New Ways of Mapping Knowledge Organization Systems: Using a Semi-Automatic Matching Procedure for Building up Vocabulary Crosswalks

Andreas Oskar Kempf*, Dominique Ritze**, Kai Eckert*** and Benjamin Zapilko*

*GESIS, Leibniz-Institute for the Social Sciences, Unter Sachsenhausen 6-8, 50667Cologne, Germany
<andreas.kempf@gesis.org>, <benjamin.zapilko@gesis.org>
**Mannheim University Library, Schloss Schneckenhof West, 68131 Mannheim, Germany
<dominique.ritze@bib.uni-mannheim.de>
***Mannheim University, Schloss, 68131 Mannheim, Germany
<kai@informatik.uni-mannheim.de>

Andreas Oskar Kempf holds a Ph.D. in sociology from Goethe University, Frankfurt am Main, Germany, and is a postdoctoral researcher at GESIS, Leibniz-Institute for the Social Sciences, Cologne. He received a master's degree in library and information science from Humboldt-University, Berlin, in 2011, with a thesis on automatic indexing of domain-specific information. Apart from automatic indexing his research interests include the interoperability of knowledge organization systems and their integration in semantic web applications.

Dominique Ritze is a Ph.D. student and project employee at the Mannheim University Library. She received her master's degree in computer science from the University of Mannheim, Germany, in 2011. Her research focuses on ontology matching as well as semantic technologies and linked data.

Kai Eckert is a research associate at the University of Mannheim, Germany, where he leads the infrastructure development of the EU-funded project DM2E (Digitized Manuscripts to Europeana). He is a computer and information scientist with master's degrees from the University of Mannheim (computer science, business informatics) and the Humboldt-University of Berlin (MA LIS). He was member of the W3C Provenance Incubator Group and the W3C Library Linked Data Incubator Group. Currently, he participates in the W3C Provenance Working Group and co-chairs the DCMI Metadata Provenance Task Group.

Benjamin Zapilko studied computational visualistics at the University Koblenz-Landau and is research associate at GESIS, Leibniz-Institute for the Social Sciences since 2007. In the department Knowledge Technologies for the Social Sciences he conducts applied research in the fields of semantic web and linked open data. In his Ph.D. thesis, he investigates methods and problems for a useful appliance of linked open data in scientific research. His research interests also include the modeling and standardization of data as well as the matching and integration of heterogeneous data sources.

Kempf, Andreas Oskar, Ritze, Dominique, Eckert, Kai, and Zapilko, Benjamin. **New Ways of Mapping Knowledge Organization Systems: Using a Semi-Automatic Matching Procedure for Building up Vocabulary Crosswalks.** *Knowledge Organization. 41(1)*, 66-75. 25 references.

Abstract: Crosswalks between different vocabularies are an indispensable prerequisite for integrated, high-quality search scenarios in distributed data environments where more than one controlled vocabulary is in use. Offered through the web and linked with each other they act as a central link so that users can move back and forth between different online data sources. In the past, crosswalks between different thesauri have usually been developed manually. In the long run the intellectual updating of such crosswalks is expensive. An obvious solution would be to apply automatic matching procedures, such as the so-called ontology matching tools. On the basis of computer-generated correspondences between the Thesaurus for the Social Sciences (TSS) and the Thesaurus for Economics (STW), our contribution explores the trade-off between IT-assisted tools and procedures on the one hand and external quality evaluation by domain experts on the other hand. This paper presents techniques for semi-automatic development and maintenance of vocabulary crosswalks. The performance of multiple matching tools was first evaluated against a reference set of correct mappings, then the tools were used to generate new mappings. It was concluded that the ontology matching tools can be used effectively to speed up the work of domain experts. By optimizing the workflow, the method promises to facilitate sustained updating of high-quality vocabulary crosswalks.

Received 31 July 2013; Revised 13 September 2013; Accepted 13 September 2013

Keywords: mapping, KOS, semi-automatic matching, vocabulary crosswalks

1.0 Introduction

For good reason, crosswalks between two or more vocabularies, also known as terminology mappings, play an important role in today´s information landscape. First and foremost, they are an essential means to achieve interoperability among different knowledge organization systems, thus overcoming problems of semantic heterogeneity. Even where a certain number of discrepancies between different terminologies prove insurmountable, crosswalks implemented in a distributed search scenario can enable an integrated search across varied information collections indexed using different subject metadata systems. In addition, alignments between different controlled vocabularies serve as a useful tool for vocabulary expansion. This is especially helpful in overcoming differences in the terminologies used in different subject disciplines. Beyond that, semantic mappings between different vocabularies can be useful for query expansion and reformulation. Automatic translation of a query into the appropriate search terms of all the different vocabularies in use enables a searcher to apply only the terminology with which he or she is familiar, while moving between different resources and databases in a collection.

