Ingetraut Dahlberg
INDEKS GmbH, Frankfurt

# Towards Establishment of Compatibility Between Indexing Languages

Dahlberg, I.: Toward establishment of compatibility between indexing languages
In: Int. Classif. 8 (1981) No. 2, p. 86–91, 31 refs.

Outlines previous work done in the field of compatibility between indexing languages (IL), and describes the scope, limitations and advantages of establishing compatibility between IL. Suggests methods for verbal comparisons between IL as well as generation of an alphabetical comparison matrix M1. Conceptual comparisons, however, demand a conceptual reorganization of M1 into a compability martrix M2 with its two alternatives, namely a system-related matrix M3 and a hierarchical matrix M4. In conclusion, the use of a compatibility matrix and organizational problems are described. (Author)

## 1. Purpose

After an abandonment of universal and special classification systems in most of the documentation and information centers for the classification of joumal literature after World War II, a prolific development of indexing langauges (ILs) (especially in the form of thesauri) has taken place, which are now recognized as a hindrance to the exchange of indexed items and to cooperation between such ILs. In this situation, however, a retum to the UDC for example or to the joint use of the classes of any other existing universal classification system as a basis for cooperation does not seem to be realistic.

In this situation a mechanism must be found to correlate the elements of the different ILs with each other in order to be able to switch between them when searching in databases with their indexations, or when exchanging indexed literature.

The following suggestions are intended to provide methods for the setting up of a compatibility instrument, a "black box" which can be used in order to enter any system of any kind by way of verbal access. This is accomplished by getting from the "black box" any descriptor, class description or notation of any system it handles.

## 2. Previous work

The topic of compatible ILs for the sharing or exchange of infonnation between different information systems (IS) was discussed as early as 1962 (1) and 1965 (2); by 1966 a literature review (3) was already published on these problems. It has received new attention through the UNISIST Report (4) which proposes among its recommendations in Section 6.2.4 (Content Analysis) to establish switching, convertibility and compatibility instruments for broad subject categorization and deep indexing. The Report defines 'compatibility' as

> a quality of systems whose products can be used interchangeably, notwithstanding differences in notation, structure, physical carriers, etc., without any special 'conversion' machinery

and 'conversion' as

> the process of transforming information records, with regard to transcription encoding, data structure, etc., so as to make them interchangeable between two or more services or systems using different conventions and media. (Glossary, p. 147).

Some empirical investigations, such as those of Wellisch (5), Agraev et al (6), Smith (7), Svenonius (8) and Wersig (9) brought insights into the nature of the ILs compared, the methodology for IL comparison, and the structure of IL elements best suited for interchangeability. Most interesting have been the results of the "intermediate lexicon" study by Horsnell (10). The idea of an intermediate lexicon had been developed by J.-C. Gardin and his Groupe d'Etude in Marseilles from 1967 to 1968[1]. Coates (11) who hoped to arrive at a model for comparison of other subject fields, took up this idea in England somewhat later, where consequently an OSTI Project (Office of Scientific and Technical Information) was started, concerned with the creation of an Intermediate Lexicon for the Information Sciences, a little different from the one that had been elaborated in 1968. The results of this new study are reported by Horsnell (10) as follows:

> The greater the relative specificity and vocabulary size of the input language, the better the switching performance. Performance was also good for the switch between the two classification systems used in precoordinate mode. Factors in the input language which adversely affected the performance were low specificity and the frequent use, in thesauri of coordination to represent concepts, especially in limited vocabulary schemes.

Soergel devoted the whole chapter K of his book on "Indexing languages and thesauri" (12) to the problems of compatibility. His solution for the present situation with its abundance of thesauri and continuous creation of new ones, envisages a Universal Source Thesaurus (UST) in computer stored form, in which the elements of all the existing thesauri could be collected, along with their relationship indications. This could then be used as a source of information on existing descriptors and relationships among their concepts, and for the creation of new systems on the basis of the existing concepts. To establish such a UST would, however, be a major undertaking. Its realization could once provide a single compatibility mechanism for all those systems deriving their elements from the common pool.

