

Consistency and Interoperability in a National Term Bank

Madsen, Bodil Nistrup; Erdman Thomsen, Hanne; Lassen, Tine; Pram Nielsen, Louise; Odgaard , Anna Elisabeth ; Hoffmann, Pia Lyngby

Document Version

Final published version

Published in:

Proceedings of the 10th Terminology and Knowledge Engineering Conference

Publication date:

2012

License

CC BY-NC-ND

Citation for published version (APA):

Madsen, B. N., Erdman Thomsen, H., Lassen, T., Pram Nielsen, L., Odgaard , A. E., & Hoffmann, P. L. (2012). Consistency and Interoperability in a National Term Bank. In G. A. de Cea, M. C. Suárez-Figueroa, R. García-Castro, & E. Montiel-Ponsoda (Eds.), *Proceedings of the 10th Terminology and Knowledge Engineering Conference: New Frontiers in the Constructive Symbiosis of Terminology and Knowledge Engineering* (pp. 242-257). Universidad Politécnica de Madrid.

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 26. Apr. 2024



Consistency and interoperability in a national term bank

Bodil Nistrup Madsen, Hanne Erdman Thomsen, Tine Lassen,
Louise Pram Nielsen, Anna Elisabeth Odgaard and Pia Lyngby Hoffmann

Dept. of International Business Communication
Copenhagen Business School,
Copenhagen, Denmark
{bnm, het, tla, lpn, aeo, plh}.ibc@cbs.dk

Abstract. In this paper we will describe some problems related to the definition of a set of data categories as well as to the import and merging of data from various resources. First, we illustrate how organizing a taxonomy of data categories is facilitated by using the principles for creating a terminological ontology (or concept system). Next, we discuss how multiple terminological entries referring to the same concept can be identified with the purpose of merging them.

Keywords: Data categories, data exchange, merging of terminological data.

1 Introduction

Projects that aim at the establishment of national and corporate term banks have to find solutions to many problems, e.g. how to increase the number of terms and at the same time to ensure the quality of the information about the terms in the term bank, how to choose, organize and present data suitable for different user groups and how to ensure interoperability with other term banks. This is also the case for the DanTermBank project, which aims at creating the foundations for establishing a national term and knowledge bank in Denmark, cf. [7], [8], [11].

In order to facilitate definition of a set of data categories, data exchange and interoperability, it is important to be able to describe elements of data collections systematically and unambiguously. This is the reason why metadata registries comprising sets of metadata categories with negotiated definitions and examples exist in many fields.

When defining a set of metadata categories, it is very useful to base it on a kind of systematization, e.g. a taxonomy, specifying main categories, categories and subcategories. Otherwise, one may end up with an incomplete and/or inconsistent set of categories that is very difficult to use and to extend. In order to obtain a well structured taxonomy it should be based on the principles used for elaboration of a terminological ontology, cf. e.g. [10]. In this way the concepts of the domain and their interrelations

are clarified. We illustrate how this is done in the DanTermBank project for some data categories in section 2.

One method of obtaining a sufficient number of terms in a national term bank is to reuse information from other terminology resources. However, this approach will often lead to problems of redundancy, since the term bank will typically contain many entries connected to the same term, but with varying formulation of the definitions and/or different translations. In section 3, we discuss these problems and sketch possible solutions.

2 Data categories

2.1 Introduction to ISO 12620:1999 and ISOcat

For many years, ISO/TC 37/SC3 has been working on defining a set of data categories for terminological resources. In the first standard, [4], terminological data categories were classified in three major groups and ten sub-groups, as shown in Fig. 1. This structure is to some degree based on the structure of data in a terminological entry, and it does not reflect a logical ordering of the data categories, cf. [15] and [9].

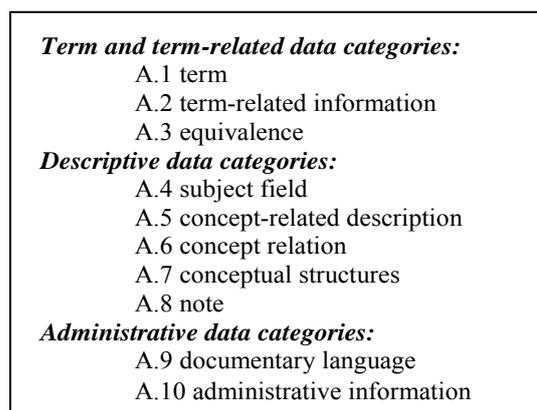


Fig. 1. Main groupings of data categories in ISO 12620:1999

In 2003, it was proposed to set up a Data Category Registry (DCR) in ISO TC 37, cf. [2], [14]. The recent version is found at www.isocat.org/ (referred to as ISOcat, c.f. [3], [6]). The terminology profile of ISOcat comprises 575 entries which are not systematized, but listed in alphabetic order. ISOcat comprises new data categories which are not in [4]. For the data categories from [4] it is possible to see the original systematic number in an explanatory comment, but it is not possible to export data comprising these numbers, and thus not possible to generate a systematic overview of all data categories in ISOcat.