Cross-concordances between controlled vocabularies usually involve three basic mapping types: equivalence, hierarchical and associative. Equivalence can be exact (between synonyms) or inexact (between quasi-synonyms). Hierarchical mappings, either broader or narrower, apply in one or the other direction between broader and narrower terms in the respective vocabularies. Associative mappings link related terms. A "null relation" describes the case where no appropriate mapping can be established for a given term. Cross-concordances are established bilaterally, i.e. cross-concordances are created from vocabulary A to vocabulary B as well as from vocabulary B to vocabulary A, and these bilateral relations are not necessarily symmetrical. Additionally, one term of vocabulary A can be mapped to a combination of terms of vocabulary B or independently to several terms of vocabulary B; both cases are known as one-to-n (1:n) term mappings.

The intellectual mapping of vocabularies done by domain experts includes a number of working steps. The first is an overall analysis of structure and topical overlap, to determine whether an alignment is possible and reasonable at all. According to Mayr and Petras (2008, 5):

> Essential for a successful mapping is an understanding of the meaning and semantics of the terms and the internal relations of the concerned vocabularies. This includes syntactic checks of word stems but also semantic knowledge to look up synonyms and other related terms.

Then, the mapping process starts. For each concept, any scope note and all its internal relationships need to be taken into account. In order to achieve overall consistency it is occasionally necessary to revise mappings already created. Finally, mappings between different vocabularies usually include retrieval tests for document recall and precision to evaluate whether the translation of search terms of one vocabulary into those of another vocabulary indeed facilitates the search across different databases and terminologies. For example, queries are translated into search terms of a controlled vocabulary A and used for keyword search in a bibliographic database which uses another controlled vocabulary B. Retrieval results can be compared by repeating this search using cross-concordances between both vocabularies which translate the original controlled vocabulary search terms into the controlled vocabulary terms of the target database.

The need for expertise and for constant consideration of the whole semantic environment of each term, make vocabulary mapping expensive and extremely time-consuming. Against this backdrop, this article seeks to examine to what extent semi-automatic matching procedures can be used to prepare vocabulary crosswalks. The results of the 2012 Ontology Alignment Evaluation Initiative (OAEI) provided basic background as to the ontology-matching approaches available. Comparing technical and intellectual evaluation results of OAEI`s most recent "Library Track" we suggest a semi-automatic method to make the intellectual evaluation of automatically-generated vocabulary crosswalks more efficient.

2.0 Related Work

Building up correspondences between vocabularies has been a crucial topic for years in library and information science. For this reason several terminology mapping projects have already addressed the issue of manual versus automatic generation of crosswalks between heterogeneous vocabularies.

A first major terminology mapping initiative was the project Multilingual ACcess to Subjects (MACS) carried out by the National Libraries of France, Germany, Switzerland and the United Kingdom. By establishing equivalences between the three indexing languages, RAMEAU for French, *Library of Congress Subject Headings* (*LCSH*) for English, and SWD[1] for German, multilingual subject access to library catalogues was made possible (Landry 2009). This led to the establishment of a link management database to create and manage links in a decentralized environment. The development of a search interface and the future and permanent management of the MACS approach are still under planning and analysis. Terminology mappings have also been created at the OCLC Online Computer Library Center, Inc. (Godby et al. 2004; Vizine-Goetz et al. 2004), where various vocabularies like the *Dewey Decimal Classification* (*DDC*), the Library of Congress *Classification* (LCC), the *Medical Subject Headings* (*MeSH*), and *LCSH* have been taken into account. Further initiatives include the High Level Thesaurus Project (HILT) (Macgregor et al. 2007) and CRISSCROSS (Panzer 2008), as well as several mapping projects from the Food and Agriculture Organization of the United Nations (FAO) (Lauser et al. 2008; Liang and Sini 2006). A manual cross-concordance between the Thesaurus for Economics (STW) and the Thesaurus for the Social Sciences (TSS) was manually created by domain experts in 2006 (Mayr and Petras 2008). All these projects have in common that they did not exploit automatic approaches systematically, due to a lack of generally available and applicable matching systems.

One impediment to the development of matching systems arises from the different formats that are used to represent knowledge organization systems (KOS). With the advent of the semantic web (Berners-Lee et al. 2001), Resource Description Framework (RDF) (Klyne and Carroll 2004), Web Ontology Language (OWL) (McGuinness and van Harmelen 2004), and Simple Knowledge Organization System (SKOS) (Miles and Bechhofer 2009), a technical basis exists that facilitates access to KOS data. Ontology matching, also called ontology alignment, is a related field where correspondences are established between ontologies that are usually represented in OWL. Ontology in this context stands for a special kind of KOS substantially differing from thesauri and classification systems. While thesauri and classifications usually apply a limited number of relationships between concepts or between terms, ontologies potentially apply an unlimited number of predicative term relations (Gietz 2001). Despite these differences between types of KOS's, however, matching approaches are to some extent transferable.

