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

Andreas Oskar Kempf [1], Dominique Ritze [2], Kai Eckert [3], Benjamin Zapilko [1]

[1] GESIS – Leibniz-Institute for the Social Sciences
[2] Mannheim University Library
[3] Mannheim University

Abstract: Crosswalks between different vocabularies are an indispensable prerequisite for integrated and high-quality search scenarios in distributed data environments. Offered through the web and linked with each other they act as a central link so that users could move back and forth between different data sources being online available.
In the past, crosswalks between different thesauri have been primarily developed manually. In the long run the intellectual updating of such crosswalks requires huge personnel expenses. Therefore, an integration of automatic matching procedures, as for example *Ontology Matching Tools,* seems pretty obvious.

On the basis of computer-generated correspondences between the Thesaurus for Economics (STW) and the Thesaurus for the Social Sciences (TheSoz) our contribution will explore cross-border approaches between IT-assisted tools and procedures on the one hand and external quality measurements via domain experts on the other hand. Thus, we will present techniques to semi-automatically perform vocabulary crosswalks. Due to intellectually evaluated results of multiple matching tools in the forerun, quality statements concerning the reliability of further computer-generated crosswalks can be made. This way, the application of various tools and procedures gradually contributes to an increase in quality. Moreover, on the long-term it facilitates a continuous update of high-quality vocabulary crosswalks.

Introduction

For good reason terminology mappings, defined as crosswalks between two or more vocabularies, play an important role in today´s information landscape. First and foremost, they are an essential precondition to achieve interoperability among different knowledge organization systems, this way being a key instrument in the treatment of semantic heterogeneity. Although insurmountable discrepancies between different terminologies need to be accepted, implemented in a distributed search scenario they enable an integrated search in varied information collections indexed on the basis of different subject metadata systems. In addition, alignments between different subject metadata schemes serve as a useful tool for vocabulary expansion. Forming a pool of fixed relations between terms of different vocabularies, mappings especially between subject metadata schemes from different disciplinary background can provide a possible route into various domain-specific languages. Beyond that, semantic mappings between different vocabularies serve as an indispensible instrument for query expansion and reformulation. Automatically translating the query in search terms of all the different vocabularies of the databases integrated within the data collection searchers, using their own vocabulary they are familiar with, could maneuver between different information resources.

Cross-concordances between controlled vocabularies usually consist of equivalence as well as hierarchy and association relations. While relations of equivalence stand for synonymous, respectively, quasi-synonymous relations, hierarchy relations include broader term and narrower term relations. Association relations represent relations between related terms. Beyond that, it might occur that a term can´t be mapped to another term. These, so-called null relations, do not form a type of relation of its own. 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, which does not imply that these bilateral relations are necessarily symmetrical. Additionally, one term of vocabulary A could be mapped to a combination of terms of vocabulary B or it could be mapped to several terms of vocabulary B, in these cases speaking of so-called one-to-n (1:n) term relations.

An intellectual mapping of vocabularies done by domain experts includes a number of working steps which build on one another starting with an overall analysis of the topical overlap as well as the structure of the different vocabularies determining in how far an alignment is possible and reasonable at all. “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 lookup synonyms and other related terms” (Mayr/Petras 2008: 5). Subsequently, starting with the mapping process, all the internal relations, including synonyms, i.e. non-descriptors, within one concept, respectively, between different concepts, need to be taken into account. The same applies to scope notes available. The mapping of terms being semantically linked to one another needs to be consistent. Therefore, occasionally, a revision of mappings already created might be necessary. Finally, mappings between different vocabularies usually include retrieval tests for document recall and precision.

It is for the need of expertise and the constant consideration of the whole semantic environment of each term which make the mapping of different vocabularies extremely resource and especially time consuming.

Against this backdrop, this article seeks to examine in how far semi-automatic matching procedures can be used for building up vocabulary crosswalks. Referring to ontology matching approaches we take results of the last Ontology Alignment Evaluation Initiative (OAEI) as base material. Comparing technical and intellectual evaluation results of OAEI`s most recent so-called Library Track we suggest a semi-automatic evaluation scenario to make the intellectual evaluation process of automatically generated vocabulary crosswalks more efficient.