All permissible values for a category are listed under the relevant category without any grouping, and each of these values also exists in separate entries, e.g. the values of the category term type (<http://www.isocat.org/datcat/DC-2677>) c.f. Fig. 2.

It is not possible to see the relation between e.g. *abbreviation* and *short form*, or *part number* and *sku* (a unique alphanumeric designation assigned to an object in an inventory control system). The lack of systematization makes it difficult to ensure completeness, consistency, user-friendliness and extensibility of categories and category values. If, for example, a user wants to introduce a new category, it may be difficult to check whether it is already in the registry, because ISOcat may use another designation for the category than the user. If a user wants to get an overview of all data categories related to for example *Usage*, which is a sub-sub group to *Term-related data categories* in[4], the user will have to go through all 575 data categories and find relevant data categories on the basis of the definitions.

- <i>abbreviation</i>	- <i>part number</i>
- <i>acronym</i>	- <i>phraseological unit</i>
- <i>clipped term</i>	- <i>transcribed form</i>
- <i>common name</i>	- <i>transliterated form</i>
- <i>entry term</i>	- <i>short form</i>
- <i>equation</i>	- <i>shortcut</i>
- <i>formula</i>	- <i>sku</i>
- <i>full form</i>	- <i>standard text</i>
- <i>initialism</i>	- <i>string</i>
- <i>internationalism</i>	- <i>symbol</i>
- <i>international scientific term</i>	- <i>synonym</i>
- <i>logical expression</i>	- <i>synonymous phrase</i>
	- <i>variant</i>

Fig. 2. Term types in ISO 12620:1999

2.2 Data categories for a national terminology and knowledge bank

In the DanTermBank project, we want the data categories of the term bank to be compatible with the ISOcat categories, to enable import from and export to other data collections by using the standard for TermBase eXchange, TBX, [5]. We will only use a subset of the ISOcat categories in the user interfaces for the national term bank, but allow for import of all ISOcat categories in the structure of the database.

In order to facilitate our work on the data categories, and the description to future users, we are developing a new logical structure in the form of a systematic taxonomy. We started by mapping all ISOcat data categories to the grouping of[4], which is much better than no structure. In order to get a better overview, we have entered all ISOcat categories into diagrams using a concept modeling tool.

In ISOcat, we found many examples of double entries with different definitions for a data category, which complicated the mapping of the categories of the national term bank to the ISOcat categories, and we still need to check the data category names, which are in some cases changed in ISOcat compared to [4]. For the purpose of the

set of categories for the national term bank, we have introduced new categories as well as excluded ISOcat categories, and introduced a new structure into the set of data categories.

Due to the large amount of categories we have created several sub-diagrams and an overview diagram with a structure corresponding to the original structure in Fig. 1. In the diagrams, it is possible to click individual categories in order to see definition, example and administrative information. These diagrams formed the basis for our work on a new structure. In Fig. 3, the new overall structure of the data categories for the DanTermBank is presented.

As already mentioned the original structure of [4], which is presented in Fig. 1, reflected the structure of a terminological entry. It is not very easy for a person who is not familiar with term bases to know where in the structure to look for a specific kind of e.g. a type of term-related information. In the DanTermBank project we have reorganized the data categories to obtain a more user-friendly structure.

As basis for the new structure we have used the Danish Standard[1] which comprises a taxonomy for the classification of lexical data, referred to as the STANLEX taxonomy. In the STANLEX taxonomy the main groups of information types are structured according to the linguistic disciplines: *graphical information*, *grammatical information*, *etymological information*, *phonetic information*, *information on usage* and *semantic information*. Furthermore it comprises the groups: *language*, *administrative information*, *structural information* and *note*. Table 1 presents a systematic list of the overall structure of the data categories for the DanTermBank project.

In the new structure it is much easier for a user to find information on the different types of data categories e.g. data categories related to examples of usage, such as the category *linguistic context*. The user does not have to know whether it is a kind of *term-related information* or a kind of *concept-related description*.

Below, we will go more into details on the differences between the original structure of data categories related to *Usage* from ISOcat, structured according to [4], c.f. Fig. 4 and the new structure for the DanTermBank, cf. Fig. 5.