In the USSR a conference was held in Riga, 6–8 Sept. 1977 on a „Unified System of Information Retrieval Languages" with a number of contributions devoted to the problems of compatibility (13–25). I would like to refer especially to the state-of-the-art report given by Ju. A. Šrejder (26) on "Interaction of classification and subject languages". One of the papers of the Edinburgh FID-Congress of 1978 summarized some of the research having been undertaken in the USSR, namely the one by S.K. Vilenskaya: On the compatibility of different information retrieval languages within the integrated information system (27). It was shown how ILs can be built-up to supplement each other hierarchically on different levels of abstraction and thus to provide compatibility through a common superstructure.

In the Federal Republic of Germany a comprehensive study was undertaken to determine the feasibility of uniting into one system all the thesauri and classification sys-

tems used by the federal ministries and agencies (9). The task, however, appeared too huge and costly, the recommendations for elaboration of a socalled Bundesdachthesaurus were never realized.

A more recent attempt to arrive at a correlation mechanism between the UDC and/or the Broad System of Ordering (BSO) and the MISON Rubricator was discussed at the Prague Meeting of the ECSSID WG 3, Sept. 9–11, 1980. The attempt was based on indexing results in a limited field of study undertaken by the Library of the CSSR Academy of Sciences on the hand, and on comparisons between the ILs themselves, done in Poland, Hungary and the USSR. No special guidelines had been used though in the elaboration of these comparisons.

## 3. Scope, limitations, advantages

The following guidelines are directed towards comparisons and the establishment of compatibility between *existing* ILs. Guidelines for procedures involving comparisons of indexed results without any prior systematic comparisons of the ILs themselves are excluded.

The empirical approach underlying these latter procedures may be of additional interest, once the comparison of the systems has led to certain results; such results, for example, might be supported by a comparison of the indexations via the juxtapositions of a compatibility matrix.

The scope involves any CS and thesaurus. The more complex such systems are, the more complex will be the necessary mechanism. Complexity is determined here by the amount of precombinations of the concepts involved. The more elementary the concepts are in a system, the easier are any systems comparisons to be carried through. The best systems for comparison are therefore the faceted ones. Another comparison problem arises whenever a class is denoted or described in a rather vague manner or when specifications or scope descriptions are missing altogether.

Other limitations are caused by lack of structural components on the part of thesauri, when compared with such classification systems which include summarizing classes of general content or the famous "other" – positions, collecting anything else not covered by any of the previous classes. All such system elements should be left out of the comparison, at least in the beginning[2].

Comparisons between ILs are always directional. This means that a system A can be compatible with a system B if the elements of A match the elements of B. But if B is a more refined system, then its more specific elements cannot find their equivalences in A; B is then only partially compatible with A. The advantages of compatible ILs are numerous, as regard input as well as search and output. Once compatibility is established one should be able

(1) to search with any term on any problem in a variety of different files.
(2) retrieve the information labelled with a specific concept from any store, indexed by any of the systems involved,
(3) inform the user of a certain IL on the availability of information on the topic of his interest, indexed by any other IL,
(4) get the equivalents of a system in one natural language (say English) in another natural language (say German), thus the switching system can also be used as a multilingual theasaurus.

Regarding the latter question, a number of solutions to problems of a multilingual thesaurus can also be applied to "compatibilize" existing ILs (see 28).

The task of establishing compatibility is a difficult, time consuming concept-analytical work. It involves a very good knowledge of the ILs under investigation and also of the subject fields covered. It is rewarding, on the other hand, with respect to the convenience and possibilities achieved for all users of any of the systems.

The systems themselves profit from such thorough comparisons with respect to depth and clarity of class descriptions, in that the necessary improvements of contents and structure can be promptly undertaken.