Related Work

Building up correspondences between vocabularies has been a crucial topic for years in the information and library sciences. It is for this reason, that several terminology mapping projects have already addressed the issue of a manual, respectively, automatic generation of crosswalks between heterogeneous vocabularies so far.

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 national indexing languages RAMEAU for French, LCSH for English, and former SWD for German multilingual subject access to library catalogues was made possible (Landry 2009). Therefore, a link management database was established 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 Online Computer Library Center (OCLC) (Godby 2004, Vizine-Goetz 2004), where various vocabularies like the Dewey Decimal Classification (DDC), the Library of Congress Classification (LCC), the Medical Subject Headings (MeSH), and the Library of Congress Subject Headings (LCSH) have been taken into account. Apart from further initiatives, like the High-Level Thesaurus Project (HILT) (Macgregor 2007) and CRISSCROSS (Panzer 2008), the Food and Agriculture Organization of the United Nations (FAO), too, has been involved in several mapping projects (Lauser 2008, Liang 2006). A manual cross-concordance between the Thesaurus for Economics (STW) and the Thesaurus for the Social Sciences (TheSoz) has been manually created by domain experts in 2006 (Mayr 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 of the main reasons why matching systems are not generally applicable are the different formats that are used to represent knowledge organization systems (KOS). With the advent of the Semantic Web (Berners-Lee 2001), RDF (Klyne 2004), OWL (McGuinness 2004), and SKOS (Miles 2009), a technical basis exists that facilitates access to KOS data. Ontology matching, also called ontology alignment, is a related field where correspondences between ontologies are established 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 are usually characterized by a well-limited amount of conceptual, respectively term relations, ontologies, potentially dispose of an unlimited number of predicative term relations (Gietz 2001). Matching approaches regarding the different types of KOS, however, to some extent are transferable.

Recently, automatic matching systems are discussed as preprocessor for manual evaluation. Involving the user into the matching process such approaches typically allow for user interaction before (To 2009), during or after the matching process (Duan 2010, Ehrig 2005). Similar to the evaluation scenario presented in this article are approaches that enable a validation of the detected correspondences after the matching process. While Paulheim (2007) enables a rating of correspondences by the user, the matching process presented by Cruz (2012) and Noy (2003) is performed iteratively. User feedback on correspondences is directly included 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 generally are used to improve matching results in various settings, we specifically focus on the task to create a manual high quality mapping, where automation is used to reduce the manual effort required.

So far, a large amount of matching techniques has already been developed (Kalfoglou 2003). Some of them take the names of the 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 evaluations have been invented.

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 participating. Taking part either at a specific track or at all tracks these matching systems and algorithms are evaluated according to special criteria, as for example time spent to build up a mapping.

In the year 2007 a so-called Library Track, dedicated to KOS specifically applied in libraries has been introduced in the OAEI conducted until 2009 (Isaac 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. The need for an updated reference alignment to evaluate the different matchers participating led to the idea to use the matching results to maintain the already existing, but outdated, alignment.

Data Set

Central prerequisite for the automatic creation of correspondences in the framework of the OAEI Library track was the disposal of two considerably overlapping domain-specific thesauri, in this case the Thesaurus for the Social Sciences (TheSoz) and the Thesaurus for Economics (STW). Both thesauri for they are commonly used for indexing by domain-specific libraries and infrastructure institutions represented a so-called real world data set.

The Thesaurus for the Social Sciences (TheSoz) serves as key indexing language for documents and research information in the German-language social sciences. Translated into English and French it contains overall about 12,000 keywords, divided into 8,000 standardized subject headings and 4,000 so-called non-descriptors. The thesaurus entirely covers topics and sub-disciplines of the social sciences. Additionally, general, non-scientific terms as well as terms from associated and related disciplines are included in order to support an accurate and precise indexing of documents from a wide inter- and multidisciplinary background. The thesaurus is owned and maintained by GESIS - Leibniz Institute for the Social Sciences[1]. Its SKOS version is published under a CC-by-NC-ND license.

