access to thesauri and simple functionalities to establish term relations for mapping.

THALEN has been developed on an MS-Access relational database and provides simple functionalities of user authentication, thesaurus loading, parallel view of two thesauri, search modalities (term browsing and searching), manual mapping implementation, summary of the established term mapping as well as exportation of mapping relations in RDF SKOS. Moreover this application can be used for human validation of the automatic mapping. In Fig. 3 a screenshot of THALEN as developed for the project is shown.

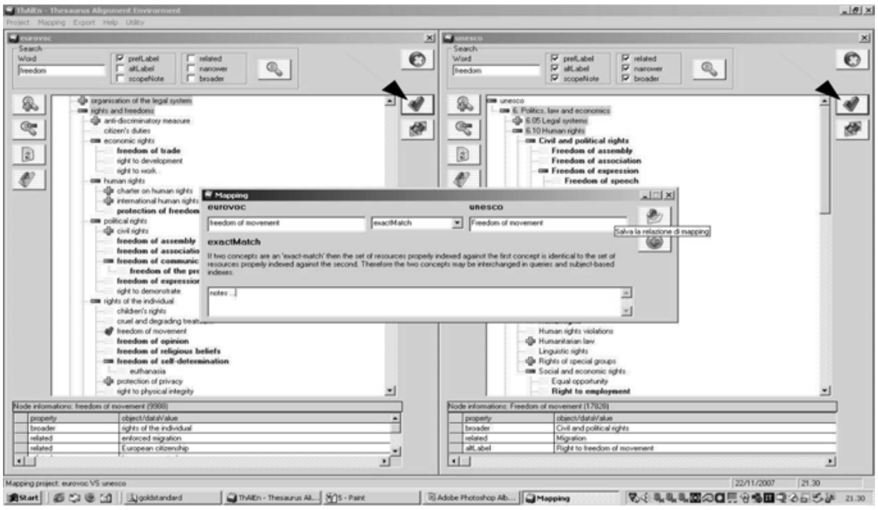

Figure 3 *A screenshot of THALEN showing a parallel view of two thesauri and a form where to establish mapping relations.*

6. Test of the proposed TM approaches

The "gold standard" produced by experts has been used to assess the proposed methodologies for automatic thesaurus mapping: it includes 624 relations, of which 346 are exactMatch relations.

The system mapping performances have been assessed with respect to the "gold standard" for each single mapping relation type, using the typical Precision and Recall measures. In particular for the project case study the system Recall has been considered since the automatic mapping is addressed to identify matching concepts within the system predictions, to be validated by humans. In our case study an important measure is represented by Recall with respect to a less specified relation (untypedMatch) able to express a generic association between concepts.

Preliminary experiments have shown satisfactory performances as regards the identification of untypedMatch relations between concepts, and, as a consequence, the identification of noMatch relations; on the contrary, good performances have been obtained as regards the selection of exactMatch relations, while the distinction between narrowMatch and broadMatch relations revealed a high degree of uncertainty. Global performances on specific types of relations are highly affected by this uncertainty, therefore meaningful

7 http://protege.stanford.edu/plugins/prompt/prompt.html

8 http://marinemetadata.org/vine

results can be given with respect to untypedMatch relations as well as the system ability to identify exactMatch relations with respect to the “gold standards”.

The proposed logical views for thesaurus concepts and the related ranking functions outperformed a simple string matching among thesaurus concept terms. In particular for the following couples of thesauri (EUROVOC vs. {ETT, ECLAS, GEMET}) the Lexical Manifestation logical view and the Levenshtein Similarity ranking function gave the best results (untypedMatch Recall = 66.2%, exactMatch Recall = 82.3%), while for the couple EUROVOC vs. UNESCO Thesaurus, the best results was obtained using the Lexical Network logical view and the Conceptual Similarity ranking function (untypedMatch Recall = 73.7%, exactMatch Recall = 80.8%).

7. Evaluation of the results

Different thesaurus concept logical views and related ranking functions to establish conceptual similarity have been proposed and tested. The preliminary results revealed that a simple Lexical Manifestation logical view and Levenshtein Similarity ranking function produced the best results on most of the matches between the thesauri couples of interest.

More complex representations (Lexical Contexts, Lexical Networks) suffer from problems of computational tractability (in particular as regards Lexical Networks), moreover the use of higher number of features for concepts description provides a higher variability of the similarity measures, affecting the algorithms performances.

To improve the performances of the system using Lexical Contexts and Lexical Networks, which usually provide a higher degree of expressiveness in concept descriptions, different criteria of features selection can be tested, aiming at reducing the computational complexity, as well as the variability of the similarity measures.

8. Conclusions and Recommendations

Providing public services at transnational levels implies processing and classifying data and information in their flow, exchange and access, making them available in a cross-border dimension and independently from the specific public administration approach as well as from user background.