## 4. Methods for the establishment of verbal comparisons between indexing languages

### 4.1 Recording of elements of IL

Whenever two or more ILs are to be compared with each other or studied with regard to their possible compatibility, it is necessary to obtain first of all an overview of the actual overlap of classes and descriptors by matching the terms *verbally*. This can be done by computer if the ILs are already machine readable or in an alphabetic file. If there is only a limited percentage of overlap, then the entire comparison should be reconsidered. If the overlap is considerable, then the contents of each of the system positions or the descriptors as such must be recorded on cards or in a computer file and their concepts must be analyzed in a so-called concept record which ought to have the following entries:

*Fig. 1: Fields of a concept record*

(1) Name of concept or class
(2) Notation
(3) Next broader concept
(4) Highest concept in hierarchy/subject category
(5) Indication of hierarchical level of concept
  (A) highest level
  (B) next lowest level
  (C) third level, etc.
(6) Number of subconcepts, if comparison only on a certain level, in brackets for each level
(7) Form category of concept
  (O) Object, entity
  (P) Process, activity, state
  (Q) Quantity, quality
  (R) Relation
  (S) Space-related concept
  (T) Time-related concept
  (W) Subject-field or discipline
(8) Definition of concept (if necessary and possible)
(9) Other names of concept or class
(10) Source of concept abbreviation of IL accord. to (29)
(11) Remarks

If an IL is already available in a machine-readable form, the record for each descriptor should be supplemented by the indications mentioned above.

### 4.2 Establishment of an alphabetical comparison matrix (M1)

Once the file of records of the ILs under investigation has been established according to 4.1 above, the concept names must be brought into alphabetical order in such a way that they form the column 2 of a matrix with further columns for each IL. Suppose, for example, that the main fields of 'Social Welfare' are to be compared, as they occur in the DDC, BBC and UNT[3], namely:

*Fig. 2: 'Social Welfare' in the UDC, BBC and UNT.*

| | | |
|---|---|---|
| DDC: | 361 | Social problems and social welfare |
| | 361.1 | Social problems |
| | 361.2 | Social action |
| | 361.3 | Social work |
| | 361.4 | Group work |
| | 361.6 | Public action |
| | 361.7 | Private action |
| | 361.8 | Community action |
| | 361.9 | Historical and geographical treatment |
| BBC: | Q | Social Welfare |
| | QAG | Social Welfare Administration |
| | QD | Social Work |
| | QE | Social Services |
| | QF | Social Security |
| | QG | Persons in need, causes of need |
| | QN | Deviants |
| | QO | Crime, criminology |
| UNT: | R85/99 | Social Welfare |
| | R86 | Social welfare philosophy |
| | R87 | Welfare policy |
| | R88 | Social welfare planning |
| | R89 | Social welfare economics |
| | R90 | Social welfare administration |
| | R91 | Social workers |
| | R93 | Social work |
| | R94/99 | Social services |
| | R95 | Social security |
| | R96 | Personal social services |
| | R97 | Health services |
| | R98 | Disaster relief work |

From these a comparison matrix will be generated as shown in Fig. 3

After establishing this matrix (see Fig. 3) the running numbers of column 1 can be added under field (12) of each record (Fig. 1). The notations of column 3–5 are those, which indicate a verbal coincidence with the linguistic form of column 2. All deviations in wording have been left out here, especially regarding the cases of a strict computer matching of terms. One can establish now a so-called verbal coincidence rate which is the sum total of the numbers under column 6 (vc = verbal coincidence), the verbal coinciding cases (here 42) divided by the sum total of possible coincidences (23 X 3 = 69). In this case the coincidence rate is 0.608, which is a rather high rate, indicating the relatively high degree of verbal compatibility between the three systems involved.

The verbal compatibility can be measured only, if ILs of one natural language are involved. Furthermore, it does not tell us whether:

(1) the terms which coincide really stand for the same concept (there may be homonyms and polysemes),
(2) the ILs may use deviations in wording (e.g. 'philosophy of social welfare' as against 'social welfare philosophy'),
(3) there may be a higher degree of compatibility between the ILs if one considers the concepts rather than the names of the system positions.
(4) concepts could be entered into the comparison which differ in specificity but may still be used as a substitute for an otherwise lacking concept.

Because of these reasons, a merely verbal compatibility matrix will not be of sufficient help in establishing compatibility, although it certainly serves as its necessary preliminary step.