Recently, automatic matching systems have been discussed as a prior step before manual evaluation. Such approaches typically enable user interaction before (To et al. 2009), during, or after the matching process (Duan et al. 2010; Ehrig et al. 2005). The most similar to the evaluation scenario presented in this article are those that enable validation of the detected correspondences after the matching process. While Paulheim et al. (2007) enable a rating of correspondences by the user, the matching process presented by Cruz et al. (2012) and Noy and Musen (2003) is performed iteratively. User feedback on correspondences is brought directly into the subsequent matching tasks. By splitting up the validation process these tools aim to reduce the manual evaluation effort. The main difference is the use case: while these approaches may be used to improve matching results in a variety of settings, we specifically address the task of creating a set of high quality mappings between vocabularies, where automation is used to reduce the manual effort required.

Many matching techniques have already been developed (Kalfoglou and Schorlemmer 2003; Aguirre et al. 2012). Some of them take the names of entities into account while others compute similarities based on the ontology hierarchy. All of them have advantages as well as disadvantages and their individual field of application. Without extensive knowledge about the systems, it is difficult to decide which system should be used for a specific matching task. That is the reason why ontology matching evaluation is needed.

3.0 OAEI library track 2012

One already established evaluation initiative is the Ontology Alignment Evaluation Initiative (OAEI http://oaei.ontologymatching.org), which started in 2004. Spanning various tracks from a wide range of different scientific disciplines, this campaign has as its main goal to improve ontology matching in general, by comparing and evaluating the different matching systems and algorithms. Taking part either in a specific track or in all tracks these matching systems and algorithms are evaluated according to special criteria, for example the time spent to build up a set of mappings. Between 2007 and 2009 the OAEI included a so-called library track, directed towards KOS's specifically applied in libraries (Isaac et al. 2009). Last year the OAEI again offered a library track focused on the automatic matching of different domain-specific thesauri, co-organized by authors of this paper. To make evaluation of the results possible, however, the organizers needed a reference set of mappings.

3.1 Data set

A key enabler for the OAEI library track was the availability of two considerably overlapping domain-specific thesauri, in this case the Thesaurus for the Social Sciences (TSS) and the Thesaurus for Economics (STW). Both thesauri are commonly used for indexing by domain-specific libraries and institutions providing information infrastructure, and so can be regarded as a real world data set.

The Thesaurus for the Social Sciences (TSS) serves as a key indexing language for documents and research information in German language social sciences. Translated into English and French it contains overall about 12,000 keywords, made up of 8,000 standardized subject headings and 4,000 non-descriptors. The thesaurus as a whole covers topics and sub-disciplines of the social sciences. Additionally some general, non-scientific terms and some terms from associated and related disciplines are included, in order to support accurate and precise indexing of documents from a wide inter- and multi-disciplinary background. The thesaurus is owned and maintained by GESIS, Leibniz-Institute for the Social Sciences (http://www.gesis.org/en/home/). Its SKOS version is published under a CC-by-NC-ND licence.

The Thesaurus for Economics (STW) provides a German and English indexing vocabulary for economics containing more than 6,000 standardized subject headings, and 19,000 entry terms. Besides subject headings used in the field of economics it includes juridical, sociological, political and geographical subject headings. The entries are richly interconnected by 16,000 hierarchical and 10,000 associative relations. An additional hierarchy of main categories provides a high level overview. The vocabulary, used for indexing purposes in libraries and economic research institutions, is maintained and further developed on a regular basis by ZBW (German National Library of Economics http://zbw.eu/index-e.html), Leibniz Centre for Economics. It is published under a CC-by-SA-NC license.

During an earlier major terminology mapping initiative conducted by GESIS, Leibniz-Institute for the Social Sciences in 2006, a bilateral reference alignment had been created manually by domain experts (Mayr and Petras 2008) between TSS and STW. It contains about 3,000 exact equivalences, 1,500 narrower and approximately 150 broader term relations in each direction. Since its initial creation in 2006, this reference alignment had not been updated. In recent years, however, the source thesauri have evolved and the changes were not reflected in the reference alignment. For the evaluation exercise, accordingly, an updated alignment would have been useful but in its absence only the established equivalence relations were used for validating the correspondences detected. This need, however, motivated subsequent investigation of whether the results could be used to update the existing alignment.

In view of the large number of concepts, semantic relations and synonyms, the overriding aim of the evaluation was to show whether and to what extent the alignment of the two thesauri could be generated automatically. The question was whether current state-of-the-art matching systems developed for ontologies would be able to deal effectively with thesauri – the so-called "lightweight ontologies" (Uschold and Gruninger 2004) that are widely used in practice.