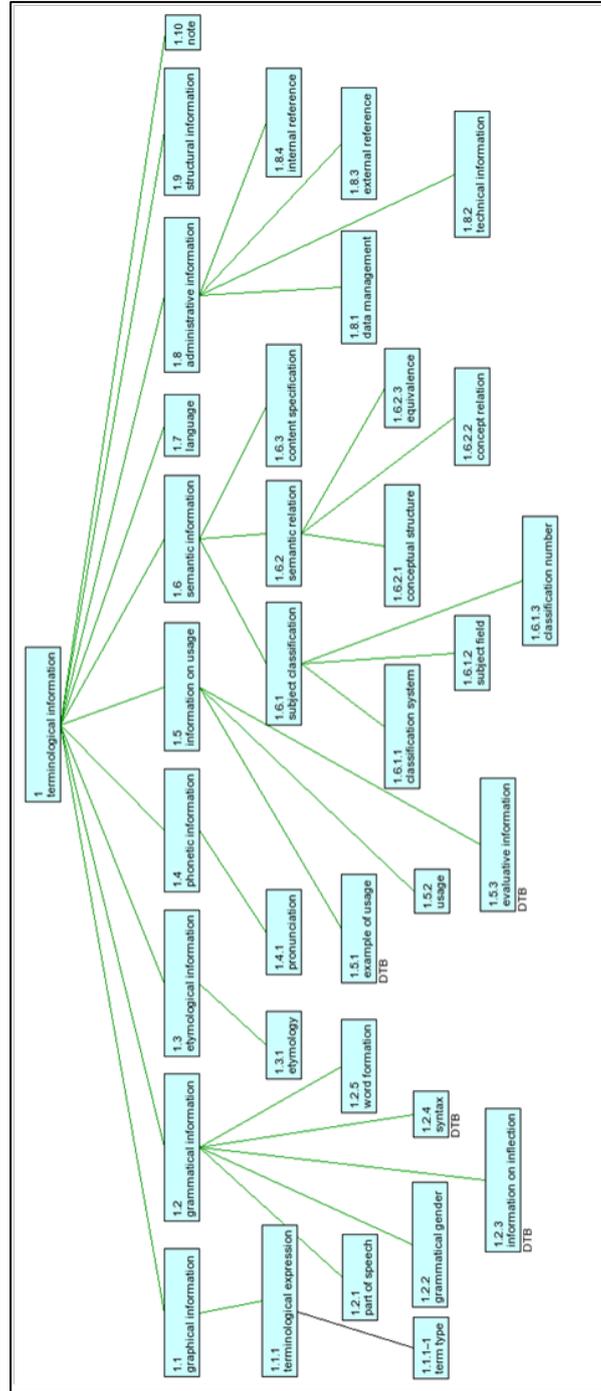


Fig. 3. Overall structure of the data categories for the DanTermBank

Table 1. Systematic list of data categories for the DanTermBank

Notation	DanTermBank category	ISO 12620:1999
1	terminological information	
1.1	graphical information	
1.1.1	terminological expression	A.1
1.1.1--1	term type	A.2.1
1.2	grammatical information	A.2.2
1.2.1	part of speech	A.2.2.1
1.2.2	grammatical gender	A.2.2.2
1.2.3	information on inflection	
1.2.4	syntax	
1.2.5	word formation	A.2.4
1.3	etymological information	
1.3.1	etymology	A.2.4.2
1.4	phonetic information	
1.4.1	pronunciation	A.2.5
1.5	information on usage	
1.5.1	example of usage	
1.5.2	usage	A.2.3
1.5.3	evaluative information	
1.6	semantic information	
1.6.1	subject classification	
1.6.1.1	classification system	A.4.1
1.6.1.2	subject field	A.4
1.6.1.3	classification number	A.4.2
1.6.2	semantic relation	
1.6.2.1	conceptual structure	A.7
1.6.2.2	concept relation	A.6
1.6.2.3	equivalence	A.3
1.6.3	content specification	
1.7	language	
1.8	administrative information	A.10
1.8.1	data management	
1.8.2	technical information	
1.8.3	external reference	
1.8.4	internal reference	
1.9	structural information	
1.10	note	