*Fig. 3: Example of verbal compatibility matrix*

| No. | Name | DDC | BBC | UNT | vc |
|---|---|---|---|---|---|
| 1 | Community action | 361.8 | QDP | R93.20.10* | 3 |
| 2 | Crime. Criminology | | QO | R75/78 | 2 |
| 3 | Deviants | | QN | | 1 |
| 4 | Disaster relief work | | QGN | R98 | 2 |
| 5 | Educat. welfare | | QEN-P | J14.15 | 2 |
| 6 | Group work | 361.4 | QDN | R93.10 | 3 |
| 7 | Health services | | | R97 | 1 |
| 8 | Personal soc. serv. | | | R96 | 1 |
| 9 | Persons in need | | QG | | 1 |
| 10 | Private action | 361.7 | | | 1 |
| 11 | Public action | 361.6 | | | 1 |
| 12 | Social action | 361.2 | | P71.20 | 2 |
| 13 | Social problems | 361.1 | | R70/84 | 2 |
| 14 | Social security | | QF | R95 | 2 |
| 15 | Social services | 361–365 | QE | R94/99 | 3 |
| 16 | Social welfare | | Q | R85/99 | 2 |
| 17 | Soc. welf. admin. | | QAG | R90 | 2 |
| 18 | Soc. welf. economics | | QAT M/Z | R89 | 2 |
| 19 | Soc. welf. philosophy | | QAE | R86 | 2 |
| 20 | Soc. welf. planning | | | R88 | 1 |
| 21 | Social work | 361.3 | QD | R93 | 3 |
| 22 | Social workers | | QB | R91 | 2 |
| 23 | Welfare policy | | | R87 | 1 |
| | | | | | 42 |

*A thesaurus without notations will be entered only through its descriptors.

## 5. Methods for the establishment of conceptual comparisons between ILs

### 5.1 Comparing concepts on the verbal level

The records of one and the same verbal form must now be compared with one another regarding the entries of fields

(3) next broader concept
(4) highest concept ("top term")
(5) level
(6) subconcepts
(7) form category
(8) definition
(9) other names

It can be assumed that there will be no complete coincidence between the contents of the fields of the different ILs under investigation. But from the comparisons one should be able to sort out those records that do not match conceptually and associate them to those concepts in the file to which they are related conceptually according to the well known concept relationships (30).

The discrepancies leading to this decision can be recorded in field (11).

The records (cards) of the other concepts having the same or a different but synonymous concepts name and similar data in fields (3)–(8) should be kept together.

### 5.2 Conceptual reorganization of matrix M1

The reorganization of the records can be done in two different ways:

(1) by designating the most detailed and most structured of all the ILs under investigation as the one to accept the funtion of a master system[4] to be placed in columns 1 and 2. In this case the equivalent classes or terms of the other IL should be juxtaposed after comparing the definitions and other entries for relevance. If it should occur that the master IL lacks a certain concept belonging to any of the other ILs in the environment of one of the concepts, a position must be left empty in columns 1 and 2 for the placement of the lacking concept in the row

under its own IL heading. Such concepts should be called "challenge concepts" since they call for the creation of a new subdivision of the master system for compatibility reasons.

(2) If no master system can be accepted, then such a system must be generated from the concepts involved in the comparison themselves, by

- sorting the material according to field (4)
- arranging the sorted material according to field (7)
- structuring according to fields (3), (5) and (6)
- comparing the results and establishing a tentative master IL on the basis of all concepts which would belong in this column.

In a further step the resulting groupings should again be checked regarding occurring super– or subordinations. Some of the groups should perhaps be brought together and some categorizations (field 7) reorganized too.

### 5.3 Establishment of a compatibility matrix M2

After a careful consideration of the steps outlined above a compatibility matrix can be generated by inserting in columns I and 2 the concepts adopted from a master system or generated through the reorganization attempts outlined, and juxtaposing in columns 3, 4 and 5 those equivalent concepts that most closely match conceptually the ones in columns 1 and 2.

It will be helpful to show the hierarchical levels of column 2 either by special print, underlining, or by indentations. The numbering in column 1 will now become a conceptual one, possibly corresponding to the hierarchical and faceted system positions placed in column 2.