For the automatic creation of cross-correspondences both thesauri needed to be available in a machine-readable format. Since OWL is used by almost all ontology matching systems, both thesauri had to be converted from their existing SKOS formats into OWL. (General differences between ontologies and thesauri and a detailed description of difficulties including the transformation from SKOS into OWL can be found in Aguirre et al. (2012)).

4.0 Automatic creation of correspondences

For the automatic creation of correspondences all matching systems participating in the OAEI 2012 were applied: AROMA, ASE, AUTOMSv2, CODI, GO2A, GOMMA, Hertuda, HotMatch, LogMapLt, LogMap, MaasMatch, MapSSS, MEDLEY, OMR, Optima, ServOMapL, ServOMap, TOAST, WeSeE, Wmatch and YAM++ (Aguirre et al. 2012). They match the ontologies and generate the resulting alignment by a fully automatic process. Our existing reference alignment made it possible to measure the quality of the alignments created. The results were evaluated by means of precision, recall and F-measure, where precision measures the correctness of the returned correspondences (i.e. the rate of all correct returned correspondences in regard to all returned correspondences), recall the completeness of the correspondences (i.e. the correct returned results in regard to all correct correspondences that should have been returned); F-measure is the harmonic mean of both.

An overview of the results can be found in Table 1 (matchers are sorted in descending order of their F-measure values). Altogether, 13 of the 21 submitted matching systems were able to create an alignment. Three matching systems (MaasMatch, MEDLEY, Wmatch) did not finish within the timeframe of one week while five exited with an error.

This evaluation is based on the original reference alignment. It can safely be assumed that if the reference alignment had been up-to-date, many more correct correspondences would have been identified by each of the matchers. GOMMA performs best in terms of F-measure, closely followed by ServOMapL and LogMap. However, the precision and recall measures vary a lot across the top three systems. The choice of matcher for a given application would depend on whether high precision or high recall is preferred. If the focus is on recall, the alignment created by GOMMA is probably the best choice, with a re-

Matcher	Precision	Recall	F-Measure	Time (s)	Size
GOMMA	0.537	0.906	0.674	804	4712
ServOMapL	0.654	0.687	0.670	45	2938
ServOMap	0.717	0.619	0.665	44	2413
LogMap	0.688	0.644	0.665	95	2620
YAM++	0.595	0.750	0.664	496	3522
LogMapLt	0.577	0.776	0.662	21	3756
Hertuda	0.465	0.925	0.619	14363	5559
WeSeE	0.612	0.607	0.609	144070	2774
HotMatch	0.645	0.575	0.608	14494	2494
CODI	0.434	0.481	0.456	39869	3100
MapSSS	0.520	0.184	0.272	2171	989
AROMA	0.107	0.652	0.184	1096	17001
Optima	0.321	0.072	0.117	37457	624

Table 1. Results of the OAEI Library Track 2012

call of about 90%. Other systems generate alignments with higher precision, e.g. ServOMap with over 70% precision, but most give lower recall values (except for Hertuda).

Concerning the run-time, LogMapLt as well as ServOMap were quite fast with a run-time below 50 seconds. These systems are even faster than a simple Java-program comparing the preferred labels of all terms. Thus, they are very effective in matching large ontologies while achieving very good results. Other matchers take several hours or even days and do not produce better alignments in terms of F-measure.

5.0 Intellectual evaluation of automatically created correspondences

The use of a partial reference alignment to identify a good matcher is interesting, but does not solve the problem of updating and extending the reference alignment in an efficient way. Manually evaluating new correspondences took up to several minutes for each mapping established. Therefore, a good strategy is needed to maximize the number of new correct correspondences while minimizing the tedium of evaluating the matcher results. Unsurprisingly, the matching tools were easily able to detect matches based on the term alone, even in cases of small variations in the character string. For example, useful matches were often found between geographical and ethnographical terms. But the tools were less effective when taking the term's context into account. Incorrect matches were often generated when:

the lexical value of the term was the same but broader and narrower terms showed the underlying concept to be different;

the lexical value of the term was the same but the scope note in one thesaurus indicated an exclusion not valid in the other;

terms in different domains looked similar, but their meanings were different;

the presence of a synonym matching a preferred term in the other thesaurus caused an incorrect equivalence to be generated.

To sum up, the overall intellectual evaluation results of the newly established vocabulary mappings vary greatly between the different matching tools. The number of successfully established equivalence mappings ranged (approximately) between 40 and 270, i.e. between 6% and roughly 54% of the total correct number.

Despite these promising results, it was judged that the alignments obtained were not precise enough for immediate use, since in a live situation every single cross-concordance has to be totally correct. Nevertheless, given the large number of matching systems and their fast, automated execution, they can be used to support domain experts in the creation of cross-concordances. Integrated in a semi-automatic workflow they can serve as a recommender system, showing a domain expert the most probable cross-concordances and hence saving a huge amount of time.