The Thesaurus for Economics (STW) provides a German and English indexing vocabulary for Economics containing more than 6,000 standardized subject headings (skos:Concepts), and 19,000 so-called entry terms (skos:altLabels). Besides terms used in the

[1] http://www.gesis.org/en/home/

field of economics it includes juridical, sociological and political as well as geographical subject headings. The entries are richly interconnected by 16,000 skos:broader/narrower and 10,000 skos:related 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 - Leibniz Centre for Economics[2]. 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 2008). The mapping from TheSoz to STW covers about 3,000 exact, 1,500 narrower as well as approximately 150 broader term relations. Since its initial creation in 2006, this reference alignment has not been updated. However, during the last years, the thesauri have been further developed. Thus, all these updates were not covered by the reference alignment. For the evaluation only the established equivalence relations were considered for validating the detected correspondences.

Taking the large amount of concepts as well as semantic relations and additional synonyms into account the overriding target of the evaluation was to show whether and to what extent the alignment of both thesauri could be generated automatically. The question was in how far current state-of-the-art matching systems were able to deal with these so-called lightweight ontologies (Uschold 2004) widely used in practice.

For the automatic creation of cross-correspondences both thesauri needed to be available in a machine-readable format. Since ontology matching systems are nearly exclusively specialized in matching OWL ontologies both thesauri, already available in SKOS, had to be transformed 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 Aguirre2012).

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 2012). They fully automatically match the ontologies and generate the resulting alignment. Based on the reference alignment the quality of the created alignments could be identified. The results are evaluated by means of precision, recall and F-measure, where precision measures the correctness, recall the completeness of the answers; 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 time frame of one week while five threw an exception.

[2] http://zbw.eu/index-e.html

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

This evaluation is based on the original reference alignment. It is for this reason it can be assumed that there are even more correct correspondences than identified by the matchers based on the outdated and incomplete reference alignment. 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. Depending on the application, an alignment either achieving high precision or recall is to be prefered. If the focus is on recall, the alignment created by GOMMA is probably the best choice with a recall about 90%. Other systems generate alignments with higher precision, e.g. ServOMap with over 70% precision, while mostly having significantly lower recall values (except for Hertuda).

Concerning the runtime, LogMapLt as well as ServOMap are quite fast with a runtime below 50 seconds. These systems are even faster than a simple Java-programm 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.

Intellectual Evaluation of Automatically Created Correspondences

Identifying a good matcher based on F-measure results on a partial reference alignment 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 relation established. Therefore, a good strategy is needed to get the most new correct correspondences out of the tedious work to evaluate the matcher results. The quality of the matching results differs: it turned out, that for terms being identical on the string base the tested matching tools achieved rather good results (see Levenshtein 1966 and the concept of Levenshtein-Distance). In cases of the same scope, these term relations (oftentimes between geographical and ethnographical terms), indeed, could be evaluated as equivalence relations. Due to additional information even in cases in which character stings are not totally identical, equivalence relations could be successfully developed. However, regarding context recognition of a term, there emerged several difficulties most matching tools could not successfully overcome. Especially, in cases where terms are identical regarding their character string but due to specified broader and narrower terms are characterized by different scopes matching tools oftentimes wrongly define them as equivalence relations. A similar problem arises in cases where scope notes exclude a certain meaning of a term included in the other thesaurus. Here again term relations are wrongly defined as equivalence relations. Additionally, matching tools are blind to different domain-specific meanings of terms which again, regarding their character string, look quite similar. Finally, the same applies for a correct context processing of indicated synonyms, i. e. non-descriptors. Again, similar synonyms, regarding their string of characters, could lead to an incorrect definition of equivalence relations.