In column 6 the conceptual coincidence number (cc) can then be inserted and summed up. The degree of conceptual compatibility of the ILs under investigation can be computed by dividing the first sum by the sum of all possible matches.

The resulting number will always be higher than the verbal coincidence rate.

In the comparison matrix (M1) it sufficed to indicate the concepts of the different ILs by their notations. In the compatibility matrix (M2) one should also add in each case the descriptor or class name for reasons of better understanding.

It is not possible to provide here a complete matrix of the entire field of Social Welfare as contained in the ILs used in the example above. In Fig. 4, however, we illustrate just one small section of such a matrix with the generic relationships of Social Welfare.

### 5.4 Additional information

It will become obvious from the display in the compatibility matrix (M2) that there may be slight or obvious differences in the juxtapositions of the concepts available. These can be made explicit by placing a mathematical sign before the concept, stating

≠ for unspecific concept or deviation of concepts
< for broader concept than the one in column 2
> for narrower concept than the one in column 2
c for concept combinations

A few examples with respect to Fig. 3 will explain this more clearly:

There is no entry at No. 2 'Crime. Criminology' under the DDC column. In the compatibility matrix one could, however, enter here 364.4 'Prevention of crime and delinquency' with a > symbol, indicating the narrower concept. Similarly, but vice versa at No. 3 the DDC class 302.5 'Relation of individual to society' with 'Deviation' mentioned expressedly in its scope in the DDC schedules could be entered with a < symbol, indicating that here the broader concept may be used, just as for 'personal social services' (No. 8) the DDC class 362 'Social welfare problems and services'. For 'disaster relief work' (No. 4) the DDC would use a combination, namely c 363.3475, combining 'disaster' and 'social action'.

A true conceptual but not verbal coincidence would be DDC 362.1–4 called "Illness and disability' under 'Social welfare problems and services' for 'health services' at No. 7, and also UNT R79/84 'People in need' at No. 9 instead of 'Persons in need'. BBC serves at No. 12 only with > 'social action programmes' and DDC at No. 14 with either > 336.249 'social securitiy taxes' or > 368.4 'social security benefits'.

The index to the DDC gives 'Welfare Economics' under 330.155, however, the schedules reveal that this is the system position for 'miscellaneous schools in economics' – the equivalence of 'social welfare' and 'welfare' in the DDC has lead thus to a homonym for 'Welfare economics' which – although apparently relevant for position 18 in Fig. 2 – is nevertheless no equivalent for the BBC and the UNT classes placed here.

### 5.5 Establishment of system-related matrices M3

As an alternative to the matrix M2 established on the basis of either a chosen or a newly created master system, any of the classification systems under comparison could play the role of the master system and be placed in columns 1 and 2. In this case, however, one would juxtapose to the classes of the master system only those classes and descriptors of the other ILs which correlate to the chosen master system. One could therefore call this matrix M3 also a 'correlation matrix'. If the master system happens to be a very broad system and the other ILs rather detailed then these latter will not be able to bring in their fully

| No. | Name | DDC | BBC | UNT | cc |
|---|---|---|---|---|---|
| 65 | Social Welfare | 361<br>Social problems and social welfare | Q<br>Social Welfare | R85/99<br>Social Welfare | 3 |
| 65.1 | Soc. welf. philos. | 361.01<br>Philos. & theory of | QAE<br>Philos. of soc. welf. | R86<br>Soc. welfare philosophy | 3 |
| 65.2 | Soc. welf. admin. | – | QAG<br>Soc. administration, Soc. welf. administration | R90<br>Social welfare administration | 2 |
| 65.3 | Welfare policy | >361.25<br>Action within establ. soc. framework (policy) | QAG P<br>Policy (in Social Welfare) | R87<br>Welfare policy | 3 |
| 65.4 | Soc. welf. planning | >361.25<br>Act. within establ. soc. framework (planning) | QAH<br>Planning for welfare, social planning | R88<br>Social welfare planning | 3 |
| 65.5 | Soc. welf. econom. | – | QAT M/Z<br>Management of soc. welf. | R89<br>Social welfare economics | 2<br>2 |

*Fig. 4: Examples of section of a conceptual compatibility matrix*

developed conceptual capacity. In oder to take care of this, it is recommended to insert after each class designation in brackets the number of its subclasses as indicated in field (6) of Fig. 1.