However, the question is how to benefit the most from the cross-concordances prepared automatically? Within an alignment, confidence values assigned to the correspondences by the matching tools indicate how trustworthy a correspondence is. Unfortunately, the confidence values are not comparable between different matchers; in particular they do not indicate how far an alignment is correct. They can only be used to order correspondences within one alignment. Traditional measures like precision, recall and F-measure do not take this ordering into account. Thus, an alignment can have a high F-measure value but if the correct correspondences are listed at the end, this alignment is not the best choice. In this case, an alignment with a low F-measure value but properly assigned confidence values is to be preferred. Thus, the domain expert gets a high amount of correct cross-concordances while verifying as few as possible.

6.0 Improving results with user interaction

Until now, the OAEI tracks have only evaluated fully automated matching systems. Similar to the library track, the results are often good, but for various applications not good enough. In these cases, it is necessary to involve domain experts, either before, during or after the matching process. Before the matching process: the expert can indicate correct and incorrect correspondences. Based on this additional source of information, the system can try to learn the perfect matching strategy. During the matching process: the matching system can ask the expert e.g. to verify or complete correspondences. Using the answer, the system can try again to adapt its strategy. After the matching process: once the alignment has been created, the expert can verify the correspondences in order to improve their quality. In this case the matching system cannot benefit from the results as they are usually not fed back into the system. Since the current state-of-the-art matching systems mostly deal with fully automated matching services, we only verified the alignments after they had been created. If the expert is interactively involved in the whole matching process, the manual effort could be further reduced. Then, of course, other measures are needed to compare the system, e.g. the number of required interactions (Paulheim et al. 2013).

7.0 Optimizing the evaluation process

In the following experiment, we investigated whether the effort of a domain expert during manual evaluation can be reduced and optimized. For our manual evaluation, we studied each alignment in isolation and checked every single correspondence. It goes without saying that this process would be quicker, if each correspondence that occurs in several alignments can be presented for checking only once. Another idea is to exploit the large number of alignments generated by the matching systems. The underlying assumption of this approach is that the more matching systems have found a certain correspondence, the more likely it seems to be correct. Additionally, we investigated whether a reorganization of the results presented for manual evaluation had an impact on the time spent by domain experts. We tested this assumption on the results of the OAEI library track 2012.

This experiment addressed the order and the number of detected correspondences the domain expert had to consider. Any duplicate correspondences (i.e. correspondences generated by more than one matcher) were removed. After de-duplication, the correspondences were grouped according to the number of matchers detecting them. This resulted in a group containing correspondences that were found by all thirteen matching systems, a group with correspondences found by twelve matchers and so on. The last group contained correspondences found by only one matcher.

In the experiment, the groups were presented to the domain expert for evaluation in descending order, i.e. the expert began with the group of correspondences found by all matching systems. From the total numbers of correspondences and of those which turned out to be correct, we can observe the rate of finding correct correspondences and compare that with the rate when no reordering of the results was done. In other words, calculation shows how many correct correspondences would be found after evaluating the same number of correspondences as before.

In Table 2, the results of the manual evaluation are summarized. For our experiment only the de-duplicated correspondences were considered.

	All correspondences (including duplicates)	De-duplicated correspondences
Total number	55466	22592
of which are correct	21541	2484 (11%)

Table 2. Number of correspondences: total; de-duplicated and correct

In Figure 1, we illustrate the percentage of correct correspondences (y-axis) found by a certain number of matching systems (x-axis). For example, x=9 means that these correspondences were identified by 9 matching systems, no matter which particular 9 systems found them. Above the graph, the total number of detected correspondences for x systems is indicated (71). Altogether, 71 correspondences were found by all matching systems, from which ~99% proved correct. Of the correspondences found by 12 matching systems (209), about 93% were found to be correct. The graph clearly shows a correlation between the number of matchers to identify a given correspondence, and the likelihood of its being correct.