To sum up, the overall intellecutal evaluation results of the newly established vocabulary mappings vary greatly between the different matching tools. Figures of successfully built-up equivalence relations range between roughly 40 and 270, respectively, six and roughly 54 %.

Altogether, the tested automatic ontology matching systems achieved quite good results, although they had to deal with several specific characteristics of this kind of thesaurus-generated ontologies. However, it turned out that the alignments obtained are not precise enough to directly use them as cross-concordances since every single cross-concordance has to be totally correct. Nevertheless, especially based on the amount of existing matching systems and their fast, automated execution, automatically established cross-concordances can be used to support domain experts in the creation of cross-concordances. Integrated in a semi-automatic workflow they could serve as a recommender system, showing a domain expert the most probable cross-concordances this way saving domain experts a huge amount of time otherwise necessary when starting from scratch.

However, the question is how to benefit the most from the cross-concordances built up automatically? Within an alignment, confidence values assigned to the correspondences indicate how trustworthy a correspondence is. With these confidence values, we can order correspondences within the 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 prefered. Thus, the domain expert gets a high amount of correct cross-concordances while verifying as few as possible.

Improving Results with User Interaction

By now, the OAEI tracks only evaluate 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 already been created, the expert can verify the correspondences in order to improve the quality. In this case the matching system cannot benefit from the results as they are usually not fed back into the systems.

Since the current state-of-the-art matching systems mostly focus on fully automated matching services, we only verified the alignments *after* they had been created. If the expert is interactively involved into 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. To set an incentive for matching systems to provide some kind of user interaction, the idea of providing an evaluation for these tools already arose (Paulheim 2013). Whenever reference alignments are available, they can be used as oracles to simulate users. Thus, it is not necessary to involve real domain experts for this kind of evaluation.

Optimizing the Evaluation Process

In the following experiment, we investigated whether the effort of a domain expert for the manual evaluation can be reduced and optimized. For our manual evaluation, we had a look at each alignment built up in isolation and checked every single correspondence. It goes without saying that this could be improved, if every correspondence, occuring in several alignments, is only checked once. Another idea is to exploit the large amount of available 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 investigate whether a reorganization of the evaluation results as input for the manual evaluation has an impact on the time spent by domain experts. We have conducted our experiment on the results of the OAEI Library Track 2012 to find out whether this assumption holds.

In this experiment, the order and the amount of detected correspondences the domain expert has to consider are changed. The amount of the correspondences is reduced to the number of unique correspondences, i.e. each correspondence is only considered once, no matter by how many matching systems it has been detected. Then, the correspondences are grouped according to the number of matchers that have detected the particular correspondences. This results in a group which contains correspondences that have been found by all thirteen matching systems, a group with correspondences found by twelve matchers and so on. The last group contains correspondences which have only been found by one matcher.

In the experiment, the groups are presented to the domain expert for evaluation in descending order, i.e. the expert starts evaluating the correspondences of the group with the correspondences found by all matching systems. By cumulating the numbers of correspondences and correct correspondences, we can observe the progress of finding correct correspondences with regard to the number of all and of all correct correspondences. Finally, we compare these numbers to those, where no reordering of the results was done, i.e. it is calculated how many correct correspondences are found after evaluating the same number of correspondences as with reordering.

In Table 2, the results of the manual evaluation are summarized. For our experiment we consider only the unique correspondences.

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

Table 2: Number of all, unique and correct correspondences

In Figure 1, we illustrate the percentage of correct correspondences (y-axis) found by a certain amount of matching systems (x-axis). For example, x=9 means that these correspondences are identified by exactly 9 matching systems, no matter which concrete 9 systems found them. Above the graph, the total number of detected correspondences for x systems is indicated. Altogether, 71 correspondences have been found by all matching systems from which ~99% are indeed correct. Having a look at the correspondences found by 12 matching systems (209 ones), about 93% are correct. Continuing this series, the less matchers find a correspondence, the less likely this correspondence is correct.