If the master system happens to be more detailed than the other ILs then the matrix will have very many empty spaces in the columns of the ILs under comparison.

### 5.6 Establishment of hierarchical matrices M4

Whenever a broader master system is to be correlated to more detailed ILs and presented in a matrix form, a special case of the system-related matrix M3 can be established by considering that all system positions which are occupied by one of the detailed ILs must be *left open* in the columns 1 and 2 of such of matrix.

If the hierarchical levels should differ between the ILs entered into this matrix it will be advisable to place them one after the other in the direction of their increase in depth of hierarchical levels. The column 1 would not have a system notation in such cases. In order to access the concepts in the matrix it will be necessary to add a consecutive numbering preferably in brackets on each level after the notations indicating the broad concept of the master system.

### 5.7 Index to the compatibility matrix

Since there is no other verbal access to any of the conceptual compatibility matrices M2—M4, it is necessary to establish an alphabetical index to all of the entries of each one, including in one alphabetical list all the descriptors and class descriptions of the other ILs, as a lead-in vocabulary for the new notation given in column 1 of M2, M3 or M4. The index entries should also include the other names of a concept or a class as given in field (9) of the concept record.

## 6. Treatment of compatibility problems

It is less cumbersome to compare two or three thesauri of one and the same discipline than a thesaurus and, say, a section of a universal classification system. Usually the thesauri have a limited number of descriptors and these are mostly single terms or combination of two terms. If there is no verbal match, the problem lies in finding the closest concept to the concept in question. A classification system, especially if it offers concept construction facilities by combinations to be expressed in combined notations too, (as e.g. the DDC, the UDC, the CC, the LCC and the UNT) must be very well known in order to construct the notations in question adequately. On the other hand, with these facilities such systems are most powerful in matching any concept of a thesaurus expressed in a controlled verbal form.

Therefore, in comparing the contents of a special thesaurus with a universal classification system one must – in each case of non-verbal-match of a possible equivalent in the index – look into the schedules and construct a notation on the basis of the rules of such a system as far as such a system allows, for the equivalent concept construction. It is advisable to do such a work together with an expert of the universal classification system in question.

On the other hand, if an IL offers a precombination, as e.g. 'social welfare administration' and the other IL tells to combine in this case 'social welfare' and 'administration' then the precombined class description determines the inclusion and the equivalent must be built up by logical addition accordingly. Such a procedure is problematical, however, if there are neither rules nor classes for such combinations. In such cases one cannot but look for the next broader concept and subsume the concept in question subsequently. If one wants to juxtapose it, then the necessary symbols must be added, (see section 5.4). All combinations occurring in one of the ILs under investigation must get an entry in the compatibility matrix 2 in order that their equivalents can be inserted in the columns of the other IL. This does not hold, however, for any possible combinations which have so far not been used as a concept in indexing.

## 7. Establishment of compatibility between ILs

Only after the preliminary work of establishing a concept record file and the two matrices M1 and M2 is it possible to establish compatibility between the individual ILs.

This process can be related to different levels and different parts of an IL and the option for the use of any of the matrices M2 to M4, depending on the ILs themselves under investigation and on an IL's administration policies. It may even lead to an entire remodelling of an IL.

As far as compatibility on the verbal level is concerned, the index of each IL can be supplemented by the term alternatives or synonymous terms of the other IL and thus improve the verbal access to its classes and concepts.

With respect to compatibility on the conceptual level in cases of missing equivalents, this can be reached by supplementing the missing concepts or classes by the equivalent of the master column and its notation. Also, an adaptation of concept contents is possible if the concept specificity is deviating.

With respect to structural problems compatibility can be established between two or more IL by using the same divisions for subject groupings and for facetting.

Finally, it will also become necessary to adapt the syntax rules of the ILs involved which are used for the formation of concept combinations within one and the same subject group or between different subject groups.