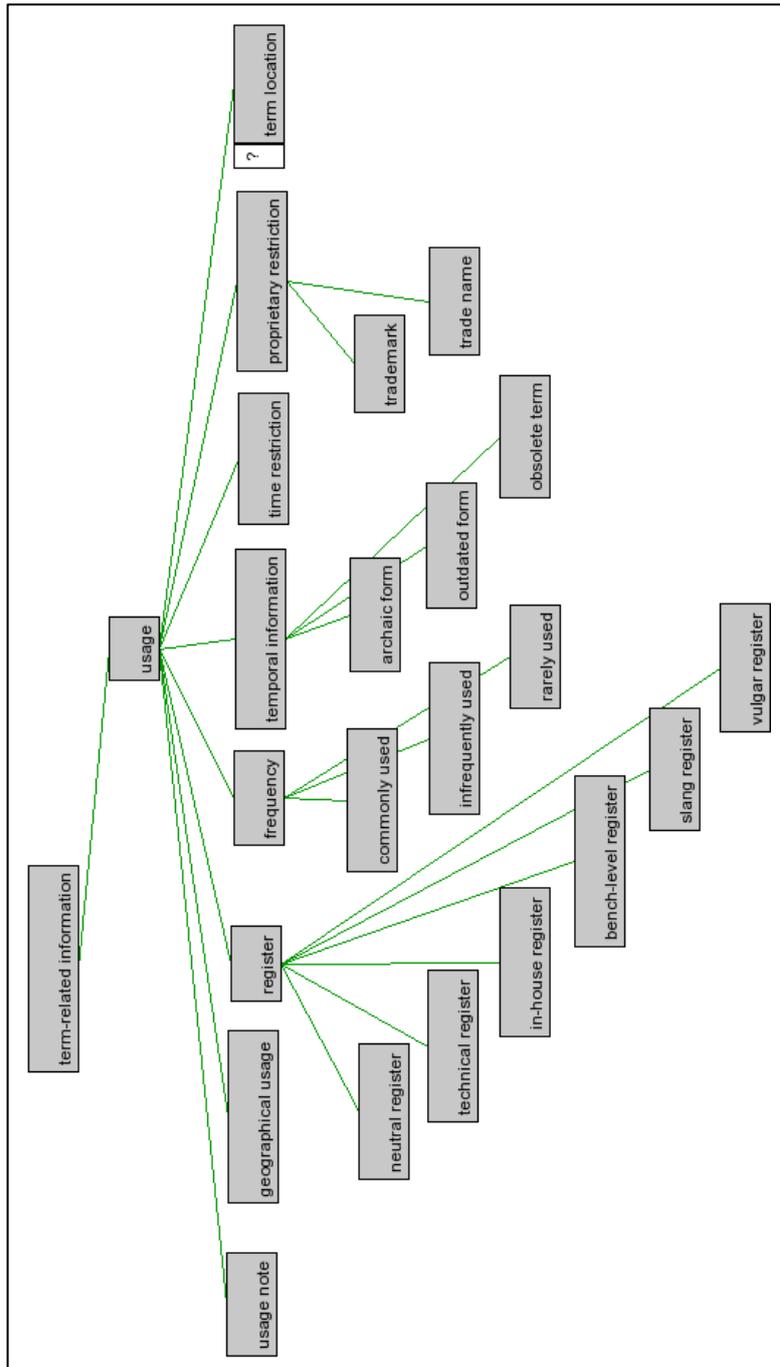


Fig. 4. Data categories related to Usage (ISOcat)