Establishing semantic interoperability between classification systems represents a crucial step in linking European Public services. Semantic tools can greatly contribute to the achievement of these goals.

This paper presented a methodological framework and a specific implementation of schema-based thesaurus mapping. The approach has been assessed on a case study focused on five thesauri of interest for the EU institutions.

Generally speaking mapping semantic resources has positive results as it allows Public adminstrations to implement interoperable services without changing legacy systems. More specifically the methodological approach provides a framework that can be adaped to different scenarios with the aim to share experiences and make comparisons.

Benefits can be derived for both Public administrations and citizens: the former can increase their efficency and promptness in providing services to citizens; the latter can access transnational public services, overcoming language and cultural barriers.

References

[1] M. Doerr, “Semantic problems of thesaurus mapping,” Journal of Digital Information, vol. 1, no. 8, 2001.
[2] A. Miles and B. Matthews, “Deliverable 8.4: Inter-thesaurus mapping,” 2004. http://www.w3c.rl.ac.uk/SWAD/deliverables/8.4.html
[3] E. Rahm and P. Bernstein, “A survey of approaches to automatic schema matching,” The International Journal on Very Large Data Bases, vol. 10, no. 4, pp. 334–350, 2001.
[4] J. Euzenat and P. Shvaiko, Ontology Matching. Springer, 2007.

[5] R. Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval. Addison Wesley, 1999.
[6] A. Miles and D. Brickley, "Skos mapping vocabulary specification," 2004.
http://www.w3.org/2004/02/skos/mapping/spec/
[7] J. Sowa, Conceptual Structures: Information Processing in Mind and Machine.
Addison-Wesley, 1984.
[8] J. Zhong, H. Zhu, J. Li, and Y. Yu, "Conceptual graph matching for semantic search," in Proceedings of 10th International Conference on Conceptual Structures (ICCS), Lecture Notes in Artificial Intelligence 2393, Springer-Verlag, 2002.
[9] M. Montes y Gomez, A. Gelhukh, A. Lopez-Lopez, and R. Baeza-Yates, Flexible Comparison of Conceptual Graphs. Lecture Notes in Computer Science 2113, Springer-Verlag, 2001.
[10] C.-A. M. Irniger, Graph Matching – Filtering Databases of Graphs Using Machine
Learning Techniques. PhD thesis, Institut fur Informatik und angewandte Mathe-
matik, Universit¨ at Bern, 2005.
[11] E. Rasmussen, "Clustering algorithms," in Information Retrieval: Data Structures & Algorithms (W. B. Frakes and R. Baeza-Yates, eds.), 1992.
[12] W.-T. Cai, S.-R. Wang, and Q.-S. Jiang, "Address extraction: a graph matching and ontology-based approach to conceptual information retrieval," in Proceedings of the Third Intemational Conference on Machine Learning and Cybernetics, pp. 1571–1576, 2004.
[13] A. C. Liang and M. Sini, "Mapping AGROVOC and the Chinese agricultural thesaurus: Definitions, tools, procedures," New Review of Hypermedia and Multimedia, vol. 12, no. 1, pp. 51–62, 2006.
[14] Trojahn, C., Moraes, M., Quaresma, P. and Vieira, R. *A Cooperative Approach for Composite Ontology Mapping*, Journal on Data Semantics, pp. 237–263, 2008.
[15] Neubert, J., *Bringing the "thesaurus for economics" on to the web of linked data*," in Proceedings of the WWW2009 Workshop on Linked Data on the Web (T. B.-L. K. I. Christian Bizer, Tom Heath, ed.), vol. 538 of CEUR Workshop Proceedings, CEUR-WS, April 2009.
[16] Isaac, A., Wang, S., Zinn, C., Matthezing, H., van der Meij, L. and Schlobach, S., *Evaluating thesaurus alignments for semantic interoperability in the library domain*," IEEE Intelligent Systems, vol. 24, no. 2, pp. 76—86, 2009.
[17] Assem, M., Malais, V., Miles, A. and Schreiber, G., *A method to convert thesauri to skos*, in Volume 4011 of Lecture Notes in Computer Science, pp. 95--109, Springer, 2006.
[18] Miles A., and Bechhofer, S., *Skos simple knowledge organization system reference*, in http://www.w3.org/TR/skos-reference, W3C Semantic Web Deployment Working Group, 2009.
[19] Shadbolt, N., O'Hara, K., Salvadores, M. and Alani, H. (2011) eGovernment. In: *John Domingue, Dieter Fensel & James Hendler (eds.), Handbook of Semantic Web Technologies*, pp. 840-900, Springer-Verlag.
[20] European Commission *Towards interoperability for European public services, doc. COM(2010) 744 final, of 16 December 2010.*