## 8. Use of a compatibility matrix

The most sophisticated use of a compatibility matrix as a "black box" will be the possible input of class descriptions in one IL with the resulting output in another IL, e.g., if the entries have an LCC notation at the input stage and appear with a DDC, or a BBC or a UDC notation in the output. Elegant solutions for such a utilization will take a long time for their elaboration, but such a capability should be easier to achieve than the translation of one natural language into another.

A compatibility matrix will however already be a helpful instrument, if it can suggest descriptors and notations for searching in differently classed files.

It will also provide a necessary tool for cooperative indexing and exchange of indexed items.

Last but not least it will demonstrate the priorities and levels of specificity of the ILs compared and will provide a better understanding of the conceptual task of structuring classification systems.

## 9. Organizational problems

The organization and administration of an IL compatibility matrix involves

– subject experts in the fields under comparison
– experts with respect to the ILs to be entered
– secretarial help
– the printed or computerized versions of the ILs and
– if possible, computer access

The time necessary to set up a single record according to section 4.1 will amount to approx. 15 minutes. This time can be minimized, if for example a printout of a computerized version of the IL with most of the information mentioned could be provided. An IL, say of 1000 concepts would demand thus less than two months for the recording of the concepts.

This work could be done by different people for the different ILs at the same time. The merging, alphabetization and setting up of the comparison matrix M1 would be a clerical job. The concept comparison must again be done by experts, probably most effectively in a common discussion of each concept.

All the decisions taken at such discussions must be recorded on the record (card) of the concept in question, field (11).

After this, the compatibility matrix M2 can be set up and printed together with its alphabetical index.

This matrix should be tested subsequently in actual indexing and retrieval cases. The results of the tests should serve as feed-back material for an improvement of the compatibility matrix or the index.

The availability of the finished compatibility matrix should be made known to the professional world.

## Notes

1 For this the elements of six thesauri in the field of information science were correlated by a matrix; a new grouping for the master thesaurus had been worked out previously by N. Gardin and F. Levy.
2 Later on one might reconsider these classes on the basis of what has been indexed under their heading.
3 Regarding these abbreviations of the Dewey Decimal Classification (DDC), the Bliss Bibliographic Classification (BBC), and the Unesco Thesaurus (UNT) please see (29).
4 This has, for example been done by Foster et al (31).