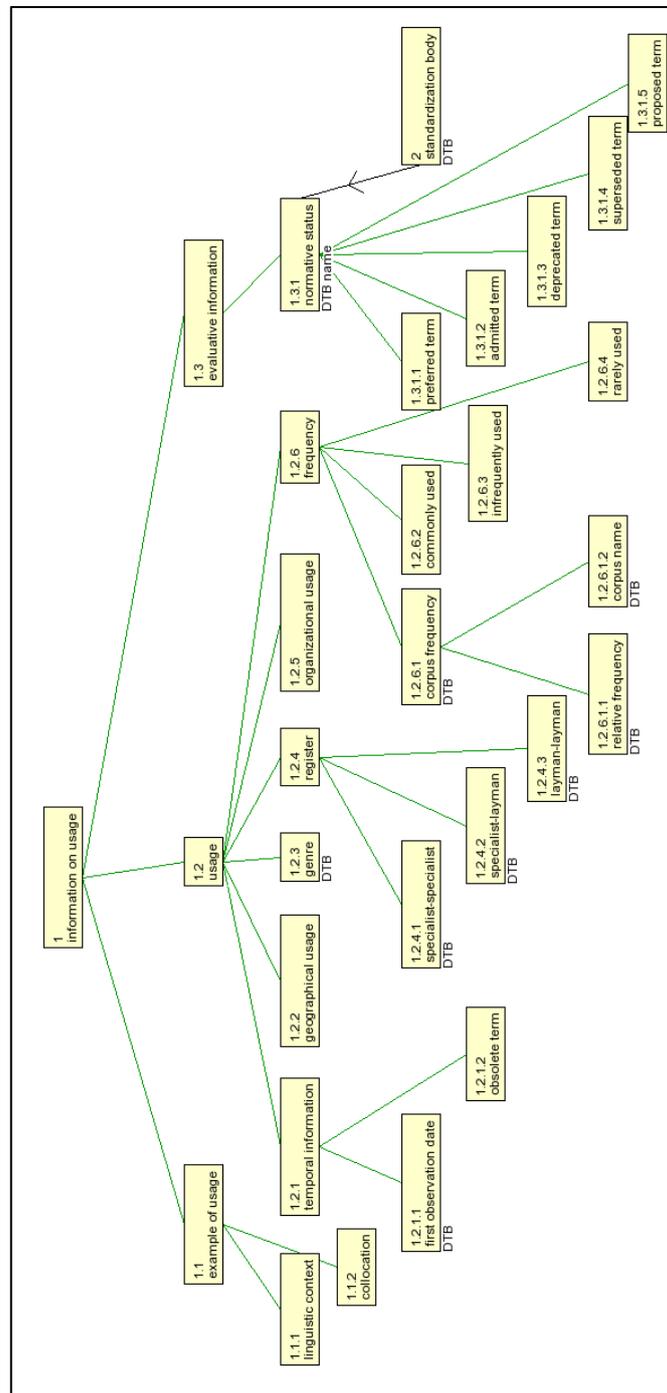


Fig. 5. Data categories related to Usage (DanTermBank)

Fig. 4 shows the data categories related to Usage from ISOcat, structured according to [4].

Fig. 5 shows the structure of data categories related to *Usage*, which has been chosen for the basic set data categories in the DanTermBank project.

As it may be seen from the sub-diagram of *information on usage*, which is presented in Fig. 5, we have introduced a new level with the categories: *example of usage*, *usage* and *evaluative information*, which makes it easier for a user to navigate in the structure. In the original structure collocation is found in group 1, subgroup 2: *Term-related data categories* under *term type*, while *linguistic context* is found in group 2, subgroup 5, *Concept-related description*. In the new structure we have placed both *linguistic context* and *collocation* under *example of usage*.

In deciding upon the new structure we have used the terminological principle of identifying the differentiating characteristics in the subgroups. For example, it is difficult to identify the subtle differences between *archaic form*, *outdated form* and *obsolete term*, cf. Fig. 4, and we doubt that terminologists and end users will be able to differentiate when working with a term base. In the new structure we have simplified this and only introduced one of these three categories, *obsolete term*. To this group we have added the data category *first observation date*, which is a very important type of *temporal information*. The data category *normative status*, placed under *evaluative information*, corresponds to the category *normative authorization* in [4], which is a *term-related category* in the sub-sub-group *term status* (A.2.9.1).

Also we doubt whether users will be able to distinguish between *temporal information* and *time restriction*, so in the new structure we have only kept *temporal information*, c.f. Fig. 5.

3 Merging of data from different sources

3.1 The problem: false doublettes and false equivalents

When integrating terminological data from existing resources, we also want to be able to merge these data properly in the term bank. Entries from different resources are likely to contain information about identical concepts which are associated with different terms and definitions. We will refer to these double entries as ‘false doublettes’. At the same time, there are bound to be entries with the same designation but clearly referring to different concepts. These we refer to as ‘false equivalents’.

False doublettes and false equivalents reduce the usefulness and the quality of the term bank considerably. If a user has to go through a number of entries that differ to some large or small degree as an answer to a given query, the reliability and usefulness of the term bank is reduced. It is not possible for a user to choose the correct term if a query returns e.g. 25 definitions in random order and a closer study reveals that there are in fact only 6 different concepts, which also means that there should only be 6 different definitions and entries. In our opinion, a query in the term bank should result in a clear and precise answer, rather than present the user with a selection of definitions from different sources which the user is left to assess, otherwise

end users may find that searching the term bank is not much better than using the ‘define’ search in google. For this reason, it is very important to merge entries with concepts having the same meaning, and, in our opinion, there is a very close connection between entry merging and quality assurance.

Merging of entries from different sources is a problem which has not been satisfactorily solved by other term banks. The Swedish term bank Rikstermbanken¹ [12], has chosen to present data from all the different sources and let the user evaluate for herself. Likewise in IATE², false doublettes have not been removed. In the version of EuroTermBank³ available in March 2011, automatic entry compounding was carried out, but the result was not always optimal since one entry may comprise definitions of completely different concepts, resulting in a presentation of false equivalents to the user, c.f. [13], [16]. In the current version of EuroTermBank this has changed considerably, we only refer to the older version to illustrate the difficulty of entry merging.

Examples.

In the following we use as an example the search results obtained from the three term banks mentioned above when searching for the term *absorption*. In Rikstermbanken we used Swedish as source language, in IATE and EuroTermBank English. In IATE we got 669 hits, in Rikstermbanken 63 and in EuroTermBank 42.

IATE combines entries from all translation related termbases in EU and contains 1,5 million entries from 5 sources and from a large number of terminology projects. In the search result, each entry is presented with its subject domain and terms, cf. Fig. 6, but no further ordering is done, not even by subject, which makes it close to impossible for the user to find the relevant entries if a very narrow field is not chosen for the search.