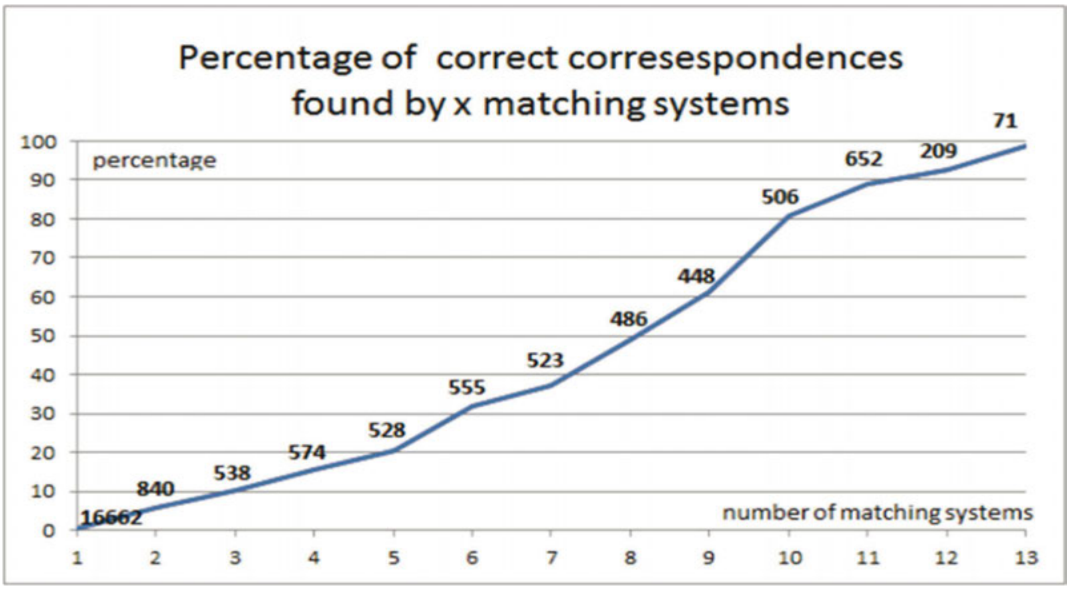

Figure 1. Percentage of correct correspondences found by x matching systems

Number of corresponding matchers	Number of all correspondences	Percentage of correct correspondences	Number of correct correspondences
1	16662	0.27007562	50
2	840	5.71428571	48
3	538	10.4089219	56
4	574	15.6794425	90
5	528	20.4545455	108
6	555	31.8918919	177
7	523	37.0936902	194
8	486	48.8659794	238
9	448	61.3839286	275
10	506	80.8300395	409
11	652	89.1104294	581
12	209	92.8229665	194
13	71	98.5915493	70

Table 3. Results of the "majority vote"

Table 3 shows the number of all correspondences and the numbers of all *correct* correspondences, grouped by the number of matchers that found these correspondences. For example, 506 correspondences were found by ten matching systems, and 409 of them (80% approximately) were correct.

These numbers confirm our assumption that the more matching systems have found a certain correspondence, the more likely it is to be correct. This "majority vote" method has already emerged as a promising technique, e.g. for combining different ontology matching systems (Eckert et al. 2009).

Regarding the time spent by users during manual evaluation, our results confirm that at least a certain number of correct correspondences can be found relatively quickly by optimizing the sequence of entries in the list of matches (see Table 4). To show the extent of the efficiency gain, the first five columns of Table 4 reverse the sequence of Table 3, beginning with those correspondences that were found by as many matchers as possible. This reveals how many correct correspondences can be found at each stage, if the list is reorganized. Percentages of correct correspondences are also shown for each group of matchers. Finally, in the last two columns of Table 4 we compare these numbers to the numbers when the evaluation is not optimized. The number of corresponding matchers (column 1) was not taken into account. The overall correctness rate of 11 % (see Table 2) was used to estimate the num-

			Optimized scenario		Normal evaluation	
Number of corresponding matchers	Number of all correspondences (corr.)	Percentage of all corr. (22592=100%)	Number of correct corr.	Percentage of all correct corr. (2484=100%)	Number of correct corr. (estimated)	Percentage of all correct corr. (2484=100%)
13	71	0.31%	70	2.82%	8	0.32%
12	280	1.24%	264	10.63%	31	1.25%
11	932	4.13%	845	34.02%	103	4.15%
10	1438	6.37%	1254	50.48%	158	6.36%
9	1886	8.34%	1529	61.55%	207	8.33%
8	2372	10.50%	1767	71.14%	261	10.51%
7	2895	12.81%	1961	78.95%	318	12.80%
6	3450	15.27%	2138	86.1%	380	15.30%
5	3978	17.61%	2246	90.42%	438	17.63%
4	4552	20.15%	2336	94.04%	501	20.17%
3	5090	22.53%	2392	96.30%	560	22.54%
2	5930	26.25%	2440	98.23%	652	26.25%
1	22592	100%	2490	100%	2485	100%

Table 4. Comparison of different evaluation strategies

ber of correct correspondences shown in column 6. This shows the number of correct correspondences that would have been found after checking the same number of candidates as were checked at the corresponding stage of the optimized process.

In summary, a critical mass of correct correspondences can be detected faster by reordering the results for manual evaluation. For example, after having evaluated 1886 correspondences a total of 1529 correct correspondences were found in the optimized scenario (i.e. 61.5 % of all correct correspondences), while only 207 correct correspondences would have been found without optimization (only 8.33 % of all correct correspondences). Nevertheless, if it is necessary to find all correct correspondences, all the results of all matchers must eventually be evaluated.

8.0 Conclusion and outlook

As is already well-known, the intellectual process of developing cross-vocabulary mappings typically requires specialist resources and can be very time-consuming. This is especially true of large-scale thesauri that cover many sub-disciplines. Our study has shown that the use of ontology matching tools can greatly speed up the process, especially if the work is organized in the most time-efficient order. This enables automatic creation of an alignment between different thesauri that are available in machine-readable format.