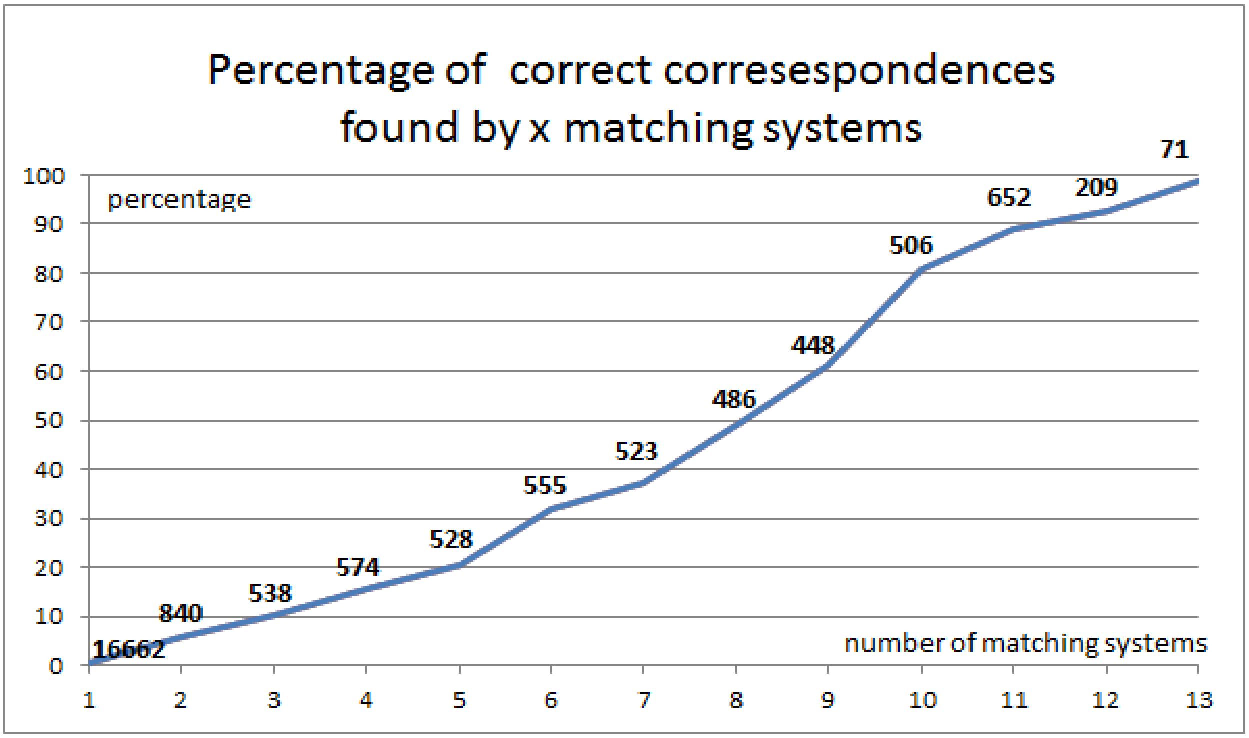

Figure 1: Percentage of correct correspondences found by *x* matching systems

Table 3 shows the number of all correspondences and the numbers of all correct correspondences grouped by the number of matchers that have found these correspondences, i.e. 506 correspondences have been found by ten matching systems. Of these correspondences are about 80 % correct, which makes 409 correct correspondences found by ten matchers.

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

By exploiting these numbers, we can verify our assumption that the more matching systems have found a certain correspondence, the more likely it seems to be correct. Using majority votes have already been proven as promising techniques, e.g. for combining different ontology matching systems (Eckert 2009).