The Swedish Rikstermbanken contains ~85.000 entries from 986 different sources of varying size, the largest contributing 4456 entries. Each entry in the search result is presented with Swedish terms, target languages in the entry (if any), the first line of the definition and the source from which the entry was imported. The search for *absorption* results in 63 hits including multiword and composite terms, refining the search to precisely the word *absorption* reduces the result to 31. A closer inspection of the 31 definitions shows that they could have been reduced to just 9 entries. For example, the two circled definitions in Fig. 8 are almost identical: “process by which a physical matter is absorbed into the interior of another physical matter” and “entrance and absorption of one physical matter into the interior of another physical matter”. They are members of a set of 7 very similar definitions .

¹ www.rikstermbanken.se

² www.iate.europa.eu

³ www.eurotermbank.com

absorption

en > en (domain: Any domain, type of search: All)

Result 21-30 of 669 for absorption

Chemistry [Council]	Full entry
EN absorption oil	☆☆☆
Technology and technical regulations, Health [Council]	Full entry
EN absorption peak	☆☆☆
Migration, Cooperation policy, FINANCE [Council]	Full entry
EN absorption capacity	☆☆☆
FINANCE, Financial institutions and credit [Council]	Full entry
EN absorption of liquidity	☆☆☆
Technology and technical regulations [Council]	Full entry
EN light absorption coefficient	☆☆☆
Technology and technical regulations [Council]	Full entry
EN sound absorption coefficient	☆☆☆
Financial institutions and credit, FINANCE [Council]	Full entry
EN structural absorption operation	☆☆☆
Technology and technical regulations [Council]	Full entry
EN radio-frequency absorption material	☆☆☆

Fig. 6. Search result in IATE

absorption (1) sv	DEFINITION: process (1) vid vilken ett ämne (1) upptas i det inre av ett annat ämne ... KÄLLA: Terminologacentrum TNC: Tekniska basord 1995 [Ordlistan är under revidering.]	Visa hela termposten
absorption (2) sv	DEFINITION: process (1) genom vilken ljudenergi minskas genom energiupptagning hos ... KÄLLA: Terminologacentrum TNC: Tekniska basord 1995 [Ordlistan är under revidering.]	Visa hela termposten
absorption (3) sv	DEFINITION: omvandling av strålningsenergi till värme KÄLLA: Terminologacentrum TNC: Tekniska basord 1995 [Ordlistan är under revidering.]	Visa hela termposten
absorption (1) sv	DEFINITION: inträngning och upptagning av ett ämne i det inre av ett annat ämne ÖVRIGA SPRÅK: de, en, fr KÄLLA: Terminologacentrum TNC: Plan- och byggtmer 1994	Visa hela termposten

Fig. 7. Almost identical definitions in Rikstermbanken

EuroTermBank is a multilingual terminology portal which gives access to a central database and 4 external terminology resources. The central database includes ~625.000 entries originating from 101 sources. Our search resulting in 42 hits was carried out in May 2011. At that time, EuroTermBank did merge entries from various sources, thus enabling the user to find translation equivalents between language pairs not originally found together in one source. In Table 2 we present an example of a merging of 4 entries from EuroTermBank. Thanks to the merging, a Russian equivalent can be found for the German term *Resorption*. In addition to identical source terms (English), the merging seems to be based on similarity in EuroVoc subject labels. In this case, however, the fact that the subject of the third entry includes both economy and health, causes entries on ONLY economy (first and second entry) to be merged with an entry on ONLY health. It is therefore very likely that above mentioned equivalence is actually a false one.

Table 2. Merged entries from EuroTermBank (visited May 2011)

EN	DE	LT	LV	RU	Subject (EuroVoc)
absorption			elastības pakāpe	степень эластичности	social affairs; miscellaneous industries
absorption	gewalttätige Hinzufügung		varmācīga pievienošana	поглощение (в экономике)	business and competition; economics; documentation
absorption		sugertis			economic analysis; financing and investment; health
absorption	Resorption				health

As mentioned, this outdated example from EuroTermBank was included to illustrate the difficulties to be faced when merging entries. EuroTermBank has now (April 2012) abandoned this type of entry merging. Instead they display search results in two ways: entries view, which gives the original entries, and translations view, which combines all the entries as illustrated in Fig. 8, allowing the user to unfold the entries for each target term. Thus, the false equivalence link between the German *Resorption* and the Russian terms in Table 2 is avoided.

The screenshot displays the EuroTermBank website interface. At the top, there is a navigation menu with links for Home, Resources, Downloads, News, Help, and About. Below this is a search bar containing the term 'absorption', with dropdown menus for 'From English (en)' and 'To DE LT', and a 'Search' button. To the right of the search bar is a link for 'Show Advanced Options'.