The most recent OAEI library track has shown significant differences between the performances of various ontology matching tools on offer. Some are rather promising. None of them, however, could alone prepare a high-quality vocabulary crosswalk. As a first conclusion, it was judged that the matching tools could be used in recommender systems. Second, the matches generated by a variety of different tools were combined and presented in the most time-efficient order, so as to speed up the intellectual evaluation of the matches. This proved highly effective.

The immediate outcome has been the development of a semi-automatic matching technique for preparing vocabulary crosswalks. Beyond that, however, more research could usefully be done into the provision of automated support for intellectually verified matching procedures. Knowledge organization systems such as thesauri are built with elaborate semantic content and structures. The challenge of achieving interoperability between them is an intellectual task that cannot easily be emulated by automatic means. That is why further research could usefully study the interplay between process-supporting technical solutions and intellectual demands.

Note

1. Schlagwortnormdatei or Subject Headings Authority File of the German National Library, has subsequently been replaced by the GND (Gemeinsame Normdatei or Universal Authority File).

References

Aguirre, José Luis, Eckert, Kai, Euzenat, Jérôme, Ferrara, Alfio, van Hage, Willem Robert, Hollink, Laura, Meilicke, Christian, Nikolov, Andriy, Ritze, Dominique, Scharffe, François, Shvaiko, Palve, Šváb-Zamazal, Ondřej, Trojahn, Cássia, Jiménez-Ruiz, Ernesto, Cuenca Grau, Bernardo and Zapilko, Benjamin. 2012. Results of the ontology alignment evaluation initiative 2012. In *Proceedings of the 7th International Workshop on Ontology Matching, OM 2012 - Collocated with the 11th International Semantic Web Conference, ISWC 2012*, pp. 73-115.

Berners-Lee, Tim, Hendler, James and Lassila, Ora. 2001. The semantic web. *Scientific American* 284: 34-43.

Cruz, Isabel F., Stroe, Cosmin and Palmonari, Matteo. 2012. Interactive user feedback in ontology matching using signature vectors. In *Proceedings of the 28th International Conference on Ocean, Offshore and Arctic Engineering-2009: presented at [the] 28th International Conference on Ocean, Offshore and Arctic Engineering: May 31-June 5, 2009, Honolulu, Hawaii, USA*, New York: American Society of Mechanical Engineers, pp. 1321-4.

Duan, Songyun, Fokoue, Achille and Srinivas, Kavitha. 2010. One size does not fit all. Customizing ontology alignment using user feedback. In Simperl, Elena, Cimiano, Philipp, Polleres, Axel, Corcho, Oscar and Presutti, Valentina, eds., *The semantic web: Research and applications: 9th Extended Semantic Web Conference, ESWC 2012, Heraklion, Crete, Greece, May 27-31, 2012. Proceedings*. Berlin: Springer, pp. 177-92.

Ehrig, Marc, Staab, Steffen and Sure, York. 2005. Bootstrapping ontology alignment methods with APFEL. In Gil, Yolanda Gil, Motta, Enrico, Benjamins, V. Richard and Musen, Mark, eds., *The Semantic Web - ISWC 2005: 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6 - 10, 2005, Proceedings*. Berlin: Springer, pp. 186-200.

Eckert, Kai, Meilicke, Christian and Stuckenschmidt, Heiner. 2009. Improving ontology matching using meta-level learning. In Aroyo, Lora, Traverso, Paolo, Ciravegna, Fabio, Cimiano, Philipp, Heath, Tom, Hyvonen, Eero, Mizoguchi, Riichiro, Sabou, Marta and Simperl, Elena, eds., *The semantic web: research and applications: 6th European Semantic Web Conference, ESWC 2009, Heraklion, Crete, Greece, May 31-June 4, 2009: proceedings*. Berlin: Springer, pp. 158-72.

Gietz, Peter. 2001. *Expertise über quality controlled subject gateways und fachwissenschaftliche portale in Europa*. Tubingen.

Godby, Carol Jean, Young, Jeffrey A. and Childress, Eric. 2004. A repository of metadata crosswalks. *DLib magazine* 10 no. 12. Available http://www.dlib.org//dlib/december04/godby/12godby.html.

Isaac, Antoine, Wang, Shenghui, Zinn, CLaus, Mattherzing, Henk, van der Meij, Lourens and Schlobach, Stefan. 2009. Evaluating thesaurus alignments for semantic interoperability in the library domain. *IEEE intelligent systems* 24 no. 2: 76-86.

Kalfoglou, Yannis and Schorlemmer, Marco. 2003. Ontology mapping: the state of the art. *The knowledge engineering review* 18: 1-31.