Regarding the time effort spent by users during the manual evaluation, the numbers indicate that at least a certain amount of correct correspondences can be found relatively fast when reorganizing the results. In order to compare the progress of detecting correct correspondences in both scenarios, we have cumulated the values of Table 3, beginning with those correspondences that were found by as many matchers as possible. This way we can observe how many correct correspondences can be found by reordering the correspondences. The numbers for correct correspondences are cumulated from Table 3 for each group of matchers. Finally, we compare these numbers to the numbers when the evaluation is not optimized. Considering the general correctness rate of 11 % (see Table 2), the number of correct correspondences for this case is estimated at 11 % of all evaluated correspondences. The results are shown in Table 4.

			optimized scenario	optimized scenario	normal evaluation	normal evaluation
Number of corresponding matchers	Number of all correspondences	Percentage of all correspondences (22592=100%)	Number of correct correspondences	Percentage of all correct correspondences (2484=100%)	Number of correct correspondences (estimated)	Percentage of all correct correspondences (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

We observed that a critical mass of correct correspondences can be detected faster when reordering the results for manual evaluation. For example, after having evaluated 1886 correspondences a total of 1529 correct correspondences have been found in the optimized scenario (i.e. 61,5 % of all correct correspondences), while only 207 correct correspondences have been found without optimization (only 8,33 % of all correct correspondences). Nevertheless, assuming that all correct correspondences should be found it is necessary that the results of all matchers are evaluated.

Conclusion and Outlook

Against the backdrop of the various working steps depicted for building up vocabulary mappings intellectually manual maintenance of vocabulary crosswalks seems rather resource and especially time-consuming. This is especially the case for large-scale thesauri including a wide range of different sub-disciplines. As turned out high-quality mapping procedures carried out intellectually could be organized more effectively referring to ontology matching instruments. This way, an alignment of different thesauri being available in a machine-readable format can be created automatically.

On the basis of evaluation results of the most recent OAEI´s Library Track it became clear that next to significant differences between the various ontology matching tools some tools provided rather promising performances. Likewise, however, it was evident that none of the different matching tools alone could ensure high-quality standards for building up vocabulary crosswalks. As an intermediary conclusion matching tools were suggested as recommender systems.

In what followed a combination of results of the different matching tools was tested in order to optimize the evaluation process. It turned out that by considering the number of accordances between the different matching tools, starting with those crosswalks detected by most of the matching tools the intellectual evaluation of vocabulary crosswalks built up automatically could be made much more time-efficient.

Using this semi-automatic matching technique for building up vocabulary crosswalks as an example it becomes obvious that more research is needed dealing with the interaction between automatically driven and intellectually done matching procedures. Increasing interoperability between different knowledge organization systems, like thesauri, is a domain which for its specific semantic structure and content is far from being imitated by automated procedures. It is for this reason, further research in this field should focus on the interplay between process-supporting technical solutions and intellectual demands.

References

Aguirre, J. , Eckert, K., Euzenat, J., Ferrara, A., Hage, W. R. van, Hollink, L., Meilicke, C., Nikolov, A., Ritze, D., Scharffe, F., Shvaiko, P., Šváb-Zamazal, O., Trojahn, C., Jiménez-Ruiz, E., Cuenca Grau and B., Zapilko. 2012. Results of the ontology alignment evaluation initiative 2012. In *Proceedings of the 7th International Workshop on Ontology Matching at ISWC'12*, 73-115.

Berners-Lee, T., Hendler and J., Lassila, O. 2001. The Semantic Web. *Scientific American* 284, 34-43.

Cruz, I. F., Stroe, C. and Palmonari, M. 2012. Interactive User Feedback in Ontology Matching Using Signature Vectors. In *Proceedings of the 28th International Conference on Data Engineering*, 1321–1324.

Duan, S., Fokoue, A. and Srinivas, K. 2010. One Size Does Not Fit All. Customizing Ontology Alignment Using User Feedback. In *Proceedings of the 9th International Semantic Web Conference*, 177–192.

Ehrig, M., Staab, S. and Sure, Y. 2005. Bootstrapping ontology alignment methods with APFEL. In *Proceedings of the 4th International Semantic Web Conference*, 186-200.

Eckert, K., Meilicke, C. and Stuckenschmidt, H. 2009. Improving Ontology Matching using Meta-level Learning. In *Proceedings of the 6th European Semantic Web Conference (ESWC)*, 158-172.