The main content area is divided into several sections:

- Search results:** A list of related terms including 'absorption', 'absorption spectrograph', 'absorption spectrometer', 'acoustic absorption', 'anomalous absorption', 'selective absorption', 'wave absorption', 'dielectric absorption', 'electric absorption', 'systemic absorption', 'absorption study', 'atomic absorption', and 'neutron absorption'.
- Translations View / Entries View:** A table showing translations for 'absorption' from English (EN) to German (DE), Lithuanian (LT), and Latvian (LV).

Language	Term	Subjects
EN	absorption	
DE	Absorption	• wood industry • natural and applied sciences • ...
DE	Aufsaugung	• natural and applied sciences • mechanical engineering
DE	gewalttätige Hinzufügung	• business and competition • economics • ...
DE	Absorbierung	• natural and applied sciences • mechanical engineering
DE	Aufnahme	• natural and applied sciences • mechanical engineering
DE	Resorption	• health
LT	sugėrimas	• natural and applied sciences • mechanical engineering
LT	sugertis	• natural and applied sciences • mechanical engineering
LT	absorbicija	• humanities
LT	sugertis	• economic analysis • financing and investment • ...
LV	absorbicija	• wood industry • natural and applied sciences
- Display options:** A panel with checkboxes for 'show source', 'show domains', and 'show definitions (29)'. All are checked.
- Filter by domain:** A panel with checkboxes for various domains such as 'business and competition', 'chemistry', 'documentation', 'economic analysis', 'economics', 'energy', 'environment', 'environmental policy', 'financing and investment', 'health', 'humanities', 'mechanical engineering', 'miscellaneous industries', 'natural and applied sciences', 'technology and technical regulations', and 'wood industry'. All are checked. At the bottom, there are links for 'Select all / Select none'.

Fig. 8. Current interface for EuroTermBank

3.2 Improved Methods

In order to improve merging of entries from different sources, we propose to refine the method based on designation and subject by including information from other data categories, cf. Fig. 9. If identical terms with the same part of speech come from the same source, there is no reason to believe that they should be merged. Likewise, if they do not belong to the same or at least comparable subject fields, given that subject classifications from different sources have been aligned. Subjects can be said to be comparable e.g. if one is a subcategory of the other one. Finally, the definitions have to be taken into consideration. Only if they are to some extent identical, should the entries be merged. The identity can be measured in terms of identical word forms or just identical or synonymous lexemes, this will have to be decided based on experiments. The same applies to the exact percentage of identity, in the diagram 50% is chosen as an example.

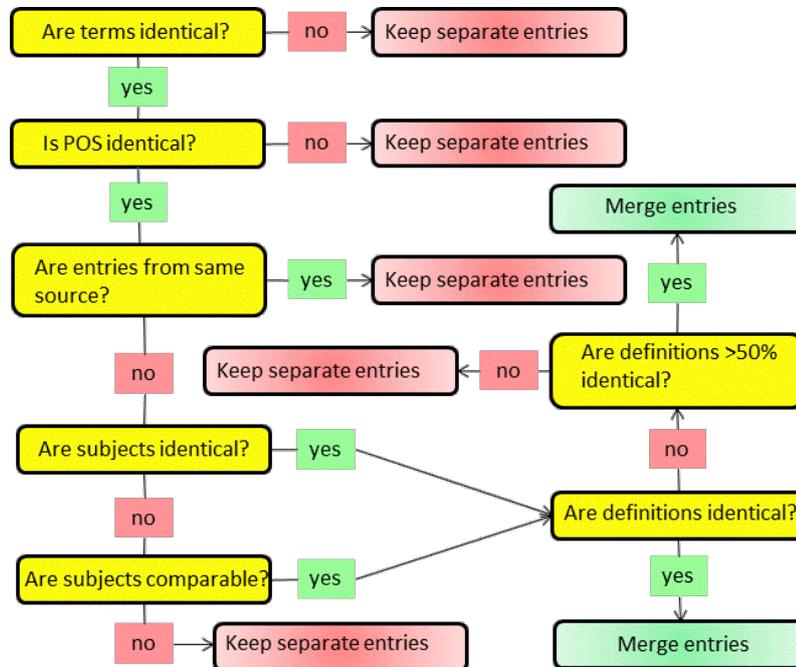


Fig. 9. Decision diagram for merging entries

In the DanTermBank project we are also developing methods for automatic construction of terminological ontologies, i.e. concept systems including delimiting characteristics. Based on the definitions in terminology resources to be included, it will be possible to extract superordinate concepts and delimiting characteristics cf. [7]. When this is implemented, it can replace the use of definition identity in the decision algorithm outlined in Fig. 9.

4 Concluding remarks

In this paper we have argued that when defining a set of metadata one should base it on a taxonomy, which allows users to get an overview and easily identify individual data categories. This is important in order to provide an overview, both if a user wants to suggest new data categories, and if the set of metadata should be used for choosing data categories for a new terminological data collection. Otherwise one may end up with an incomplete and inconsistent set of categories that is very difficult to use and to extend.