Klyne, Graham and Carroll, Jeremy J.. 2004. Resource description framework (RDF): Concepts and abstract syntax – W3C recommendation. Available http://www.w3.org/TR/rdf-concepts/.

Landry, Patrice. 2009. Multilingualism and subject heading languages: how the MACS project is providing multilingual subject access in Europe. *Catalogue & index: Periodical of CILIP cataloguing & indexing group* 157: 9-11.

Lauser, Boris, Johannsen, Gudrun, Caracciolo, Caterina, van Hage, Willem Robert, Keizer, Johannes and Mayr, Phillip. 2008. Comparing human and automatic thesaurus mapping approaches in the agricultural domain. In Greenberg, Jane and Wolfgang, Klas, eds., *Metadata for semantic and social applications: Proceedings of the International Conference on Dublin Core and Metadata Applications: Berlin, 22-26 September 2008: DC 2008: Berlin, Germany.* Göttingen: Universitätsverlag Göttingen, pp. 43-53.

Liang, A. C. and Sini, M. 2006. Mapping AGROVOC and the Chinese Agricultural Thesaurus: Definitions, tools, procedures. *New review of hypermedia and multimedia* 12: 51-62.

Macgregor, George, Joseph, Anu and Nicholson, Dennis. 2007. A SKOS core approach to implementing an M2M terminology mapping server. In *Proceedings of the International Conference on Semantic Web and Digital Libraries (ISCD-2007), 21-23 February 2007, Bangalore, India.* Bangalore: Documentation Research & Training Centre, Indian Statistical Institute, pp. 109-20. Available http://strathprints.strath.ac.uk/2970/1/strathprints002970.pdf

McGuinness, Deborah L. and van Harmelen, Frank. 2004. *OWL web ontology language – W3C recommendation*. Available http://www.w3.org/TR/owl-features/.

Mayr, Philipp and Petras, Vivien. 2008. Building a terminology network for search: The KoMoHe Project. In Greenberg, Jane and Wolfgang, Klas, eds., *Metadata for semantic and social applications: Proceedings of the International Conference on Dublin Core and Metadata Applications: Berlin, 22-26 September 2008: DC 2008: Berlin, Germany.* Göttingen: Universitätsverlag Göttingen, pp. 177-82.

Miles, Alistair and Bechhofer, Sean. 2009. SKOS simple knowledge organization system reference - W3C recommendation. Available http://www.w3.org/TR/skos-reference/.

Noy, Natalya F. and Musen, Mark A. 2003. The PROMPT suite: Interactive tools for ontology merging and mapping. *International journal of human-computer studies* 59: 983-1024.

Panzer, Michael. 2008. Semantische Integration heterogener und unterschiedlichsprachiger Wissensorganisationssysteme: CrissCross und jenseits. In Mitgutsch, Konstantin, Netscher, Sebastian and Ohly, H Peter, eds., *Kompatibilität, medien und ethik in der wissensorganisation : Proceedings der 10. Tagung der Deutschen Sektion der Internationalen Gesellschaft für Wissensorganisation Wien 3. - 5. Juli 2006*. Würzburg: Ergon, pp. 61-9.

Paulheim, Heiko, Rebstock, Michael and Fengel, Janina. 2007. Context-sensitive referencing for ontology mapping disambiguation. In Bouquet, Paolo, Euzenat, Jérôme, Ghidini, Chiara, McGuinness, Deborah L., Serafini, Luciano, Shvaiko, Pavel and Wache, Holger, eds., *C&O:RR-2007 Contexts and Ontologies: Representation and Reasoning: Proceedings of the International Workshop on Contexts and Ontologies: Representation and Reasoning.* CEUR, pp. 47-56. Available http://ceur-ws.org/Vol-298/paper5.pdf.

Paulheim, Heiko, Hertling, Sven and Ritze, Dominique. 2013. Towards evaluating interactive ontology matching tools. In Cimiano, Philipp, Corcho, Oscar, Presutti, Valentina, Hollink, Laura, Rudolph, Sebastian, eds., *The semantic web: Semantics and big data: 10th International Conference, ESWC 2012, Montpellier, France, May 26-30, 2013: Proceedings*. Heidelberg: Springer, pp. 31-45.

To, Hoai-Viet, Ichise, Ryutaro and Le, Hoai-Bac. 2009. An adaptive machine learning framework with user interaction for ontology matching. In *Proceedings of the International Joint Conferences on Artificial Intelligence, Workshop on Information Integration on the Web*, pp. 35-40.

Uschold, Michael and Gruninger, Michael. 2004. Ontologies and semantics for seamless connectivity. *ACM SIGMOD record* 33 no. 4: 58-64.

Vizine-Goetz, Diane, Hickey, Carol, Houghton, Andrew and Thompsen, Roger. 2004. Vocabulary mapping for terminology services. *Journal of digital information* 4 no. 4. Available http://journals.tdl.org/jodi/index.php/jodi/article/view/114/113.