## References


(1) Hammond, W., Rosenborg, S.: Experimental study of convertibility between large technical indexing vocabularies. Silver Springs, Md.: Datatrol 1962.
(2) Newman, S. M. (Ed.): Information systems compatibility. Washington, D. C.: Spartan Books 1965.
(3) Henderson, M. M., Moats, J. S., Stevens, M. E., Newman, S. M.: Cooperation, convertibility and compatibility among information systems: a literature review. Washington, D. C.: Government Printing Office 1966.
(4) Unesco: UNISIST Study Report on the feasibility of a world science information system. Paris: Unesco 1971.
(5) Wellisch, H.: A concordance between UDC and Thesaurus of Engineering and Scientific Terms (TEST). Results of a pilot project. In: Proc. Intern. Symp. UDC in Relation to Other Indexing Languages. Herceg Novi, 1971. Belgrad: Yugoslav Center for Techn. &Scient. Doc. 1972.
(6) Agraev, V. A., Kobrin, R. Ju., Sul'c, M. M.: Information retrieval system compatibility. (Transl. from Russian). In: Automatic Doc. &Math. Linguistics 8 (1974) No. 2, p. 29–37.
(7) Smith, L. C.: Systematic searching of abstracts and indexes in interdisciplinary areas. In: J. Amer. Soc. Inform. Sci. 25 (1974) p. 343–353.
(8) Svenonius, E.: Translation between hierarchical structures: an exercise in abstract classification. In: Neelameghan, A. (Ed.): Ordering systems for global information networks. Proc. 3rd Intern. Study Conf. on Class. Res., Bombay 1975. Bangalore: DRTC 1979. p. 204–211.
(9) Wersig, G.: Experiences in compatibility research in documentary languages. (same source as (8), p. 423–430.
(10) Horsnell, V.: Intermediate lexicon for information science: a feasibility study, London: Polytechnic of North London, School of Librarianship 1974.
(11) Coates, E. J.: Switching languages for indexing. In: J. Doc. 26 (1970) p. 102–110.
(12) Soergel, D.: Indexing languages and thesauri: construction and maintenance. Los Angeles, CA: Melville 1974.
(13) Včerasnij, R. P., Bunova, M. A.: Compatibility among classificatory retrieval languages. In: Edinaja sistema informacionno-poiskovych jasykov. (Unified system of information retrieval languages: summaries of papers presented at an All-Union Conference held in Yurmala, Latvian SSR, 6–8 Sept. 1977. (In Russian). Riga: Latv. Universitet 1977. p. 170.
(14) Giljarevskij, R. S.: A note on the compatibility between retrieval language types. In same source as (13) and in: Naučno-tech. inform., Ser. 2 (1978) No. 1, p. 11–13.
(15) Gluškov, V. M., Skorochodjko, E. F., Stognij, A. A.: Measuring the compatibility level of document retrieval languages. In same source as (13) and in: Naučno-Techn. Inform., Ser. 2 (1978) No. 1, p. 14–19.
(16) Gruzitis, B. E., Sover, N. B., Morozov, V. A.: Some design concepts and the structure of a compatible language system of a republic computer-based STI system. In same source as (13).
(17) Konšin, G. I.: Contribution to the problem of compatibility between classificatory and descriptor languages. In same source as (13).
(18) Konšin, G. I., Moskalenko, Z. D.: Information language compatibility problems in the context of an evolving information network in a field: with special reference to the "Geology" automated STI system. In same source as (13).
(19) Kosmačeva, N. V., Makurina, R. G.: Vocabulary compatibility of different IR systems in a field with special reference to the "Electrical Engineering" and "Source" computer-based IR system. In same source as (13).
(20) Kocetova, S. M., Murinson, E. A., Šenderov, V. Z.: The philosophy of interfacing classificatory and descriptor languages. In same source as (13).
(21) Orlov, A. N.: Social science-related problems of compatibility between a subject heading language and a descriptor language. In same source as (13).
(22) Pšeničnaja, L. E.: On the compatibility between thesaurus vocabularies in an IR system. In same source as (13).
(23) Margulis, A. M., Vychodcev, L. A., Šuchman, E. S.: Automated maintenance of interrelated classifications. In same source as (13).
(24) Margulis, A. M., Vychodcev, L. A., Questions of compatibility of MIS linguistic means. In same source as (13).
(25) Černyi, A. I.: Information retrieval languages: types, design principles, compatibility. (In Russian). In: Naučno-techn, inform., Ser. 2 (1978) No. 1, p. 1–10.
(26) Šrejder, Ju. A.: Interaction of classification and subject languages. (In Russian). In: Naučno-techn. inform., Ser. 2 (1978) No. 1, p. 20–23.
(27) Vilenskaja, S. K.: On the compatibility of different information retrieval languages within the integrated information system. In: Taylor, P. J. (Ed.): New trends in documentation and information. Proc. 39th FID Congress, Edinburgh 1978. London: Aslib 1980. p. 315–325.
(28) Unesco: UNISIST Guidelines for the establishment and development of multilingual thesauri. Rev. text. Paris: Unesco 1980. 85 p. (Text in English and in French available). PGI/80/WS/12.
(29) Dahlberg, I.: Citation codes in classification and terminology. In: Intern. Classificat. 5 (1978) No. 2, p. 91–92.
(30) Dahlberg, I.: A referent-oriented, analytical concept theory for Interconcept. In: Intern. Classificat. 5 (1978) No. 3, p. 142–151.
(31) Forster, W. R., Beckhard, E. J., Hart, C. S.: The coordinated development of health-related clearinghouse vocabularies. In: Tally, R. D. (Ed.) et al: Information choices and policies. Proc. 42nd ASIS Ann. Meeting. Minneapolis, Minn. 1979. Washington D. C.: Amer. Soc. Inform. Sci. 1979, p. 160–167.