Gietz, P. 2001. *Expertise über Quality Controlled Subject Gateways und fachwissenschaftliche Portale in Europa*.

Godby, C. J., Young J. A. and Childress, E. 2004. A Repository of Metadata Crosswalks, D-*Lib Magazine* 10(12). Available http://www.dlib.org//dlib/december04/godby/12godby.html

Isaac, A., Wang, S, Zinn, C., Mattherzing, H., van der Meij, L. and Schlobach, S. 2009. Evaluating Thesaurus Alignments for Semantic Interoperability in the Library Domain. *IEEE Intelligent Systems*, 24(2), 76-86.

Kalfoglou, Y. and Schorlemmer, M. 2003. Ontology mapping: the state of the art. *The Knowledge*

Engineering Review, 18(1), 1-31.

Klyne, F. and Carroll, J. 2004. Resource Description Framework (RDF): Concepts and Abstract Syntax - W3C Recommendation. Available: http://www.w3.org/TR/rdfconcepts/

Landry, P. 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*.

Lauser, B. Johannsen, G., Caracciolo, C., Keizer, J., van Hage, W. R. and Mayr, P. 2008. Comparing human and automatic thesaurus mapping approaches in the agricultural domain. In *Proceedings of the 8th International Conference on Dublin Core and Metadata Applications*, 43-53. Available http://edoc.hu-berlin.de/docviews/abstract.php?lang=ger&id=29135

Liang, A. C. and Sini, M. 2006. Mapping AGROVOC and the Chinese Agricultural Thesaurus: Definitions, tools, procedures. *New Review of Hypermedia and Multimedia*, 12 (1), 51-62.

Macgregor, G, Joseph, A. and Nicholson, D. 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)*, 109-120. Available http://strathprints.strath.ac.uk/2970/1/strathprints002970.pdf

McGuinness, D. and van Harmelen, F. 2004. OWL Web Ontology Language - W3C Recommendation. Available: http://www.w3.org/TR/owl-features/

Mayr, P. and Petras, V. 2008. Building a Terminology Network for Search: The KoMoHe Project. In *Proceedings of the 2008 International Conference on Dubline Core and Metadata Applications (DC)*, 177-182.

Miles, A. and Bechhofer, S. 2009. SKOS Simple Knowledge Organization System Reference - W3C Recommendation. Available: http://www.w3.org/TR/skos-reference/

Noy, N. F. and Musen, M. A. 2003. The PROMPT suite: interactive tools for ontology merging and mapping. *International Journal of Human-Computer Studies*, 59(6), 983–1024.

Panzer, M. 2008. Semantische Integration heterogener und unterschiedlichsprachiger Wissensorganisationssysteme: CrissCross und jenseits. *Fortschritte in der Wissensorganisation*, Band 10. Kompatibilität, Medien und Ethik in der Wissensorganisation, 61-69.

Paulheim, H., Rebstock, M. and Fengel, J. 2007. Context-Sensitive Referencing for Ontology Mapping Disambiguation. In *Proceedings of the 2007 Workshop on Context and Ontologies Representation and Reasoning*, 47–56.

Paulheim, H., Hertling, S. and Ritze, D. 2013. Towards Evaluating Interactive Ontology Matching Tools. *In Proceedings of the 10th Extended Semantic Web Conference*, 31-45.

To, H.-V., Ichise, R. and Le, H.-B. 2009. An Adaptive Machine Learning Framework with User Interaction for Ontology Matching. In *Proceedings of the International Joint Conferences on Artifical Intelligence, Workshop on Information Integration on the Web*, 35–40.

Uschold, M. and Gruninger M. 2004. Ontologies and semantics for seamless connectivity. *ACM SIGMOD Record* 33(4), 58-64.

Vizine-Goetz, D., Hickey, C., Houghton, A. and Thompsen, R. 2004. Vocabulary Mapping for Terminology Services. *Journal of Digital Information* 4(4), Article no. 272. Available http://journals.tdl.org/jodi/index.php/jodi/article/view/114/113