We have also argued that the reliability and usefulness of a national term bank depends on the quality of the data. Users will be frustrated, if they have to go through a large number of entries as an answer to a given query, if the entries comprise false doublettes and false equivalents, and we have outlined a method for merging entries.

Acknowledgements

The DanTermBank project, which aims at creating the foundations for establishing a national term bank in Denmark, is funded by the VELUX foundation⁴.

References

1. DS 2394-1:1998, *Leksikalske datasamlinger, Indholds- og strukturbeskrivelse (STANLEX)*. Dansk Standard.
2. Ide, N., Romary, L. (2004). A Registry of Standard Data Categories for Linguistic Annotation. In: *Proceedings of the Fourth Language Resources and Evaluation Conference (LREC)*, Lisbon, pp. 135-39.
3. ISOCat. ISOCat Data Directory registry. <http://www.isocat.org>.
4. ISO 12620:1999, *Computer assisted terminology management — Data Categories*. International Organization for Standardization.
5. ISO 30042: 2008. *Systems to manage terminology, knowledge, and content – TermBase eXchange (TBX)*. International Organization for Standardization.
6. Kemps-Snijders, M., Windhouwer, M., Wittenburg, P., Wright, S.E. (2008). 'A Revised Data Model for the ISO Data Category Registry'. In: *Proceedings from the 8th International Conference on Terminology and Knowledge Engineering, TKE 2008. Managing Ontologies and Lexical Resources*. Editors: Madsen, B.N., Thomsen, H.E. Institut for Internationale Sprogstudier og Vidensteknologi (ISV), 2008, Litera, ISBN: 978-87-91242-50-9, pp. 137-148.
7. Lassen, T.: 'A Corpus Compilation and Processing Prototype for Terminology Work'. In *Proceedings from TKE 2012* (this volume).
8. Lassen, T., Madsen, B.N., Thomsen, H.E.: "Automatic Knowledge Extraction and Knowledge Structuring for a National Term Bank" In: *Proceedings of the Workshop CHAT 2011: Creation, Harmonization and Application of Terminology Resources*. Editors: Tatiana Gornostay and Andrejs Vasiljevs. NEALT Proceedings Series, Vol. 12 (2011), 23-26. © 2011 The editors and contributors. Published by Northern European Association for Language Technology (NEALT) <http://omilia.uio.no/nealt> . Electronically published at Tartu University Library (Estonia) <http://hdl.handle.net/10062/16956> . Pp 23-26.
9. Madsen, B.N., Thomsen, H.E. (2008). 'A Taxonomy of Lexical Metadata Categories'. In: *Proceedings of the 6th International Conference on Language Resources and Evaluation - LREC 2008*. ISBN: 2-9517408-4-0, pp 3402-3408. (Also at <http://www.lrec-conf.org/proceedings/lrec2008/>).
10. Madsen, B.N., Thomsen, H.E., Vikner, C- (2004). 'Principles of a system for terminological concept modelling'. *Proceedings of the 4th International Conference on Language Resources and Evaluation, Vol. I*, pp. 15-18. Lisbon.
11. Nielsen, L.P.: 'Target user oriented knowledge dissemination'. In *Proceedings from TKE 2012* (this volume).

⁴ <http://www.veluxfoundation.dk>

12. Nilsson, H. (2009). 'The realisation of a national term bank – how and why?'. In: *ELETO – 7th Conference "Hellenic Language and Terminology"*, Athens, Greece, 22-24 October 2009. <http://www.eleto.gr/en/papers.htm>.
13. Rirdance, S., Vasiljevs, A. (2006). Towards Consolidation of European Terminology Resources - Experience and Recommendations from EuroTermBank Project. Tilde, Riga, Letland. <http://www.eurotermbank.com/downloads.aspx>.
14. Wright, S.E.. (2004). A Global Data Category Registry for Interoperable Language Resources. In: *Proceedings of the Fourth Language Resources and Evaluation Conference (LREC)*, Lisbon.
15. Wright, S.E.. (2005). 'A Guide to Terminological Data Categories - Extracting the Essentials from the Maze'. In: *Proceedings from the 7th International Conference on Terminology and Knowledge Engineering, TKE 2005. Terminology and Content Development*. Editors: Madsen, Bodil Nistrup & Hanne Erdman Thomsen. Association for Terminology and Knowledge Transfer (GTW) 2005, Litera, ISBN: 87-91242-46-0, pp. 63-75.
16. Vasiljevs, A., Rirdance, S., Liedskalnins, A. (2008). 'EuroTermBank: Towards Greater Interoperability of Dispersed Multilingual Terminology Data'. *Proceedings of the First International Conference on Global Interoperability for Language Resources ICGL 2008* (2008), p 213-220.